



Figure S8. Multiple sequence alignment of CX family proteins. These proteins are all from nematodes. The six conserved cysteines are highlighted in **yellow** (changes of cysteines to other residues are in **black** background). Predicted transmembrane segments are underlined. Cysteines (possible palmitoylation sites) and positively charged residues (helping determining the topology of the transmembrane proteins) near the C-terminal end of the Predicted transmembrane segment are in **red** letters and **blue** letters respectively. PY motifs (PPxY or LPxY) are highlighted in **cyan**. The N-terminal segments before the cysteine-rich domain are not shown. The first residue numbers are shown before the sequences. The sequence with three random cysteine-rich domains (gi: 308468441) is marked by **green** names. Sequence names are followed by species abbreviations: Brm, *Brugia malayi*; Cb, *Caenorhabditis briggsae*; Ce, *Caenorhabditis elegans*; Cr, *Caenorhabditis remanei*; Ll, *Loa loa* and Ts, *Trichinella spiralis*.

```

86565585_ref_NP_001033410.1_Ce 115 SYCKRPDE-----PFRPDILCPRES---MFFFYKCCPSSQDHEKSECCAKIRIWLMLAIVCVVLSCLVGICYSIFHYFCRHCRLPSWKSSWKSTRLLRSDV
308460146_ref_XP_003092380.1_Cr 115 SYCKRPDE-----PFHPSVLCPRES---MFFFYKCCPSSADPEKSECCAKIRIWLMLAIVCVVLSCLVGICYSIFRYFCRHCRLSAWKPTRLRSHV
268572725_ref_XP_002649032.1_Cb 115 SYCKRPEE-----PFRPDILCPRES---MFFFYKCCPSSLDAEKSECCAKIRIWLMLAIVCVVLSCLLGISYFFRYFCRHCRLSHHWKPTRLRSDV
170582513_ref_XP_001896163.1_Bm 72 TWCVPV-----VIGTQCPASS---IFHYKCCGTA---NKECCFNLTQWVIVVLAIVAVMMLASFVLSVLRCLFCRR
312078855_ref_XP_003141921.1_Ll 72 TWCVPV-----VIGTQCPASS---IFHYKCCGTA---NKECCFNLTQWVIVVLAIVAVMMLASFVLSVLRCLFCRR
268572629_ref_XP_002641370.1_Cb 74 TYCPVP-----LVGTRCGTSS---IFHYWKCCGDL---NKECCFNLTQWVIVVLAIVAVMMLASFVLSVLRCLFCRR
17555852_ref_NP_499520.1_Ce 74 TYCPVP-----LVGTRCGTSS---IFHYWKCCGEL---NKECCFNLTQWVIVVLAIVAVMMLASFVLSVLRCLFCRR
308483826_ref_XP_003104114.1_Cr 74 TYCPVP-----LVGTRCGTSS---IFHYWKCCGEL---NKECCFNLTQWVIVVLAIVAVMMLASFVLSVLRCLFCRR
312094490_ref_XP_003148040.1_Ll 75 LACPPI-----VVGEOCPDET---PLYYRCCGDL---NSSCCFRLQEWVIVLVIMGILIGLSLVNFIRYLFCIR
17507487_ref_NP_492071.1_Ce 82 LACPLP-----VVGQAQPEEN---VFYFKCCGQI---ADQCCFRLQDWVIVVLLFLAVCTILSIIINCVRCFCCA
268560470_ref_XP_002646218.1_Cb 82 LACPLP-----VVGSAQPEEN---VFYFKCCGQI---ADQCCFRLQDWVIVVLLFLAVCTILSIIINCVRCFCCA
308469178_ref_XP_003096828.1_Cr 82 LACPLP-----VVGSAQPEEN---VFYFKCCGQI---ADQCCFRLQDWVIVVLLFLAVCTILSIIINCVRCFCCA
308462989_ref_XP_003093773.1_Cr 85 VFCPSKE-----NQVTDGEGS---FAHTHKCCGGD---EKECCFALQVWVIVLAVLVLCIIISTVIGIICCCCKK
268572633_ref_XP_002641371.1_Cb 86 AYCPCG-----NGKFCGSTS---FAHYNCGGDN---NABCCFALQVWVIVLAVLVLMIIISTFIGIICCSVRSQSTS
