Research article

Open Access TMC and EVER genes belong to a larger novel family, the TMC gene family encoding transmembrane proteins

Gabor Keresztes^{†1,2}, Hideki Mutai^{†1,2} and Stefan Heller^{*1,2}

Address: 1Department of Otolaryngology and Program in Neuroscience, Harvard Medical School and 2Eaton Peabody Laboratory, Massachusetts Eye & Ear Infirmary, Boston, MA

Email: Gabor Keresztes - gabor@epl.meei.harvard.edu; Hideki Mutai - mutai@epl.meei.harvard.edu; Stefan Heller* - hellers@epl.meei.harvard.edu

* Corresponding author †Equal contributors

Published: 17 June 2003

BMC Genomics 2003, 4:24

This article is available from: http://www.biomedcentral.com/1471-2164/4/24

© 2003 Keresztes et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Received: 27 March 2003 Accepted: 17 June 2003

Abstract

Background: Mutations in the transmembrane cochlear expressed gene I (TMCI) cause deafness in human and mouse. Mutations in two homologous genes, EVER1 and EVER2 increase the susceptibility to infection with certain human papillomaviruses resulting in high risk of skin carcinoma. Here we report that TMC1, EVER1 and EVER2 (now TMC6 and TMC8) belong to a larger novel gene family, which is named TMC for trans membrane channel-like gene family.

Results: Using a combination of iterative database searches and reverse transcriptase-polymerase chain reaction (RT-PCR) experiments we assembled contigs for cDNA encoding human, murine, puffer fish, and invertebrate TMC proteins. TMC proteins of individual species can be grouped into three subfamilies A, B, and C. Vertebrates have eight TMC genes. The majority of murine TMC transcripts are expressed in most organs; some transcripts, however, in particular the three subfamily A members are rare and more restrictively expressed.

Conclusion: The eight vertebrate TMC genes are evolutionary conserved and encode proteins that form three subfamilies. Invertebrate TMC proteins can also be categorized into these three subfamilies. All TMC genes encode transmembrane proteins with intracellular amino- and carboxyltermini and at least eight membrane-spanning domains. We speculate that the TMC proteins constitute a novel group of ion channels, transporters, or modifiers of such.

Background

The complete sequencing of the human genome led to the conclusion that the number of human genes is approximately 30,000 - 40,000 [1,2]. This initial estimate was proven to be incorrect because a comparison of the predicted genes that were identified in public and commercial sequencing projects and inclusion of a third cluster of known genes revealed little overlap among the three groups [3]. Although many of the discrepancies between the different human genome assemblies are caused by fundamental disparities in the compilation of the sequence information [4], it is still evident that the human genome harbors probably significantly more than 40,000 genes and that these genes are not easily identified. This is particularly true for genes that are either expressed transiently, at low levels, or in very few specific cell types. Such elusive genes are underrepresented in cDNA libraries and consequently also in expressed sequence tag (EST) databases. Computational annotation of such genes from genomic DNA becomes additionally challenging when the genes consist of many scattered small exons.

Two pairs of previously un-notated genes encoding transmembrane proteins have recently been identified based on their linkage with inherited disorders. The first set of genes, transmembrane cochlear expressed genes 1 and 2 (*TMC1*) and *TMC2* [5] were found in a search for mutations causing dominant and recessive deafness in human and mouse [5,6]. Mutations in the second pair of genes, *EVER1* and *EVER2*, are linked to epidermodysplasia verruciformis that is associated with an increased susceptibility to infection with some human papillomaviruses causing a high risk of skin carcinoma [7]. We noticed that the proteins encoded by these genes are homologous, which raised our interest in testing whether the genome accommodates additional related genes.

Results

The mammalian TMC gene and protein family

We used human TMC1 and TMC2, murine Tmc1 and Tmc2, and human EVER1 and EVER2 sequences in conjunction with database search algorithms to identify homologous sequences in genomic and expressed sequence tag (EST) databases. We iteratively assembled contiguous theoretical coding sequences for individual human and mouse proteins. We verified regions of individual murine cDNAs that were not unequivocally predictable with database information by RT-PCR and by sequencing. This strategy resulted in the identification of the murine orthologues of EVER1 and EVER2, and four hitherto uncharacterized human and mouse proteins (Figure 1, Table 1). All identified proteins have eight predicted membrane-spanning domains (TM1 - TM8) and share the completely conserved amino acid triplet C (cysteine) - W (tryptophan) - E (glutamic acid), predicted to be located in the extracellular loop upstream of TM6 (Figure 2). The presence of this hereby named TMC signature sequence motif CWETXVGQElv(K/R)LtvXD is our defining criterion for this novel TMC protein family (Figures 1, 2, additional file 2).

The eight mammalian TMC proteins can be grouped into three subfamilies A, B, and C, based on sequence homology and on similarities of the genomic structure of their respective genes (Figures 1, 3, and 4). The murine TMC protein subfamily A consists of three proteins, Tmc1, Tmc2, and the novel Tmc3. All TMC subfamily A proteins are between 757 and 1130 amino acids in length. Tmc3 bears a long carboxyl-terminal tail that is unlike all other TMC proteins (Figure 1). The overall identity within the TMC subfamily A is 36–56%; the positions of >73% of the genes' introns within the conserved core region are conserved (Figure 3A, 3C, additional file 1).

The murine TMC subfamily B consists of Tmc5 and Tmc6 (mouse orthologue of EVER1), proteins of 810 and 757 amino acid residues that are 31% identical and share a

>92% conservation of the corresponding genes' intron locations within the conserved core region (Figure 3A, 3C, additional file 1). A significant structural difference between subfamily B and subfamily C proteins is that the long presumptive extracellular loop of TMC subfamily A proteins between TM5 and TM6 is much shorter in subfamily B proteins and mainly consists of the TMC signature sequence motif (Figures 1, 2).

Finally, the three members of the murine TMC subfamily C, Tmc4, Tmc7, and Tmc8 (mouse orthologue of EVER2) are the shortest TMC proteins with 694, 726, and 722 amino acid residues. The overall identity within the murine TMC protein subfamily C is 29–33% with a common gene structure of >92% conserved intron locations within the conserved core region (Figure 3A, 3C, additional file 1).

Analysis of the human TMC gene and protein family yielded principally identical results, likely because of the high degree of conservation between human and murine TMC proteins [5] (Figure 4). The new TMC classification now designates *TMC6* for *EVER1*, and *TMC8* for *EVER2* (Table 1).

The TMC genes map to six chromosomal locations in the human and mouse (Table 1). Two chromosomal locations in both species harbor two neighboring TMC genes, *Tmc5* and *Tmc7* on murine chromosome 7, and *Tmc6* and *Tmc8* on murine chromosome 11; the human orthologues are located on the syntenic regions of chromosomes 16 and 17, respectively [7]. An additional murine locus on chromosome 8 represents a partial gene fragment, likely a pseudogene of *Tmc2*.

Expression of murine TMC transcripts

To demonstrate that all predicted murine TMC genes are transcribed, we performed RT-PCR experiments. Primer pairs specific for each TMC transcript amplified products of predicted length and sequence (Figure 5).

Tmc1 and Tmc3 mRNAs were detectable in most neuronal organs and we also found expression in some non-neuronal organs. Tmc2 transcripts were only detectable in testis; we did not reveal by RT-PCR Tmc2 expression in cochlea. However, we were able to verify that Tmc2 is expressed in the cochlea by using an organ of Corti cDNA library as a template for PCR, which corroborates the results presented by Kurima et al. (2002).

We observed that mRNAs encoding Tmc5, Tmc6, Tmc4 and Tmc7 are expressed in most murine organs tested. Tmc8 mRNA is detectable in thymus and lung; we also found expression of Tmc8 mRNA in spleen (not shown).

Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Und QUEEKEENERSSEE Disc. Und QUEEKEENERSSEE Und QUEKEENERSSEE Und QUEKEENERSSEE Und QUEKEENERSSEE Und QUEKEENERSSEE Disc. Und QUEKEENERSSEE Und QUEKEENERSSEE Und QUEKEENERSSEE Und QUEKEENERSSEE Und QUEKEENERSSEE				
The State is a second sta	Tmc1	(Mm)	1	
The Description Description <thdescription< th=""></thdescription<>	Tmc2	1		
The Control Image: Additional program and the program			- C.	
The Constraint Description Description Description The Constraint Independence in the set of				
The Tory International Control State				
<pre>11 1 INCOMENT IN</pre>				
The Composition 94 AMPROXIMATION CONCLUMENTS AND ADDRESS ADDR	Tine /	(Mm)	1	MSEGSAGDPGHGSSKQKAVIPENLSLGSS
The Composition 94 AMPROXIMATION CONCLUMENTS AND ADDRESS ADDR	Tmc1	(Mm)	42	INEDDPEPEPEDEETRKAREKERRRRLRRGAEEE EIDEEELERLKAL DENROMIATVRCKPWKMEKKIEVLKEAKKFVSENEGALGKGKKKW
The Composition 9 HTMSERBAGCTLOQUELLASSESFINATION CONFIGURATION HTMSERBAGCTLOQUELLASSESFINATION CONFIGURATION 10 HERDERBAGGTLOQUELLASSESFINATION CONFIGURATION HTMSERBAGGTLOQUELLASSESFINATION CONFIGURATION 11 INFORMATION CONFIGURATION CONFIGURATION CONFIGURATION HTMSERBAGGTLOQUELLASSESFINATION CONFIGURATION 11 INFORMATION CONFIGURATION				
The S 101 THETYLERASSEPT.LAUVYSARELLY CLU OFT.LEDDERGEL, VELO GAAGGEDWYSHELLGERC, LEEDERGER, UGG, C.A.M. THE THETYLERASSEPT.LAUVYSARELLY TOLL OFT.LEDDERGEL, VELO GAAGGEDWYSHELGERC, CLEAR ST. PERIL PHYLOLAUVES AND TOLL OFT.LEDERGER CLEAR ST. THE THETYLERASSEPT.LAUVYSARELLY TOLL OFT.LEDERGER CLEAR ST. PERIL PHYLOLAUVES AND TOLL OFT.LEDERGER CLEAR ST. THE THE THE THE THE THE THE THE THE THE THE	Tmc3	(Mm)	9	RYRSIRRNASQCYLYQDSLLLGNSDDSFNADETGDSSDPEQIFQNIQFQKDLMANIRCRPWTMGQKLRALRRAKEIVLKFEGRLTRTRG
The description 15 KARTERSELAUMADLESALTINGETURPHOLEDDDIOGRALDATATAQUESH FEBRUARIESALTINGETURPHOLEDDIOGRALDATATAQUESH Test	Tmc5	(Mm)	58	NPGNDRRHNRLLRFSSLNESISQIYHGSECLVTDESCTFHETVQGQKLIASLIPMTTRDRIKTIRNQPRTMQEKRELRKIVDKEKNKQSHGILEA-NC
The Tory (1) 30 OPSPYPHYLQLPPERSYNARUTILISENCEGOLULTERING THELANG STREENE DIG CONTROL NUMBER OF DISCUSSED (1) Tory (1) 10 PATOGENESS (2)	Tmc 6	(Mm)	101	YNRTVRLRRRSSRPLLGNVVPSARPSLRLYDLELDSTILEEDEKRSLLVKELQGLSAAQRDHMVRNMPLSLGEKRCLREKSWSPKGKRRHLQGRSGAFSC
Table 1 Description Table 1 Description Table 1 Table 2 Table 2 <thtable 2<="" th=""> <thtable 2<="" th=""> Tabl</thtable></thtable>				
 11.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1				
The Control 189 TATMONDARK-VEYERDED FUNCTION FUNCTION TO THE ADDRESS OF THE ADDR	Tmc8	(Mm)	1	MFRQWSVQSGPAPRRPESQAASEELWEQEVERLCASRTPVRMLPYAMADKRF RELREPEGVETTF
The Control 189 TATMONDARK-VEYERDED FUNCTION FUNCTION TO THE ADDRESS OF THE ADDR	Tmc 1	(Mm)	137	PAPKMMAKK-WAKFLRDFENFKAACUPWENKIKATESOFGSSVASVFLFLRNWYGUNMULFVLTFSLTMLPEVINGLPVGSL
 Theo J (m) 98 - TQAAAAAL-IMERAAAAAUNYUT JITUREE UKI SEPT-GOVAANT TJUREE UKI JITUREE UKI STUPPE - VALUE UKI SEPT-SEPT SERVICE UKI SEVEN SE		(/		
The S (m) 15 CACLOSCISTITEREBRICALISTIC LANGE THE LAY FILL OF THE TABLE OF THE TABL				
 Tincel (m) 9. SESTANTTROLARTKORDERGY ZIG (WARTER KI LOGG) FOATERS YELLAR LIJZ LUKUNG YELLOW LIJZ (LIDVA) FLIJT KI KI				
The T Init INIT </td <td>Tmc6</td> <td>(Mm)</td> <td></td> <td></td>	Tmc6	(Mm)		
The Dire 61 62 NUMBER AND ADDRESS AD	Tmc4	(Mm)		
10 21 DET/INFORME/DATE/DATE/DATE/DATE/DATE/DATE/DATE/DAT				
The C Number C TORK HILLING TORK HILLING <thtork hilling<="" th=""> TORK HILLING</thtork>	Tmc8	(Mm)	67	WQRWHRPRRVARQHLREAEQRLARGFGLWEGALYEIGGLFGLGIQSYFTFLRFLLLLNLLTMLLTAC.VLLPL/WLRPPELGPALKLSSLQCSS
The C Number C TORK HILLING TORK HILLING <thtork hilling<="" th=""> TORK HILLING</thtork>	Tree 1	(Mm)	220	RETURDA FRASANFOUL VERICE AVSUL FYGYVENKR
The 1 The 1 The 1 The 2 The 2 <td< td=""><td></td><td></td><td></td><td></td></td<>				
Thick Number 2015 FUNCTION FUNCTION FUNCTION <				
The G (M) 24 -VPFFSGL				
Time 1 19 21 YTENTIGUARPTQUFHIGURATION FOR FULL SEGUENCE LUNG FULL SEGUENCE LARGE SEGUENCE AND				
Thm 7 (im) 214 YEISSGUTYTSTICLAGGOVLEETSLYCHTTODVEOGST- UDL/ALICETY ALSL/UNTVESTOR MURPHYRING 1000 SPLADGUTHENTMINICAGUAL CONTRAST SINCE SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES 2101 1001 STATUPENNYCKENDLIGGER CONTRAST SINCE SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES 2101 1001 STATUPENNYCKENDLIGGER CONTRAST SINCE SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES 1002 1001 SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES 1001 1102 SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES 1001 1102 SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES SSLADGUTHENTVESTOR SALES 1001 1112 SSLADGUTHENTVESTOR SALES SS			192	YIPHTHGLVAFPTQLFNLLSGEGYLEWSPLFYGFYPARSNLAITYLCSVFAISVIYLLCILRRSVSGLKETLLAESD
Time (W) 160 SPLEQCOLFPRINTPUNITCIPAL-INTELYCATRAGEESSE		(Mm)	214	YPISSSGLIYFYSYIIDLLSGTGFLEETSLFYGHYTIDGVKFQSFTYDLPLAYLISTIAYLALSL.WIVKRSVEGFKINLIRSEE
<pre>TmC (W) 35 DETYPERMETENCT.LOSSET.DARKYCHITTEFF</pre>	Tmc8	(Mm)	160	SPLPQSDIPRFHNPLWNILTGRAF-NNTYLFYGAYRAGPESSSEYSIRLAYLLSPMVCLLLCFCGILQRMAEGLPQQTLLGQR
<pre>TmC (W) 35 DETYPERMETENCT.LOSSET.DARKYCHITTEFF</pre>				
Tmc1 (m) 261 PHTTPCMWYCAMOULOUPEAAESKYARIAUSIE Pmc3 (m) 315 SKGLA-LTYCENOFYTHEKKURKUKAUSTUETA				
<pre>Tmc5 (mp) 11 yENGLA-KLIPCMPTYTERKUKLEVELMELTEIRELAISLERGEVELTPRQUTEFSAUVAUKUTSTVIALCUMYTLEYNBERLETREIP Fmc6 (mp) 25 KIISLA-TUFCMPTYTERKUKLEVELMELTEIR</pre>				
<pre>TmG (m) 37 KOINAL-TYPESDORTWTTERSERVEDUKELLABMILEREREVEDUKE_VVERSERVEDUKE_UVERSERVEDUKEUVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUVERSERVERSERVEDUKEUTUKEUVERSERVERSERVERSERVEDUKEUTUKEUVERSER</pre>				
 Tmc4 (m) 269 ILISTSSERVFANMPGLCGDURVELAGILITELQUDLEERVVRAAAE/CLSGRAVVBAAE/CLSGRAVVBAAE/CLSGRAVVLAAE/CAAFYGIVAAEFYLLULGEFYLVRQPF MMC9 (M) 242 WEPTSAAVTSSWDCICKWEAATIKKELINELKKELEGGAVELAQUTAAGKACLLTYLKULVAAUAATIKATYJQDNEESUF MM 384 WENEH				
Tmc? (m) 299 HPSSCHKTPAGNPCTT HENDELENERLEAPERERTEGUTERTTERTTERTTERTTERTTERTTERTTERTTERTTER				
Tmc8 (Mm) 242 YRTPLSAKUYESSINPCTRVREAATIKKEEIMELKEEGRAVIELAQUTRACKACKLITYIKTNII, IVILUVVAISAITIAATIKYSQDNKEESIF				
<pre>Tmc1 (m) 396 VERNEN</pre>				
 Tmc2 (km) 449 YERREYE_WESLGHCCPFLFTLALNN-HERTCLINGLORJFALLONUTFY-LALADOVUSINULASIERTINTLYNY Tmc3 (km) 55 WIKENSVVXSUVVENUASSAPOLIAAUNA-UPRTCLINGUSULANUTSIITALDOVUSINUEAUNEENTENTYPEEG Tmc5 (km) 407 GAVLLLPFVVSULLAVRFYFLGSAPOLIAAUNA-TERTEPPEGUVVLLUNGUTFYLLADOVUSINUEAUNEENTENTYPEGU Tmc4 (km) 36 LEFLLUDULPSIFJSUFHULPVYFFLGSUGUVULUNGUTFULISITUG_LYTUAL-VALS- Tmc6 (km) 37 GENLUDULPSIFJSUFHULPVYFFLGSUGUVULUNGUTSILLARVFYLLANSUVF_UVSUSGICCOGOMELAECKAGGYNK Tmc7 (km) 38				
