



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Genomics 82 (2003) 300–308

GENOMICS

www.elsevier.com/locate/ygeno

Characterization of the transmembrane channel-like (*TMC*) gene family: functional clues from hearing loss and epidermodysplasia verruciformis☆

Kiyoto Kurima,^a Yandan Yang,^a Katherine Sorber,^a and Andrew J. Griffith^{a,b,*}

^a Section on Gene Structure and Function, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Rockville, MD 20850, USA

^b Hearing Section, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Rockville, MD 20850, USA

Received 10 March 2003; accepted 15 May 2003

Abstract

Mutations of *TMC1* cause deafness in humans and mice. *TMC1* and a related gene, *TMC2*, are the founding members of a novel gene family. Here we describe six additional *TMC* paralogs (*TMC3* to *TMC8*) in humans and mice, as well as homologs in other species. cDNAs spanning the full length of the predicted open reading frames of the mammalian genes were cloned and sequenced. All are strongly predicted to encode proteins with 6 to 10 transmembrane domains and a novel conserved 120-amino-acid sequence that we termed the TMC domain. *TMC1*, *TMC2*, and *TMC3* comprise a distinct subfamily expressed at low levels, whereas *TMC4* to *TMC8* are expressed at higher levels in multiple tissues. *TMC6* and *TMC8* are identical to the *EVER1* and *EVER2* genes implicated in epidermodysplasia verruciformis, a recessive disorder comprising susceptibility to cutaneous human papilloma virus infections and associated nonmelanoma skin cancers, providing additional genetic and tissue systems in which to study the *TMC* gene family.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Deafness; Epidermodysplasia verruciformis; EVER1; EVER2; Hearing; Ion channel; Papillomavirus; Receptor; TMC domain; Transmembrane

The *TMC* gene family was recently discovered through positional cloning of the gene underlying both dominant and recessive nonsyndromic sensorineural hearing loss at the DFNA36 and DFNB7/B11 loci, respectively, on chromosome 9q13–q21 [1]. These were shown to be allelic disorders caused by mutations of *TMCI*, a novel gene of unknown function. The mouse ortholog, *Tmc1*, also has dominant and recessive mutant alleles that cause hearing loss in the Beethoven (*Bth*) and deafness (*dn*) mutant mouse strains, respectively [1,2]. The dominant mutations are mis-

sense substitutions that likely act via dominant negative or gain-of-function mechanisms, since recessive mutations appear to be functional null alleles of *TMCI* or *Tmc1*, and heterozygous carriers of recessive mutations are unaffected [1].

TMCI was initially identified among unassembled genomic sequence fragments with similarity to a second predicted gene of unknown function, *TMC2*, on chromosome 20 [1]. Neither *TMCI* nor *TMC2*, nor their mouse orthologs *Tmc1* and *Tmc2*, has nucleotide or amino acid sequence similarity to any known genes or domains. All four genes are strongly predicted to encode at least six conserved transmembrane domains, raising the possibility that TMC proteins may function as ion channels, pumps, or transporters [1]. The cochlear electrophysiologic phenotype of homozygous deafness mice, which segregate a partial genomic deletion of *Tmc1* [1], is consistent with an ion channel defect [3].

☆ Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries under Accession Nos. AY236490 (*TMC3*), AY236491 (*Tmc3*), AY236492 (*TMC4*), AY236493 (*Tmc4*), AY236494 (*TMC5*), AY236495 (*Tmc5*), AY236496 (*TMC6*), AY236497 (*Tmc6*), AY236498 (*TMC7*), AY236499 (*Tmc7*), AY236500 (*TMC8*), and AY236501 (*Tmc8*).

* Corresponding author. Fax: +1-301-402-7580.