193206295_ref_NP_501165.3_Ce 93 IWCPTH-----RVGNCPPEE---LLSFYKCCGHL---NKECCFHLRVVIVIMIVLPICLVLPASTFLIRILCSSSQPTGRAPSAADRVRQLAEDSTTFRRSDDMVML
268531584_ref_XP_002630919.1_Cb 105 LHCPSIE-----NAVSTCPKDS---NWVYTCGGA---NMYCCEHIQTWLLSALAVIAVFLTLLIGCCVRCCCSYKKKTKQQSYSPDK
308502874_ref_XP_003113621.1_Cr 120 LHCPSIE-----NSVSTCPKDS---NWVYTCGGA---NMYCCEHIQTWLLSALAVIAVFLTLLIGCCVRCCCSYKKKTKQQSYSPDK
115532536_ref_NP_001040783.1_Ce 108 LHCPSIE-----NAVSSCPKDS---KWVYTCGGA---NMYCCEHIQTWLLSALAVIAVFLTLLIGCCVRCCCSYKKKTKQQSYSPDK
86563167_ref_NP_001033369.1_Ce 59 ETCY-----PDTADPDLDTFFSFGTCCD-----NECCQWRIRYVYVPASIAIGFVAGAFFALCFQCR
312084143_ref_XP_003144154.1_Ll 82 EICI-----PQDCPEET---NVVYTCGD-----DCCQQRWIVVLPALFASTAFTMGAFFAFCFHCR
193206848_ref_NP_001122817.1_Ce 71 EKNDTS-----GNPLNCPSNF---IYFECGA-----NECCLRTQVVPAVIIAIGAIMTGFCLVSCVAYCCCGE
309359596_emb_CAR99881.1_Cb 70 EPCQDV-----GKELHCPSNF---IYFECGS-----NECCLRTQVLPAVIAVAFGAIVMTLFCVSCVAYCCCGD
170590574_ref_XP_001900047.1_Bm 88 QWCRNRT-----ISKYIQPMGS---ALHYSCCGET---DTECCFRLQPTVVAVLLSIQCFLRYSNSIIRL
268529264_ref_XP_002629758.1_Cb 83 QWCRNFT-----VNEHTCQPAS---AFHQDCCGQH---ETECCFALQGWVIVILIVGTCASIVLIFSTLLKLNLICPLPYSTKPIPTXGTIHD
308459747_ref_XP_003092187.1_Cr 82 QWCRNFT-----VNEHTCQPRGS---AFHQDCCGQH---ETECCFALQGWVIVILITVGTGCAILIFSALLKFNLICPLPYSSKPAPTXGTIQD
71981497_ref_NP_001021939.1_Ce 82 QWCRNFT-----VNEHTCQPAS---AFHQDCCGHL---ETECCFALQGWVIVILITVGTGCAILIFSVLLKMNLICMPYTTKQTPSTYTGIQE
309355322_emb_CAS00965.1_Cb 59 WSCPFINPMGQ-----MGMPVGEDCPDTS---FIHYSCDDN---PFQCCFHFETWAIVIFGIVGTIVVGSLVFVAGLLMTPNYSKNPPNGRV
17556895_ref_NP_498853.1_Ce 45 GNCGGGILGSLVGMAGGAGGYGGYGGYGRCGADN---IFYRWRCCDYS---PYECCIQLETWVIVFLVIFIGFFVCLCACLAGCVWSARNRE
308483972_ref_XP_003104187.1_Cr 44 GSCGGGILGSLVGMAGG-YGGYGGYGGYGGYGRCGADN---IFYRWRCCDYS---PYECCIQLETWVIVFLVIFIGFFVCLCACLAGCVWSARNRQ
268574408_ref_XP_002642181.1_Cb 44 GSCNGGILGSLVGMAGG---GGYGGYGGYGGYGRCGADN---IFYRWRCCDYS---PYECCIQLETWVIVFLVIFIGFFVCLCACLAGCVWSARNRQ
312081944_ref_XP_003143239.1_Ll 62 MVEDGAGNEFHS-----LLYGLGYGLAQICPPDN---FFFSYICDQD---YFECIRFELWFMYSIRKKSSGKVISLTVSPSL
312091987_ref_XP_003147177.1_Ll 100 LECN-----GKKCPNNN---IFVYCCGENV---PKECCWHIRYWLTKLVDHDFIFFPDDKWKRYVYV
115534231_ref_NP_499169.3_Ce 64 VICRRPWFYN-----SPRPAPMRCPPTAS---FTTYECCGEF---LENCCWRFRQEPIIIGVILLVLLLVCCCCCVSVIFVGRKKKSSNIEPEKPTKSDSIQITISSSTIDSGTQWELRRSYEETDKRRSYAAARDQELDYQYFQ
308502285_ref_XP_003113327.1_Cr 105 VICRRPWFYN-----SPRAPLRCPTPS---FTTYECCGEF---LENCCWRFRQEPIIISVILLVLLLVCCCCCIAWLAFGRKKKSTNVEPEKQLTTTKSDSIQITISSSTIDSGTQWELRRSYEETERRRSYAAARDRELDYQYFS
309361634_emb_CAF29516.4_Cb 58 VICRRPWFYN-----SPRAPLRCPTPS---FTTYECCGEF---LGNCCWRFRQEPIICVILVLLLVCCCCCIAWCVGRKKKSTSVAPERKTTTTTKSDSIQITISSSTIDSGTQWELRRSYEETERRRSYAAARDRELDYQYFS

