

Evolutionary selection pressure and family relationships among *connexin* genes

V. Cruciani and S.-O. Mikalsen

Biol. Chem. 388 (2007), pp. 253-264.

SUPPLEMENTARY FIGURES

Suppl. Fig. 1 <i>Connexin</i> sequences from chicken (<i>Gallus gallus</i>).....	2
Suppl. Fig. 2 <i>Connexin</i> sequences from <i>Xenopus tropicalis</i>	10
Suppl. Fig. 3 <i>Connexin</i> sequences from zebrafish (<i>Danio rerio</i>).....	20
Suppl. Fig. 4 <i>Connexin</i> sequences from pufferfish (<i>Fugu rubripes</i>).....	33
Suppl. Fig. 5 <i>Connexin</i> sequences from green spotted pufferfish (<i>Tetraodon nigroviridis</i>).....	51
Suppl. Fig. 6 <i>Connexin</i> sequences from tunicates (<i>Ciona intestinalis</i> and <i>Halocynthia pyriformis</i>).....	70
Suppl. Fig. 7 Coincidence of SNPs in human Cx43 with sequence differences of Cx43 and Cx43p1.....	75
Suppl. Fig. 8 Schematic explanation of terms used in the text when describing a branch from a phylogenetic tree.....	77
Suppl. Fig. 9 Detailed branches of the vertebrate <i>connexin</i> phylogenetic tree.....	78
Suppl. Fig. 10 Phylogenetic tree illustrating long-branch attraction of tunicate connexins with vertebrate Cx40.1 orthologs.....	90
Suppl. Fig. 11 Phylogenetic tree including the tunicate <i>connexins</i> , but without vertebrate Cx31.3, 31.9 and 40.1 orthologs.....	92
Suppl. Fig. 12 The grouping of some <i>Ciona connexins</i> together with vertebrate sequences.....	93

SUPPLEMENTARY TABLES

Suppl. Table 1 Genomic positions of chicken <i>connexins</i>	96
Suppl. Table 2 Databank entries of chicken <i>connexins</i>	97
Suppl. Table 3 Database positions of <i>Xenopus tropicalis connexins</i>	98
Suppl. Table 4 <i>Xenopus tropicalis</i> and <i>Xenopus laevis</i> genome databank entries.....	100
Suppl. Table 5 Genomic positions of zebrafish <i>connexins</i>	101
Suppl. Table 6 Zebrafish genome databank entries.....	102
Suppl. Table 7 Databank entries and database positions of <i>Fugu rubripes connexins</i>	105
Suppl. Table 8 Databank entries and positions of <i>Tetraodon nigroviridis connexins</i>	108
Suppl. Table 9 d_N and d_S values for full-length mammalian <i>connexins</i>	109
Suppl. Table 10 Selection pressures (d_N/d_S) on full-length mammalian <i>connexins</i>	111
Suppl. Table 11 d_N and d_S values for the conserved domains of mammalian <i>connexins</i>	112
Suppl. Table 12 Selection pressures (d_N/d_S) on the conserved domains of mammalian <i>connexins</i>	114
Suppl. Table 13 d_N and d_S values for the variable domains of mammalian <i>connexins</i>	115
Suppl. Table 14 Selection pressures (d_N/d_S) on the variable domains of mammalian <i>connexins</i>	117

Suppl. Table 15 Nucleotide distances for full-length coding sequences, introns and 3'-UTRs for some mammalian <i>connexins</i>	118
Suppl. Table 16 Nucleotide distances for <i>Cx43</i> - and <i>Cx31.3</i> -like (pseudo-)genes.....	118

Abbreviations:

ic: Follows a gene identification, and indicates that the sequence has an incomplete conserved domain.

iv: Follows a gene identification, and indicates that the sequence has an incomplete variable domain.

Suppl. Figure 1 *Connexin* sequences from chicken (*Gallus gallus*, Gg). The sequences are found in the Ensembl database or GenBank, and are updated as of July 1 2006. If the sequences previously have been named within the connexin nomenclature, this is normally the name that is used here (see Suppl. Table 2). The Ensembl gene identification is given in abbreviated form. E.g., for *Gallus gallus Cx31* (*GgCx31*) the abbreviated form *Gg17136* should be read as ENSGALG00000017136. If the sequence only is found in GenBank, the accession number is given. Other information is given in short form where needed. If our interpretation of the most likely connexin sequence differs from the predictions found in Ensembl (or GenBank), the database sequence is given first, and immediately followed by our suggested sequence, which is the one used in our analyses. The conserved domains are marked in yellow (unmodified sequence) or green (modified sequence). Alternating black/blue sequences are exon predictions copied from Ensembl. Other colors are explained for each of the relevant sequences.

>GgCx31 Gg17136 Ensembl prediction in 3 exons (black/blue/black letters). Removal of 96 nt at the 5' end (marked gray) to reach initiation codon consistent with other connexins. Addition of 18 nt at 3' end to reach stop codon (marked blue). Modelled according to NM_204931 (there called Cx30).

GTGATGGGGAATTGTCTGTCTAGCTCTGCCACTAAGGACAAGGTATCCTTTAGACTTC
CTAAGGCAGGCTGTGCAGGCCTCCTCATACAAGACAATGGATTGGGGAGCTGTGCAGACC
ATCTTAGGAGGTGTAATAAACACTCCACCAGCATTGGGAAGATATGGCTCACAGTCCTG
TTCATCTTCCGTATCATGATCCTGGTTGTGGCTGCAGAGAGAGTCTGGGGAGATGAACAA
GATGACTTTATCTGCAACACTTTGCAACCTGGTTGCAAAAATGTTTGCTATGATCACTTT
TTCCCCATCTCTCACATCAGACTCTGGGCCCTGCAGCTGATCTTCGTCTCTACCCCTGCG
CTGCTGGTGGCCATGCACGTGGCTTACAGGAGACACGAGAAGAAAAGGCAGTTCAGAAAAG
GGAGACCAGAAATGCGAGTATAAGGACATTGAAGAAATCAGAACCCAGAGGTTTCGTATT
GAGGGCACCCCTGTGGTGGACGTACACTTGCAGCATCTTCTTCAGATTGGTCTTTGAAGCA
GTCTTCATGTATGGCTACATTTTATTTTCATGTATGATGGGTTCCGAATGCCTCGCTTAATGAAA
TGAGTGCTTGGCCCTGCCCAACAGTAGACTGCTTTGTTTCTCGACCTACTGAAAAG
ACAGTGTTCACTATTTTTCATGATTGCTGTGTCCAGCATTTCATTCTTTAAATGTGGCT
GAATTATGTTATTTACTGACAAAATTTTTCTCAGAAGGTCTAGAAAAGCTGGAAATCAG
AAACACCAACCCCAACCATGAGAATAAAGAAGAAACCAACCAAAATGAAATGAACGAGTTA
ATATCTGATAGCTGTGACAAACAGTTATA

>GgCx31 Gg17136 Our suggested modification according to NM_204931

ATGGATTGGGGAGCTCTGCAGACCATCTTAGGAGGTGTAATAAACACTCCACCAGCATT
GGGAAGATATGGCTCACAGTCCTGTTTCATCTTCCGTATCATGATCCTGGTTGTGGCTGCA
GAGAGAGTCTGGGGAGATGAACAAGATGACTTTATCTGCAACACTTTGCAACCTGGTTGC
AAAAATGTTTGCTATGATCACTTTTCCCCATCTCTCACATCAGACTCTGGGCCCTGCAG
CTGATCTTCGTCTCTACCCCTGCGCTGCTGGTGGCCATGCACGTGGCTTACAGGAGACAC
GAGAAGAAAAGGCAGTTACAGAAAGGGAGACCAGAAATGCGAGTATAAGGACATTGAAGAA
ATCAGAACCCAGAGGTTTCGTATTGAGGGCACCCCTGTGGTGGACGTACACTTGCAGCATC
TTCTTCAGATTGGTCTTTGAAGCAGTCTTCATGTATGCATTTTATTTTCATGTATGATGGG
TTCCGAATGCCTCGCTTAATGAAATGTAGTGCTTGGCCCTGCCCAACACAGTAGACTGC
TTTGTTTCTCGACCTACTGAAAAGACAGTGTTCACTATTTTTCATGATTGCTGTGTCAGC
ATTTGCATTCTTTAAATGTGGCTGAATTATGTTATTTACTGACAAAATTTTTTCTCAGA
AGGTCTAGAAAAGCTGGAAATCAGAAACACCAACCCCAACCATGAGAATAAAGAAGAAACC
AAACAAAATGAAATGAACGAGTTAATATCTGATAGCTGTGACAAACAGTTATAAGGATTC
ACAAGTAGCTAA

>GgCx31sim XM_425641 When comparing the XM sequence with the genomic sequence in Ensembl, it is evident that the XM sequence predicts an intron (4 first nt of exon 2 marked in dark green /white letters). By extending exon 1 with 4 nt into the predicted intron (marked in blue), a stop codon is apparent.