E-mail address: griffita@nidcd.nih.gov (A.J. Griffith).

```

human TMC1 MSPKKVOIKVE-----EKEDETEESSEEE 24
mouse TMC1 MLIQIIVE-----EKEEETEESSSEE 20
human TMC2 MSHOVKGLKEEARGVVKGRVKS GSPHTGDRLGRRSSSKRALKAEGT PRRGAQRSOKERAGGSPSPG SPRRKTQTRRRRHE 81
mouse TMC2 MSPQLKSLDEE-----GDKSARRPTRKQTSRAACPQDGHRAQSSRKDPAGKSPRP GSSRRKQMEHGSYHK 65
human TMC3 -----MEEN-PTLESEAWG 13
mouse TMC3 -----MEAWG 5
human TMC4 MLSD-----DHVNEITIQ--VENVSSGVQSHPS-----NOIFQEKVL--LDSSINMVLISID 49
mouse TMC4 MQSD-----DOVDEITIE--VENVPSGVQNHLS-----SGLALRKS--ANPFA-CVLSSSA 48
human TMC5 MAQPLAFLILDVVPETPGDOGQGSPYPDESEVHDSFOOLIQEQSOCTAQEGLELQOREREVTGSSQOTLWRPFG--TQTATLRLILAM 85
mouse TMC5 MTQSLALALADVPETLGDDEGLESPYVESEVHDSFHQLIQEQSLRVAAEGLELELPLG--LGRGDTLPGLEGAPALSSATLRLILAM 84
human TMC6 -----MSESSGALQFGR-- 13
mouse TMC6 -----MSE--GSAGDPGHG 12
human TMC7 -----MLLPRSV 7
mouse TMC7 -----MFRQSV 7
human TMC8 ----- 0
mouse TMC8 ----- 0

human TMC1 EEEVEDKLPRESLRPK-RKRTRD-----VINEDDPEPEPEDEETRKARE----KERRRRLKRGAEKEE- IDEEELERLKAELDE 98
mouse TMC1 EE--DKLPRRESLRPK-RKRTRD-----VINEDDPEPEPEDEETRKARE----KERRRRLKRGAEKEE- IDEEELERLKAELDE 92
human TMC2 ELGEOERGAERTCEGRKRDERA-----SFOERTAAPKREKIEPRKEEKSKROKPKRSSLSASGGSLSEEBLAQILVEE 162
mouse TMC2 GLQGGPKPRKVERSLOGR-KKDRRT-----SLKQRASPKKEREALRKEAG-KOLRKPRTSLGSSVSTGSLSEEBLAQILVEE 144
human TMC3 MKTSKASORYRSGIRRNASOC-----YLVOESLILNSLD-----DSFADTETGDSNDPEQIFQIF 56
mouse TMC3 MKTSKASORYRSGIRRNASOC-----YLQDSLILNSLD-----DSFNADTETGDSNDPEQIFQIF 56
human TMC4 SSRGWLAPREARGAPSSPGPSLS-----SVLNLEPSAATLRYRDPGVLP-----WGALEEBEDGSRSKAFETVQT 82
mouse TMC4 QSPACSSSRKAR-----TGP SLA-----SVLNLEPSAATLRYRDPGVLP-----WGLTDEEDGEGRSQAFAETA 68
human TMC5 IDVIDSOTVSKRNDQKGNVLRFS-----SLNESMSOATLHSLRCMGIDTP-----GSSHETVGO--KIATSLIPMTSR 117
mouse TMC5 ADRVDCOTFNPGRNRHRNLRFS-----SLNESISOTIYHSECLVTDES-----CFHETVGO--KIATSLIPMTTR 115
human TMC6 PRTIGRSRGAIIISQYYNRTVOLRRSSRPLLGNVFVRSAPSLRLYDLELD-----PTALEEEKOSLLVKELQSLSAA 159
mouse TMC6 PRTIGRSRGAIIISQYYNRTVOLRRSSRPLLGNVFVRSAPSLRLYDLELD-----STILEEKEKRLSVKELQSLSAA 158
human TMC7 PSRQAVHPENLSLGGSCFS-PPV-----NFLQELPSYRSIARRRTVHS-----RDKOSGTLKPTDSYSSOLEBRIA 82
mouse TMC7 PSRQAVHPENLSLGGSCFS-PPV-----NFLQELPSYRSIARRRTVHS-----RDKOSGTLKPTDSYSSOLEBRIA 82
human TMC8 SERRAPGVPE-----BELWEAEMERL-----SSG-----TP 35
mouse TMC8 QSGPAPRRPESQ-----ASEELWEQEVERLC-----ASR-----TP 39

human TMC1 KRQIATVKCKPWKMEKKIEVLKBAKPFVSENEGA-----LGKGGKRWKFAFKMMMAKKWAKFLRDFENFKAACVPWENKIKATESQ 180
mouse TMC1 NRQMIATVKCKPWKMEKKIEVLKBAKPFVSENEGA-----LGKGGKRWKFAFKMMMAKKWAKFLRDFENFKAACVPWENKIKATESQ 174
human TMC2 KKLIATVWRKSWMAKKLTELREAQEFVEKYEGA-----LGKGGKQLYAYKMLMAKKWKFKRDFDNFKTCIPWEMKIKATESH 244
mouse TMC2 KKLIATVWRKSWMAKKLTELREAQEFVEKYEGA-----LGKGGKQLYAYKMLMAKKWKFKRDFDNFKTCIPWEMKIKATESH 226
human TMC3 OKDLMANIRCPWTGOKLRAKRALKRNVLKFEGR-----LTRTREG-----QAAGAELEWKKFARLACNFVIFIPWEMKIKATESH 133
mouse TMC3 OKDLMANIRCPWTGOKLRAKRALKRNVLKFEGR-----LTRTREG-----QAAGAELEWKKFARLACNFVIFIPWEMKIKATESH 133
human TMC4 ELQDPHSRELPPWQAR-RAHRK-----NASRDQ-----VYVGSQTKTDRMARLLR-----RKKGKMEGFPOTLPWAWTKIKTGG 195
mouse TMC4 ES--HPSRELPPWQAR-RAHRK-----NASRDQ-----VYVGSQTKTDRMARLLR-----RKKGKMEGFPOTLPWAWTKIKTGG 137
human TMC5 DR--IKTRNORPTWEEK-RNTERKI-----VDKEKSK-----OTHRILQMNCCIOCLNSI SRAYYRKNMSEILENLSISLWAKLTKGGK 195
mouse TMC5 DR--IKTRNORPTWEEK-RNTERKI-----VDKEKSK-----OTHRILQMNCCIOCLNSI SRAYYRKNMSEILENLSISLWAKLTKGGK 193
human TMC6 QR--DHMLRGMPLSLAEK-ESLREK-----SDTTPKRWGPESGGVCS-CCGRLLRYACVLALHSLGLALLSALQALMPWYALRKG 240
mouse TMC6 QR--DHMLRGMPLSLAEK-ESLREK-----SDTTPKRWGPESGGVCS-CCGRLLRYACVLALHSLGLALLSALQALMPWYALRKG 240
human TMC7 ENLSSYSRLNYSALNISEK-RRLE-----DIOETC-----MKYLS-----EWDONKRYSSKSKWRFLBKAREMTTHLELWRKDRSIEGK 155
mouse TMC7 ENLSSYSRLNYSALNISEK-RRLE-----DIOETC-----MKYLS-----EWDONKRYSSKSKWRFLBKAREMTTHLELWRKDRSIEGK 153
human TMC8 VRL--GLPYAMMDK-RFLW-----ELREPA-----GVQTL--RWRORRRRQTVERRLEAQRARLGLWEGALYETGGL 101
mouse TMC8 VRL--GLPYAMMDK-RFLW-----ELREPA-----GVQTL--RWRORRRRQTVERRLEAQRARLGLWEGALYETGGL 101

human TMC1 FGSSVASYELFLRWYGVNMLVFILTFSLIMLPEYLWGLPYG-----SLPRKTVPRAEAS-----AANFGVLYDFNGLA 252
mouse TMC1 FGSSVASYELFLRWYGVNMLVFILTFSLIMLPEYLWGLPYG-----SLPRKTVPRAEAS-----AANFGVLYDFNGLA 246
human TMC2 FGSSVASYELFLRWYGVNMLVFLIIFGLVIIPEVLMGMPYG-----SIPRKTVPRAEBEK-----AMDFSVLWDFEYIKY 316
mouse TMC2 FGSSVASYELFLRWYGVNMLVFLIIFGLVIIPEVLMGMPYG-----SIPRKTVPRAEBEK-----AMDFSVLWDFEYIKY 298
human TMC3 FGSSVASYELFLRWYGVNMLVFLIIFGLVIIPEVLMGMPYG-----SIPRKTVPRAEBEK-----AMDFSVLWDFEYIKY 298
mouse TMC3 FGSSVASYELFLRWYGVNMLVFLIIFGLVIIPEVLMGMPYG-----SIPRKTVPRAEBEK-----AMDFSVLWDFEYIKY 205
human TMC4 FGAGTESYSLRFLLLNLVSLVAMACMLLTPWLGAPPG-----PPGPDISSPCGSYNPHSGLVTFATQLFNLLSAGEYLEW 236
mouse TMC4 FGAGTESYSLRFLLLNLVSLVAMACMLLTPWLGAPPG-----PPGPDISSPCGSYNPHSGLVTFATQLFNLLSAGEYLEW 218
human TMC5 FGTSLVSYENFLRWLLKFNIFSFILNFSFIIIPQTVFA-----KNTLQFTG-----LEFFTGVGVRFD 254
mouse TMC5 FGTSLVSYENFLRWLLKFNIFSFILNFSFIIIPQTVFA-----KNTLQFTG-----LEFFTGVGVRFD 252
human TMC6 FGSSVLSYELFLKTLAFAFNALLLPLVAFLVGVQAAFPP-----ALPGPAPVCTG-----LELLTGACGFT 302
mouse TMC6 FGSSVLSYELFLKTLAFAFNALLLPLVAFLVGVQAAFPP-----ALPGPAPVCTG-----LELLTGACGFT 301
human TMC7 FGTGQSYEFSLRFLVLLNLVFLIIFMLVLVLLTKYKITNSFVLVLPFKMDKQCTVYPVSSGLIYFYSYIDLLSGTGLPE 242
mouse TMC7 FGTGQSYEFSLRFLVLLNLVFLIIFMLVLVLLTKYKITNSFVLVLPFKMDKQCTVYPVSSGLIYFYSYIDLLSGTGLPE 240
human TMC8 FGTGQSYEFSLRFLVLLNLVFLIIFMLVLVLLTKYKITNSFVLVLPFKMDKQCTVYPVSSGLIYFYSYIDLLSGTGLPE 181
mouse TMC8 FGTGQSYEFSLRFLVLLNLVFLIIFMLVLVLLTKYKITNSFVLVLPFKMDKQCTVYPVSSGLIYFYSYIDLLSGTGLPE 185

human TMC1 SVLFYGYDNDK-RITGW-----LNFRPLPSYFLVGMCIQYFLVVLKAMTKNIGDDGG--GDDNTFNFSSKVFTSWD 322
mouse TMC1 SVLFYGYDNDK-RITGW-----LNFRPLPSYFLVGMCIQYFLVVLKAMTKNIGDDGG--GDDNTFNFSSKVFTSWD 316
human TMC2 SALFYGYNNQ-RITGW-----LYRLPMAYFLVGVSVFVGLMIVIRSMASNTQGTSEGDSNDFPSFKMTSWD 387
mouse TMC2 SALFYGYNNQ-RITGW-----LYRLPMAYFLVGVSVFVGLMIVIRSMASNTQGTSEGDSNDFPSFKMTSWD 369
human TMC3 SVLFYGYGREG-RKITGR-----AGYRLPLAYFLVGMVAFYFIVLLKMKAKNSRISLASASNYTFCWWRVCAWD 276
mouse TMC3 SVLFYGYGREG-RKITGR-----AGYRLPLAYFLVGMVAFYFIVLLKMKAKNSRISLASASNYTFCWWRVCAWD 276
human TMC4 SPLFYGFYPPR-----PRLAVYLCWAFAVYICLILLHRSVSGLKQTLAESEALTYSHRVSADW 300
mouse TMC4 SPLFYGFYPPR-----SNLAITYLCSVFASIVYLLCILLHRSVSGLKQTLAESEALTYSHRVSADW 282
human TMC5 TVMYGYFTNSTIHOHN-----SGASYNMOLAYIFTIGAICLVCFSSLFSPMAKYPFRNNFINPHYSRGVAKLICW 327
mouse TMC5 TVMYGYFTNSTIHOHN-----GGASYNMOLAYIFTIGAICLVCFSSLFSPMAKYPFRNNFINPHYSRGVAKLICW 325
human TMC6 TVMYGYHNSNATLNGPGLDGSQCTPRVGLPYNMLAYLSTVGVSFITCTILVYSMAHSGESYRVG--STSGHAIITVCSWD 388
mouse TMC6 TVMYGYHNSNATLNGPGLDGSQCTPRVGLPYNMLAYLSTVGVSFITCTILVYSMAHSGESYRVG--STSGHAIITVCSWD 387
human TMC7 TSLFYGHYTIIDVQFQNS-----FTYDLPAYLLSTIATLALSLVIVKRSVEGPKINLIRSEEHFQSYCNKIIFAGW 314
mouse TMC7 TSLFYGHYTIIDVQFQNS-----FTYDLPAYLLSTIATLALSLVIVKRSVEGPKINLIRSEEHFQSYCNKIIFAGW 312
human TMC8 TYLFYGAIRVGPES--S-----SVSTRAYLSPSLACLILCFGCTLRMRMKGLPKQKTLRGQYQAPLSAKVSSWD 251
mouse TMC8 TYLFYGAIRVGPES--S-----SEYSTRAYLSPSLACLILCFGCTLRMRMKGLPKQKTLRGQYQAPLSAKVSSWD 255

human TMC1 YLIGNPETADNKFNSTIMNFKHAITEEK-----AAOveenVHLIRFLRFLANFVFLTLGGSGYLI FWAVKRSOEFQAQDPT 400
mouse TMC1 YLIGNPETADNKFNSTIMNFKHAITEEK-----AAOveenVHLIRFLRFLANFVFLTLGGSGYLI FWAVKRSOEFQAQDPT 394
human TMC2 YLIGNSETADNKYSITTSFKESIVDEO-----ESNKENIHLTRFLRVLANFLIICL CGSGYLIFVVKRSOEFKSMON 463
mouse TMC2 YLIGNSETADNKYSITTSFKESIVDEO-----ESNKENIHLTRFLRVLANFLIICL CGSGYLIFVVKRSOEFKSMON 445
human TMC3 YLIGNPEAAESKTAATVNSTREAILDEO-----EKKSKNLAVITCLRIITANLVLVLSLAGSYLTYFVVDRSOKLEQSKKE 353
mouse TMC3 YLIGNPEAAESKTAATVNSTREAILDEO-----EKKSKNLAVITCLRIITANLVLVLSLAGSYLTYFVVDRSOKLEQSKKE 353
human TMC4 FGLCGDVHLRQRITLLELVKELETVVRRQAARVTLGQAVRVLVRLNLLVALLGAAFVGVYVAWGCTVLETQEMPLVQELPL 387
mouse TMC4 FGLCGDVHLRQRITLLELVKELETVVRRQAARVTLGQAVRVLVRLNLLVALLGAAFVGVYVAWGCTVLETQEMPLVQELPL 369
human TMC5 FTVTHEKAVLKKKNLSTETRENLSLRL--QENSKLTFNQLTRFSAYMVAWVSTGVAIACCAVYVLAEBNLEFKTH 405
mouse TMC5 FTVTHEKAVLKKKNLSTETRENLSLRL--QENSKLTFNQLTRFSAYMVAWVSTGVAIACCAVYVLAEBNLEFKTH 403
human TMC6 VKVTOKRASRLOODNTRTOKELLAEWV--LRKSPSVCGLRQAOAVLGLWLLCLGTALGCAVAHVHVSFMIORPAS 466
mouse TMC6 VKVTOKRASRLOODNTRTOKELLAEWV--LRKSPSVCGLRQAOAVLGLWLLCLGTALGCAVAHVHVSFMIORPAS 465
human TMC7 FCHITNRSMDLKHSSIRYELRADLEBERMROKTAERTSETIRIYSLRFLNCTIVLAVGACFVAIYVAVFVSOHHMKKEIDKMVF 401
mouse TMC7 FCHITNRSMDLKHSSIRYELRADLEBERMROKTAERTSETIRIYSLRFLNCTIVLAVGACFVAIYVAVFVSOHHMKKEIDKMVF 399
human TMC8 FCIRVQEAATIKHETSNELKVCLEGRRVLEMOQQTACRLLSYLRNVLNGLLVVGAISAFIWAATKYSODNKEE--S 331
mouse TMC8 FCIRVQEAATIKHETSNELKVCLEGRRVLEMOQQTACRLLSYLRNVLNGLLVVGAISAFIWAATKYSODNKEE--S 335