```

Figure S9. Multiple sequence alignment of DUF2650 family proteins. These proteins are all from nematodes. The six conserved cysteines are highlighted in yellow. Predicted transmembrane segments are underlined. Cysteines (possible palmitoylation sites) and positively charged residues (helping determining the topology of the transmembrane proteins) near the C-terminal end of the predicted transmembrane segment are in red letters and blue letters respectively. PY motifs (PPxY or LPxY) are highlighted in cyan. The N-terminal segments before the cysteine-rich domain are not shown. The first residue numbers are shown before the sequences. Sequence names are followed by species abbreviations: Bm, *Brugia malayi*; Cb, *Caenorhabditis briggsae*; Ce, *Caenorhabditis elegans*; Cr, *Caenorhabditis remanei* and Ll, *Loa loa*.

269954664_ref_NP_694961.2_Hs 59 AKCGLILACPKGFKCGDSCCQENELFPQVRI~~VFVIFLVLISL~~FCICGLAKCFCRNCRPEPEPDSVDCRGPLELPSIIPPERVRLSAPPPYSEVILKPSLGPPTPEPPYSFRPEEYTGQQRGIDNPAF

139948842_ref_NP_001077120.1_Bt 56 ATCG-LLTCPKGFKCGSDCCQEVYQLEQYDFFSHPRLTSVIFLIIILLLCIYGVTKHLYLNLCKRSEQEAAPTALPEQPPIAVERVRLTAPISEPPYSEIILKPVGLPLPEPPYSFRPEEYAGFCRGINDSF

297716026_ref_XP_002834334.1_Pa 59 AKCGLIFACPKGFKCGDSCCQENELFPQVRI~~VFVIFLVLISL~~FCICGLAKCFCRNCRPEPEPDTMDCRGPLELPSIIPPERVRLSAPPPYSEVILKPSLGPPTPEPPYSFRPEEYTGQQRGIDNPAF

296202492_ref_XP_002748479.1_Cj 59 ARCGLSFACPKGFKCGDSCCQENDIFPGVRI~~VFVIFLVLISL~~FCICGLAKCFCRNCRPEPEPDTLDCRRPELPSIIPPERVRLSAPPPYSEVILKPTLGPPTPEPPYSLRPEEYTRDQRGIDNPAF

297694989_ref_XP_002824748.1_Pa 58 AKCGLSFTCPKGFKCGDSCCQENELFPGPLRIFVIFLVLISLFCICGLAKCFCRNCRPEPEPDLDPALDTCRGPQELPSIIPPERVRLSAPPPYSEVILKPTLGPPTPEPPYSFRPEEYTGQQRGIDNPAF

301776715_ref_XP_002923774.1_Am 59 ARCGLFFTCPKGFKCGDSCCQEVYELFSGPLRVFVITVLLIILPLLCICGLAKCFCRNCRSEGEQEQDPPVDHQGLPEPPIVPPERVRLTAPISEPPYSEIILKPVLPPEPPYSFRPEEYTGQQRGIDNPAF