Thm 3 17 WENNEYSYWELVTELAPESPTCLEALENT-PETTLEPCLAVVULTURUUTLEALLENTERTEPEQUAVVULTURUUTLEALLENTERTEPEQUAVVULTURUUTLEALLENTERTEPEQUAVVULTURUUTLEALLENTERTEVEQUAVULTURUUTLEALLENTERTEVEQUAVULTURUUTLEALLENTERTEVEQUAVULTURUUTLEALLENTERTEVEQUAVULTURUUTLEALLENTERTEVEQUAVULTURUUTLEALLENTERTEVEQUAVULTURUUTLEALLENTERTEVEQUAVUULTURUUTLEALLENTERTEVEQUAVUULTURUUTLEALUPVEVENTURUUTLEATURUUTULEATURUUTULEATURUUTULEATURUUTULEATURUUTURUUTURUUTURUUTURUUTURUUTURUUTUR		()		
 TIMES (MI) 407 GAVLILPFVVSCINLAUPPRYSHPRLVERYEPPROEVVLLVENHITPLKISJVGILTVINITVALS TIME 407 GAVLIL				
 Tmc6 (Mn) 449 VIALAPLVWSVLMIGGSLIFRGG_TLERRDSPUEVUMICENT.TUNWAVG/UVUMLCRWAT				
<pre>Tmc4 (Mm) 369 LFKLUPDLPSTJSLFNPVLPPVRFLSLKVC-TQSRCJVLLLLRVYFLRLASLW/IVSKNSQI</pre>		(
<pre>Tmc? (Mm) 399 GINLLITURESTVITLARY TEPTIFAKT INTERD: SPGPETRATING/VPRILARY CUTYLIGS(INTEGESGINAC</pre>		1		
 Tmc8 (Mm) 338LULQULPPGUISLUNPLGPQLPTVLIQLENY-PPGTEVRLTLINCUVLKLASIGNESPSIGQTVLUIGREKTS-CESYGYENAC				
<pre>Tmc1 (M) 485NESLSGLSGNTTGAPFPVHPADVPROPCNETWVGGEFVKLTVSDVLTIYTTLIGDPLRACFVRPCNTCHCHDLEYGYP Tmc2 (M) 533NSSGGNESVPRPPHPADVPROSCNETTVGGEFVKLTVSDULTIYTTLIGDPLRACFVRPNNCKCHDLEAGFP Tmc5 (M) 473GECNETTVGGFVLTVUGGEFVKLTVGGEFVKLTSDULTIYTTLIGDPFGFVKLNSSKYPENKLSSKPP Tmc5 (M) 473GECNETTGAPFVVEETSFTETHTKANKTVPUQGFQCVENTTVGGEFUKLTIDPFFNLTUFTSDVECHDLESKPP Tmc5 (M) 474GECNETTGAPFVVEETSFTETHTKANKTVPUQGFQCVENTTVGGEFUKLTIDPFFNLTUFTSDVECHDLESKPP Tmc5 (M) 473GECNETTGAPFFVVEETSFTETHTKANKTVPUQGFQCVENTTVGGEFUKLTIDPFFNLTUFTSDVECHSESKPP Tmc5 (M) 454GECNETTGAPFVETTFLTURTUFGGEFUKLTIPFFLTURGFFVERSTESKPP Tmc5 (M) 451DCDNETTGAPFVETTFLTURTUFGGEFUKLTIPFFILTURGFFVERSTESKPP Tmc5 (M) 554 STEFPISGNTLALFPNQGHIMNGSFTAPGLOPHILLITSNTPQCNAVGSNVFEARVFXASRSNNFTLGLLLIFFSTEFVKKLRKG Tmc6 (M) 555 EIGFFNIGGETUKLLIFSTEFFTIGGLIFFNIGGEFUKLTIPFFILTURGFFVERSTEGF Tmc2 (M) 564 STEFPISGNTLGLFRQGHIMNGSFTAPGLOPHILLITSNTPQCNAVGSNVFEARVFXASRSNNFTLGLLLIFFSTEFFTIGGT Tmc2 (M) 555 EIGFFNIGGETUKGETUKGLIFFFILTFFILTFFUKTULGGHINGSFTAPGLOPHILTFFILTFFILTFFILTFFILTFFILTFFILTFFILTF</pre>			338	
 Tmc2 (Mm) 533		()		
 Tmc3 (Mm) 452 QMPTEGGAELRRITSTMYVERTSFLTSTMYVERTSFLTSTMYVERTSGQCMATYUGGPEGGAELRSVTENGPTVASILLDFPFRGLFVRILDSTVENDUESSFPF Tmc5 (Mm) 536 Tmc7 (Mm) 541 QQCCMEDYORGLYRPHVUPFTMILDSFLGGLVRRLIGM	Tmc1	(Mm)		
<pre>Tmc5 (Mm) 473 Tmc6 (Mm) 536 Tmc6 (Mm) 536 Tmc6 (Mm) 536 Tmc7 (Mm) 451 Tmc1 (Mm) 54 Tmc7 (Mm) 451 Tmc1 (Mm) 564 SYTEPDISGNVLALIPNQCHINNOSFYAPSLPGINILALHTSNIPQCMAVMSCNVPEMUTPDINIL/VTLPVDPPRULUSELFCEUVRUSLIPS Tmc1 (Mm) 564 SYTEPDISGNVLALIPNQCHINNOSFYAPSLPGINILALHTSNIPQCMAVMSCNVPEMUTVAPAFUSLSEPRLUVERFSGM</pre>				
Tmc6 (Mm) 536				
Tmc4 (Mm) 454			473	GEECHETLIGODIVRLLLANDFVFSLADSLLGEPLARLIGHKFTSLS
 Tmc7 (Mm) 481 Tmc6 (Mm) 419 Tmc7 (Mm) 481 Tmc8 (Mm) 419 Tmc8 (Mm) 419 Tmc8 (Mm) 564 SYTEPDISONULALIPNOGHIMMOSPAPSLOPINILRIHTSMIPQCHAVMCSVPEARUPKASRSNNFYLAULURLPSTAFUVSRSHPPSPCG Tmc1 (Mm) 564 SYTEPDISONULALIPNOGHIMMOSPAPSLOPINILRIHTSMIPQCHAVMSSVPEARUPKASRSNNFYLAULURLFISLEVRA-TYMSLPPSPCG Tmc3 (Mm) 552 SUGEFKIARNULHUVNOGUINMOSPYAPSLOPANULKILISMIPQCHAVMSSVPERVFKASRSNNFYLAULURLFISLEVA-TYMSLPPSPCG Tmc3 (Mm) 554 SUGEFKIARNULHUVNOGUINMOSPYAPSLOPANULKILISMIPQCHAVMSSVPERVFKASRSNNFYLAULURUPTPAIVMXPSLOPSPCG Tmc3 (Mm) 554 SUGEFKIARNULHUVNOGUINMOSPYAPSLOPANULKILISMIPQCHAVMSSVPERVFKASRSNNFYLAULLUVPTYAIVMXPSLOPSPCG Tmc4 (Mm) 564 SUGEFKIARNULHUVNOGUINMOSPYCHULPUNTURUPTVINVSLAMNOQAPRENHASRMNTYPILLUVEFSFTGVLCI.INTITMELKPSALGG Tmc4 (Mm) 563 SUGEFKIARNULDITGOTICHIGAPFSPLLPAUQITRHLF:FHIKKASLANNOQAPRENHASRMTYPFLIVLLGLLAAVUVPL-VSIFSHSKLCG Tmc6 (Mm) 663 PFSURMKIPEVIGETLEHDPFSM:AKILRQLSNPGLVIAVILVNULTIYLN-ATAGQKAANLDLKKKKQQALENKKKMAARAAAAAG Tmc1 (Mm) 663 PFSUGEKIPOTULWVINDUSUSPVULLPILINSVFLFTFYTIKYTLINSTASEFRFFHANGQKAANLDLKKKKQQALENKKKMAARAAAAAG Tmc3 (Mm) 663 PFSUGEKIPOTULWVINDUSUSPVULPILIPTITKYTLINSTSSTSSALGUNGQINAGINAKKKAQAVEALIPSDA Tmc5 (Mm) 664 PFRILNTUREAGTVNVRRLEHAG-SGASH.PHILHFLVENTYFFF1LIVLITYIN-ATAGQKAANLDLKKKKQQALENKKKMAARAAAAAG Tmc6 (Mm) 663 PFRILNTUREAGTVNVRRLEHAG-SGASH.PHILHFLVENTYFF1LIVLIVUNUTUTIVINGUNKGUNVRILLSSGQIRMGINGARSINKKKAQUVEALAIPSDA Tmc6 (Mm) 664 PFRILNTUREAGTVNVRRLEHAG-SGASH.PHILHFLVENTYFF1LAUNTUNUTUGUNKURURURURURURURURURURURURURURURURURURU			330	CONTRACT AND A CONTRA
Tmc8 (Mm) 419			481	
Tmcl (Mm) 564 SYTEPDISGNVLALIFNQCHINNGSPFAPSLPGINILRLHTSNYPQCNAVNCCNVPEARVFKARSNNFYLGHLLLVESTNPVL-VHIVSLPPSPCG Tmc2 (Mm) 609 SYAEPDISGNVLGLIFNQCHINNGSPFAPSLPGINILRLHTSNYPQCNAVNCSNVPEARVFKARSNNFYLGHLLLVESLEVA-YTVNSLPPSPCG Tmc3 (Mm) 552 SYGEFALRAVLHLVINQUHINNGAPFSCLPAPNULKLIGHTLASNNFPGNARSNNFYLGLLLVESLEVA-YTVNSLPPSPCG Tmc5 (Mm) 552 SYGEFALRAVLHLVINQUHINNGAPFSCLPAPNULKLIGHTLASNNSTVENTASSNNFYLGLLLVESLEVALLET-FALVARLEPSLCG Tmc6 (Mm) 552 QRPEFDIARNVLBLINQUILNQUFSULTAVLUVINVSLGNNCOPPSKARASQNTFPILLCPSFLGAAVFLCVANVQVRPSSTCG Tmc4 (Mm) 663 PFDIARNVLBLINQUITNAULTVQUISPCLLPLINTARFLIFFILSNSTFFASTANTFPILVLLUGLAISNVFV-VISTSTESSNCG Tmc1 (Mm) 663 PFSUKRNIFEVIGETLEHDPFSNLALITKFVI FYVRELSLITCRPSPROFRASSNFPFHLVLLIGLAISNVFV-VISTSTSSNCG Tmc1 (Mm) 663 PFSUKRNIFEVIGETLEHDPFSNLALITARFVI IPTVRELSLITCRPSPROFRASSNFPFHLVLLIGLAISNVFL-VISTSTSSNCG Tmc1 (Mm) 663 PFSUKRNIFEVIGETLEHDPFSNLALITARFVI IPTVRELVINVUSLISSNFFFHANSANAPARSNEFFFHANKKNAAARAAAAAG Tmc3 (Mm) 663 PFSUKRNIFEVIGETLEHDPFSNLALISSNUTULINGUSKFFFHITTUNSTNSASPRFFHANSANAPURKKNAAAARAAAAAG Tmc5 (Mm) 661 PFSUKRNIFEVIGETLEHDPFSNLAKINVUNUSLSSNFFFHITTUNSTNSASPRFFASSTFFFHANKKNAAAARAAAAAA Tmc5 (Mm) 661 PFSUKRNIFEVIGETLEHDPFFLALISSNFFFFITTULITUNUSTNENSASPRFFFFFHITUNULLUKKKNKAAAANAAAAAAAAAAAAAAAAAAA				
Tmc2 (Mm) 608 SYAEFDISGNULGLIFNQGULUNGSYLAPGULGINVLELLISMUFPCCAUVERSNUPERNYEASRSNNPYLAULLILULFLSLLUVA-YTTMSLEPSFDCG Tmc3 (Mm) 552 EXGEFFISAENVLHUVNGGULUNGGPFSPCLPAPNVLKLIGINVLKSNAVLTCNVPHQVFRASRSNNPYLAULLFULFULCMLPT-FAIVHYKPSLUCG Tmc5 (Mm) 552 CKPEFDIANNVLELINGGILVAQUEPSCLPAPNVLKLIGINVLKSNAVLTCNVPHQVFRASRSNNPYLAULFULFUCMT-FAIVHYKPSLUCG Tmc6 (Mm) 582 CKPEFDIANNVLELINGGILVAQUEPSLPAVQILRLJFFHIKNSLAGARPOPPSKANRASQUITYFIFLUCHOLLSNVPL-VISTELIPSSTGULCIAITINKEN Tmc4 (Mm) 564 GILEGUVDEVLGUNGVEPSFCLELLINTKKFLIFFIVENUNSLAGNCQAPRAPHLASHNSTVFILLCPFSFLGAAVFLCYAVWQNPSSTGG Tmc4 (Mm) 663 QPEFDIANNVLGUNQVENUNSSFCCLELLINTKKFLIFFIVENUNSLAGNCQAPRAPHASHNSKFPFLVLLULGULAISNVPL-VISTENIPSKLCG Tmc1 (Mm) 663 PFS/GKNNVDVLGETIENDPFNFLORLINSVFLFTYVKKYLLINKSRAPRFPASSTFPFHLVLLIGULAISNVPL-VISTENSSTCG Tmc3 (Mm) 663 PFS/GKNNVDVLGETIENDPFNFLAXINFYUTVVKUTVINUNUNUNUNUNULSSGURAGIQALANNARARAAAAG Tmc5 (Mm) 661 PFS/GKNNVDVLGETIENDPFNFLAXINVUNUNUSUNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUN				