ATGGATTGGGGAACCTCTGCAGGCTGTTTTGGGAGGTGTAATAAACACTCCACCAGCATT
GGGAAGATATGGCTCACAGTCCTGTTTCATCTTCCGTATCATGATCCTGGTTGTGGCTGCG
GAGAGAGTCTGGGGAGATGAACAGCAAGATTTTGTCTGCAATACGCTTCAACCTGGGTGT
AGAAATGTTTGCTATGATCACTTTTCCCCATCTCTCACATCAGACTCTGGGCCCTGCAG
CTGATCTTCGTCTCTACCCCTGCACTGCTGGTGGCCATGCACGTGGCTTACACCAGGCAC
GAGAGGAAAAGGCGATTAGAAAGTGGTGATAAAATTAATATTGAAGAGCTGAAAAATGAA
AAGATCCACATTCGAGGTCCCCTGTGGTGGACGTATACCTGCAGCATCTTCTTCAGGATC
GTCTTTGAAGCAGTCTTCATGATCGTGTTCTACTACATGTACGACGGGTACCATGATGCT
CGTCTGGTGAAGTGTGATGCGTGGCCCTGCCCAATGTCGTGGATTGTTTCGTGTCTCGG
CCTACTGAGAAAACCACTTTACTATTTTATGCTTGCTGATCTGGGATCTGCATGATG
TTGAATCTGGCTGAGTTGTTACTTAGTGATAAAAGTTTGCCTGAAAGATTCTGGGAAA
ACAACAGTTTTTAAACAAACGAGGAAGGAAGAGAGACTTCAGTTCTTTTGTGTTGACCCG
CTGCCCCCTGGCAGTTCGCTGTCTCAGCATTTAAATATCTCTAGTGCCATCTTCCGTAGC
ACAGCGGCATGTCAAGCGAAGTTCTTACCGTTCCTGACGACGAGGGCGGTTGCCCAA
CGCACTGAAGTACAGCAGTCCTTTACTCGGGGCACTTGCCTCTCTGCTTCTGAGTCTT

CTGCAGCCATCTCAGAGATTTGTTGACAGTCCAAAACCTGTTCTTCCCTTCTAGAAGGGAA
TTCTTGGAAAGGAATTTATTAGTGGATAGACGTAGTCTGATGTGGGCACCCACAGCATC
AGATTATTCTATTTTGAAGATATTGAATTCTATCTTAAAAACCTTCTGTGTCAAACCC
AACCTCTCCGGCCGCTGCGCGCGGGTACGCGGGGCCGCCACTGGCGCGCGTGTCTCCGC
CCCGCCAGGGCCGGGTGCCCGCGCCGCCAAAGAGCCGTGCCGGGGCGCGCCGTGGGC
AGCGCCCGGCTGAGGGGAGCCGGGCTGCCGGCAGGGAGCGCCCGCCCGCGGGCACCC
GCAGCTCTCCCGCTCCCGAGCGAAGCTTTGCGGGCTTGTTCCTCCAGAGTTAAACGA
GCTATGAGTGAACGGTGTCCGGACTACCCGAACGAGCTGAGCTTTTGCTACTGCCGATA
GCACTGTCTCCCGCAACGCGTAAGTAGAGCGCCCGGGCGGGTGCCTGCTGGGGGACGG
AGCGGGCAGCTCAGCCCTCAGCCGTGCCCGTGTCTAGGGGGGCATCGCGGGCCGGGT
CGGTCCCTGCGCTACCCGGCGGGACGGAGTCTGGCACGGCGAGAAGCGCCAGACCTCCC
CCAGCACCTCGGGAGTGCAGCGGAGGGACGCACCGGGTGTCTCCGGGATGGTCACGGCTGC
TGCCCGCCCACTGTGCACGTGTAGGGATGGCATTCTCTCCAAAGGGTGCCACTCAC
ATTGA

>GgCx31sim XM_425641 Our suggested modification

ATGGATTGGGGAACCTCTGCAGGCTGTTTTGGGAGGTGTAAATAAACACTCCACCAGCATT
GGGAAGATATGGCTCACAGTCTCTTCCGTATCATGATCCTGGTTGTGGCTGCG
GAGAGAGTCTGGGGAGATGAACAGCAAGATTTTGTCTGCAATACGCTTCAACCTGGGTGT
AGAAATGTTTGTCTATGATCACTTTTCCCATCTCTCACATCAGACTCTGGGCCCTGCAG
CTGATCTTCGTCTCTACCCCTGCACTGCTGGTGGCCATGCACGTGGCTTACACCAGGCAC
GAGAGGAAAAGGCGATTTAGAAGTGGTGATAAAATTAATATTGAAGAGCTGAAAAATGAA
AAGATCCACATTCGAGGTCCTGTGGTGGACGTATACCTGCAGCATCTTCTTCAGGATC
GTCTTTGAAGCAGTCTTCACTGTACGTGTTCTACTACATGTACGACGGGTACCAGATGCCT
CGTCTGGTGAAGTGTGATGCGTGGCCCTGCCCAATGTCGTGGATTGTTTCTGTCTCGG
CCTACTGAGAAAACACCGTTTACTATTTTATGCTTGTGTATCTGGGATCTGCATGATG
TTGAATCTGGCTGAGTTGTGTTACTTAGTGATAAAAGTTTGCCTGAAAGATTCTGGGAAA
ACAACAGTTTTAAAGTAA

>GgCx31.1sim XM_425784 When comparing the XM sequence with the genomic sequence in Ensembl, it is evident that the XM sequence predicts an intron (4 first nt of exon 2 marked in dark green/white letters). By extending exon 1 by 154 nt into the predicted intron (marked in blue), a stop codon is apparent.

ATGAACCTGGTGGTGTGTTGAGGGGCTGCTCAGTGGGGTCAACAGATACTCCACAGCCTTC
GGCCGCATCTGGCTCTCCCTCGTCTTCACTTCCGCTTCTGGTCTACGTGGTGGCAGCC
GAGCGCGTCTGGAGCGAGCCAGCAAAAGACTTCGACTGCAACACGCGGACGCCAGGCTGC
ACCAACGTCTGTTATGACCATTCTTCCCATCTCCACATCCGCTGTGGGCTCTGCAG
CTCATCTGGTCACTGTCCCTCGTGGTATCATGCACGTGGCTTACCGCCAAGCC
AAGGCGCTGCAGCACCGCGCGAGAGGGGGACAGTACCGCTCCCGCTTCCCAACCCC
GGCAAGAAGCGAGGCGGGCTGTGGTGGACGTACCTGTTCAAGCTCATCGTCAAGGCTGGC
GTGGACATGGTTTTCTCTACATCTTCTACCACATCTACAGGAACACACCCTGCCCGC
TTGGTGAAGTGCAGCTGCCCGCTGCCCAACATCGTGGACTGCTTCATCTCCCGGCC
ACTGAGAAGACTATATTCACCCTCTTATGGTGGTCAACCTGCATCTGCATCGTGCTC
AGCTGGTAGAGGCCGCTACCTGATTGGCAAGCGGTGTGGGAGCGCTGCAGGCGGC
ACTGGGAGAGCCACTGCCACAGCCATGACTGTGCCGAGTCTGCAAGCACAGGGGGACAG
GTGTTCCATGGGGTGGGATGGGACCGTCCACGACCCCATCTTGGGAGCCTGGGGTGGC
CCAACCATGGAGACGGGACTCGGCCTCTCTCTCAAGGCAAGGAGTTTGCATCGTT
CTGGTGGATTGGGAGGGGACGAAGGGCTGTGCTGCCCTGGAAAGGCACAGAGCAGCT
CTGGAGCCGGTTGTGAGCGATGCTGCTTTGGTACAGCAGCCACAGCCACGAAGCAAAAG
GGTGAGGAGCCTGCTGTGGGCTGGAGAAAAACACCCGAGCAAAAGCAGCAAATAAAGGAG
AGGAAAAACACCCAGCCATTGATTAA

>GgCx31.1sim XM_425784 Our suggested modification

ATGAACCTGGTGGTGTGTTGAGGGGCTGCTCAGTGGGGTCAACAGATACTCCACAGCCTTC
GGCCGCATCTGGCTCTCCCTCGTCTTCACTTCCGCTTCTGGTCTACGTGGTGGCAGCC
GAGCGCGTCTGGAGCGAGCCAGCAAAAGACTTCGACTGCAACACGCGGACGCCAGGCTGC
ACCAACGTCTGTTATGACCATTCTTCCCATCTCCACATCCGCTGTGGGCTCTGCAG
CTCATCTGGTCACTGTCCCTCGTGGTATCATGCACGTGGCTTACCGCCAAGCC
AAGGCGCTGCAGCACCGCGCGAGAGGGGGACAGTACCGCTCCCGCTTCCCAACCCC
GGCAAGAAGCGAGGCGGGCTGTGGTGGACGTACCTGTTCAAGCTCATCGTCAAGGCTGGC
GTGGACATGGTTTTCTCTACATCTTCTACCACATCTACAGGAACACACCCTGCCCGC
TTGGTGAAGTGCAGCTGCCCGCTGCCCAACATCGTGGACTGCTTCATCTCCCGGCC
ACTGAGAAGACTATATTCACCCTCTTATGGTGGTCAACCTGCATCTGCATCGTGCTC
AGCCTGGTAGAGGCCGCTACCTGATTGGCAAGCGGTGTGGGAGCGCTGCAGGCGGC
ACTGGGAGAGCCACTGCCACAGCCATGACTGTGCCGAGTCTGCAAGCACAGGGGGACAG
GTGTTCCATGGGGTGGGATACAAAGCCCCACAGCCTCCATCCCCAGTGCCTCCATCCCT
ACAACCTCCATCCCCACTGCCTCCATCCCTACAACCTCCATCCCCACCGCTCCATCCCT
ACAACCTCCATCCCCACCACTGCCCCGAGCAGGAGAAGACTCCTTCTAG

>GgCx32 Gg05541 Ensembl prediction. Runs into Ns at the 5'-end. NM_204371 contains the full-length sequence. The 156 5'-end nt are taken from the NM sequence (marked in green).