```

Fig. 1. ClustalW alignment of human and mouse TMC amino acid sequences. Amino acid residues that are identical or conservatively substituted among more than 50% of the homologs are shaded. Amino acid positions are indicated at right. Black bars are shown over six regions of TMC1 that are predicted by TMHMM2.0 to span membranes. A striped bar is shown over the 120-amino-acid TMC domain. A TMC2 splice isoform from the inner ear contains a previously unreported exon encoding 16 in-frame amino acids near the N-terminus; the corresponding isoform and exon could not be detected in the mouse.

human TMC1 LGWWEKNEMNVMVSLGMCFCPTLFDLFAELEDYHPLIALKWLGRIFALLGNLYVFLALMDEINNKIEE EKLVKANITLWEANM 486
mouse TMC1 LGWWEKNEMNVMVSLGMCFCPTLFDLFAELEDYHPLIALKWLGRIFALLGNLYVFLALMDEINNKIEE EKLVKANITLWEANM 480
human TMC2 VSWYERNEVEI VMSLLGMFCPPLEETIAALENYH PRGTGLKWLGRIFALLGNLYVFLALMDDVHLKLANBETIKN - -ITHWTL 546
mouse TMC2 VSWYERNEVEI VMSLLGMFCPPLEETIAALENYH PRGTGLKWLGRIFALLGNLYVFLALMDDVHLKLANBETIKN - -ITHWTL 528
human TMC3 LTLWEKNEVSVVSLVMTIAPSADFLIAALEMYH PRRTTRFOLARVFLVYLGNLYSLI IALLDKVNSMSIEBAMTKN -NTSHWID 437
mouse TMC3 LTLWEKNEVSVVSLVMTIAPSADFLIAALEMYH PRRTTRFOLARVFLVYLGNLYSLI IALLDKVNSMSIEBAMTKN -NTSHWID 437
human TMC4 LKLGVNYLPPPI FAGVNFVLPVPEKLIAPLEGYT RSRQVFLRLTRVFLRSLASLVVLLFSLWNOITCGGDSEA - - -EDCKT 465
mouse TMC4 FKLGLVNYLPSIFALFNFLVLPVSKFIASLEGYTS OSRQVFLRLTRVFLRSLASLVVLLFSLWNOITCGGDSEA - - -EGCKA 447
human TMC5 SNPGAVLLLPVVSCTINLAVPCVSMFRFLVREYEMERHEVYVLLTRNITFKISITIGILCYWLNITVALSGE - - - - - 476
mouse TMC5 RNPQAVLLLPVVSCTINLAVPCVSMFRFLVREYEMERHEVYVLLTRNITFKISITIGILCYWLNITVALSGE - - - - - 474
human TMC6 GQEAVALLPLVVSGLNLGAPYICRVLAALRPHSDSPVLEYVYATCRNLIKLAALGTLCYHWLGRVGVLOG - - - - - 538
mouse TMC6 GQEAVALLPLVVSGLNLGAPYICRVLAALRPHSDSPVLEYVYATCRNLIKLAALGTLCYHWLGRVGVLOG - - - - - 537
human TMC7 ENLFFLILPSIVITLANFITPMIPAKIIRYEDYS PGFEHLTLRLCVFMRLATL CVLVFTLGSKITSCDD - - - - - DTCDL 476
mouse TMC7 ENLFFLILPSIVITLANFITPMIPAKIIRYEDYS PGFEHLTLRLCVFMRLATL CVLVFTLGSKITSCDD - - - - - STCEL 474
human TMC8 LFLLLQYLPFGVIVLNFGLPLLFTFLVLENYN PNTENLTLTWCVVVKLASLGMFSVSLGQTFLCIGR -DK - - - - - SSCES 408
mouse TMC8 LFLLLQYLPFGVIVLNFGLPLLFTFLVLENYN PNTENLTLTWCVVVKLASLGMFSVSLGQTFLCIGR -NK - - - - - TSCES 412

human TMC1 IKAYNASFS - - - - - EN - - - - - STGPPFVHPADVPRGPCWEYVMGOEFVRLTVDLVITHYVTI 539
mouse TMC1 IKAYNESL SGLS - - - - - GN - - - - - TTGAPFVHPADVPRGPCWEYVMGOEFVRLTVDLVITHYVTI 536
human TMC2 FNMYNSSGG - - - - - N - - - - - EBSVPRPLHPADVPRGSCWEYVMGOEFVRLTVDLVITHYVTI 598
mouse TMC2 FNMYNSSGG - - - - - N - - - - - EBSVPRPPHPADVPRGSCWEYVMGOEFVRLTVDLVITHYVTI 580
human TMC3 STTFFAIRTAPDEHEKWSISRPGMGLRRNR - TWALEETSISAYTMLIKANKTSLHTSQSODOCWEYVMGOEMKLSIIDLMLFTVASI 523
mouse TMC3 APTFSAARTVPEBGQWPVPSGGAELRRNSTWVVEETSIFLTSITHTKANKTVPYMQGPOOCWEYVMGOEMKLSIIDLMLFTVASI 524
human TMC4 CGYNYKQL - - - - - - - - - - PCWETVLGOEMKLLFDLITVLAVA 499
mouse TMC4 CGYNYKEI - - - - - - - - - - PCWETRLGOEMKLLFDLITVLAVA 481
human TMC5 - - - - - - - - - - ECWETLLIGODIYRLLMDFVFLVNS 502
mouse TMC5 - - - - - - - - - - ECWETLLIGODIYRLLMDFVFLVNS 500
human TMC6 - - - - - - - - - - QCWEDFVGOEYRFLVGMGFVLLDLD 564
mouse TMC6 - - - - - - - - - - QCWEDFVGOEYRFMVVDVFI MLLDS 563
human TMC7 CGYNOKLY - - - - - - - - - - PCWETOVGOEMKLMDFDITILAVT 510
mouse TMC7 CGYNOKLY - - - - - - - - - - PCWETOVGOEMKLMDFDITILAVT 508
human TMC8 YGYNVCDY - - - - - - - - - - QCWENSVGEELYKLSIFDFTLVAFA 442
mouse TMC8 YGYNACDY - - - - - - - - - - QCWENSVGEELYKLSIFDFTLVAFA 446