 Tmc3 (Mm) 552 EVGEFKIAENVLHUVENQOHTENGAPFSPELPAPVUKLIGLMYLAGUNELGUVENQOVEPAARSENNEYLAMLIFULFLELLPT-FAIVUKRESLIGG Tmc5 (Mm) 519 LO-FPITAENVLHUTVAQTENLAGUFPSCLPAPNUKLIGLMYLNEVSKIAGAQPREMIASBNSTVFITLLCPPSFTGVLC1.TITTINELKSASLGG Tmc6 (Mm) 582 QRPEFDIARNVLDLIVQQTLTHLGYLFSPLLPAVQILHLFLFHNKKASLAMNCQAPREMIASBNSTVFITLLCPPSFTGVLC1.TITTINELKSASLGG Tmc7 (Mm) 582 QRPEFDIARNVLDLIVQQTLTHLGYLFSPLLPAVQILHLTART.LIPCLKNTLFS ISSASHTFMASTANFPFLVLLUGLAISAVEVL-YSIFLIPPSKLGG Tmc7 (Mm) 582 QRPEFDIARNVLDUVQQTCTNUGHTSPFLLPLATTARV, IFVKELSLIVTCRFSFTGVLC1.BASINTSVFLLVLLOLLIAAVFVL-YSIFLIPPSKLGG Tmc8 (Mm) 469 DREEFLVFKNVLDIVAQTVTNMGLFYCPLLFLLNSVFLFLTFYIKKTYLLNSRASPRFPRQFASSSTFFFLVLLLGLLAAVFLA-YVISSTHSSKLGG Tmc1 (Mm) 663 PFSCKNRHFEVIGETLEHDPFSFNGATIANSPGLVIAVILVMVLTIYLN-ATAKGQFAANLDLKKKOKQQALBKNKRKKNAARAAAAAG Tmc3 (Mm) 618 PFSCKNRHDVULESTIENDPFFNGATIANGDYKIPTIATLANGULANAQUKANAQUVENLAITPSDA Tmc3 (Mm) 619 PFSCLSTINNVLDIVSTSTENOPFTNENVVCHISSPVTLPAULLHFULTVUL-SVSISLSTANAQURKKI-QUNERKKKNAQUVENLAITPSDA Tmc6 (Mm) 631 PFRLNTVEGATUVNRLEHAG-SGASULFHLHHIVSNTFFILASALLIAVITYNQVKQULVILLHSQINNGGKKNFLINKLKVAQUVENLAITPSDA Tmc7 (Mm) 631 PFRLNTVEGATUVNRLEHAG-SGASULFHLHHIVSNTFFILASALLIAVITYNVLOGISLSKKKNQVVSLALAITPSDA Tmc6 (Mm) 631 PFRLNTVEGATUVNRLEHAG-SGASULFLHHIVSNTFFILASALLIAVITYNVLOGISLSKKRQIENSGENKKLINKLAUAUSSKKKSQUVENLAINESSQITTENSKNGSTS Tmc7 (Mm) 566 LFTNYSAPWQVVPELVALQLPLPSQAILXVLSSRQAPFFHILTINLANVCHISSEKKRQPFH4KKVAQUVENKVFLQAVALASSRGST Tmc1 (Mm) 756 GQ* 757 Tmc2 (Mm) 798 EDTIKNSSKNATQIHLKEEGSSSQIQTLDKKKQQPHTSSTEGGASPSTSHHHVGSQPPRGRDSGQPQSQTYTGRSPSGKRTQRPH* 888 Tmc3 (Mm) 744 RQAGASTEAESSENSK KKLQARQTHESSSQIQTLDKKAQQPHTSSTEGGASPSTSHHHVGSQPPRGRDSGQPQSQTYTGRSPSGKRTQRPH* 888 Tmc6 (Mm) 661 PGSPHSRASR RSPCPGPPCPGSPGPTFRSSSSLGAPSAVASSRPHFPSRTEL* 722 <ll>Tmc3 (Mm) 847 EQDTNRHPRHRPCSSTSLHKKNSCSSVTQQPLKKVRSEPLSRKDSVY</ll>	Tmc1	(Mm)		
 Tmc5 (Mm) 519 LQ-EFDIARNULEIIYAQTITULGIFYCELLFFQMITLFLIFUFYVKNSLANNCQAPREWLASGMITYFYELLFPFSTGULCT AITIMELKPSALGG Tmc6 (Mm) 532 QKPEFDIARNULEIIYAQTITULGYLFSELPAVQITELFFFUKKSLANNCQAPREWLASHNSTYFILTLECPSFLGALVFLCTANNQVKSSTGG Tmc4 (Mm) 540 GTLEFQVPDEVLGLIYQQTITULGYLFSELPAVQITELFLFFIKKSLANNCQAPREWLASHNSTYFILTULGEFSFLGALVFLCTANNQVKSSTGG Tmc7 (Mm) 532 GQGEFAIPDNVLGIYGQTGUTGAPFSFLFAIATLKYIIFYVKSLLIXTCRFSPRGFASNSNFFFLVLLIGLLAIPLT-ISHARIPSKACG Tmc1 (Mm) 663 PFS/KNRHFEVIGGTLEHDPFSHVGLVIAVILYTYTIN-ATAKGQKAANLDLKKKOKQQALDKKRIKKAAAARAAAAG Tmc2 (Mm) 663 PFS/KNRHFEVIGGTLEHDPFSHVGKILLFLINSJSVETELFFYIKKYTLLNSSKSLGAANAQURKKI-QAVESKKKKAAAARAAAAG Tmc3 (Mm) 661 PFS/KNRHPEVIGGTLEHDPFSHVGKIDSVILPTYTIN-ATAKGQKAANLDLKKKOKQQALDKKKAAAARAAAAAG Tmc3 (Mm) 663 PFS/KNRHPEVIGGTLEHDPFSHVGKIDSVILPAVLLAFNSTKSSKSCG Tmc3 (Mm) 664 PFRINTERGETUNKKUTAARAAAAAG Tmc3 (Mm) 665 PFS/KNRHPEVIGGTLEHDPFFWFIGKIPFICTUTYINULIGUNGVKKITINTLN-SSKSLSSAANQURKKI-QAVESKIKSKKAAAVESKAAAAAAG Tmc3 (Mm) 847 EQDTNRHTKEAGTUNVRELENG-SGASULPHANLNKUTYITYTUVVTUTVILVNUTTINTLNVGTEGKKVKITELBKSKKUTAAVESKKAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA				
Tmc6 (Mm) 582 QKPEFDIARNVLDLIVGOVICHLAVLEYELLPAVQILELIFLETIKKASLANCQAPRAPHLASHNSTVPILLLCPESFLGAAVFLCXAVMQVRPSSTG Tmc4 (Mm) 504 GTLERQVPDEVLGIVXQCVIVMGDFPCLLLLIFLETIKKSLSLYTCRPSPRQFRASNSNFPFLLVLLGLLATSAVFL-YSIFLIPSKSLGG Tmc7 (Mm) 532 GQQEFAIPDNVLGIVXQCVIVMGLFYCLLFLLNSVFLFLTFYIKSYLLSSASTFRASASANFPFPLLVLLIGLLATSAVFL-YSIFLIPSKSLGG Tmc8 (Mm) 663 PFSGKRR!FEVIGETLEBDFPSM:AKILRQLSNPGLVIAVILVMVLTIYIN-ATA/GQLANLDLKKKKQQALENKKKMAARAAAAAG Tmc1 (Mm) 663 PFSGKRR!PEVIGETLEBDFPSM:AKILRQLSNPGLVIAVILVMVLTIYIN-ATA/GQLANLDLKKKKQQALENKKKMAARAAAAAAG Tmc2 (Mm) 661 PFS/GKRIPEVIGETLEBDFPSM:AKILRQLSNPGLVIAVILVMVLTIYIN-ATA/GQLANLDLKKKKQQALENKKKMAARAAAAAG Tmc5 (Mm) 661 PFS/GKRIPEVIGETLEBDFPSM:AKILRQLSNPGLVIAVILVMVLTIYIN-ATA/GQLANLDLKKKKQQALENKKKMAARAAAAAG Tmc5 (Mm) 618 PFS/GKRIPGVILWVINUTUNIGSWFFFILTUNIVIN-SVS/SLS/SNAQLCKKVLQVVENKSIKKKA/AQVVEALAIPSDA Tmc6 (Mm) 682 PFRILNTWEAGTVMVRLEHAG-SGASH.PHILHFILVENTFFILTUVIITYLN-SVS/SLS/SNAQLCKKVLLKQIRNEGEKKIFLELKLAQUAVKALAIPSDA Tmc6 (Mm) 631 PFRKLSI <wqvipelaie-slpqzaqmflv lgqatvpllissil="" mvtv-alancygliselkrqikvllaquavalssrgst<="" td=""> Tmc1 (Mm) 756 GQ* 757 Tmc2 (Mm) 756 GQ* 757 Tmc2 (Mm) 756 GQ* 757 Tmc3 (Mm) 744 RQAGSATEAESSENSK KTLQARIQTHEESSK_LLKDSDLISQLSSVMATSPNNGHMLNFDSLSSKSLAMAAITRSQSPGQGSRDFCSPLLDGSRSRP Tmc3 (Mm) 756 T</wqvipelaie-slpqzaqmflv>				
Tmc4 (Mm) 504 GTLSFQUPDEVLGLIVACYUWVGSFFCELLPLINTARF. ILFCLKKTIFLSISPASHTFPASTANFYFFLVLLVGLAISAVEVL-YSIFLIPPSKLGG Tmc7 (Mm) 532 GQQSFAIDDNYLGUYQQTICHIGAFYSPLLPAIATLKFYIIFYVKELSILYTCRPSPRQFASHSNHFYFLLVLLGLLAIYPL-YSIFLIPPSKLGG Tmc8 (Mm) 469 DREEFLVFRNULDIVAQTYCHIGAFYCLLIALSKYFLFTFIKKYTLLRNSRASPRAFASSSHFFFHLVLLGLLAAVPLA-YVISSTHSSKCG Tmc1 (Mm) 663 PFSCKNRKFEVIGETLEEDPFSN-AKILRQLSNPGLVIAVILVNULTIYLN-ATA/GQKAANLDLKKKURQQALENKRNKKMAAARAAAAG Tmc2 (Mm) 707 PFSCKNRKFEVIGETLEEDPFSN-AKILRQLSNPGLVIAVILVNULTIYLN-SVSSLSS-ANAQLARKKIQALREVEKNHKSIKGKAITYS Tmc3 (Mm) 651 PFSCKNRKPGVIGETLEEDPFSN-AKILRQLSNPGLVIAVILVNULTIYLN-SVSSLSS-ANAQLARKKIQALREVEKNHKSIKGKAITYS Tmc5 (Mm) 651 PFSCKNRKPGVINVVRULTSSVUTVOKISSSVUTJAVILVNUSSGQLENGKKKNQQVERLAINSPEDA Tmc5 (Mm) 651 PFSCKNRK				
 Tmc7 (Mm) 532 GQQEFAIPDNVLGIVYQQTCMIGAFFSPLLPAIATLKFYIFIYVKEJLLTCRPSPRQFAASNSNFPFLLVLLIGLCLAIPLA-ISNARIPSSKACG Tmc8 (Mm) 469 DREEFLVPKNVLDIVAAQTVTNNGLFYCPLLPLLNSVFLFTFYTKKYTLLTCRPSPRQFAASNSNFPFLLVLLIGLLAAVPLA-YVTSSTHSSHCCG Tmc1 (Mm) 663 PFSUKRNIFDVIGETLENDPFNFUKTILANDEDGIPTYTKKYTLANSRASPRPRASSSTFPFHLVLLGLLAAVPLA-YVTSSTHSSHCCG Tmc2 (Mm) 663 PFSUKRNIFDVIGETLENDPFNFUKTILANDEDGIPTYLN-SVSUSLSNANQURKKUKQQAL NKKNKKNAAARAAAAG Tmc5 (Mm) 661 PFSUKRNIFDVIGETLENDPFNFUKTIANSRYFFFLVLVITUG-SLASSAKKKQUKEKNKSIKKKNAARAAAAG Tmc5 (Mm) 618 PFSUKRNIFDVILWETTENDPFNFUKTIVNLONDEGIPTYLIN-SVSUSLSNANQURKKU-QUKEKNKSIKKKNAARAAAAG Tmc5 (Mm) 618 PFSUKRNIPGVLWUVNUTUNI.GSWFFFLTLVINUTGCRKVKUTURGURAVITUGCURAVITURGKUKKU-LENKUTURGKAKKU-LENKUTUNI.GSUSFFLTLNUTURGKUKKU-LENKUTURGKUKKU-LENKUTURGKUKKU-LENKUTUNI.GSUSFFLTLNUTURGKUKKU-LENKUTURGKUKKU-LENKUTURGKUKKU-LENKUTURGKUKKU-LENKUTUNI.GSUSFFLTLNUTURGKUKKU-LENKUTURGKUKKU-LENKUTURGKUKKU-LENKUTURGKUKKU-LENKUTURGKUKKU-LUKUTUNI.GSUSFFLTLNUTUGUTUKUTURGURKUTURGURKUTURGKUKKU-LUKUSTANGKUKU-LENKUTUNI.GSUSFFLTLNUTUGUTUKUTURGURKUTURGURKUTURGURKUTURGURKUTUNGKUKKU-LANGKUKU-LENKUTUNI.GSUSFFLTUNGUNUTUNI GUKUKUTURGURKUTURGURKUTURGURKUTURGURKUTURGURKUTURGURKUTUNI.GSUSFFLTUNITUS-LENKUTURGURKUTURGURKUTURGURKUTURGURKUTUNI.GSUSFFLTUNITUS-SISUSIANAUAGURKUTURGURKUTURGURKUTURGURKUTUNI.GSUSFFLTUNITUS-GUKUKUTURGURKUTURGURKUTURGURKUTUNI.GSUSFFLTUNITUS-GUKUKUTURGURKUTURGURKUTURGURKUTUNI.GSUSFFLTUNITUS-SISUSTANATURGURUKUTURGURKUTURGURKUTUNI.GSUSFFLTUNITUS-SISUSTANATUNITUSCUKUKUVUNUTUS-TPSSIGTITUSKITUTUG-SISUSANATURGURUNGULSLSSRIKUTURGURKUTUNI.GSUSFFLTUNITUS-SISUSANATUNITUSTUNITUS-SISUSANATUNITUSCUKUKUVUNUTUSLSSRUSTUNATUNUTUS-SISUSANATUNITUSCUKUKUVUNUTUSLSSRUSTUNATUNITUSCUKUKUVUNUTUSLSSRUSTUTUSTINTUSTUSTUSTUNATUNITUSTUNTUSTUNITUSTUNITUSTUNTUSTUNITUSTUSTUSTUSTUSTUNTUSTUNITUSTUNTUSTUNITUSTUSTUSTUSTUSTUSTUSTUSTUSTUSTUSTUSTUST				