TGCAACACGCAGCAGCCGGGCTGCAACAGCGTGTGCTACGACCACTTCTTCCCATCTCT
CACATCCGCTGTGGGCCCTGCAGCTCATCTGGTCAACAGCGCGGCTGTCTGTGGCC
ATGCACGTGGCTACCAAGCAGCAGCAAGAGAGAGAAAGCTGCTGGTGTATGACGGGCCACGGG
GACGCCAAGCACCTGGAGGAGGTGAAGAAGCACAAGATGCGCATCGCGGATCGCTGTGG
TGGACGTACGTCTGCAGCGTGGTCTTCCGGCTGATCTTCGAGGCCGTCTTCATGTACCTC
TTCTACATGATCTACCCGGCTACCAAGATGGTGGCTGGTCAAGTGCAGGGCTACCCC
TGCCCCAACACCGTCTGACTGCTTCTGCTCTCGCGGCCACCGAGAAGACCATCTTACCGTC
TTCATGCTGGTCACTCCAGCGTCTGCATCGTCTCAACATGGCCGAGTTGGTCTACCTG

GTGGTGCGGGCTGCACGCGCCACAGCCACCACAACCTCAACCCTTCCTCGGGGAAAGGC
TCCTTCTACGGCCACAAGATCTCCTCTGAGTACAAGCAGAACGAGATCAACCAGCTGCTG
AGCGAGCAGGACGATCCCTCAAGGACATGCTGCGCCGAACCCGGCCTGCAGGAGAAG
GGCGACCGCTGCTCCGCCTGCTGA

>GgCx32 Gg05541 Modified according to NM_204371

ATGAACTGGGCGGCTCTACGCGGTGCTGAGCGGGGTGAACCGCCACTCCACCGCCATC
GGGCGCATCTGGCTCTCCGTCATCTTCATCTTCAGGATCATGGTGCTGGTGGTGGCGGGC
GAGAGCGTCTGGGGGACGAGAAGTCCGCCCTTACCCTGCAACACGCAGCAGCCGGGCTGC
AACAGCGTCTGCTACGACCACTTCTTCCCATCTCTCACATCCGCCTGTGGGCCCTGCAG
CTCATCTCTGGTCAACACGCCGGCCCTGCTCGTGGCCATGCACGTGGCCCTACAGCAGCAC
CAGGAGAAGAAGCTGCTGGTGATGACGGGCCACGGGGACGCCAAGCACCTGGAGGAGGTG
AAGAAGCACAAGATGCGCATCGCGGGATCGCTGTGGTGGACGTACGTCTGCAGCGTGGTC
TTCCGGCTGATCTTCGAGGCGCTCTTCATGTACCTCTTCTACATGATCTACCCCGGCTAC
CAGATGGTGCGGCTGGTCAAGTGCGAGGCGCTACCCCTGCCCAACACCGTGCAGTCTTC
GTCTCGCGGGCCACCGCAGGAAGACCATCTTACCCTCTTCATGCTGGTCACTCCAGCGTC
TGATCTGCTCAACATGGCCGAGTTGGTCTACCTGGTGGTGCGGGCCTGCACGCGCCAC
AGCCACCACAACCTCAACCCCTTCCTCGGGGAAAGGCTCCTTCTACGGCCACAAGATCTCC
TCTGAGTACAAGCAGAACGAGATCAACCAGCTGCTGAGCGAGCAGGACGGATCCCTCAAG
GACATGCTGCGCCGAACCCCGCCTGCAGGAGAAGGGCGACCGCTGCTCCGCCTGCTGA

>GgCx36 Gg09845 Ensembl prediction. No modification

ATGGGGGAATGGACTATTCTAGAGAGGCTACTGGAAGCTGCCGTGCAGCAGCACTCTACT
ATGATAGGGAGGATCCTGCTACCGCTGCTGGTGATCTTCAGGATCCTCATTGTGCCATT
GTAGGGGAGACGGTGTACGATGACGAGCAAACTATGTTTGTCTGTAATACGCTGCAGCCA
GGCTGCAATCAGGCTTGTACGACCAAGCTTTCCCTATTCTCACATAAGGTACTGGGTG
TTCCAGATCATCATGGTGTGCACTCCCAGCCTGTGCTTCATAACGTACTCTGTTACACAG
TCTGCTAAGCAGAGGGAACGAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTA
GATTCAATGAAGCGTGAGGACAGTAAGAAAATCAAGAACAGGATTGTCAATGGGGTGCTG
CAGAACACTGAGAAGTCCACCAAGAGGAGCAGAACAGACTGCTTAGAAGTGAAGGAGATC
CCCAATCCTGCGATCAGAAGTACAAAATCAAAGATGAGGAGGCAAGAAAGGCAATTTCTCGA
TTTATATCATCCAAGTGGTCTTTGAAATGCCCTAGAGATTGGATTCTTAGTCGGGCAG
TATTTTCTGTATGGATTCAATGTCCCTTCCATGTACGAATGTGACAGATACCCTTGCATT
AAAGAAGTAGAGTGCTATGTCTTAGACCCACTGAGAAGACTGTGTTCTTGGTGTTCATG
TTTGCTGTCAAGTGGCATTGTGTGGTGTCAATTTGGCAGAACTGAACCACTTGGGCTGG
AGAAAGATCAAAATGCGAGTGAGAGGAGTACAGGCAAAAAGGAAATCCATATACGAAATC
AGAAATAAGGACCTGCCAAGAATGAGCATGCCTAACTTTGGCAGGACTCAGTCAAGTGAC
TCAGCTTATGTGTA

>GgCx39 Gg02532 Ensembl prediction in 2 exons (black/blue letters). Modified according to
MN_206984. The 16 nucleotide intron regarded as a part of the reading frame (marked blue).
154 nt added at the 3'-end to reach stop codon (marked blue).

ATGGGCGACTGGGGATTCTGGACCATAAGCTGGACCAAGTCCAGGAGCACTCCACCGTG
ATCGGGAAGATCTGGCTACCGTCTCTTCATCTTCCGTATCCTCATCCTGGGCTTGCT
GGTGAGTCCGTGTGGGGGACGAGCAGTCCGACTTCGTGTGCAACACCAAGCAGCCGGGC
TGACCAATGTCTGCTATGACAAAGCCTTCCCATCTCCACATCCGCTACTGGGTGCTG
CAGTTCTCTTTGTGACACCCCGACCCCTCATTTACCTGGGCCACGTTGTCTATCTCTCG
CGGAAAGAGGAGAAGCTGAAACAGCAGGAGAATGAGCTCCGGGCTATTACAGCAAGGAC
CCAAAGATTGAGCAAGCATTTGGCAGCTGTAGAAAAGAAGATGTCCAAGATCTACGTGACA
GAAGATGGGCGGCTCAAGATCCGAGGGGCACTGATGTGGACGTACGTATCAGTGTGATC
TGCAAGAGCATCTTTGAGGCAAGCTTCCCTCGTTGGCCAGTGGTACCTGTATGGCTTCTCC
ATGGTGCCCCGCTACGTGTGACGAGGAGGCCCTGCCCCACCAAGTGGGACTGCTTCATC
TCCGCCCCACCGAGAAGACCATCTTCATCATCTTCATGATGGTGATGGGCTCATTTCC
TTAATCTCAACCTCCTGGAGCTCTTCCACCTCTGCTGCAATCAAGAAAGGTGCCGGTGC
CGGTTGGCCGAGCCAGAGTGCCCTTGTGATGACATGGGCTCGGGTCCCTACCCAGCCA
AGCATTACCCCTTCTGCCATGGCTGAGAGCCACACACCGCCCTACCAGGCTTACAACA
AGCTCTCCAGCGAGCAGAACTGGGCCAACT