human TMC1 LTGDFLRACFVRFCNYCWWDLEYGYPSYTFEDTSGNVLALIFNOGMIWMSGFFAFSLPGINILRHDTSMYFQCWAVMCCNVBEARV 626
mouse TMC1 LTGDFLRACFVRFCNYCWWDLEYGYPSYTFEDTSGNVLALIFNOGMIWMSGFFAFSLPGINILRHDTSMYFQCWAVMCCNVBEARV 623
human TMC2 LTGDFLRACFVRFCNYCWWDLEYGYPSYTFEDTSGNVLALIFNOGMIWMSGFFAFSLPGINILRHDTSMYFQCWAVMCCNVBEARV 685
mouse TMC2 LTGDFLRACFVRFCNYCWWDLEYGYPSYTFEDTSGNVLALIFNOGMIWMSGFFAFSLPGINILRHDTSMYFQCWAVMCCNVBEARV 667
human TMC3 LLIIDDFRGFLVRYLSDYCWDLSEKFLMGEFKEAENVHLVYNOGMIWMAFFSPCLPAPNVLKILGLMLYLSWALTGCVNHQOV 610
mouse TMC3 LLIIDDFRGFLVRYLSDYCWDLSEKFLMGEFKEAENVHLVYNOGMIWMAFFSPCLPAPNVLKILGLMLYLSWALTGCVNHQOV 611
human TMC4 LLIIDDFRGFLVRYLSDYCWDLSEKFLMGEFKEAENVHLVYNOGMIWMAFFSPCLPAPNVLKILGLMLYLSWALTGCVNHQOV 581
mouse TMC4 LLIIDDFRGFLVRYLSDYCWDLSEKFLMGEFKEAENVHLVYNOGMIWMAFFSPCLPAPNVLKILGLMLYLSWALTGCVNHQOV 563
human TMC5 LLVQFPKLLCGLCPG-ALGRSLGT - - - - - EFQDDEVLGLIYAQTVVWVGSFFCCPLLPLINTAKVILFLCKKILFHSIYFASPT 580
mouse TMC5 LLVQFPKLLCGLCPG-ALGRSLGT - - - - - EFQDDEVLGLIYAQTVVWVGSFFCCPLLPLINTAKVILFLCKKILFHSIYFASPT 577
human TMC6 LLVQFPKLLCGLCPG-ALGRSLGT - - - - - EFQDDEVLGLIYAQTVVWVGSFFCCPLLPLINTAKVILFLCKKILFHSIYFASPT 564
mouse TMC6 LLVQFPKLLCGLCPG-ALGRSLGT - - - - - EFQDDEVLGLIYAQTVVWVGSFFCCPLLPLINTAKVILFLCKKILFHSIYFASPT 641
human TMC7 LFVDFPRKLLVYCYSSCKLIQCWGOO - - - - - EFAIPDNVLDIVYQOTICWIGAFFSPLLPAIATLKFIIIFYKWSLILYTCRSPSP 593
mouse TMC7 LFVDFPRKLLVYCYSSCKLIQCWGOO - - - - - EFAIPDNVLDIVYQOTICWIGAFFSPLLPAIATLKFIIIFYKWSLILYTCRSPSP 591
human TMC8 FLVTLPRLLDFRFSG-RFWAWLERE - - - - - EFLVPKNVLDIVAGQTVTWMLGYCPLLPLNLSVFLFITFYIKKYLTLRNSRASRRP 524
mouse TMC8 FLVTLPRLLDFRFSG-RFWAWLERE - - - - - EFLVPKNVLDIVAGQTVTWMLGYCPLLPLNLSVFLFITFYIKKYLTLRNSRASRRP 528

human TMC1 PKASR-SNNFYLGMLLLLIIFLSTMPVLYMIVSLPSFDCGPPFSGKN - - - - - RMFEVIGETLEHDFPSWMAKILROLSNPGLVIAVI 706
mouse TMC1 PKASR-SNNFYLGMLLLLIIFLSTMPVLYMIVSLPSFDCGPPFSGKN - - - - - RMFEVIGETLEHDFPSWMAKILROLSNPGLVIAVI 703
human TMC2 PKASR-SNNFYMGMLLLLIIFLSTMPVLYMIVSLPSFDCGPPFSGKN - - - - - RMVDVLOETIENDPEFTFLGRIFAFLANPGLII PAI 765
mouse TMC2 PKASR-SNNFYMGMLLLLIIFLSTMPVLYMIVSLPSFDCGPPFSGKN - - - - - RMVDVLOETIENDPEFTFLGRIFAFLANPGLII PAI 747
human TMC3 PKASR-SNNFYLAMLLFMFLCMLPTIIPAVRYKPSLNCGPPFSGOE - - - - - KIYDVISETEIENDPEFTFVWFGSVVGHISPPVILPAV 690
mouse TMC3 PKASR-SNNFYLAMLLFMFLCMLPTIIPAVRYKPSLNCGPPFSGOE - - - - - KIYDVISETEIENDPEFTFVWFGSVVGHISPPVILPAV 691
human TMC4 FRASA-ANFFPLVLVLLGLAITSVPLLVSIFLPSKLCGPPFRGQS - - - - - SIWAQIPESISS-LEPATIONFLFFLGTQAFAVPLL 660
mouse TMC4 FRASA-ANFFPLVLVLLGLAITSVPLLVSIFLPSKLCGPPFRGQS - - - - - SIWAQIPESISS-LEPATIONFLFFLGTQAFAVPLL 642
human TMC5 WRASOVITFFIFLFFSFTGVICTLAIITWRLRKPSADCGPFRGLPIFIHSIYSWIDTILSR-PCWLVVWVYRNLIGSVHFFILT 666
mouse TMC5 WRASOVITFFIFLFFSFTGVICTLAIITWRLRKPSADCGPFRGLPIFIHSIYSWIDTILSR-PCWLVVWVYRNLIGSVHFFILT 663
human TMC6 WLASHMSTVFLTLLCFPSFLGAAVFLCMVAVOVRPSSTCGPFRGLPIFIHSIYSWIDTILSR-PCWLVVWVYRNLIGSVHFFILT 729
mouse TMC6 WLASHMSTVFLTLLCFPSFLGAAVFLCMVAVOVRPSSTCGPFRGLPIFIHSIYSWIDTILSR-PCWLVVWVYRNLIGSVHFFILT 728
human TMC7 FRASN-SNFFFLVLLVLLGLAIIPLTISIRSRPSSKACGPFITNPN - - - - - TTWEVIPQVST-FFSSLQSIHTWVISEAFAVFF 672
mouse TMC7 FRASN-SNFFFLVLLVLLGLAIIPLTISIRSRPSSKACGPFITNPN - - - - - TTWEVIPQVST-FFSSLQSIHTWVISEAFAVFF 670
human TMC8 FRASS-STFFFHLVLLVLLGLLAAVPLVYVSSIHSSWDCGLFTNYS - - - - - APWQVPELVALGLPIGORALHYLGHAFSPFL 604
mouse TMC8 FRASS-STFFFHLVLLVLLGLLAAVPLVYVSSIHSSWDCGLFTNYS - - - - - APWQVPELVALGLPIGORALHYLGHAFSPFL 608