Tmc8 (Mm) 469 DREEFLVPKNVLDIVAAQTVTHMGLFYCPLLPLLNSVFLFLTFYTKKYTLLRNSRASPRRPRASSSTFFFHLVLLLGLLAAVPLA-YTISSTHSSHCG Tmc1 (Mm) 663 PFSGKRRIFEVIGETLEHDPPSM-GKIFA-LANPGLIPALLSNPGLVIAVILTYYLN-ATAKGQKAANLDLKKKGKQQAL MKMRNKKMAAARAAAAG Tmc3 (Mm) 651 PFSGKRRIFEVIGETLEHDPPSM-GKIFA-LANPGLIPALLSNPGLVIAVILTYLN-SVSSLSFANQLGKKTQALREVEKHKSIKGKAIVTYS Tmc3 (Mm) 651 PFSGKRRIFEVIGETLEHDPPKF-GKIFA-LANPGLIPALLIPALLYLN-SVSSLSFANQLGKKTQALREVEKHKSIKGKAIVTYS Tmc5 (Mm) 618 PFRILPSFIGSTYSNIDTLSRRPGKLWVVIVQHISSPVUIPLAVLLIPALLYLN-SVSSLSFANQLGKKT-QARSEDKKKAQUVEALLIPSDA Tmc6 (Mm) 618 PFRILPSFIGSTYSNIDTLSRRPGKLWVVIVQNLGSVEFPF ITLJVIVG-SIAVSLKKSQGLKVICLLKGQINKSGUKKPLLEKLTKLQDMEKKV Tmc6 (Mm) 618 PFRILPSFIGSTYSNIDTLSRRPGKLWVVIVQNLGSVEFPF ITLJVITTV-GIGLKGUKVGLKVICLLKGQINKSGUKKPLLQUAVLSKKVFLGQUAVLSKKKVFLGQUAVLSKKVFLGQUAVLSKKKVFLGQUAVLSKKVFLGQUA				
<pre>Tmc1 (Mm) 663 PFSGKNRMFEVIGETLEHDPFSHGKILRQLSNPGLVIAVILVHVLTIYIN-ATAKGQKAANLDLKKKGKQQALBKKKHAAARAAAAAG Tmc2 (Mm) 707 PFSGKNRMYDVLHETIEHDPFSHGKIFAFLANPGLIIPAILLAPLAIYYLN-SVSGLSBAAAQLKKIQALREVEKNHKSIKGKAIVYS Tm3 (Mm) 661 PFSGQKIYDVLHETIEHDPFKPIGKIFAFLANPGLIIPAILLAPLAIYYLN-SVSGLSBAAAQLKKI-QALKKKGKQALBKKKAAQAVRALAIPSDA Tm55 (Mm) 618 PFRGLGFFIGSIKSVIDTISRRPGVLWVHVIGULGSVFFFILTUIVLGTTUTVNGTEGGKVVHILLNGGUNAGESDKKKAAQAVRALAIPSDA Tm65 (Mm) 618 PFRGLFSFIGSIKSVIDTISRRPGVLWVHVIGULGSVFFFFILTUIVLGTTUTVNGTEGGKVVHILLNGGUNAGESDKKKAAQAVRALAIPSDA Tm66 (Mm) 628 PFRLHTWYERAGTUHVRRLEHAG-SGASH_PHLHHT_VENTFFFFILTUIVLGTUTUVNITLLKEQTINSGGEKFFL INKLHSVEEEGGS Tm64 (Mm) 631 PFTNFNTMQUTPAIL-SLPGTAQHFYLTUGUAGYPHILLSSILAUTYFNIQVVGQLKVICLLSELKRGIESTSVQNKVPLQUAVALSSRGTS Tm67 (Mm) 631 PFTNFNT</pre>				
Tmc2 (Mm) 707 PPS%KRN#YDVLBETIENDPPFW[GKIFATLANPG_IIPALIYIN-SVS%SLS%ANAQLEKKIQALEVEKNHKSIKGKIIYIS Tmc3 (Mm) 651 PPS%QEKIYDVLSETIENDPPFW[GKIFATLANPG_IIPALIYIN-SVS%SLS%ANAQLEKKIQALEVEKNHKSIKGKIIYIS Tmc5 (Mm) 651 PPS%QEKIYDVLSETIENDPPFW[GKIFATLANPG_IIPALIYIN-SVS%SLS%ANAQLEKKI-QALARS%DKKK%AQNVEALAIPSDA Tmc5 (Mm) 661 PPRCLMPYEAGTWWRRLEHAG-SGASH_PKLHH-LVSNTFFFIITLIVL.ITTLYNQTWGORKVITLLERQIINSGEKKFLIEKLTKLQNEKKV Tmc4 (Mm) 603 PPRKIKSIWQIPQITS-SLPG7AQHVIRLEHAG-SGASH_PKLHH-LVSNTFFFIFIASLLAVITYNQVWGORKVICLISELKSQIESSKKFLIEKUTKQANAVASKROTS Tmc4 (Mm) 631 PPTKPKTWQIPQITS-SLPG7AQHVITLYLGCGATVPIJ_LILSSILMQTY-ALLANCYGALLSELKSQIESSKKKTLQKLTERQARVERSQ Tmc3 (Mm) 566 GQ* 757 Tmc3 (Mm) 766 TMSASKSKIQHLIKEEPTSHSSSQIQTLDKKAQGPHTSSTEGGASPSTSMHHVGSQPPRGRDSGPQSQTYTGRSPSGKRTQRPH* 888 Tmc3 (Mm) 716 NPSALDEREVEPQIPILIKEECAAPDLERKSAQEENPIA* 757 Tmc4 (Mm) * 694 Tmc7 (Mm) 731 RGRTQARTEPTAHAPGDARDERSKULKDSDLISQLSSVJAATSPNNGHUNFDSLSSKSLRMAAITRSQSPGQGSRDPCSPLLDGSRSRP Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNSCSSVTQPPLKDVRSEPLSRKDFPPSREL* 722				
 Tmc3 (Mm) 651 PFS/QERIYDIVSETIENDPTMEHAVVGHISSPVVILPAULLIPHLIYTQ-SIANSLKLSSQDLRMQIQNASELDKKKVAQVVERLAIFSDA Tmc5 (Mm) 618 PFRULPSPIQSIYSKIDTLSRRPGYLMVVHYQLIGSUFFFILITYLNYDITEGRKWHILLLEQIINEGUKKFLEKUKAQVVERLAIFSDA Tmc6 (Mm) 620 PFRULPSPIQSIYSKIDTLSRRPGYLMVVHYQLIGSUFFFILITYLNYDITEGRKWHILLLEQIINEGUKKFLEKUKAQVVERLAIFSDA Tmc6 (Mm) 621 PFRULPSPIQSIYSKIDTLSRRPGYLMVVHYQLIGSUFFFILITYLNYDITEGRKWHILLLEQIINEGUKKFLEKUKAQVVERLAIFSDA Tmc7 (Mm) 631 PFRUKLSIWQVPELIELS-LPQTAQWFIXT/GQATYPILILSIINMYTY-ALANCYGLISELSRKWFILSSRKWFLAQWAVLSSRKGTS Tmc7 (Mm) 756 GV*757 Tmc2 (Mm) 756 GQ*757 Tmc3 (Mm) 744 RQAGSATEAESSENSK KTLQARIQTHEESSK LLKKSDISSTEGGASPSTSWHHVGSQPPRGRDSGQPQSGYTYGRSPSGKRTQRPH* 888 Tmc4 (Mm) 744 RQAGSATEAESSENSK KTLQARIQTHEESSK LLKKSDISQLSVMATSPNNGHLMFDSLSSKSLMMAAITRSGSPGQGSRDPCSPLLDGSRSRP Tmc5 (Mm) 718 RPGRTQATEPTAWHEDGGDQEVPISS*800 Tmc6 (Mm) 611 PGSPHSRARTRSCPFPLSTBSSSLGKTPPFS*810 Tmc6 (Mm) 624 Tmc7 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNSSSSTTOPPESSTEGGRSPSSSSLGAPSASVPASRPHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPRPRCSSTSNLHKNSSSSTTOPPLKDVRSEPLSRKDPQPISPPCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVESPUSSFEDQLERPFFVBRRP 	Tmc1	(Mm)		
Tmc5 (Mm) 618 PFRILPSFIQSIYSHIDTLSRRPGLUVVHIQNLIGSVEFFF ILT.IV.IT.TT.WQITEGREVVHIRLHEQIINEGGIKHFLIEKLTKLQDMEKKV Tmc6 (Mm) 682 PFRIKLSI MILLEAG-SGASH_PHILHFLVENTFFF ILSSILAUIYINGVKGQEKVICLLKEQIRNEGGIKIFI INKLKSVEEEGRS Tmc4 (Mm) 633 PFRKKLSI MILLEAG-SGASH_PHILHFLVENTFFF ILSSILAUIYINGVKGQEKVICLLKEQIRNEGGIKIFI INKLKSVEEEGRS Tmc4 (Mm) 631 PFRKKLSI MILLEAG-SGASH_PHILHELVENTFFF ILSSILAUIYINGVKGQEKVICLLKEQIRNEGGIKIFI INKLKSVEEEGRS Tmc7 (Mm) 631 PFRKKLSI MVDIFEAGUTMKVLGULSERGEKKVILGKLTEAGREKVEGGISS Tmc7 (Mm) 631 PFRKKSI MVDIFEGREVKULVGULSERGEKKVILGKLTEAGREKVEGGISS Tmc7 (Mm) 756 GG* 757 Tmc2 (Mm) 756 Tmc5 (Mm) 766 GG* 757 Tmc6 (Mm) 761 RPGRTQATCHTEESSKULLKOGDISGLSSVHAATSPNNGHLNFDSLSSKSLRMEAITRSQSPGQSRDPCSPLLDGSRSRP Tmc5 (Mm) 761 RPGRTQATCHTEAGUTMKEGGIGKEPFN* 810 11 11 Tmc6 (Mm) 781 RPGRTQATCHTEAGUTMKEGGIGKEPFN* 810 11 123 ASA* 726 Tmc6 (Mm) 661 PGSPHSRASR*RSFCPGFPCPGSPGPTFRELSSSSLGAPSASVPASRFHFPSRTEL* 722 124 Tmc3 (Mm) 847 EQDTNRHPHRPCSSSTSNLHKNRSCSSVTQTQPLKLVRESPLSRKDF	Tmc2	(Mm)		
Tmc6 (Mm) 682 PPRCINTYFERAGTWAVRRLEHAG-SGASHLPHLHHTLVENTPFLFLASLLAVITFNIQVVROGMAVICLERGIENGEGEKTFLINKLHSVEEEGGS Tmc4 (Mm) 603 PPRINKISINAQIPEATE-SLPGTAGHTVJTUGGARVYVFLILDSILHAVITY-ALANCYGLISELKEGIENGEKTFLINKLHSVEEEGGS Tmc7 (Mm) 631 PPTNFNTNQVPELVALQLPLPSGRALRISSHAPSPFLILLSIVJTYV-ALANCYGLISELKEGIENSKERVLIGKLTEAQREVRSQ Tmc8 (Mm) 568 LFTMYSAPNQVVPELVALQLPLPSGRALRISSHAPSPFLILLSIVLTVCI-SGS*ANAFAIQGLRKQLVHQVQLKNHLVDDLSRLLPELSPE Tmc1 (Mm) 756 GQ* 757 Tmc2 (Mm) 798 EDTIKNSSKNATQIHLIKEEPTSHSSSQIQTLDKKAQGPHTSSTEGGASPSTSWHHVGSQPPRGRDSGQPQSQTTTGRSPSGKRTQRPH* 888 Tmc3 (Mm) 744 RQAGSATEAESSENSK KKILQARIQTHESSKLLKDSDLISQLSVMATSPNNGHHLNFDSLSSKSLRMEAITRSQSPGQGSRDPCSPLLDGSRSRP Tmc4 (Mm) * 694 Tmc7 (Mm) 611 PGSPHSRASR"RSPCPGSPCPGSPGPCFGSPGSVSLTMDASSKSLGAPSASVPASRPHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQOPLKDVRSEPLSRKDFQFISPFCGSGVSTLTHDHSPRARYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQVMSVVPNQPCSPUSEEEEELRRDLVQMSIFASSLGAPSASVPASRTHFPSRTEL* 722				
Tmc4 (Mm) 603 PPRIKLSIWQUPEATE-SLPQTAQMFLYLGTQATTUPLILSSILMMTY-ALANCYGHLISELKRQIETEVQMKVFLAQIAVALSSRNGTS Tmc7 (Mm) 631 PPTNRTTWQUPPCTS-TPPSSLCULTANTSERA-WPFFMIICL/WFYT-ALANCYGHLISELSSRNKTLGLISSRNKFULGKIETEVQMKVFLAQIAVALSSRNGTS Tmc8 (Mm) 568 LPTNYSAPWQUVPELVALQLPLPSQRALRYLSSHAPSFPILTLSIVLTVCI-SQSRANAFAIQGLRKQLVMQVQLKWHLVDLSRLPELSPE Tmc1 (Mm) 756 GQ* 757 Tmc2 (Mm) 798 EDTIKNSSKNAQIHLTKEEPTSHSSSQQIQTLDKKAQGPHTSSTEGGASPSTSWHHUGSQPPRGRRDSGQPQSQTYTGRSPSGKRTQRPH* 888 Tmc3 (Mm) 744 RAGASATEAESSENSK/KTLQARIQTHEESSKLLKDSDLISQLSSVMATSPNNGHLNFDSLSSKSLMMEAITRSQSPGQGSRDPCSPLLDGSRSRP Tmc5 (Mm) 716 NFSALDLERREVEPQIT/HLEELGAPDLRERSQEEPFIA* 757 Tmc6 (Mm) 781 RGRTQDATEPFAMHEDGGDKEPCPPISF* 810 Tmc7 (Mm) 723 PASA* 726 Tmc8 (Mm) 661 PGSPHSRASR'RSPCPGSPGPGSPGPFTPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQPLKDVRSEPLSRKDPQDFISPPFCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKLDAL GDIVELYPRNVQQVMSVVPNQPCSPOLSEEEEELRRDLVQMSIFMSDFLSPLSCFTTGDSNNTHDPKYQRRDVYNERDSHKKTHRAFWPERHFKLDAL GDIVELYPRNVQQVMSVVPNQPCSPOLSEEEEELRRDLVQMSIFMSESCFTTGDSNNTHDPKYQRSDNSFEDGLERPFTYURRP	Tmc5	(Mm)	618	PFRGLPSFIQSIISWIDTLSRRPGXLWVVVIYQNLIGSVHPFFILTIVLIITYLYWQITEGR«VMIRLLHEQIINEGKOKMFLIEKLTKLQDMEKRV