>GgCx39 Gg02532 Ensembl prediction. No modification

ATGGGCGACTGGGGATTCTGGACCATAAGCTGGACCAAGTCCAGGAGCACTCCACCGTG
ATCGGGAAGATCTGGCTACCGTCTCTTCATCTTCCGTATCCTCATCCTGGGCTTGCT
GGTGAGTCCGTGTGGGGGACGAGCAGTCCGACTTCGTGTGCAACACCAAGCAGCCGGGC
TGACCAATGTCTGCTATGACAAAGCCTTCCCATCTCCACATCCGCTACTGGGTGCTG
CAGTTCTCTTTGTGACACCCCGACCCCTCATTTACCTGGGCCACGTTGTCTATCTCTCG
CGGAAAGAGGAGAAGCTGAAACAGCAGGAGAATGAGCTCCGGGCTATTACAGCAAGGAC
CCAAAGATTGAGCAAGCATTTGGCAGCTGTAGAAAAGAAGATGTCCAAGATCTACGTGACA
GAAGATGGGCGGCTCAAGATCCGAGGGGCACTGATGTGGACGTACGTATCAGTGTGATC
TGCAAGAGCATCTTTGAGGCAAGCTTCCCTCGTTGGCCAGTGGTACCTGTATGGCTTCTCC
ATGGTGCCCCGCTACGTGTGACGAGGAGGCCCTGCCCCACCAAGTGGGACTGCTTCATC
TCCGCCCCACCGAGAAGACCATCTTCATCATCTTCATGATGGTGATGGGCTCATTTCC
TTAATCTCAACCTCCTGGAGCTCTTCCACCTCTGCTGCAAGAGCCTGCTCAGCAGCATC
AAGAAGGTGCCGGTGCCGGTTGGCCCGAGCCAGAGTGCTTTGCTGATGACATGGGCTCG
GGTCCCTACCCAGCCAAGCATTACCCCTTCTGCCATGGCTGAGAGCCACACACCGCCC
TACAGGCTTACAACAAGCTCTCCAGCGAGCAGAACTGGGCCAACTTCCGCAACGAGGAG
AACCTGGCACTGGGCAAGCGGCCACAGGCCCTCTCGGACCCCTACGCCCCAGGGCTCCA
GATGCCCCCATCCCTGAGGAGAAGCAGTGAAGCGGCCAGCAGCTCAGCATCAAGAAG
CAGTATGTTAG

>GgCx40.1 Gg07065 Ensembl prediction. Named according to XM_418567, which predicts intron starting approx at TM3/EC2. Ensembl prediction does not reach a stop codon. By extending the reading frame with 288 nt at the 3'-end (marked in blue), a stop codon is reached.

ATGGAGCACTGGGACTCGCTGGGATTTTTGGTTGTCACTCTCAACTATAATGTGACAATT
ATAGGAAAGATCTGGCTAATGCTGGTGGTTCTGCTGCGAATGGCAGTGGTGGTGTAGCA
GGCTATCCGCTCTATCAGGACGAACAGGAGCGCTTCGTCTGCAACACCCTGCAGCCTGGA
TGCTCCAATGTTTGTCTAGCTTGTCTCTCCCATATCTCACTTCAGGTTCTGGCTCATT
CACACTGTGTCTATCTTGTACCTTACACTGCATTGAGCATTATGTTTTGCACAAGGTA
GCAACGCACATTGTAAGAAAGCACTGTTTGGTGTATGGATGCAGAGGGAATAAGGGTTTA
TCAAGTCCCAGAGACCTGAAGGAACTCTGTAGGAGTGATCTTGTCATAGATTGGATTGT
GATGTGGACAGCCTAAGTGTCTCAATTTTTCTGGGGCATATACTGTTCACTATTTTTTT
AGGACACTGATTGAGGCTGCTTTTGCAGCTGCACAGTATTTCTCTTTGGATTTTTTGT
CCCGAGCGCTTTTCTGTACCATTCACCTTGTACAAGCACAGTTGATTGCTACATCTCC
CGGCCCCTGAGAAATCCATCATGATGGTTTTCTGAGTGTCAACAGTCTGTCTTT
CTGCTCAGCCTTTGCTAGCTTGTCTGTCTCTCGGCAAATGACAGCAAGAAGCCAGAGG
AACAAGCAGCTGGCCAACCCCCACAGAGAGAATGAGGGCCTTTAAATAATCCCCAGCA
CAGTCTGGTAGCTCTTCCCCACCCCAACTCAAGACTGCCAGCATTGAATGGCAGCCAG
ACCAAGTGACGGCTCTGCTCACTTCTCTCTGAAGAGGAAGAGGAGGCTACTCTTCATCCT
GAAGAGGTCTCTCAGCAAGCTGCCAGCACTAACCTTAACAGCAACAGCCATAAGCCTTGT
GCG

>GgCx40.1 Gg07065 Our suggested modification

ATGGAGCACTGGGACTCGCTGGGATTTTTGGTTGTCACTCTCAACTATAATGTGACAATT
ATAGGAAAGATCTGGCTAATGCTGGTGGTTCTGCTGCGAATGGCAGTGGTGGTGTAGCA
GGCTATCCGCTCTATCAGGACGAACAGGAGCGCTTCGTCTGCAACACCCTGCAGCCTGGA
TGCTCCAATGTTTGTCTAGCTTGTCTCTCCCATATCTCACTTCAGGTTCTGGCTCATT
CACACTGTGTCTATCTTGTACCTTACACTGCATTGAGCATTATGTTTTGCACAAGGTA
GCAACGCACATTGTAAGAAAGCACTGTTTGGTGTATGGATGCAGAGGGAATAAGGGTTTA
TCAAGTCCCAGAGACCTGAAGGAACTCTGTAGGAGTGATCTTGTCATAGATTGGATTGT
GATGTGGACAGCCTAAGTGTCTCAATTTTTCTGGGGCATATACTGTTCACTATTTTTTT
AGGACACTGATTGAGGCTGCTTTTGCAGCTGCACAGTATTTCTCTTTGGATTTTTTGT
CCCGAGCGCTTTTCTGTACCATTCACCTTGTACAAGCACAGTTGATTGCTACATCTCC
CGGCCCCTGAGAAATCCATCATGATGGTTTTCTGAGTGTCAACAGTCTGTCTTT
CTGCTCAGCCTTTGCTAGCTTGTCTGTCTCTCGGCAAATGACAGCAAGAAGCCAGAGG
AACAAGCAGCTGGCCAACCCCCACAGAGAGAATGAGGGCCTTTAAATAATCCCCAGCA
CAGTCTGGTAGCTCTTCCCCACCCCAACTCAAGACTGCCAGCATTGAATGGCAGCCAG
ACCAAGTGACGGCTCTGCTCACTTCTCTCTGAAGAGGAAGAGGAGGCTACTCTTCATCCT
GAAGAGGTCTCTCAGCAAGCTGCCAGCACTAACCTTAACAGCAACAGCCATAAGCCTTGT
GCGTCAAGGAGATCTTGCTGTTAAGCAAGATAGCATTGAAGAACCCTGTGTGCTGATGAC
CAGCAAGGAAGCTGATGAGCAAGTGAAGCTGGGCTCCAGCAAGAGCTTTGTTAAAGAT
ACTGCCCTGACTTTGAGACCTCAGATCAAGTCTCACCTTGGAGGTTCTTCTCTGTAGTT
CACAGCAAGCTTTTAGGATACTACCTTTCAGCTGAGCTAAAAAATCCCGATGCACAATCA
AATTATAGCAGTACTAGTTGCTTGAGATCGAAAAAGTCTGAATGGGTATAG

>GgCx42 Gg15487 Ensembl prediction in 3 exons (black/blue/black letters). Modelled according to NM_205504. Predicted introns of 30 nt and 45 nt regarded as part of reading frame (marked in blue).

ATGGGGGAGTGGAGTTTCTGGGAGAGTTTCCTTGAGGAGGTCCACAAGCACTCCACAGTG
GTGGGGAAGGTCTGGTTGACTGTGCTGTTTCATCTTCCGGATGCTGGTGTGTTGACAGCA
GCTGAGTCTCTTGGGGGATGAGCAGTCCGATTTTCATGTGTGACACCCAGCAGCCTGGC
TGCGAAACAGTCTGCTACGACAAAGGCTTCCCCATTTCCACGTCCTGCTTTGGGTCCTT
CAGATCATCTTTTGTCTCACTCCATCCCTGGTGTATATGGGCCACGCGATGATACGGTG
CGCATGGAGGAGAGAGGAAGATGAAGGAGGCAGAAAGAGAGGCCAGGAGCAGAAAGTGC
CCCGTGGCAGAGAAGACTGAGTTGTCTTGTGGGATGAGTCAGGAGGCAAAATTATACTC
AGAGGCACTGCTGCTAAACACCTATGCTACAGTATTTTATTGCTGCACTGCCATGGAAATC
GCCTTCATTGTGGGCAATACATCTTGTATGGGATCTTCTGGAGACCCTGTATATCTGC
CAGCGGGGACCTTTGCCCCACCCCTGTCAACTGCTACGTTTCCCGTCCCACAGAGAAGAAC
GTGTTTCATCATCTTCATGCTTGTGTGGCTGTGCTCTCCCTGTTTCTCAGCCTGGCTGAG
TTATACCAGCTGGGCTGGAAGAAAGCCAAAGAGAGGTGCTCCCGTCTTACAAGCCAGC
CTAGCACAGCTCCTCGAAGGCTGGAGTCCGCCCCACAGGTAGAAAGGGCCAGATGTAC
ACTCCCCCACCAGATTTTAATCAGTGTGCAAGTCCCAATGGGAAGTTTCATCAGCCCC
TTCAGCAACAAGATGGCCTCGCAGCAGAACACAGCCAACCTTGGCACTGAGAGGTCCAGC
TACATGGAAAGCCAGAGGTGGCCAGTGAATGTGCGGCCACCCGCCCTCCTGAGAGCTAC
TTCAATGAGAAACGCCGCTTCAGCAAGGCCAGCCGGGCCAGCAGCAAGGCCAGGTGAGC
GATTTGTGCGGTGTGA

>GgCx42 Gg15487 Ensembl prediction.