human TMC1 LVMVLAITYLNAKQKAANL DLKMKMKALENKMRN - - - - - KKMA 749
mouse TMC1 LVMVLAITYLNAKQKAANL DLKMKMKALENKMRN - - - - - KKMA 746
human TMC2 LLMFLAITYLNSWSKSLSRANAOLRKKIQVLRVEKSHK VVGKATVYSEDTPKSSSKNATQLOLTKEEPSPSASQSOAMDKKAQ 852
mouse TMC2 LLMFLAITYLNSWSKSLSRANAOLRKKIQVLRVEKSHK VVGKATVYSEDTPKSSSKNATQIHLTKEEPSPSSHSSQIOTLDKKAQ 834
human TMC3 LLLFMLIYYLQSLARSLSLNHOLKMOIQNASIS - - - - - 724
mouse TMC3 LLLFMLIYYLQSLARSLSLNHOLKMOIQNASIS - - - - - 724
human TMC4 LISSILMAYTVLANSYGRLLISELKROROTEAQNKVFLA - - - - - RRAV 703
mouse TMC4 LISSILMAYTVLANSYGRLLISELKROROTEAQNKVFLA - - - - - ORAV 685
human TMC5 LIVLITITLYLYOITEGRKIMIRLLHEOIIINEGDKMFLI - - - - - EKLI 709
mouse TMC5 LIVLITITLYLYOITEGRKIMIRLLHEOIIINEGDKMFLI - - - - - EKLT 706
human TMC6 ALLLAVIYFNIOVVRGQKRVICLLKEOISNEGEDIKFLI - - - - - NKHL 772
mouse TMC6 ALLLAVIYFNIOVVRGQKRVICLLKEOISNEGEDIKFLI - - - - - NKHL 771
human TMC7 MIIICLIMFYFIALAGAKRVVIOLEOOLSLESRDKCYLI - - - - - OKLT 715
mouse TMC7 MIIICLIMFYFIALAGAKRVVIOLEOOLSLESRDKCYLI - - - - - OKLT 713
human TMC8 MIIISLVITVCSVTQANARAIHRLRKLQVWQVEKWWHL - - - - - EDLS 647
mouse TMC8 MIIISLVITVCSVTQANARAIHRLRKLQVWQVEKWWHL - - - - - DDLS 651

human TMC1 AARAAAAGRO - - - - - 760
mouse TMC1 AARAAAAGGQ - - - - - 757
human TMC2 GPGTNSASRSTTPASGHLPISRFPFGIGPDSGHAPSQTHPWRSSASGKSARPPH - - - - - 906
mouse TMC2 GPHSTSTEGGAS PSTSWHHVGSQPPRRGRDSGQPOSQTYTGRSPSGKRTRQPPH - - - - - 888
human TMC3 - - - - - 724
mouse TMC3 - - - - - 724
human TMC4 ALTSTKPAL - - - - - 712
mouse TMC4 ALSSRNGTS - - - - - 694
human TMC5 KLQDMKANKPS SILVLERREVEOQFFLHGEHDSGLDRSRRSVOEGNPRA - - - - - 760
mouse TMC5 KLQDMKERVDPALDLERREVEPOIPLHELELGAAPDLRLRRSAQENPIA - - - - - 757
human TMC6 STYERKERERESRVGTTEAAAP - ALLTDEQDA - - - - - 805
mouse TMC6 SVYE - - - - - BEGRSRPGRTOATEPE - AWHEDGDOGKEPCNPR - - - - - 810
human TMC7 EAQRDMRN - - - - - 723
mouse TMC7 EAQREVRSO - - - - - PASA - - - - - 726
human TMC8 RLLPEPGPDSGPKYPASQASR - QSFCPCGCPGSPGHQAPRPGPSVDAAGLRS - PCPCOHGAPASARRRFPPSGAEL - - - - - 726
mouse TMC8 RLLPELSE - - - - - PHSRASR - RFSFCGFPCCPSPPGRTPR LAP - - - - - SNRLS - SSLGAPASVAPASRREHFPSRTEL - - - - - 722

Fig. 1 (continued)

Table 1
Pair-wise amino acid sequence conservation among human TMC paralogs

	TMC1	TMC2	TMC3	TMC4	TMC5	TMC6	TMC7	TMC8		
TMC1		57	41	25	23	24	26	25		
TMC2	73		45	25	22	22	23	25		
TMC3	61	65		25	24	23	27	24		
TMC4	43	43	42		26	26	37	35	% amino acid sequence identity	
TMC5	42	40	38	37		36	27	25		
TMC6	40	38	37	36	49		25	28		
TMC7	45	42	43	56	47	43		34		
TMC8	42	43	41	50	45	43	54			
% amino acid sequence similarity ^a										

^a Summed percentage of identical and conservatively substituted residues.

Tmc1 mRNA is specifically expressed in the neurosensory hair cells of the mouse cochlea [1,2], and therefore this gene and its ortholog were named transmembrane cochlear-expressed gene 1. The development and subsequent rapid degeneration of cochlear hair cells in affected Beethoven and deafness mice indicate a direct and essential role for *Tmc1* in normal hair cell physiology [2,4], but the low levels and restricted tissue distribution of its expression are significant obstacles for in situ studies of its function. We thus sought to identify and clone additional *TMC* genes to identify additional genetic models, tissues, and molecular reagents with which to explore the molecular and physiologic function(s) associated with this gene family. Our results show that mammalian *TMC* genes are expressed with a broad tissue distribution and some underlie mutant phenotypes that do not include hearing impairment, so they have been renamed transmembrane channel-like genes.

Results

Mammalian paralogs and orthologs

A tBLASTn search of genomic and expressed sequence tag (EST) databases with human and mouse *TMC1* and *TMC2* sequences identified six additional human *TMC* paralogs and six corresponding mouse orthologs (Fig. 1). Their corresponding cDNA nucleotide sequences each contained large open reading frames predicted to encode polypeptides with 36–73% pair-wise amino acid sequence similarity and 22–57% pair-wise amino acid sequence identity among

TMC family members within a species (human comparisons shown in Table 1; mouse comparisons not shown). The human and mouse ortholog pairs had higher (75–95%) pair-wise amino acid sequence identity (Table 2), which, in combination with the syntenic relationships among their human and mouse genomic locations (Table 3), confirms our assignment of orthologous relationships. Each gene comprises 14 to 24 exons whose organization is nearly identical within each pair of human and mouse orthologs (not shown).

None of the mammalian *TMC* genes have significant sequence similarity to any other genes or motifs, although they all encode a conserved 120-amino-acid domain that we have named the TMC domain (Fig. 2). This 120-amino-acid sequence was used in a PSI-BLAST query to identify additional, more distantly related, genes from the NCBI non-redundant sequence database. Despite a low *E*-value threshold of 10, the only additional genes that were identified were *Chlamydomonas reinhardtii* adenosine triphosphatase and Crimean–Congo hemorrhagic fever virus envelope glycoprotein precursor. Direct inspection revealed that the similarity to *TMC* genes was not significant (not shown).

One large predicted exon on human chromosome 1 (GenBank Accession No. AL512353) contained sequence homologous to the entire *TMC6* open reading frame but interrupted by numerous frameshifts and termination codons. There are no corresponding ESTs and an orthologous locus is not present in the mouse genome. This processed pseudogene, *TMC9P*, probably arose through a retrotransposition event after the evolutionary divergence of mice from humans.

Table 2
Pair-wise amino acid sequence conservation between human and mouse TMC orthologs

	TMC1	TMC2	TMC3	TMC4	TMC5	TMC6	TMC7	TMC8
% identity	95	83	90	75	80	75	91	78
% similarity ^a	96	88	93	85	86	83	95	83

^a Summed percentage of identical and conservatively substituted residues.

Table 3
Human and mouse *TMC* chromosomal locations, mutant phenotypes, and positional candidate phenotypes

Gene(s)	Location	Mutant phenotypes (MIM No.)	Positional candidate phenotypes (MIM No.)
<i>TMC1</i>	9q13	Dominant deafness DFNA36 (606705) Recessive deafness DFNB7/B11 (600974)	
<i>Tmc1</i>	19 (15 cM)	Dominant deafness: Beethoven Recessive deafness: deafness	
<i>TMC2</i>	20p13		Harboyan syndrome (217400) Corneal endothelial dystrophy 2 (217700)
<i>Tmc2</i>	2 (78 cM)		Dominant deafness: Tailchaser Susceptibility to Sindbis virus replication Nocturnal frontal lobe epilepsy 2 (603204)
<i>TMC3</i>	15q25.3		
<i>Tmc3</i>	7 (41 cM)		
<i>TMC4</i>	19q13.4		Primary ciliary dyskinesia 2 (606763) Spinocerebellar ataxia 14 (605361) Hydatidiform mole (231090)
<i>Tmc4</i>	7 (4 cM)		
<i>TMC5, TMC7</i>	16p12.3		Medullary cystic kidney disease 2 (603860) Dominant deafness DFNA40
<i>Tmc5, Tmc7</i>	7 (53 cM)		
<i>TMC6, TMC8</i>	17q25.3	Epidermodysplasia verruciformis (226400)	
<i>Tmc6, Tmc8</i>	11 (75 cM)		

Protein sequence motifs

TMHMM2.0 strongly predicts [5] the presence of 6 to 10 membrane-spanning domains in each of the *TMC* proteins (Fig. 3). No N-terminal signal peptide sequences or other trafficking signals were identified by PSORT II, and all of the proteins were predicted to reside in the plasma membrane. The C-termini of all *TMC* proteins, with the exception of human *TMC6*, were predicted to be cytoplasmic by TMHMM2.0. The predicted topologic orientation of N-termini was not uniform among the different genes, even using other algorithms (TMPred and PSORT).