<pre>Tmc7 (Mm) 631 PFTHFNTTWEVIPQTVS-TFPSSLQTLIHAVTSEAFAVPFFHIICL/WFYFI-ALAGAH/QVAAQLREQLSLESROKRYLLQKLTEAQREVRSQ Tmc8 (Mm) 568 LFTNYSAPKQVVPELVALQLPLPSQRALRYLSSHAFSFPLILLSIVLTVCI-SQSsAAABAIQGLRKQLVMQVQLKWHLVDDLSRLLPELSPE Tmc1 (Mm) 756 GQ* 757 Tmc2 (Mm) 798 EDTIKNSSKNATQIHLTKEEPTSHSSSQIQTLDKKAQGPHTSSTEGGASPSTSWHHVGSQPPRGRRDSGQPQSQTYTGRSPSGKRTQRPH* 888 Tmc3 (Mm) 744 RQACSATEAESSENSK KTLQARIQTHEESSKLLKOSDLISQLSSVTMATSPNNGHMLNFDSLSSKSLRMEAITRSQSPGQGSRDPCSPLLDGSRSRP Tmc5 (Mm) 716 NFSALDLERREVEPTOI FULLEELGAAPDLRLR#SQDEENPIA* 757 Tmc6 (Mm) 781 RPGRTQDATEPFANHEDGGDQKEPCNPISF* 810 Tmc4 (Mm) * 694 Tmc7 (Mm) 661 PGSPHSRASR*RSFCPGFPCPGSPGPTTPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQTQPLKDVRSEPLSRKDFQPISPPCGSGVSTLTHDHSPRAPRYYVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSVVPNQPCSPQLSEEEEMLRRDLVQMSIFASSLTDLPRSSCFTTGDRSNNTRDFKYQRRVYYRSGDNSFEDQLERPTFVRRP</pre>	Tmc 6	(Mm)	682	PFRINTELEGYWYKKLEHAG-SGASWLFWLHIFLVENTFFLFLASALLGAVIFNIQVVKGQMKVICLLKEQIRNEGEKIFLINKLESVYEEEGRS
Tmc8 (Mm) 568 LFTNYSAPWQVVPELVALQLPLPSQRALRYLSSHAPSPPLITLISTVLTVCI-SQSFANAMAIQGLRKQLVWQVQEKMHLVDDLSRLLPELSPE Tmc1 (Mm) 756 GQ* 757 Tmc3 (Mm) 798 EDTKNSSKNATQIHLTKEEPTSHSSSQIQTLDKKAQGPHTSSTEGGASPSTSNHHVGSQPPAGRRDSGQPQSQTYTGGSPSGKRTQRPH* 888 Tmc3 (Mm) 744 RQAGSATEAESSENSK KTLQARIQTHEESSK LLKDSDLISQLSSVYMATSPNNGHMLNFDSLSSKSLRMEAITRSQSPGQGSRDPCSPLLDGSRSRP Tmc5 (Mm) 744 RQAGSATEAESSENSK KTLQARIQTHEESSK LLKDSDLISQLSSVYMATSPNNGHMLNFDSLSSKSLRMEAITRSQSPGQGSRDPCSPLLDGSRSRP Tmc6 (Mm) 781 RPGRTQATEPFAMHEDGGDQKEPCNPISP* 810 Tmc7 (Mm) * 694 Tmc8 (Mm) 661 PGSPHSRASR RSPCPGFPCPGSPGPTPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQTQPLKDVRSEPLSRKDFQPISPPCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQVMSVVPNQPCSPQLSEEEEMLRRDLVQMSIPASSLTDLPRSSCFTTGDRSNNTRDPKYQRAVYYRGDSNSFEDQLERPTFVBRKP	Tmc4	(Mm)	621	PPKKALSI
Tmc1 (Mm) 756 GQ* 757 Tmc2 (Mm) 798 EDTIKNSSKNATQIHLTKEEPTSHSSSQIQTLDKKAQGPHTSSTEGGASPSTSWHHVGSQPPRGRRDSGQPQSQTYTGRSPSGKRTQRPH* 888 Tmc3 (Mm) 744 RQAGSATEAESSENSK KTLQARQTHESSKLLKOBOLISQUSSYVATATSPNNGHLNFDSLSSKSLRMEAITRSQSPGQGSRDPCSPLLDGSRSRP Tmc5 (Mm) 716 NFSALDLERREVERD[TLKLEELGAAPDLRLR+SAQDENPIA* 757 Tmc6 (Mm) 781 RFGRTQDATEPFAWHEDGGDQKEPCNPISP* 810 Tmc7 (Mm) * 694 Tmc8 (Mm) 661 PGSPHSRASRFRSCPGFPCPGSPGPTTPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQPLKDVRSEPLSRKDFQPISPPCGSGVSTLTHDHSPRAPRYYVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEMLRDLVQMSIFASSLTDLPRSSCFYTGDGSNNTRDPKYQRAVYYRSGDNSFEDQLERPTFVBRKP	Tmc8	(Mm)	568	LPTMYSAPWOVVPELVALOLPLPSORALRYLSBARATOBARSPILLLSTVLTVCI-SOSEANA ATOCIA KOLWOVOVALUKANI. UDAL DEL SALADE
Tmc2 (Mm) 798 EDTIKNSSKNATQ:HLTKEEPTSHSSQ1QTLDKKAQCPHTSSTEGGASPSSWHHUGSQPPGGRDSGQPQSQTTTGRSPSGKHTQRP#* 888 Tmc3 (Mm) 744 RQAGSATEAESSENSK*KLQARIQTHESSK*LLKDSDLISQLSSVYMATSPNNGHMLNFDSLSSKSLRMEAITRSQSPQGGRDCSPLLDGSRSRP Tmc5 (Mm) 716 NFSALDLERREVEPQ1"LILEELGAAPDLRLR:SQEENF1A* 757 Tmc6 (Mm) 781 RFGFNQDATEP*ANHEDGGDQKEPCNP:SP* 810 Tmc7 (Mm) 723 PASA* 726 Tmc8 (Mm) 661 PGSPHSRASR*RSFCPGFPCPGSPGP*TPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQTQPLKDVRSEPLSRKDFQPISPPCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEEMLRRDLVQMSIPASSLTDLPRSSCFYTGDRSNNTRDPKYQRAVYYRSGONSFEDQLERPTFVRRP		()		
Tmc2 (Mm) 798 EDTIKNSSKNATQ:HLTKEEPTSHSSQ1QTLDKKAQCPHTSSTEGGASPSSWHHUGSQPPGGRDSGQPQSQTTTGRSPSGKHTQRP#* 888 Tmc3 (Mm) 744 RQAGSATEAESSENSK*KLQARIQTHESSK*LLKDSDLISQLSSVYMATSPNNGHMLNFDSLSSKSLRMEAITRSQSPQGGRDCSPLLDGSRSRP Tmc5 (Mm) 716 NFSALDLERREVEPQ1"LILEELGAAPDLRLR:SQEENF1A* 757 Tmc6 (Mm) 781 RFGFNQDATEP*ANHEDGGDQKEPCNP:SP* 810 Tmc7 (Mm) 723 PASA* 726 Tmc8 (Mm) 661 PGSPHSRASR*RSFCPGFPCPGSPGP*TPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQTQPLKDVRSEPLSRKDFQPISPPCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEEMLRRDLVQMSIPASSLTDLPRSSCFYTGDRSNNTRDPKYQRAVYYRSGONSFEDQLERPTFVRRP				
Tmc5 (Mm) 716 NPSALDLERREVEPQIFLELELGAAPDLRLRESAQEENPIA* 757 Tmc5 (Mm) 716 NPSALDLERREVEPQIFLELELGAAPDLRLRESAQEENPIA* 757 Tmc6 (Mm) 781 RPGRTQDATEPFANHEDGGDQKEPCNPRSP* 810 Tmc7 (Mm) 723 PASA* 726 Tmc8 (Mm) 661 PGSPHSRASRFRSFCPGFPCPGSPGPFTPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQTQPLKDVRSEPLSRKDFQPISPPFCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELTPRNVQQYMSWVPNQPCSPQLSEEEEMLRRDLVQMSIPASSLTDLPRSSCFTTGDRSNNTRDFKYQRAVYYRGGONSFEDQLERPTFVBRAP	Tmc2	(Mm)	798	EDTIKNSSKNATQIHLTKEEPTSHSSSQIQTLDKKAQGPHTSSTEGGASPSTSWHHVGSQPPRGRRDSGQPQSQTYTGRSPSGKRTQRPH* 888
Tmc5 (Mm) 781 RPGRTQDATEPFANHEDGGDQXEPCNP#SP* 810 Tmc4 (Mm) * 694 Tmc7 (Mm) 723 PASA* 726 Tmc8 (Mm) 661 PGSPHSRASR#RSFCPGFPCPGSPGP*TPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQPLKDVRSEPLSRKDFQPISPPFCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEEMLRADLVQMSIPASSLTDLPRSSCFTTGDRSNNTADFKYQRAVYYRSGONSFEDQLERPTFVBRAP				
Tmc3 (Mm) * 694 Tmc7 (Mm) 723 PASA* 726 Tmc8 (Mm) 661 PGSPHSRASR RSPCPGFPCPGSPGP TPRLAPSNRLSSSSLGAPSASVPASRPHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQPLKDVRSEPLSRKDFQPISPPFCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSKVPNQPCSPQLSEEEEEMLRADLVQMSIPASSLTDLPRSSCFTTGDRSNNTRDPKYQRRVYYRSGDNSFEDQLERPTFVBRKP				
Tmc7 (Mm) 723 PASA* 726 Tmc8 (Mm) 661 PGSPHSRASR RSFCPGFPCPGSPGP TPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQTQPLKDVRSEPLSRKDFQPISPPFCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSNVPNQPCSPQLSEEEEELRRDLVQMSIPASSLTDLPRSSCFYTGDRSNNTRDPKYQRAVYYRGGDNSFEDQLERPTFVBRKP			781	RFGKTQDATEF#AWHEDGGDQKEPCNP#SF* 810
Tmc8 (Mm) 661 PGSPHSRASR*RSPCPGFPCPGSPGP*TPRLAPSNRLSSSSLGAPSASVPASRFHFPSRTEL* 722 Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQPLKDVRSEPLSRKDFQPISPPFCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEEMLRADLVQMSIPASSLTDLPRSSCFYTGDRSNNTRDPKYQRAVYYRSGDNSFEDQLERPTFVBRKP			700	- 074 DACA + 776
Tmc3 (Mm) 847 EQDTNRHPHRPCSSTSNLHKNRSCSSVTQTQPLKDVRSEPLSRKDFQPISPPFCGSGVSTLTHDHSPRAPRYYVVNERDSHKKTHRAFWPERHFKIDAL GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEEMLRRDLVQWSIPASSLTDLPRSSCFYTGDRSNNTRDPKYQRRVYYRSGDNSFEDQLERPTFVHRKP				
GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEEMLRRDLVQWSIPASSLTDLPRSSCFYTGDRSNNTRDPKYQRRVYYRSGDNSFEDQLERPTFVHRKP	1000	(man)	001	
	Tmc 3	(Mm)	847	GDIVELYPRNVQQYMSWVPNQPCSPQLSEEEEEMLRRDLVQWSIPASSLTDLPRSSCFYTGDRSNNTRDPKYQRRVYYRSGDNSFEDQLERPTFVHRKP