ATGGGGGAGTGGAGTTTCTGGGAGAGTTTCCTTGAGGAGGTCCACAAGCACTCCACAGTG
GTGGGGAAGGTCTGGTTGACTGTGCTGTTTCATCTTCCGGATGCTGGTGTGTTGACAGCA
GCTGAGTCTCTTGGGGGATGAGCAGTCCGATTTTCATGTGTGACACCCAGCAGCCTGGC
TGCGAAACAGTCTGCTACGACAAAGGCTTCCCCATTTCCACGTCCTGCTTTGGGTCCTT
CAGATCATCTTTTGTCTCACTCCATCCCTGGTGTATATGGGCCACGCGATGATACGGTG
CGCATGGAGGAGAGAGGAAGATGAAGGAGGCAGAAAGAGAGGCCAGGAGATGAAAAAC
AGTGGTGACACATACTATCAGCAGAAAGTGCCCGTGGCAGAGAAAGACTGAGTTGTCTTGC
TGGGATGAGTCAGGAGGCAAAATTATACTCAGAGGCAGTCTGCTAAACACCTATGTCTAC
AGTATTTTATTGCTGCACTGCCATGGAAATCGCCTTCATTGTGGGGCAATACATCTTGTAT
GGGATCTTCTGGAGACCCTGTATATCTGCCAGCGGGACCTTGCCCCACCCCTGTCAAC
TGCTACGTTTCCCGTCCCACAGAGAAGAACGTGTTTCATCATCTTCATGCTTGTGTGGCT
GTGCTCTCCCTGTTTCTCAGCCTGGCTGAGTTATACCACCTGGGCTGGAAGAAAGCCAAA

GAGAGGTGCTCCCGTGCTTACAAGCCCAGCCCTAGCACAGCTCCTCGAAGGCTGGAGTCC
GCCCCACAGGTAGAAAGGGCCAGATGTACACTCCCCACCAGATTTTAAATCAGTGTTTG
GCAAGTCCCAATGGGAAGTTTCATCAGCCCTTCAGCAACAAGATGGCCTCGCAGCAGAAC
ACAGCCAACTTTGCCACTGAGAGG**GTCCACAGCCAGGAGGATGCTGCCGGGAAGGGCCC**
TTTCATGAAGTCCAGCTACATGGAAAGCCAGAGGTGGCCAGTGAATGTGCGGCACCCGCC
CTCCCTGAGAGCTACTTCAATGAGAAACGCCGCTTCAGCAAGGCCAGCCGGGCCAGCAGC
AAGGCGAGGTCAGACGATTTGTGCGGTGTA

>GgCx43 Gg14873 Ensembl prediction. No modification

ATGGGTGATTGGAGTGCCCTGGGAAAACCTTCTTGACAAGGTTCAAGCTTATTCTACTGCA
GGAGGAAAAGTGTGGCTGTCTGCTCTTTATTTCCGAATCCTGCTATTGGGGACAGCA
GTGAGTCTGCTTGGGGAGATGAACAGTCTGCTTTTCGGTGCAATACTCAACAGCCTGGT
TGTGAGAACGTCTGCTATGACAACTCTTCCCATCTCCACAGTGCCTTCTGGGTTCTG
CAGATCATATTTGTGTCTGTACCAACCCCTCTTGACCTGGCGCACGTGTTCTACGTGATG
AGGAAAGAGAGAAGCTGAACAAAAGAGAGGAAGAGCTCAAGGTGGTCCAGAATGATGGT
GTGAATGTGGATATGCACCTCAAACAAATAGAAATTAAGAAATTCAGTATGGAATTGAA
GAGCACGGCAAGGTAAAGATGCGTGGAGGACTGCTTCGTACT**TACATCATCAGCATCCTG**
TTTAAGTCTGTCTTCGAGGTGGCTTTCTGCTGATACAGTGGTACATTTATGGGTTTAGC
CTGAGTGCCATGTACACCTGTGAGCGAGACCCATGCCCTCACAGGGTGGACTGTTTCCTG
TCCCGTCCAACAGAGAAAACCATTTTCATCGTCTTCATGCTGGTAGTGTCTTTGGTGTCT
CTTGCCCTTGAACATCATTGAGCTTTTCTACGTGTTCTTCAAGGGTGTCAGGATCGTGTG
AAAGGAAAAACCGACCCCTACTCCACAGTGGCAGCATGAGCCCTTCAAGGACTGTGGT
TCCCCCAAATAGAGCTGTACCAATGGCTGTTCTCACCAGCCGCCCTTGTCTCCCATG
TCTCCCCGGGGTACAAGCTTGTACCGGAGACAGGAACAATTCCTCTGTCTGAACCTAC
AATAAGCAAGCCAGCGAGCAAACTGGGCCAATACAGTGCAGGAGCAGAATAGGATGGGG
CAGGCTGGCAGCACCATCTCAACTCTCATGCCAGCCCTTCGACTTCGCCGATGAGCAC
CAGAACACGAAAAAAGCTGGCGTCGGGACACGAGCTGCAACCCCTCACCATTGTGGACCG
AGGCTCCAGCAGAGCCAGCAGCCGAGCCAGCAGCAGGCTCGACCTGATGACCTGGAG
ATCTAA

>GgCx45 Gg00997 Ensembl prediction. No modification

ATGAGTTGGAGCTTTCTGACCCGTCTATTAGAGGAGATCCACAACCACTCCACCTTCGTT
GGCAAAATCTGGCTGTGCGGTGCTGATCGTCTTTCGGATCGTTCTGACTGCTGTGGGAGGC
GAATCCATTTATTACGATGAGCAAAAGCAAGTTTGTCTGCAACACAGAGCAGCCCGGCTGT
GAGAACGTTTGTGCTGAGCTGCTTTTGTCTCTCTTCTCACGTGAGATTTTGGGTGTTCCAG
ATCATTCTCGTCGCCACTCCATCCGTCTGTATTTGGGCTACGCCATCCATAAGATTGCC
CGCATGGTGGAGCACAGTGACGTTGACAGAAGGTTCAGAAGCAAAAGCTTTCCACGCGC
TGGAAACAGCACCGCGGCTTGGAGGAGGCCGAGGATGACCACGAGGAGGACCCGATGATG
TACCCAGAGATAGAGCTGGAAGCGAAGCGGAGAACAAAGGAGCAGCAGCCCGCCCAAG
GCCAAGCAGCAGCGGCGGCGGCGGATCCGCGAGGACGGGCTGATGAGGATC**TATGTGCTG**
CAGCTGCTGGTGAGGGCGACGTTTCGAGGTCGGCTTTCTGATCGGTCACTACCTCCTGTAT
GGCTTTGAGGTGAGCCCGTGTGTTGTGTGCAGCAGGAAGCCCTGCCCCATAAGATAGAT
TGCTTCATTTCAAGCCGACCGAAGAACCATTTTCTGCTCATAATGTACGGGGTGAGC
TGTATGTGTTTGTGTTGAATGTGTGGGAGATGCTCCATTTAGGGTTTGGCACGATCCGG
GACACGTTGAACAACAAGAGGAAAGAGCTGGAGGACTCTGGTACCTACAACCTACCCCTT
ACCTGGAACACACCCCTCTGCCCTCCTGGCTACAACATTGCGGTCAAACCCGATCAGATG
CAATACACAGAAGTGTCAACGCCAAGATGGCTACAAGCAGAACAAAGCCAACATCGCC
CAGGAGCAGCAGTACGGCAGCAACGAGGAGAACATCCCGCCGACCTGGAAAACTGACG
AGGGAAATCAAAGTGGCTCAGGAACGCCTGGACATGGCCATCCAGGCGTACAACAACCAA
AACAACCCCGGACGAGCTCCAGAGAGAAGAAGTCCAAGGCGGGCTCCAACAAAAGCAGC
GCCAGCAGCAAATCGGGGACGGGAAGAACTCTGTCTGGATTTAA

>Gg05331 Ensembl prediction. No modification

ATGACCAACATGAGCTGGAGCTTCTCACCCGCTGCTAGAAGAGATCCACAACCACTCC
ACCTTCGTGGGGAAGTCTGGCTCACTGTGCTCATCGTCTTCGCATCGTCTTAACGGCG
GTGCGTGAGAAATCCATCTATTCTGATGAGCAAAAGCAAGTTCACTTGCAACACCAAGCAG
CCCGGCTGCGACAACGTCTGTTACGATGCCTTCGCCCGCTGTACACGTTAGGTTTTGG
GTCTTCCAGATCATGATATCCACGCTTCAGTCATGTACCTGGGCTATGCCATCCAC
CGGATCGCCCGGTGAGCGAGGAGGAGAAGAAGTTCAAGGGCTTCAAGAAGAAGAAGCAG
TTGCTTTGAACTGGCAGGCGGTGCGCAACATGGAGGATGCGATGGAGGCGGATGAAGAG
GAGCCCATGATCTCCGATGACGCAGCAGAAAATGAGAAAGCCAAAGCCAAGCCAAAGAGC
AAAGAGCAGCAAAAGCAGATGGGAGGAGGCGAATCCAACAGGAAGGACTGATGAAAATC
TACGTCTTCCAGCTGCTTACCAGAGCTTCATTTGAAGTTTGTGTTTGTATAGGGCAGTAT
TTGCTCTATGGTTTTGAGGTAGAAAGCTTATTTTGTCTGCAACAGACTGCCTTGTCTCAC
ACGGTGGACTGCTTTGTGTCCCGGCCAACAGAGAAGACAATCTTCTGCTGGTGTATGTAC
GTGGTGAGCTGTCTGTGCTTGTGCTGAACATGTGCGAGATGTTCCATCTGGGCTTCGGG
ACCATCCGAGATGCCATCCGCAACCGGAAAAATCAACAGCTTCAGGCAGCCCCCTACAAC
TACGCTTACCCTAAGACATCTCTGCCCGCCGAGTACAACCTGGTGGTGAATCAGAG
AAGTCCACCAAGATCCCCAACAGCTGTATGGCTCACGAGCAGAACTTGGCTAACGTTGCT
CAGGAGCAGCAGTGACACAGCCCGGATGAGAACATCCAGCAGACTTGTCCACCTCCAC
AAACACCTGCGAGTGGCCAGGAGCAGCTGGACATAGCATTTTCAAGTTATGGCAGCAGC
CAGGGCAATGCACAGCCCTCCCGGACCAAGCAGCCCTGCCTCGGGTGGCAGGGTGGTAGAG
CAGAATAGGGCCAACACCGCCAGGAGAAGCAAGCGTAAGCCCAAGGCGTCTTGGAG
AAAGGCAGCTCCAGCAGTAAGGATGAAAGACTTCTGTGTGGATATAG