PROSITE identified numerous potential sites in each human and mouse *TMC* amino acid sequence for amidation, glycosylation and myristoylation, as well as phosphorylation by protein kinase C, casein kinase II, and cAMP- and cGMP-dependent protein kinases (not shown). Some *TMC* sequences were also predicted to have potential tyrosine kinase phosphorylation sites and leucine zipper structures (not shown). No potential binding sites for ATP or GTP were identified. The C-termini contained potential class I PDZ ligand sequences in mouse *TMC3* (SII) and mouse *TMC8* (TEL), and potential class II PDZ ligand sequences in mouse *TMC7* (ASA) and human *TMC8* (AEL). The N- and C-terminal regions of all mammalian *TMC* proteins were enriched in charged amino acid residues, including approximately equal proportions of basic and acidic side chains. The predominance of charged residues in the N- and C-termini of *TMC1* [1] was also observed for the N- and C-termini of *TMC2* and *TMC3* (Fig. 1). However, the proportions of charged residues in the N-termini of the latter four genes are lower, and the residues do not occur in the same pattern of regularly alternating clusters of similarly charged side chains observed in the N-termini of *TMC1* [1].

Tissue distribution of expression

PCR amplification products corresponding to each of the *TMC/Tmc4*, -5, -6, -7, and -8 genes were detected in each of the placenta, prostate, and testis cDNA libraries. Moreover, numerous EST clones corresponding to *TMC4*, *TMC5*, *TMC6*, *TMC7*, and *TMC8* have been isolated from a broad array of tissue sources from mice and humans (not shown). Whereas Unigene assemblies contained 14 to 80 ESTs each for *TMC4*, *Tmc4*, *TMC5*, *Tmc5*, *TMC7*, *Tmc7*, *TMC8*, and *Tmc8*, they had more than 200 each for *TMC6* and *Tmc6*. In contrast, there are only 5 or fewer ESTs for each of *TMC1*, *Tmc1*, *TMC3*, and *Tmc3*, and no ESTs have been deposited for *TMC2* or *Tmc2*. PCR amplification products corresponding to *TMC3/Tmc3* were detected only in brain and pituitary cDNA libraries. The full range of tissues expressing *TMC1* to *TMC3* may not have been adequately sampled if their mRNAs are expressed at very low levels.

Nonmammalian homologs

Expressed sequence tags with homology to the *TMC* genes were also identified from other mammals, including *Rattus norvegicus* (rat) and *Bos taurus* (cow), and nonmammalian species including *Gallus gallus* (chicken), *Xenopus laevis* (frog), *Danio rerio* (zebra fish), *Anopheles gambiae* (mosquito), *Ciona intestinalis* (sea squirt), *Strongyloides stercoralis* (nematode), and *Molgula tectiformis* (sea shell). A tBLASTn search of the *Fugu rubripes* (Japanese pufferfish) genomic DNA database identified eight distinct genomic loci homologous to mammalian *TMC* sequences (not shown).

tBLASTn searches of the BDGP (Berkeley *Drosophila* Genome Project) database identified a clone (GenBank Ac-

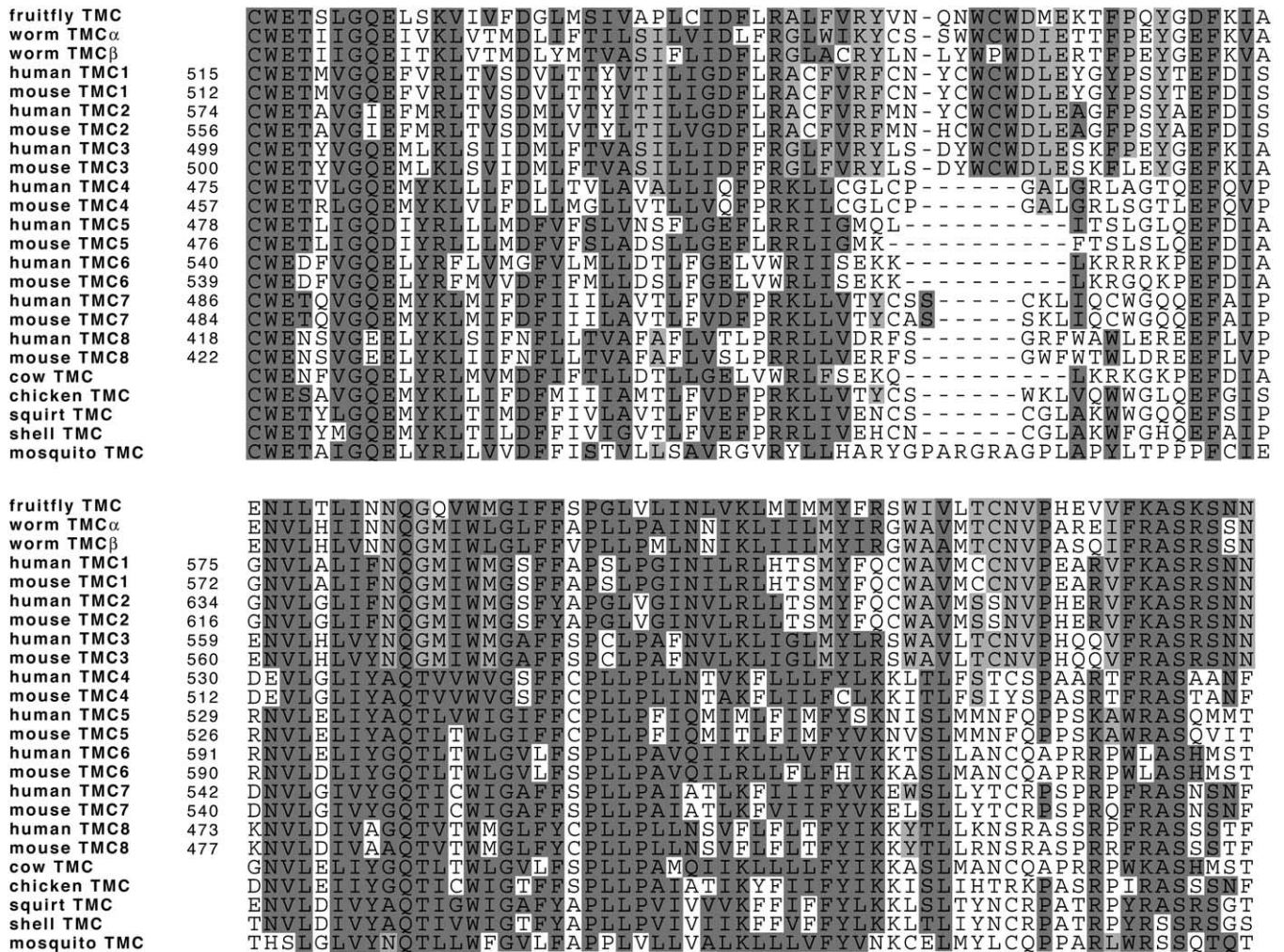


Fig. 2. ClustalW alignment of conserved TMC domains. Amino acid positions of the human and mouse sequences are indicated. Residues that are identical or conservatively substituted among more than 50% (>12/24) of the homologs are darkly shaded. Lightly shaded residues are additional residues that are conserved or conservatively substituted among seven or more of the following homologs: *Drosophila melanogaster* TMC (FlyBase CG3280), *Caenorhabditis elegans* TMC α and TMC β (WormBase T13G4.3 (α) and B0416.1 (β)), and human and mouse TMC1, TMC2, and TMC3. The species origins of additional TMC sequences are indicated: cow, *Bos taurus* (GenBank Accession No. BM255599); chicken, *Gallus gallus* (BU440863); sea squirt, *Ciona intestinalis* (BW054501); shell, *Molgula tectiformis* (AU282683); mosquito, *Anopheles gambiae* (BM636384). Homologous ESTs from *Rattus norvegicus* (BQ199107), *Danio rerio* (AI353074), and *Xenopus laevis* (BI442346) did not fully span this region and were not included in the alignment.

cession No. AE003551) predicted to encode a 1937-amino-acid protein (FlyBase Annotation CG3280) with significant sequence similarity to mammalian TMC sequences, including the highly conserved TMC domain (Fig. 2). tBLASTn searches of *Drosophila melanogaster* genomic DNA confirmed the presence of just one genomic locus with homology to TMC genes. A tBLASTn search of WormBase (<http://www.wormbase.org/db/searches/blast>) identified at least two distinct *Caenorhabditis elegans* genes predicted to encode proteins with significant sequence similarity to the mammalian TMC domains. The deduced amino acid sequences of these two *C. elegans* homologs and the single *D. melanogaster* homolog have 14 to 38% pair-wise amino acid identity with mammalian TMC sequences, with the highest sequence similarities observed with TMC1 to TMC3 (Fig. 2). There were no mutant phenotypes colocal-

izing with the Tmc homologs in *C. elegans* or *D. melanogaster*. Due to the lack of complete genomic sequence information for the other homologs shown in Fig. 2 (cow, chicken, sea squirt, sea shell, and mosquito), the TMC sequences were derived from ESTs and probably do not represent the full complement of Tmc genes in those species.