Figure I

The murine TMC protein family Murine (Mm) TMC amino acid sequences were aligned using the MultAlin sequence comparison algorithm. Positions that are conserved (identical or similar) in all three subfamilies are marked orange (present in at least 4 of the 8 sequences) and red (conserved in at least 6 of the 8 sequences). Conserved positions within individual TMC protein subfamilies are depicted in blue. Amino acids that were considered similar based on the physicochemical features of their side chains: L, I, V and M; F and Y; E, D, N and Q; T and S; K and R. The TMC motif is denoted with a green bar. Presumptive transmembrane regions are underlined. Indicated with bold letters is experimentally verified TMC sequence.

Gene	Chromosome (strand) or framework	GenBank Accession number
Mm Tmc1 19(-)	AF417579, AY263155	
Mm Tmc2	2 (+)	AF417581, AY263156
Mm Tmc3	7 (+)	AY4263157
Mm <i>Tmc</i> 5	7 (+)	AY4263159
Mm <i>Tmc</i> 6	II (-)	AY4263158
Mm Tmc4	7 (-)	AY4263162
Mm Tmc7	7 (-)	AY4263161
Mm <i>Tm</i> c8	(+)	AY4263160
Hs TMCI	9q21.13 (+)	AF417578
Hs TMC2	20p13 (+)	AF417580
Hs TMC3	l 5q23 (+)	AY4263163
Hs TMC5	16p12.3 (+)	AY4263164
Hs TMC6 (EVERI)	17q25.3 (-)	AY057379, AY099356
Hs TMC4	19q13.42 (-)	AY4263166
Hs TMC7	16p12.3 (+)	AY4263165
Hs TMC8 (EVER2)	17q25.3 (+)	AY057380, AY099358
Fr Tmc2-rs1	S2870	AY4263167
Fr Tmc2-rs2	SI4II	AY4263168
Fr Tmc3	S2699	AY4263169
Fr Tmc5	S1110	AY4263172
Fr Tmc6-rs1	S1355	AY4263170
Fr Tmc6-rs2	S2029	AY4263171
Fr Tmc4	S3252	AY4263173
Fr Tmc7	S1110	AY4263174
Ce TmcAh l	B0416.1	NM_077168
Ce TmcAh2	T13G4.3	NM_075820
Ag TmcAh	2	AY4263175
Ag TmcBh	Х	AY4263176
Ag TmcCh	3	AY4263177
Dm TmcAh	3L (CG3280)	NM_140079

Table I: TMC genes

We did not detect Tmc8 transcripts in any other organs investigated.

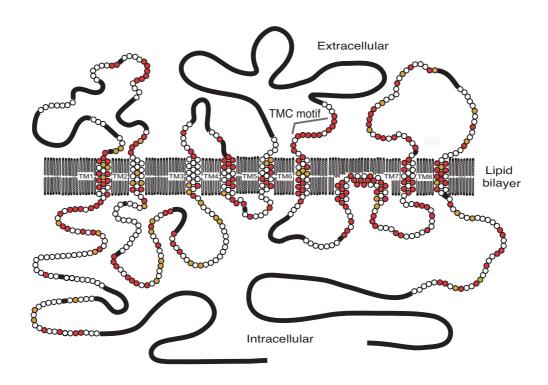
Our expression analysis results are corroborated by an analysis of applicable TMC ESTs obtained from GenBank <u>http://www.ncbi.nlm.nih.gov/dbEST/</u> (Figure 5).

Non-mammalian vertebrate and invertebrate TMC genes and proteins

The high degree of similarity between the corresponding human and murine TMCs suggests a conserved function of these proteins. This role of TMC proteins may also be conserved among other species. We therefore decided to investigate the TMC genes of other vertebrates and invertebrates.

We identified eight TMC loci in the Japanese pufferfish (Torafugu, *Fugu rubripes*, Fr) genomic database <u>http://fugu.hgmp.mrc.ac.uk/</u>. Whereas the genome of *Fugu rubripes* is one order of magnitude shorter than the length of the human genome, the total number of genes is estimated to be approximately the same [8,9]. Because of the

low homology of amino- and carboxyl-terminal sequences, we were not able to determine the complete coding sequences of the eight pufferfish TMC proteins unequivocally; nevertheless, we obtained sufficient sequence information of the central parts of the proteins bearing the transmembrane domains to classify the eight pufferfish TMCs into the three subfamilies. This subfamily assignment of individual TMC proteins was further substantiated by an analysis of the degree of conservation of intron positions within the Fugu rubripes TMC gene family (Figure 3B). The Fugu rubripes genome contains three TMC subfamilyA genes Tmc2-rs1 (Tmc2-related sequence1), Tmc2-rs2 (Tmc2-related sequence2), Tmc3, three TMC subfamilyB genes Tmc5, Tmc6-rs1 (Tmc6-related sequence1), Tmc6-rs2 (Tmc6-related sequence2), and two subfamilyC genes Tmc4 and Tmc7 (Table 1). The nomenclature of the pufferfish TMC genes and proteins is derived from the phylogenetic relation of the corresponding sequences with the mammalian TMCs (Figure 4). It is interesting that the pufferfish genome lacks orthologues of mammalian TMC1 and TMC8. Fugu rubripes Tmc5 and Tmc7 are



Proposed structure of TMC proteins Schematic illustration of the presumptive TMC protein topology. Each circle represents an amino acid. Amino acid positions conserved in at least 50% and 75% of murine TMC proteins are shaded orange and red, respectively. Black lines represent regions of variable length within the TMC family. The position of the TMC signature sequence motif is indicated.

clustered, equivalent to the clustering of their mammalian orthologues [7].

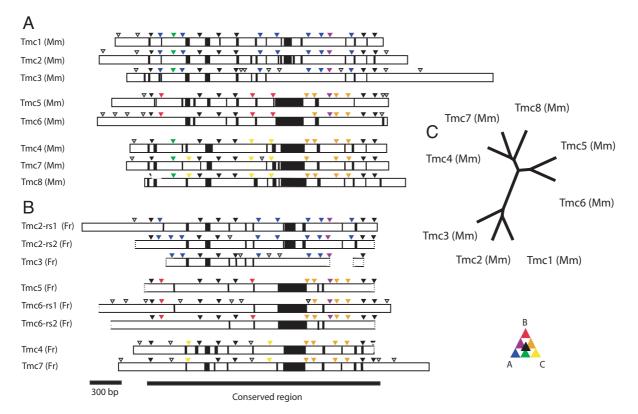
TMC genes also exist in invertebrates. In GenBank, we identified two mRNA sequences encoding *Caenorhabditis elegans* (Ce) TMC proteins. These mRNAs are transcribed from *TmcAh1* (*Tmc subfamily A homologue1*) and *TmcAh2* (*Tmc subfamily A homologue2*) (Table 1).

Whereas the *C. elegans* genome appears to lack TMC subfamily B and C genes, some insects have genes for all three TMC subfamilies. For example, the mosquito *Anopheles gambiae* has three TMC genes, *TmcAh* (*Tmc subfamily A homologue*), *TmcBh* (*Tmc subfamily B homologue*), and *TmcCh* (*Tmc subfamily C homologue*) (Figure 4, Table 1). We did not find deposited cDNA *TmcAh* and *TmcCh*, but *TmcBh* appears to be transcribed (ESTs BM645887, BM621478, BM605758, and BM636384). A search of the *Drosophila melanogaster* genome database revealed only a single TMC gene, *TmcAh*, (Figure 4, Table 1).

We did not find evidence for TMC genes in genomes and cDNA databases of yeast and plants.

Discussion

In an effort to define a novel gene family, we set out to identify genes related to *TMC1*, *TMC2*, *EVER1*, and *EVER2* [5–7]. We obtained the coding sequences of additional homologues in human and mouse, which form the TMC protein family (Figures 1, 2). We subdivided the TMC protein family into three subcategories A, B, and C. This subfamily-classification is based on two major observations. First, the sequence homologies among the different TMC protein sequences of individual species' do cluster into three groups (Figures 1, 4, and additional file 1). Second, our analysis of the organization of vertebrate TMC

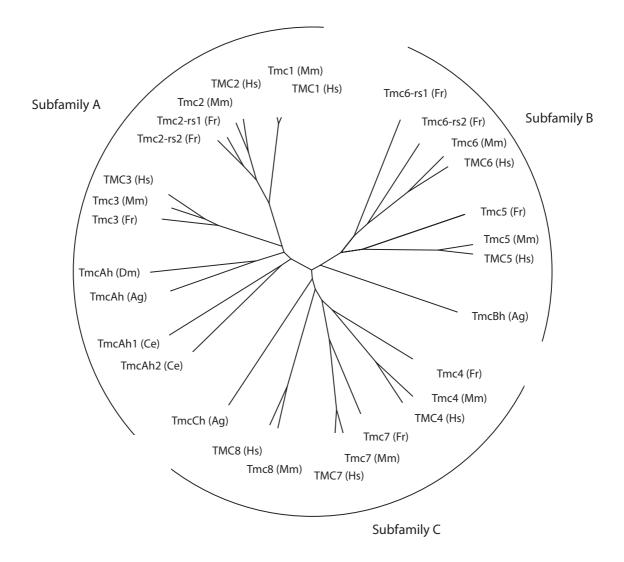


Comparison of murine and pufferfish TMC gene structures Comparison of murine (Mm)(A) and pufferfish (Fr)(B) TMC gene structures. To compare and visualize the gene structures of TMC genes, the predicted coding sequences (depicted here with white boxes) were aligned. Gaps were introduced when required for appropriate comparison (illustrated as black boxes). Triangles mark positions where the coding sequence is interrupted by introns in the genome (intron locations). The triangles are colour-coded to reflect the degree of conservation of the intron locations among TMC subfamilies A, B, and C. Positions of introns conserved in all members of the TMC family are denoted by black triangles; blue, red and yellow triangles label positions of introns conserved within individual TMC subfamilies; orange, magenta and green triangles denote intron positions shared by members of two subfamilies; intron positions unique to a single family member are indicated by open black triangles. (C) Phylogenetic tree drawn based on the similarity of the gene structures of murine TMC genes restricted to the conserved region of the TMC proteins. The corresponding matrix of distances is shown in additional file I.

genes implied identical allocation of the individual TMCs (Figure 3 and additional file 1).

On the basis of comparative structural predictions of all mammalian TMCs we propose a basic topology of the generic TMC protein with variable intracellular aminoand carboxyl-termini, a conserved core with eight membrane-spanning domains, and some variability in the length of several intra- and extracellular loops (Figure 2). This prediction further refines the previously published TMC protein features [5,7]. The structural presumptions for the region flanked by TM6 and TM7 were somewhat ambiguous, despite the region's high conservation among all TMC proteins with a high proportion of apolar amino acids (Figure 1). We hypothesize that it is unlikely that some members of the TMC family display an atypical topology, thus we propose that the lipophilic intracellular loop between TM6 and TM7 bears some flexibility, which may enable this domain to integrate into the inner surface of the phospholipid bilayer or even to pass through the plasmamembrane (Figure 2).

The high amino acid sequence conservation of 75–96% identity among the individual human and mouse TMC proteins implies that mutations that alter the coding sequence in the corresponding genes are subjected to significant selective pressure; thus advocating that TMC proteins have important cellular roles. Mutations in *TMC1* are responsible for the autosomal dominant human hearing disorder DFNA36 and for the recessive



Evolutionary relationship of vertebrate and invertebrate TMC proteins Phylogenetic tree, based on comparison of TMC protein sequences, illustrating the evolutionary relationship of TMC proteins of *Homo sapiens* (Hs), *Mus musculus* (Mm), *Fugu rubripes* (Fr), *Drosophila melanogaster* (Dm), *Anopheles gambiae* (Ag), and *Caenorhabditis elegans* (Ce).

	Cortex Cerebellum	Pituitary gland	Eye	Cochlea	Salivary gland	Small intestine	Colon	Lung	Thymus	Liver	Kidney	Urinary bladder	Testis	Prostate	Ovary	Heart	No. of cycles
Tmc1	accord and	-	-												-		32 42
Tmc2													Enny			1	32 42
Tmc3			-	(Internet	- Sector				-	ere Bee	-		-		-	nin. Mene	32 42
Tmc5		-	-	-	-	_	E	-		-	-	6			-	-	32 42
Tmc6			-				-	-						-		-	32 42
Tmc4		-				-						-					32 42
Tmc7				- Second			-	-		-	-	-	Ē		-		32 42
Tmc8	Bionara							and a									32 42
GAPDH			•	ŀ	-	-	ŀ	-	-	-	-	-	-	-	•	-	25

TMC expression analysis RT-PCR analysis of expression of TMC family members in mouse organs as judged by 32 and 42 cycles of amplification. Expression analysis of the ubiquitously expressed glyceraldehyde 3-phosphate dehydrogenase (GAPDH) is shown for reference. Dots indicate that one or two (small dot) or three and more (large dots) ESTs representing the respective organ are present in public databases.

deafness DFNB7/11 [5]; the corresponding murine gene *Tmc1* also causes deafness in the mutant mouse strains *Beethoven* (*Bth*) and *deafness* (*dn*) [6]. *Bth/Bth* mice display altered potassium currents in early postnatal cochlear hair cells, in particular Ik,n and Ik,f currents appear to be depressed [10]. This observation led to the hypothesis that Tmc1 may participate directly or indirectly in regulating the permeability of potassium channels [10]. Because of the high level of conservation of the amino acid sequences among the eight mammalian TMC proteins (Figure 1, Additional file 1), we speculate that other TMC proteins may as well be modifiers of ion channels or transporters.