>Gg11233 Ensembl prediction only includes first conserved domain. The sequence is extended by 583 nt according the genomic sequence, but then runs into a row of Ns. However, a reasonable

second conserved domain (green) is included in the extended sequence. The 902 5' nt then correspond to the 5'-part of XM_422042 (Hypothetical protein MGC75965)

ATGAGCTGGAGCTTCTGACGCGGCTCCTGGAGGAGATCAACAACCACTCGACCTTCGTG
GGCAAGATCTGGCTTAGCGTCTCATCATCTTCAGGATCGTGTGACGCGCGTGGGGGGC
GAGTCCATCTACTACGACGAGCAGAGCAAGTTCGTCTGCAACACGACGAGCGCGGCTGC
GAGAACGCTCTGTACAACGCCTTCGCCCCGCTCTCCACGTCGCTTCTGGTTTTCCAC
ATCATCATGGTGGCCACCCCTCGGTGCTCTACCTGGGCTTCGCCATGCACCGCATCGCC
CGCATGCCCGAGTCTCGCGGCGCGGGCGCCGGCAGCTACGCGC

>Gg11233iv Our suggested modification

ATGAGCTGGAGCTTCTGACGCGGCTCCTGGAGGAGATCAACAACCACTCGACCTTCGTG
GGCAAGATCTGGCTTAGCGTCTCATCATCTTCAGGATCGTGTGACGCGCGTGGGGGGC
GAGTCCATCTACTACGACGAGCAGAGCAAGTTCGTCTGCAACACGACGAGCGCGGCTGC
GAGAACGCTCTGTACAACGCCTTCGCCCCGCTCTCCACGTCGCTTCTGGTTTTCCAC
ATCATCATGGTGGCCACCCCTCGGTGCTCTACCTGGGCTTCGCCATGCACCGCATCGCC
CGCATGCCCGAGTCTCGCGGCGCGGGCGCCGGCAGCTACGCGC
GTGCGGCGCGGAGCGGGGCGGACTACGAGGAGGCGGAGGACGATAACGAAGAGGACCCC
ATGATCTTCGAGGAGATCGAGGTGGAGAAGGAGAAGGGCGCCGAGGGCGCGAGAAAGCAC
GACGGCGCGCGCCGATCAAGCGGGACGGGCTGATGCGCGCC
TGCCGATGCTGGAGTCTGGGCTTCTCTTCGCGGAGTACCTGCTGTTCCGCTTTAAG
GTCAGCCCCCTCGTACGTGTGACGCCACAGCCCCCTGCCGACATGGTCGACTGCTTCGTC
TCGCGCCCCACCGAGAAGACCATCTTCTGCTCATCATGTACGCCGTCAGCGGGCTCTGC
CTCTTCTCAACATCTGCGAGCTCTGACCTGGGCGTG
GGGCGCATCCGCGACGCGCTG
AGCCAGCCCCACCGGCCCGCCGCGGAGCCCGCGCCGCACTACGCCAAGAAGCCCCC
AGCGCGCGCCACCTACCCTCCCTGAAGAAGGAGCCGCTGAAGACCCGATGACCGAC
GGCAAGATGGACTACAGGAGAACCTGNNN

>GgCx56 Gg17137 Ensembl prediction in 2 exons (black/blue letters). Predicted intron of 48 nt regarded as part of reading frame. Modified according to XM_425642.

ATGGGTGACTGGAGCTTTCTGGGGAGGCTGCTAGAGAACGCGCAGGAGCACTCCACAGTT
ATTGGCAAGGTTTGGCTGACAGTACTGTTTATCTTCAGGATACTGGTGTGGGGGCGCT
GCTGAGGAAGTCTGGGGAGACGAGCAGTCAAGCTTTACATGCAACACTCAGCAACCTGGT
TGCGAAATGTTTGCTATGACAAAGCCTTCCCCATTTCTATATCCGCTTCTGGGTGCTG
CAGATCATTTTTGCTCCACTCCAACCTCATATACCTGGGCCAGTGTGCACATTGTA
CGCAGTGTCAAAGACAACCACTACCCAGGAGCAGCAACGTCTGGTGGTGGTAGTGGAGGT
GGCAATAACTCTGAGGAGTCTTCTTCAAAATGGGAAAGGAGAAGCTTCGATCCGTGAT
GAGCGTGGTAGAATCCGTATGGGAGGTGCTCTGCTCCGTACCTACATCTTCAACATCATT
TTCAAGACACTGTTTGGGTAGGCTTCATTGTGGGCCAGTACTTCTTTATGGCTTTGAA
CTAAAGCCAGTCTACCACTGACGCCGTCACCTTGTCCACACACTGTGGACTGCTTCATC
TCACGGCCCACTGAGGAAGCAATCTTCATCATCTTCATGTTGGTGGCTCTGTCTCC
CTGCTGCTGAACATGCTTGGATATATCACTTGGGATGGAAGAAGCTTAAGCAGGGCATG
ACAAGCCAGTACAGCCTTGGATGCCTGTTACAACACTGACACCGGTATGGTAACAGGG
GAATCCAACACAGTTCCTTGGCACCAGCCAGCACCCTGTGGTGGTAACGACCACTGCA
CCTGCTCTGTTCTGACGCGGTGCTGTACACCGCTGCTGGCTCCAGTGACCATG
GCACCCTACTATGCTGCGGCTGCTCCTAGGACACGACCCCTTCCAACACTGCCTCCATG
GCAAGCTACCCTGTTGCTCCACAGTTCAGAGAACAGGACCGGTGCTGTACCCCTACA
CCGCTCTCCACTCCTGTACCATCCCAACGCCCATCCCAACCCACACCAAGCCATCATC
AACTACTTCAACAGCAAAAGCAATGCCCTGGCAGCCGAGCAGAAGTGGGTCAACATGGCA
GCTGAGCAGCAAGGCAAGGCACCTCCAGCTCAGCAGGTTCTCCACGCTAGCAGTGTG
CGGCATCCCTTCCAGAGCAGGAAGAGCCCTGGAGCAGCTTCTCCCACTCCAGCTGGG
CCGCCCATCACCAACAACAGTGGCAGCAGCAGCTTGGAGCGGAGCAAGTGGCAGC
AAATGGGATGTGGAAGGCGAGGAGGAGTGGCAGAGGAACGCCCATCTCAGCCACCTGC
ACCACCGTGGAGATGCATGAGCCGCGCTTACTTGTAGACACACGCGTTTGGAGAGGGCC
AGTAAGTCGAGCAGCAGCAGAGCCAGGTGATGACTTGGCAGTGTAG