Evolution of TMC genes

The *C. elegans* and *D. melanogaster* homologs were used to root a phylogenetic tree of the human and mouse TMC genes (Fig. 4). The results indicate that TMC1, -2, and -3 comprise an evolutionarily distinct subfamily, while TMC4, -5, -6, -7, and -8 form another subfamily. TMC5 and TMC7 are separated by approximately 400 kb in a tandem

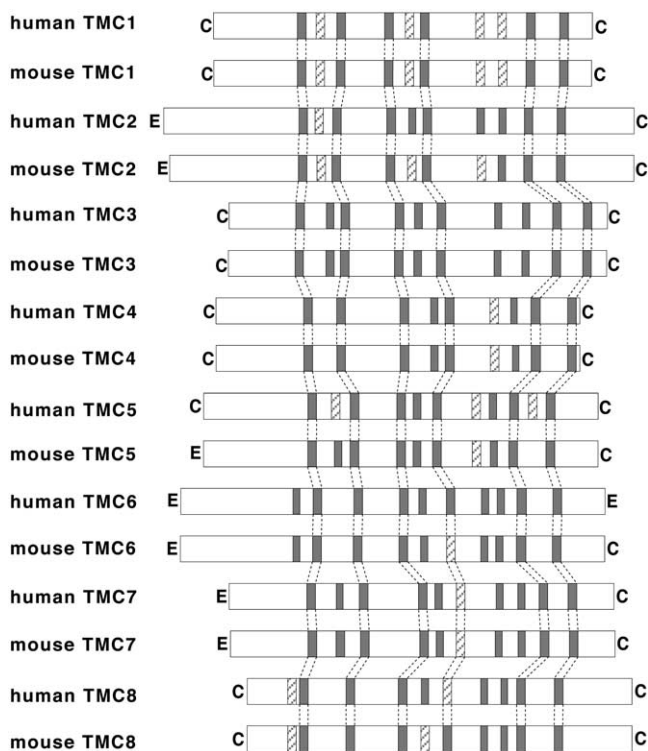


Fig. 3. Transmembrane topologies of TMC proteins predicted by TMHMM2.0. Shaded regions are predicted to span membranes, and striped regions have a lower, subthreshold posterior probability of spanning membranes. Sequences homologous to the six transmembrane domains of TMC1 are aligned with dotted lines. Predicted topologic orientations of N- and C-termini are indicated with E (extracellular) or C (cytoplasmic).

head-to-tail configuration with at least two intervening genes on chromosome 16p12, while *TMC6* and *TMC8* are arranged in a head-to-head configuration separated by only 4 kb on chromosome 17q25.3. The mouse orthologs *Tmc5* and *Tmc7* are oriented head-to-head and separated by 50 kb on chromosome 7, whereas *Tmc6* and *Tmc8* are separated by only 2 kb in a head-to-head configuration on chromosome 11 (not shown). Given the high sequence similarity between *TMC5* and *TMC6* and between *TMC7* and *TMC8* (Table 1), these four paralogs likely arose from a duplication of two tandemly arranged ancestral homologs. However, we could not detect shared linkage of these two gene pairs with paralogs of other gene families to confirm this hypothesis. Moreover, the TMC sequences have diverged to an extent that precludes inference of previous evolutionary events.

TMC6/TMC8 mutant phenotype

Four different recessive, truncating mutations of *EVER1* or *EVER2* have recently been reported to cause epidermodysplasia verruciformis (EV; MIM 226400), characterized by susceptibility to cutaneous human papilloma virus (HPV) infections and associated nonmelanoma skin cancers [6]. The long splice isoforms of *EVER1* and *EVER2* are

identical in sequence and genomic location to *TMC6* and *TMC8*, respectively. EV-based nomenclature is not applicable to most of the *TMC* family members, and it is unknown if *TMC3* to *TMC8* are specifically expressed in cochlea. Therefore, according to HUGO gene nomenclature guidelines [7], the *TMC* locus prefix is retained for all of these genes but their name has been changed to transmembrane channel-like genes.

Discussion

There are 10 published mutant alleles of *TMC1/Tmc1* that include genomic deletions, frameshift, nonsense, and splice site mutations and three different missense substitutions: D572N in dominant DFNA36 hearing loss, M654V in recessive DFNB7/B11 deafness [1], and M412K in the dominant hearing loss mutant Beethoven [2]. All three substituted residues are conserved between human and mouse *TMC1*, whereas only D572 and M412K are also conserved in *TMC2*, and none of them are uniformly conserved among all of the remaining family members. It is possible that one or more of the missense substitutions are nonpathogenic, rare polymorphisms in linkage disequilibrium with undetected mutations elsewhere in *TMC1* or a closely linked gene. Alternatively, these residues may underlie a critical function or structure that is unique to *TMC1* and, possibly, *TMC2*.

The amino acid sequences of the predicted transmembrane domains and intervening polypeptide segments are generally the most conserved among the TMC proteins (Fig. 1), including the 120-amino-acid TMC domain, which immediately precedes the fifth transmembrane domain (Fig.

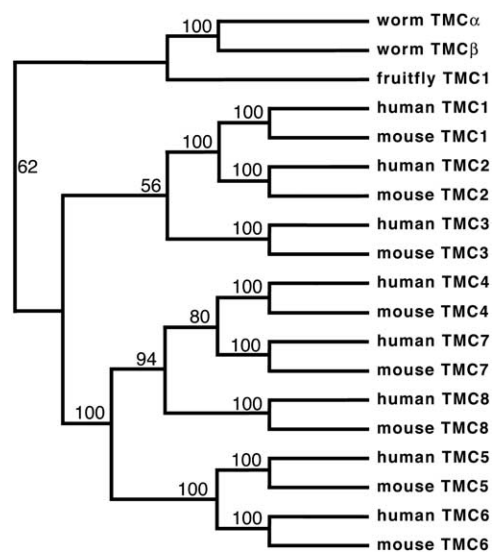


Fig. 4. Phylogeny of human, mouse, *C. elegans* (worm), and *D. melanogaster* (fruitfly) TMC genes predicted by PAUP* 4.0. Bootstrap values for 100 replicates of each parsimony analysis are indicated at nodes of the phylogenetic tree, which was rooted with the three invertebrate homologs.

2). However, the amino acid sequences between the fourth transmembrane domain and the TMC domain are not conserved, with several different gaps and insertions among the family members (Fig. 1). Similarly, the N- and C-termini of the TMC proteins are highly divergent, which is observed in some receptor and ion channel gene families such as the Kir3 (GIRK) inwardly rectifying potassium channel family [8]. By analogy, the conserved transmembrane domains and intervening polypeptide regions of TMC proteins may mediate potential core molecular properties shared among some or all of the TMC family members, such as signal transduction, transmembrane ion conductance, or homo- or heterotypic assembly into higher order multimers.

The recent description of truncating mutations of *TMC6* and *TMC8* that cause EV [6] identifies new experimental models for investigating their function. A typical feature of EV is decreased cell-mediated immunity [9,10], and the identification of *TMC6* and *TMC8* ESTs and cDNA from lymphoid tissue (T lymphocytes and palatine tonsil) is consistent with a model in which EV mutations exert their direct effects in the immune system. Perhaps *TMC6*, *TMC8*, or both underlie signal transduction or ion channel activities in the immune system whose genes have not been cloned, such as Ca²⁺ release-activated Ca²⁺ channels or Ca²⁺ release-activated nonselective cation channels in T lymphocytes [11,12]. The localization of transiently expressed *TMC6* and *TMC8* proteins to the endoplasmic reticulum of a human keratinocyte cell line [6] may not reflect their subcellular distribution in situ, although it does provide experimental evidence that *TMC* genes encode integral membrane proteins. A molecular pathogenesis similar to that in EV may also underlie the observed association of HPV with oropharyngeal (tonsillar) and cervical carcinomas [13,14].

TMC2 is a positional candidate for Harboyan syndrome (CDPD1; MIM 217400; Table 3), an autosomal recessive disorder comprising corneal dystrophy and postlingual progressive sensorineural hearing loss [15], and *Tmc2* is located within the critical interval for the Tailchaser mutation (*Tlc*), which causes progressive hearing loss, vestibular dysfunction, and abnormal hair cell stereocilia development in affected *Tlc/+* mice [16]. However, no mutations have been detected in *Tmc2* exons and adjacent splice sites in Tailchaser genomic DNA (personal communication, Ronna Hertzano and Karen Avraham, Tel Aviv University, March 5, 2003). *Tmc2* is also located in the region of a quantitative trait locus for susceptibility to neuroadapted Sindbis virus replication with paralysis and mortality in female mice (*Nsv1*) [17].