311267573_ref_XP_003131634.1_Ss 59 ATCAPFHTCPKGLKCGDTCCPEYQREPPGFFSGPFRIFVIFLIFLIPLLCICGLAKCFCRNCRPEPEPTEHEGPELPSIAPPEERVIAPISDSPPYSEVILKPELGLPPTPEPPYSCRPEEYTRACRGINDPAF

109114310_ref_XP_001097096.1_Mam 59 AKCGFIFNCPKGFKCGDSCCQENELFPQVRI~~VFVIFLVLISL~~FCICGLAKCFCRNCRPEPEPAPMDCRGPLELPSIIPPERVRLSAPPPYSEVILKPSLGPPTPEPPYSFRPEEYTGRRRGINDPAF

253683499_ref_NP_001156644.1_Mm 63 NTCD-ILNCCPKGFTCCVKECCPERKVVWDANDRRFRFLVILACIIFPILFICALVSLFCPCNCTELQHDVRRVDHQTPIEPPIAPLESIWVTSLDPPYSQVVMPTTPEPPYSLRPEGPAGQMRGRAYATL

157738690_ref_NP_001099031.1_Mm 63 NTCD-ILNCCPKGFTCCDKDCCPERKVVWGPESWPKNEPFKIVVILLVMLPLLCICGLVRRFCPCNRELQHNFRADHQTSPLSPVALENIWVTSLDPPYSQVVLKPTTPEPPYSLRPEGTSGQRGHAYATL

51766385_ref_XP_484066.1_Mm 63 NTCD-ILNCCTKGFTCCHNGCCPERELSDLANDHFRIVDILLWALLFVICGTIRHFFYKRTYLLHGVRTSDHQTPPEPPIAPLESIRVTTLDPPYSQVVMPTTPEPPYSLRPEGPAGQMRGRAYATL

309262391_ref_XP_003085797.1_Mm 63 NTCD-NLNCCTKGFTCCHNGCCPERELSDLANDHFRIVDILLWALLFVICGTIRHFFYKRTYLLHGVRTSDHQTPPEPPIAPLESIRVTTLDPPYSQVVMPTTPEPPYSLRPEGPAGQMRGRAYATL

309262389_ref_XP_997064.2_Mm 63 NTCD-ILNCCTKGFTCCDNGCCPERKFWDLNENRFLVNIILYAILVLSFICVLLRHFFYLCRDLQHDVVRTSDHQTPPEPPIAPLESIWVTSLDPPYSQVVMPTTPEPPYSLRPEGPAGQMRGRAYATL

309272522_ref_XP_001474210.2_Mm 63 NTCD-ILSCCPKGFTCCDNECCPERKVVWDSANPPFRILSIIICTILPVLFCALVRCIYSKCRDLQHDVVRTSDHRAPPEPPIAPLESIRVTSFDPPYSQVVMPTTPEPPYSLRPEGPAGQMRGRAYATL

63617748_ref_XP_618920.1_Mm 63 NTCD-ILSCCPKGFTCCDNECCPERKVVWDSANPPFRILSIIICTILPVLFCALVRCIYSKCRDLQHDVVRTSDHQAPPEPPIAPLESIRVTSFDPPYSQVVMPTTPEPPYSLRPEGPAGQMRGRAYATL

293340380_ref_XP_002724681.1_Rn 63 NNCE-DLHCCPKGFGCCGNECCQESKILDPENDPLKIVFITLLVIVLLLCICGLVMSKCKRLHHDRAADHQTPPDAPSIAPLESIWVTSLDPLPTYSQVVLNSTRTEPPYSLRPERPTGRRGTRYATI

293340376_ref_XP_002724680.1_Rn 63 NNCE-DLHCCPKGFGCCGNECCHESKILDPENDPLKIVFITLLVIVLLLCICGLVMSKCKQLHHDRAADHQTPPDAPSIAPLESIWVTSLDPLPTYSQVVLNSTRTEPPYSLRPERPTGRRGTRYATI

293351785_ref_XP_001075148.2_Rn 63 NNCE-DLHCCPKGFGCCGNECCQERKILDPENDPLKIVFITLLVIVLLLCICGLVMSKCKRLHHDRAADPAPSIAPLESIWVTSLDPLPTYSQVVLNSTRTEPPYSLRPERPTGRRGTRYATI