>GgCx56 Gg17137 Modified according to XM_425642

ATGGGTGACTGGAGCTTTCTGGGGAGGCTGCTAGAGAACGCGCAGGAGCACTCCACAGTT
ATTGGCAAGGTTTGGCTGACAGTACTGTTTATCTTCAGGATACTGGTGTGGGGGCGCT
GCTGAGGAAGTCTGGGGAGACGAGCAGTCAAGCTTTACATGCAACACTCAGCAACCTGGT
TGCGAAATGTTTGCTATGACAAAGCCTTCCCCATTTCTATATCCGCTTCTGGGTGCTG
CAGATCATTTTTGCTCCACTCCAACCTCATATACCTGGGCCAGTGTGCACATTGTA
CGCATGGGAAGAGAAGAGGAAAAAGAGGAAGAGCTGAAAAAGAGGGGAAAGTGTCAAA
GACAACAACCTACCCAGGAGCAGCAACGCTCTGGTGGTGGTAGTGGAGGTGGCAATAACTTC
AAGGATCCTCTATCAAAATGGGAAAGGAGAAGCTTCCGATCCGTGATGAGCGTGGTAGA
ATCCGATGAGGAGTGTCTGCTCCGTACCTACATCTTCAACATCATTTCAGAGCACTG
TTTGAGGTAGGCTTCATTGTGGGCCAGTACTTCTTTATGGCTTTGAACTAAAGCCAGTC
TACCACTGACGCCGTCACCTTGTCCACACACTGTGGACTGCTTCATCTACGGCCCACT
GAGAAGACAATCTTCATCATCTTCATGTTGGTGGTGGCCTCTGCTCCCTGCTGCTGAAC
ATGCTTGAGATATACCTTGGGATGGAGAAGCTTAAGCAGGGCATGACAAGCCAGTAC
AGCCTTGAGATGCTGTTTACAACACTGACACCGGTGATGGTAACAGGGGAATCCAAACCA
GTTTCTTGGCACCGCCAGCACCCTGTGGTGGTAACGACCACTGCACCTGCTCCTGTT
CTGCTGACACGCGTGTGTACACCGCTGCTGGCTCCAGTGACCATGGCACCCTACTAT
GCTGCGGCTGCTCTAGGACACGACCCCTTCCAACACTGCCTCCATGGCAAGCTACCT
GTTGCTCCACCACTTCCAGAGAACAGGACCGGTGCTGTACCCCTACACCGCTCCACT
CCTGTACCATCCCAACGCCATCCCAACCCACACCAAGCCATCATCACTACTTCAAC
AGCAAAAGCAATGCCCTGGCAGCCGAGCAGAAGTGGGTCAACATGGCAGCTGAGCAGCAA
GGCAAGGACCCCTCAGCTCAGCAGGTTCTCCACGCCTAGCAGTGTCCGGCATCCCTT
CCAGAGCAGGAAGAGCCCTGGAGCAGCTTCTCCCACTCCAGCTGGGCGGCCATCAC
ACAACCAACAGTGGCAGCAGCAGCTTGGAGCGGAGCAAGTGGCAGCAATGGGATGTG

GAAGGCGAGGAGGAGCTGGCAGAGGAACGGCCCATCTCAGCCACCTGCACCACCGTGGAG
 ATGCATGAGCCGCCGTTACTTGTAGACACACGGCGTTTGAGCAGGGCCAGTAAGTCGAGC
 AGCAGCAGAGCCAGGTCAGATGACTTGGCAGTGTAG

>Gg39.9 XM_426260 When comparing the XM sequence with the genomic sequence in Ensembl, it is evident that the XM sequence predicts an intron (4 first nt of exon 2 marked in dark green/white letters). An analysis of the XM sequence suggests that exon 1 is the 5'-part of the first conserved domain, and exon 2 is the 3'-part of the second conserved domain plus the cytoplasmic tail. The Ensembl genomic sequence reveals that there is an apparent unsequenced area in-between the two domains, but the conserved domains can be extended by a few tens of nt (in green). The remaining unsequenced area include the cytoplasmic loop (marked in blue).

ATGGGGGACTGGAGCCTGCTGGGGAGGCTGCTGGAGAACGCCAGGAGCACTCCACGGTG
 GTGGGGAAGGTTTGGCTCACTGTACGCTTCGTCTTCCGCATCCTGGTGCTGGGGGCGGCC
 GCTAAACGAGTGTGGTACGCTTGTACGGGCTTGACCTGAAACCCCTTACACCTGCAGC
 CGCTGGCCCTGCCCAACACCGTCAACTGCTACATCTCCGGGCCACCGATAAGACCATC
 TTCATCTCTTTCATGCTGGGGGTGGCCTGCGTGTCTCTGCTGCTCAACCTGGTAGAGATT
 TACCACCTGGGCCCTACCAAGTGCCGGCAGCGGCCACGTCCCAGATCCCGCATCTGCC
 GGCCCTGGTGGCCTTGCGGGGTACACGGTGCCCGCTTACCTTGCCTGGTGGTGGAAC
 CCGGTGGATGCCAGCACCCACCTGCGACCCCTGGGGAAGGCTGGTGGCTGGGG
 GTGGCCAGCGGGTCCCGTGGGAAGGCGCGAGCTGGGGACCTGGCAGTGTGA
 >Gg39.9ic XM_426260 Our suggested modification.

ATGGGGGACTGGAGCCTGCTGGGGAGGCTGCTGGAGAACGCCAGGAGCACTCCACGGTG
 GTGGGGAAGGTTTGGCTCACTGTACGCTTCGTCTTCCGCATCCTGGTGCTGGGGGCGGCC
 GCTAAACGAGTGTGGTNGACGAGCTGGCCGGCTTCTCCTGCGACACGAGCAGNNNNNN
 NNN
 NNN
 NNN
 NNN
 NNN
 NNN
 NNNNNNNNNNNNGGTGGATTCTGAGGANGTACGCTCTGCAGCGTCGTCTTCAGGNTCTCTC
 TTTGACCGTGACTTCNTTGGCGGGGCACTACGCTTGTACGGGCTTGACCTGAAACCCCTC
 TACACCTGCAGCCGCTGGCCCTGCCCAACACCGTCAACTGCTACATCTCCCGGCCACCG
 GATAAGACCATCTTCATCTCTTCATGCTGGGGGTGGCCTGCGTGTCTCTGCTGCTCAAC
 CTGGTAGAGATTTACCACCTGGGCCCTACCAAGTGCCGGCAGCGGCCACGTCCAGATCC
 CGCATCTCTGCCCGCCCTGGTGGCCTTGCGGGTACACAGGTGCCGCTTACCTTGCCT
 GGTGGTGGCAACCCCGTGGATGCCAGCAGACCCACCTGGCACCCCTGGGGAAGGCTGGT
 GCGTGGCTGGGGGTGGCCAGCGGTCCCGTGGGAAGGCGGAGCTGGGGACCTGGCAGTG
 TGA

>GgCx45.6 Gg15488 Ensembl prediction into two exons (black/blue letters). The 60 nt predicted intron is included in the sequence (marked in blue), which then becomes consistent with NM_204997.

ATGGGTGACTGGAGTTTCTTGGGGAACATTTTAGAGCAGGTGAACGAGCAGTCCACTGTC
 ATCGGGAGAGTTTGGCTCACGGTGCTCTTCATTTTCCGCATCCTGATCCTGGGAACAGCT
 GCTGAAGTAGTATGGGGAGATGAACAGTCAGACTTTGTGTGCAACACCCAGCAACCTGGT
 TGTGAGAACGCTGCTATGATGAGGCCCTTCCCATCTCCACATCCGGCTCTGGGTCCTA
 CAGATCATTTTGTATCCACGCCTTCGCTAGTGTACTTTGGGCATGCGGTGCACACGTC
 CGCATGGAGGAGAAGAGGAAAGAGAGGGAGGAAGCTGAGAGGCGTCAGCAAACTGAGGTG
 GATGAAGAGAAGCTGCCCTAGCTCAAATCAAACAAGGGCAACAACCCAGATGGGACC
 AAGAAGTTTCGCTGGAGGGGACCCTCCTGAGAACCACATCTCCACATCATTTTCAA
 ACTCTCTTTGAAGTGGGATTCTTGTGGCCAGTATTTCTGTATGGTCTCCGATTCTC
 CCCCTTTACCGCTGTGGGCGGTGGCCCTGTCCCAACCTAGTGGACTGTTTGTCTCCAGG
 CCCACAGAGAAGACCATCTTTATTATGTTTCATGCTTGTGGTGGCTGCTGTGTCCCTCTTC
 CTCAACCTGGTGGAGATCAGTCACTTGTATCCTGAAAAGGATCCGGAGGGCTCTGAGAAGA
 CCAGCAGAGAGCAGATGGGGAGGTGCCAGAGAAGCCCTCCATGCTATTGCAGTGTCC
 TCCATCCGAAGGCCAAAGGCTACAAGCTGCTAGAAGAAGAAAAGCCAGTGTCCCACTAT
 TTCCCTCTCACGGAAGTAGGGGTTGAGCCAGTCCCCTCCATCAGCCTTCAATGAGTTT
 GAGGAGAAGATTGGGATGGGGCACTGGAAGATCTCTCCAGGGCATTTGATGAGAGGTTA
 CCATCGTATGCACAAAGCAAGCACCTCAGGAAGAGCCAGGGGTGAAGAAAGCAGAGGAG
 GAGGTGGTGAAGCAGTGAAGTGGAGGGCCTTCAGCACCTGCTGAACCTGCCACCGATGTA
 AGATCCCTCAGCAGGCTAAGTAAAGCCAGCAGCCGGGCCAGGTGACGATCTGACTGTA
 TGA