TMC5 and *TMC7* are located on chromosome 16p12, where nonsyndromic dominant hearing loss DFNA40 is reported to be located (Hereditary Hearing Loss Home page), although there are no published genetic map data for this locus. *TMC5* and *TMC7* are also located within the critical interval for autosomal dominant medullary cystic kidney disease 2 (MCKD2; MIM 603860) on chromosome

16p12.3 [18], and both of these genes or their mouse orthologs have ESTs derived from kidney. The phenotypically similar polycystic kidney disease is known to be caused by mutations in polycystin-1 or -2 [19,20], which comprise a nonselective cation channel implicated in intracellular release of Ca²⁺ [21] and mechanosensation of fluid flow by primary cilia of renal tubule cells [22]. The latter phenomenon is notable in its similarity to mechanotransduction of auditory stimuli by cochlear hair cell stereocilia [23]. Another potential link to cilia function may be provided by *TMC4*, which is located within the interval for primary ciliary dyskinesia 2 (CILD2; MIM 606763) [24]. Interestingly, mice that are homozygous for a functional null allele of polycystin-2 have lateralization defects [25], a phenotypic hallmark of primary ciliary dyskinesia. Our study now provides a theoretical foundation and experimental tools to address the roles of *TMC* genes in these and other disorders.

Materials and methods

In silico analyses

TMC1 and *TMC2* amino acid sequences (GenBank Accession Nos. AAL86399, AAL86400, AAL86401, and AAL86402 for human and mouse *TMC1* and human and mouse *TMC2*, respectively) were used as query sequences for tBLASTn analyses with the Celera Discovery System, NCBI BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>), and Ensembl Genome Browser (<http://www.ensembl.org/>). ClustalW (MacVector 7.0, Genetics Computer Group, Madison, WI, USA) was used to align human and mouse amino acid sequences.

Deduced amino acid sequences of human and mouse *TMC1* to *TMC8*, two *C. elegans* *TMC* homologs (T13G4.3 and B0416.1), and a *D. melanogaster* homolog (CG3280) were aligned with ClustalX version 1.8 [26] using Gonnet Protein Matrix with a gap opening penalty of 5.0 and gap extension penalty of 0.1. Nonhomologous regions, including the N- and C-termini as well as sequences with gaps in one or more homologs, were deleted for analysis by the parsimony method with PAUP* 4.0 (Sinauer Associates, Inc, Sunderland, MA, USA). The phylogenetic tree was rooted with the three nonmammalian homologs. Transmembrane topology and domain search analyses of mammalian *TMC1* to *TMC8* were performed with PROSITE (<http://us.expasy.org/cgi-bin/prosite-list.pl>), PSORT II (<http://psort.ims.u-tokyo.ac.jp/>), TMHMM2.0 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>), and TMPred (http://www.ch.embnet.org/software/TMPRED_form.html).

cDNA cloning and sequence analysis.

Genomic DNA, Unigene, and EST sequence information was used to construct composite predicted sequences and design PCR primers for amplification of *TMC/Tmc3*, *-4*, *-5*,

-6, -7, and -8 from placenta, prostate, testis, brain, or pituitary Marathon-Ready cDNA libraries (Clontech, Palo Alto, CA, USA). PCRs were performed in 25 μ l with 2.5 μ l cDNA library, 20 pmol forward and reverse primers, 200 mM each dNTP, 1 \times PCR buffer, and 2.5 U of Takara LA-*Taq* polymerase (PanVera LLC, Madison, WI, USA). Cycling conditions were 95°C for 1 min followed by 35 cycles of 96°C for 10 s and 68°C for 3 min. 5'- and 3'-RACE were also performed as needed to amplify cDNA ends. Nucleotide sequence analysis of amplification products was performed as described [27]. Primer sequences are available from the authors upon request.

Acknowledgments

We thank Ed Wilcox (NIDCD, NIH) for advice on PSI-BLAST analysis, Ronna Hertzano and Karen Avraham (Tel Aviv University) for communicating unpublished data on *Tailchaser*, and Tom Friedman and other Laboratory of Molecular Genetics (NIDCD, NIH) colleagues for helpful discussions and critical comments on the manuscript. This study was supported by NIDCD/NIH Intramural Research Fund 1-Z01-DC-000060-02.

References

- [1] K. Kurima, et al., Dominant and recessive deafness caused by mutations of a novel gene, *TMC1*, required for cochlear hair-cell function, *Nat. Genet.* 30 (2002) 277–84.
- [2] S. Vreugde, et al., Beethoven, a mouse model for dominant, progressive hearing loss *DFNA36*, *Nat. Genet.* 30 (2002) 257–8.
- [3] K.P. Steel, G.R. Bock, The nature of inherited deafness in deafness mice, *Nature* 288 (1980) 159–61.
- [4] G.R. Bock, K.P. Steel, Inner ear pathology in the deafness mutant mouse, *Acta Otolaryngol.* 96 (1983) 39–47.
- [5] S. Moller, M.D. Croning, R. Apweiler, Evaluation of methods for the prediction of membrane spanning regions, *Bioinformatics* 17 (2001) 646–53.
- [6] N. Ramoz, et al., Mutations in two adjacent novel genes are associated with epidermodysplasia verruciformis, *Nat. Genet.* 32 (2002) 579–81.
- [7] H.M. Wain, et al., Guidelines for human gene nomenclature, *Genomics* 79 (2002) 464–70.
- [8] D. Ma, et al., Diverse trafficking patterns due to multiple traffic motifs in G protein-activated inwardly rectifying potassium channels from brain and heart, *Neuron* 33 (2002) 715–29.
- [9] K.D. Cooper, E.J. Androphy, D. Lowy, S.I. Katz, Antigen presentation and T-cell activation in epidermodysplasia verruciformis, *J. Invest. Dermatol.* 94 (1990) 769–76.
- [10] S. Majewski, et al., Natural cell-mediated cytotoxicity against various target cells in patients with epidermodysplasia verruciformis, *J. Am. Acad. Dermatol.* 22 (1990) 423–7.
- [11] R.S. Lewis, Calcium signaling mechanisms in T lymphocytes, *Annu. Rev. Immunol.* 19 (2001) 497–521.
- [12] Z. Su, et al., A store-operated nonselective cation channel in lymphocytes is activated directly by Ca(2+) influx factor and diacylglycerol, *Am. J. Physiol. Cell. Physiol.* 280 (2001) C1284–92.
- [13] J.M. Walboomers, et al., Human papillomavirus is a necessary cause of invasive cervical cancer worldwide, *J. Pathol.* 189 (1999) 12–9.
- [14] E. Ringstrom, et al., Human papillomavirus type 16 and squamous cell carcinoma of the head and neck, *Clin. Cancer Res.* 8 (2002) 3187–92.
- [15] M.J. Abramowicz, J. Albuquerque-Silva, A. Zanen, Corneal dystrophy and perceptive deafness (Harboyan syndrome): *CDPD1* maps to 20p13, *J. Med. Genet.* 39 (2002) 110–2.
- [16] A.E. Kiernan, et al., *Tailchaser* (*Tlc*): A new mouse mutation affecting hair bundle differentiation and hair cell survival, *J. Neurocytol.* 28 (1999) 969–85.
- [17] D.C. Thach, S.R. Kleeberger, P.C. Tucker, D.E. Griffin, Genetic control of neuroadapted Sindbis virus replication in female mice maps to chromosome 2 and associates with paralysis and mortality, *J. Virol.* 75 (2001) 8674–80.
- [18] F. Scolari, et al., Identification of a new locus for medullary cystic disease, on chromosome 16p12, *Am. J. Hum. Genet.* 64 (1999) 1655–60.
- [19] The European Polycystic Kidney Disease Consortium, The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16, *Cell* 77 (1994) 881–94.
- [20] T. Mochizuki, et al., *PKD2*, a gene for polycystic kidney disease that encodes an integral membrane protein, *Science* 272 (1996) 1339–42.
- [21] P. Koulen, et al., *Polycystin-2* is an intracellular calcium release channel, *Nat. Cell Biol.* 4 (2002) 191–7.
- [22] S.M. Nauli, et al., *Polycystins 1 and 2* mediate mechanosensation in the primary cilium of kidney cells, *Nat. Genet.* 33 (2003) 129–37.
- [23] J. Garcia-Anoveros, D.P. Corey, The molecules of mechanosensation, *Annu. Rev. Neurosci.* 20 (1997) 567–94.
- [24] M. Meeks, et al., A locus for primary ciliary dyskinesia maps to chromosome 19q, *J. Med. Genet.* 37 (2000) 241–4.
- [25] P. Pennekamp, et al., The ion channel *polycystin-2* is required for left–right axis determination in mice, *Curr. Biol.* 12 (2002) 938–43.
- [26] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The *CLUSTAL_X* windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.* 25 (1997) 4876–82.
- [27] J.M. Bork, et al., 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 (2001) 26–37.