Our assembly of TMC protein sequences allowed us to analyze their phylogenetic relationships (Figure 4). Because our analysis did not reveal any TMC genes in yeast and plants, we conclude that TMC proteins are a specific trait of animals. The *C. elegans* genome harbors two genes encoding presumptive TMC subfamily A-like proteins, both of which branch off close to the center of the phylogenetic tree, which may indicate similarity with the primordial TMC protein sequence. It is interesting that some of the intron locations of the two *C. elegans* genes are conserved when compared with vertebrate TMC genes, in particular with the TMC subfamily A (Additional file 3).

It is likely that the TMC family has diversified into three subfamilies before the *Protostomia* and the *Deuterostomia* diverged because the genome of the mosquito *Anopheles gambiae* contains one putative member of each TMC subfamily. Interesting in this regard is that the fruit fly's genome only has a single TMC subfamily A-like gene. The vertebrate TMC gene family is diversified, with each subfamily represented by multiple members.

Conclusions

The recently identified genes encoding the cochlear transcripts for TMC1 and TMC2, and encoding the proteins EVER1 and EVER2, belong to a novel gene family that is conserved in animals. The TMC protein family has eight members in vertebrates and forms three subfamilies, A, B, and C. TMCs are proteins with a conserved core of eight membrane-spanning domains. Most murine TMC genes are widely expressed at relatively low transcription levels; some TMC genes, however, display more restricted expression patterns.

Methods

Database-aided and experimental assessment of TMC cDNA and protein sequences

We created and continuously refined contigs of members of novel TMC cDNAs by aligning fragmented sequence information obtained from public and commercial EST databases (GenBank/NCBI – <u>http://</u> www.ncbi.nlm.nih.gov, Celera Discovery System - http:// www.celeradiscoverysystem.com). We started this procedure by using sequence fragments homologous to the known mammalian TMC proteins. Individual fragments were assembled to longer contigs using multiple such as ClustalW [11] http:// alignment tools www.es.embnet.org/Doc/phylodendron/clustalform.html Multalin [12] http://prodes.tou or louse.inra.fr/multalin/multalin.html. Our sequence determination was continuously refined with the results of repeated BlastN and tBlastN database searches with the updated contigs, and finally we extended our search by using genomic databases. The majority of Blast searches were conducted with the default parameter settings. With this strategy were able to generate presumptive "framework" contigs for the in databases less abundantly represented TMC genes. Because of the inter-species conservation of protein-coding sequences we were able to predict the presumptive DNA sequence encoding aminoand carboxyl-termini of various TMC proteins by comparing human and murine genomic TMC sequence.

We constantly refined individual TMC sequences by using genomic sequence information and comparing gene structures. In addition to iterative database searches, we also took into account the results of gene prediction algorithms using genomic DNA harboring TMC genes as inputs (GeneScan [13], GenomeScan [14] <u>http://genes.mit.edu</u>). Some invertebrate TMC genes were already annotated as hypothetical proteins. We used human and murine genome databases (GenBank/NCBI – <u>http://www.ncbi.nlm.nih.gov</u>) to determine the chromosomal locations of all mammalian TMC genes.

In a second phase, we used polymerase chain reaction and 5' and 3' rapid amplification of cDNA ends (RACE) experiments to obtain experimental proof for our cDNA predictions and to determine murine TMC sequence that was only ambiguously predicted.

Secondary structure predictions were done with PSORTII [15] <u>http://psort.ims.u-tokyo.ac.jp</u>. The phylogenetic distances between various TMC protein sequences were calculated with ClustalW [11] <u>http://www.es.embnet.org/</u><u>Doc/phylodendron/clustal-form.html</u> and the resulting distance matrix Additional file: 1 (A) was visualized as a tree diagram with centered node position using the Phylodendron phylogenetic tree printer available at <u>http://iubio.bio.indiana.edu/treeapp</u>.

Evolutionary distance between two gene structures was defined as the reciprocal of the ratio of conserved intron locations with respect to the coding sequences of the TMC genes. The distances were determined for each pair of murine TMC genes (additional file 1 (B)). We compared the relationship of the TMC gene structures by performing cluster analysis of distances using the kitsch algorithm of the PHYLIP Phylogeny Inference Package [16] <u>http://bio</u> <u>web.pasteur.fr/seqanal/phylogeny/intro-uk.html</u>. The results of the cluster analysis were visualized as a tree diagram with centred node position using the Phylodendron phylogenetic tree printer.

Accession numbers for all TMC cDNAs reported in this paper are listed in Table 1.

mRNA expression analysis

Organs we dissected from eight-week-old mice and quickly frozen in liquid nitrogen. Specifically, we used cortex, cerebellum, eye, cochlea, pituitary gland, salivary gland, small intestine, colon, lung, thymus, liver, kidney, heart, urinary bladder, testis, prostate, and ovary. We also used in a control experiment aimed to verify Tmc2 expression in the cochlea, plasmid DNA obtained from an organ of Corti cDNA library that was kindly provided by Dr. B. Kachar, NIDCD/NIH. Total RNA was extracted using RNeasy Mini and Midi kits (Qiagen, Valencia, CA). For RT-PCR, 1 µg of total RNA was reverse-transcribed into cDNA using SuperscriptII reverse transcriptase (Invitrogen, Carlsbad, CA). One tenth of the resulting cDNA was used as template for each PCR amplification with primers specific for each member of the TMC gene family (Tmc1: 5'-GAAACAATGGTGGGGGCAGGAA and 5'-GATGGCG-GGAGGGAGACGAT, Tmc2: 5'-CCCCGGCCACCAC-CACAC and 5'-GAAGGGGCCACAGTCAAACGAG, Tmc3: 5'-TCTGGACAAAGTGAACAGCATG and 5'-GAACATT-GAATGCTGGCAGACA, Tmc5: 5'-GGGCGGGTTATTTT-GGGGACAC 5'and ACAACAAAGGGCAGCAACAACACC, Tmc6: 5'-CTAT-5'-GCAGCCGGA-CACTGGCTGGGCCGCAGG and GAATCTGTACTGC, Tmc4: 5'-GTGGAGTCAGATCACGTGTGGG and 5'-CAGGCA-GAACAGTATGAGGAAC, Tmc7: 5'-GGGCTCCAAGATCA-CATCCTGT and 5'-CGAATTTCAAGGTGGCAATGGC, and Tmc8: 5'-GGGTCAGACGGTGCTGTGCATA and 5'-AGAGGAAGACACTGTTGAGTAG). Care was taken to select primer pairs that amplify TMC transcripts encoded in multiple exons to rule out the amplification of genomic DNA contamination. Specific products obtained with each primer pair from various organs and in several independent experiments were verified by sequencing. To estimate the quality of the template and to compare expression levels, we amplified mouse glyceraldehyde 3phosphate dehydrogenase (GAPDH), a ubiquitously expressed abundant transcript with specific primers (5'-AACGGGAAGCCCATCACC and 5'-CAGCCTTGGCAG-CACCA). Cycling parameters were 94°C for 3 min, 8 cycles of (94°C20 s, 60°C20 s -1°C/cycle, 68°C45 s), followed by 17, 24 or 34 additional cycles of (94°C30 s, 52°C30 s, 68°C45 s) and 68°C for 5 min.

Authors' contributions

KG carried out iterative database searches to identify mammalian members of the TMC family, he did PCR and RACE experiments to obtain TMC cDNA sequence, and he performed part of the transcript expression analysis. HM identified non-mammalian and invertebrate TMC genes, he assembled their coding sequences from public databases, and he performed part of the transcript expression analysis. KG and HM drafted the manuscript. All authors conceived of the study as result of discussions during the group's laboratory meeting. SH participated in design and coordination of the study and finalized the manuscript, which has been read and approved by all authors.

Additional material

Additional File 1

Identity matrices of murine TMCs. (A) Percentage of amino acid residue identity among individual murine TMC proteins. (B) Evolutionary distance ratios of the gene structures of individual murine TMC genes. We calculated the ratio of conserved intron locations within the conserved region of the eight murine Tmcs (see Figure 2). The reciprocal values of the distance ratios represent the degree of conserved intron locations in percent. The three subfamilies are shaded in green (TMC Subfamily A), red (TMC Subfamily B), and blue (TMC Subfamily C). Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-4-24-S1.pdf]

Additional File 2

TMC signature sequence motif consensus sequence. Comparison of the conserved TMC signature sequence motif among vertebrate and invertebrate TMC proteins. The consensus sequence is based on a 100% conservation of the three leading amino acid residues. All other capitalized amino acids are conserved in more than 70% of all TMCs, and lower case letters indicate the most probable amino acid residue.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-4-24-S2.pdf]

Additional File 3

Comparison of invertebrate TMC gene structures. Comparison of the intron locations projected onto a schematic comparison of invertebrate TMC proteins. A representation of the conserved intron locations of the vertebrate TMC subfamilyA is displayed as a reference. For details on the color-coding of intron locations, see figure 3, with the addition that open blue triangles refer to intron positions conserved within invertebrate cladi. Each black box indicates a gap to adjust the alignment. Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-4-24-S3.pdf]

Acknowledgements

We thank Drs. E. Corrales and W.F. Sewell for comments on the manuscript. In particular, we thank Drs. K. Kurima and A. Griffith for sharing the HUGO and mouse nomenclature committee-approved TMC nomenclature, which has been submitted elsewhere by their research team in a related context. This work was supported by grant DC04563 to SH from the NIH/NIDCD and in part by Basil O'Connor Starter Scholar Research Award Grant No. 5-FY00-528 from the March of Dimes Birth Defects Foundation.

References

- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA and Holt RA *et al.*: The sequence of the human genome. *Science* 2001, 291:1304-51.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M and FitzHugh W et al.: Initial sequencing and analysis of the human genome. Nature 2001, 409:860-921.
- Hogenesch JB, Ching KA, Batalov S, Su Al, Walker JR, Zhou Y, Kay SA, Schultz PG and Cooke MP: A comparison of the Celera and Ensembl predicted gene sets reveals little overlap in novel genes. *Cell* 2001, 106:413-5.
- Li S, Liao J, Cutler G, Hoey T, Hogenesch JB, Cooke MP, Schultz PG and Ling XB: Comparative analysis of human genome assemblies reveals genome-level differences. *Genomics* 2002, 80:138-9.
- Kurima K, Peters LM, Yang Y, Riazuddin S, Ahmed ZM, Naz S, Arnaud D, Drury S, Mo J and Makishima T et al.: Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function. Nat Genet 2002, 30:277-84.
- Vreugde S, Erven A, Kros CJ, Marcotti W, Fuchs H, Kurima K, Wilcox ER, Friedman TB, Griffith AJ and Balling R et al.: Beethoven, a mouse model for dominant, progressive hearing loss DFNA36. Nat Genet 2002, 30:257-8.
- Ramoz N, Rueda LA, Bouadjar B, Montoya LS, Orth G and Favre M: Mutations in two adjacent novel genes are associated with epidermodysplasia verruciformis. Nat Genet 2002, 32:579-81.
- 8. Hedges SB and Kumar S: Genomics. Vertebrate genomes compared. Science 2002, 297:1283-5.
- Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P, Christoffels A, Rash S, Hoon S and Smit A et al.: Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. Science 2002, 297:1301-10.
- Steel KP, Marcotti W, Erven A, Johnson SL and Kros CJ: Cochlear function in Beethoven mouse mutants. In: Association of Research of Otolaryngology Winter Meeting Daytona Beach, Florida; 2003:1063.
- Thompson JD, Higgins DG and Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994, 22:4673-80.
- 12. Corpet F: Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 1988, 16:10881-90.
- 13. Burge C and Karlin S: **Prediction of complete gene structures in** human genomic DNA. J Mol Biol 1997, 268:78-94.
- Yeh RF, Lim LP and Burge CB: Computational inference of homologous gene structures in the human genome. Genome Res 2001, 11:803-16.
- Nakai K and Horton P: PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 1999, 24:34-6.
- Felsenstein J: PHYLIP Phylogeny Interference Package (Version 3.2). Cladistics 1989, 5:164-6.