293340378_ref_XP_001081299.2_Rn 63 NNCE-DLHCCPKGFGCCGNECCQERKILDPENDPLKIVFITLLVIVLLLCICGLVMSKCKRLHHDRAADPAPSIAPLESIWVTSLDPLPTYSQVVLNSTRTEPPYSLRPERPTGRRGTRYATI

293340374_ref_XP_002724679.1_Rn 63 YKCE-DIHCCPKGFGCCDNECCLERKIWNWDNEPFRILFIIILLVMLPLLCICGLVRRFCPCNRELQHEVTRADHQSPDPPIAPLESIWVTSLDPPYSQVVMPTTPEPPYSLRPEPDSFCPMRETSHETL

Figure S10. Multiple sequence alignment of TMEM92 family proteins. These proteins are all from mammals. The six conserved cysteines are highlighted in yellow. Predicted transmembrane segments (TMs) are underlined. Cysteines (possible palmitoylation sites) and positively charged residues (helping determining the topology of the transmembrane proteins) near the C-terminal end of the predicted TMs are in red letters and blue letters, respectively. PY motifs (PPxY or LPxY) are highlighted in cyan. The N-terminal segments before the cysteine-rich domain are not shown. The first residue numbers are shown before the sequences. Sequence names are followed by species abbreviations: Am, *Ailuropoda melanoleuca*; Bt, *Bos taurus*; Cj, *Callithrix jacchus*; Hs, *Homo sapiens*; Mam, *Macaca mulatta*; Mm, *Mus musculus*; Pa, *Pongo abelii*; Rn, *Rattus norvegicus* and Ss, *Sus scrofa*.

16418447_NP_443186.1_Hs 29 AQCGKDKSYCDGGTTPYCCSYAYIGNILSGTAIAGIVFGVIFMGVIAGIAICICMCKMNNRATRVGLRTRTHINTVSSYPGPPYGHDEMEYCADLPPYSPTQGPQRSPPPYPGNARK

21450067_NP_659102.1_Mm 29 AQCGKECHSYCDGSGTPYCCSYAYIGNILSGTAIAGIVFGVIFMGVIAGIAICICMCKMNNRGT RVGVIRAAHINAISSYPAPPYTYDHEMEYRTDLPPYSAAPQASQRSPPPYPGNPRKYSSSQNRIRDN

118083813_XP_001233379.1_Gg 227 AQCDNDCKAYCDGTAPYCCSYAYIGNVLSGTAIAGIVFGVIFMGVIAGIAICICMCKMNNRGT RVGVIRTRTHINTISTYPTVPPYSYEHMEYVVDLPPYTPTPTSLQYPPYPGYSKG

62858163_NP_001016493.1_Xt 27 AYCEADCKSYCDGTPPPYCCSYAYIGNVLSGTAIAGIVFGVIFMGVIAGIAICICMCKMNSRGT RVGVIRTRTHINAISSYPAPPYSYEYEMDFPLDLPPYTPTPTISHYPSPPYPGPSRK

47086271_NP_998047.1_Dr 31 AQCDG-IEYCDGVPPFCCCSYAYVGDVLSGTAISGIVFGVIFMLGAVAAVFLCVCMCKNSRGRSVGVFSSTYINTVTQYGGPPPPYSYDHEMFPPDLPPYTPTVPRSANYSPPYPGFSRK

47226858_CAG06700.1_Tn 18 AQEG-LEYCDASPPFCCCSYAYVGDVLSGTAISGIVFGVIFMLGAVAAMCLCCMCKMNGR GARVGVFSTCINSVTVQYGGPPPPYTPAQSQPANYSPPPYPGCK

Figure S11. Multiple sequence alignment of representative CYR1 family proteins. These proteins are from vertebrates. The six conserved cysteines are highlighted in yellow. Predicted transmembrane segments (TMs) are underlined. Cysteines (possible palmitoylation sites) and positively charged residues (helping determining the topology of the transmembrane proteins) near the C-terminal end of the predicted TMs are in red letters and blue letters, respectively. PY motifs (PPxY or LPxY) are highlighted in cyan. The N-terminal segments before the cysteine-rich domain are not shown. The first residue numbers are shown before the sequences. Sequence names (gene identification number and accession number) are followed by the species abbreviations as follows: Dr, *Danio rerio*; Gg, *Gallus gallus*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Tn, *Tetraodon nigroviridis*; and Xt, *Xenopus tropicalis*.