>GgCx45.6 Modified according to NM_204997

ATGGGTGACTGGAGTTTCTTGGGGAACATTTTAGAGCAGGTGAACGAGCAGTCCACTGTC
 ATCGGGAGAGTTTGGCTCACGGTGCTCTTCATTTTCCGCATCCTGATCCTGGGAACAGCT
 GCTGAAGTAGTATGGGGAGATGAACAGTCAGACTTTGTGTGCAACACCCAGCAACCTGGT
 TGTGAGAACGCTGCTATGATGAGGCCCTTCCCATCTCCACATCCGGCTCTGGGTCCTA
 CAGATCATTTTGTATCCACGCCTTCGCTAGTGTACTTTGGGCATGCGGTGCACACGTC
 CGCATGGAGGAGAAGAGGAAAGAGAGGGAGGAAGCTGAGAGGCGTCAGCAAACTGAGGTG
 GATGAAGAGAAGCTGCCCTAGCTCAAATCAAACAAGGGCAACAACCCAGATGGGACC
 AAGAAGTTTCGCTGGAGGGGACCCTCCTGAGAACCACATCTCCACATCATTTTCAA
 ACTCTCTTTGAAGTGGGATTCTTGTGGCCAGTATTTCTGTATGGTCTCCGATTCTC
 CCCCTTTACCGCTGTGGGCGGTGGCCCTGTCCCAACCTAGTGGACTGTTTGTCTCCAGG
 CCCACAGAGAAGACCATCTTTATTATGTTTCATGCTTGTGGTGGCTGCTGTGTCCCTCTTC
 CTCAACCTGGTGGAGATCAGTCACTTGTATCCTGAAAAGGATCCGGAGGGCTCTGAGAAGA
 CCAGCAGAGGAGCAGATGGGGAGGTGCCAGAGAAGCCCTCCATGCTATTGCAGTGTCC
 TCCATCCGAAGGCCAAAGGCTACAAGCTGCTAGAAGAAGAAAAGCCAGTGTCCCACTAT

TTCCCTCTCACGGAAGTAGGGGTTGAGCCAGTCCCCTTCCATCAGCCTTCAATGAGTTT
 GAGGAGAAGATTGGGATGGGGCCACTGGAAGATCTCTCAGGGCATTGATGAGAGGTTA
 CCATCGTATGCAAACCGAAGGAACCGGAAGAGGAGAAGGTAAAGCAGAGGAGGAAGAG
GAACAAGAAGAGGAGCAGCAAGCACCTCAGGAAGAGCCAGGGGTGAAGAAAGCAGAGGAG
 GAGGTGGTGAGCGATGAAGTGGAAAGGGCCTTCAGCACCTGCTGAACTTCCACCGATGTG
 AGATCCCTCAGCAGGCTAAGTAAAGCCAGCAGCCGGGCCAGGTGACACGATCTGACTGTA
 TGA

>Gg15606 Ensembl prediction into 2 exons (black/blue letters), without the inclusion of a stop codon. Modified by inclusion of the 57 nucleotide intron into the coding part (marked in blue), and adding 222 nt at the 3'-end to reach a stop codon (also in blue).

ATGGGCGACTGGAACCTGTTGGGCAGCATCCTTGAAGAAGTGCACATCCACTCCACTATA
 GTTGGCAAAATCTGGCTCACGATCCTGTTTCATATTCGGGATGCTAGTGCTGGGAGTGGCT
 GCTGAGGATGTCTGGGATGACGAGCAGTCGGAGTTCATCTGCAACACAGAGCAACCTGGC
 TGCAGCAATATATGTTATGACAAGGCCCTTCCCTATCTCTTTGATCAGATACTGGGTATTG
 CAGATCATATTTGTATCTTCTCCATCTCTTGTTTACATGGCCCATGCACTGTACAGATTA
 AGGGCTCTAGAGAAAGAGCGACAGAAGAGGAAAGCCCACTTGCGGGCTCAGCTAGAAGAT
 CTGGAGCCCATGCCTGAAGAACACCGGAGAGTGGAGAGGGAGCTGCGAAAGCTGGAAGAA
 CAGAAGAAAGTGAATAAGGCACCCCTTGAGAGGGTCTCTGCTGCGCACCTATGTCTACAT
 ATCCTGACCCGCTCAGTGGTTGAAGTGGGCTTTATGATAGGTCACTATCTTTTGTATGGA
 TTTTCACATGTACCCCTTTTACAATGTACTCGGCCCTTGCCTAACACAGTGGATTGT
 TTTGTGTCCCGACCCACAGAGAAGACCATCTTTATGGTCTTATGCACAGCATTGCTGCA
 GTCTCCCTATTCTCAACATCCTAGAAATTGCCACCTGGGCCCTCAAGAAGATCCACAAG
 ACCCTCTACGGGCGACTGCGGCAGCCCGTGGGCCCTGAGTATGACGAGGCCAGTCCTTAC
 AACTCCAAGAAGAACTCCGTGGTCCCACCGGCCCTGCCCGGCCGCGACACCTCCCTTCA
 CGCTCTGCCGCGGCCCGCCGCCCTCCCAACGCTCCCTCGGCCACCCCGCCCTCCCTCC
 CCTCTGGGACAGCCGGGACCTCAGCACAGGGAGAGCTCCGAGCCACCGCCCGCCGCGC
 CGCCGGCACAGCGCGGCTCAGCACACGGACAGCAGCCACCGCGTCTCCAGCAGCGAG
 GGAGTGTCTCGTAAGCAGAGCCGCGTGAGCGTCTGTGCGACCTGGAGGAGAACCAGCGGG
 GAATCCCGGACAGCGGGCACTGCCCGGGCACCCGCAATCCAGCTTCCTCTCCCGGGTG
 CTCTCCGAGAGCCGGGCGGGCAGCGACAGCGAGAGCGCGCTCCCGCCGAGGCTCCGCT
 CCCGGCTCCGCTCTGTTCCGGCTCCGAGGGGAGCGCCACGAGGAGGGCAGCCCGGCC
 GGCAGCCCGCGCGCCCGCGGGCATGGGACGCGCGTGTGATGGCAAGTAGTGGC

>Gg15606 Our suggested modification. Position corresponding to potential intron start in mammalian Cx62 orthologs is underlined.

ATGGGCGACTGGAACCTGTTGGGCAGCATCCTTGAAGAAGTGCACATCCACTCCACTATA
 GTTGGCAAAATCTGGCTCACGATCCTGTTTCATATTCGGGATGCTAGTGCTGGGAGTGGCT
 GCTGAGGATGTCTGGGATGACGAGCAGTCGGAGTTCATCTGCAACACAGAGCAACCTGGC
 TGCAGCAATATATGTTATGACAAGGCCCTTCCCTATCTCTTTGATCAGATACTGGGTATTG
 CAGATCATATTTGTATCTTCTCCATCTCTTGTTTACATGGCCCATGCACTGTACAGATTA
 AGGGCTCTAGAGAAAGAGCGACAGAAGAGGAAAGCCCACTTGCGGGCTCAGCTAGAAGAT
 CTGGAGCCCATGCCTGAAGAACACCGGAGAGTGGAGAGGGAGCTGCGAAAGCTGGAAGAA
 CAGAAGAAAGTGAATAAGGCACCCCTTGAGAGGGTCTCTGCTGCGCACCTATGTCTACAT
 ATCCTGACCCGCTCAGTGGTTGAAGTGGGCTTTATGATAGGTCACTATCTTTTGTATGGA
 TTTTCACATGTACCCCTTTTACAATGTACTCGGCCCTTGCCTAACACAGTGGATTGT
 TTTGTGTCCCGACCCACAGAGAAGACCATCTTTATGGTCTTATGCACAGCATTGCTGCA
 GTCTCCCTATTCTCAACATCCTAGAAATTGCCACCTGGGCCCTCAAGAAGATCCACAAG
 ACCCTCTACGGGCGACTGCGGCAGCCCGTGGGCCCTGAGTATGACGAGGCCAGTCCTTAC
 AACTCCAAGAAGAACTCCGTGGTCCCACCGGCCCTGCCCGGCCGCGACACCTCCCTTCA
 CGCTCTGCCGCGGCCCGCCGCCCTCCCAACGCTCCCTCGGCCACCCCGCCCTCCCTCC
 CCTCTGGGACAGCCGGGACCTCAGCACAGGGAGAGCTCCGAGCCACCGCCCGCCGCGC
 CGCCGGCACAGCGCGGCTCAGCACACGGACAGCAGCCACCGCGTCTCCAGCAGCGAG
 GAGGCGCGCGGGCGGGCGGGGCGGGGACGGCGGGGCGGGAACGGGGGCTGGCCGCCGA
 GTGCTCCGTAAGCAGAGCGCGTGAGCGTCTGTGCGACCTGGAGGAGAACCAGCGGGGAA
 TCCCGGACAGCGGGCACTGCCCGGGCACCCGCAATCCAGCTTCCTCTCCCGGGTGCTC
 TCCGAGAGCCGGGCGGGCAGCGACAGCGAGAGCGCGCTCCCGCCGAGGCTCCGGTCCC
 GGCTCCGCTCTGTTCCGGCTCCGAGGGGACGCGCCACGAGGAGGGCAGCCCGGCCGCGC
 AGCCCGCCGCGCCCGCGGCCATGGGACGCGCGTGTGATGGCAAGTAGTGGCCCCGGG
GAGCGGGCGGGGGTACTGACAGCCGGGGCAGCTCTTCTAAGTTCTTGGCACCCACG
GCTATGATCGTGCCCGCACCTTCTGCGCTTCGGCGGGGAAAGCAGAAGCGCTGCAGA
ACTGCGTTTCCATCAGCTGTTTCATGTCGTTTGGCTGGATGTTCTAAACTTGCTTATACTC
AGTGTTAATGTCATGCTTCCCTCCCTGAACAAATGA

Suppl. Figure 2 *Connexin* sequences from *Xenopus tropicalis* (Xt). The sequences are found in the Ensembl database, GenBank or the *Xenopus tropicalis* database (<http://genome.jgi-psf.org/Xentr3/Xentr3.home.html>). The sequences are updated according to the Ensembl database as of July 1 2006. In Suppl. Table 4, information on *Xenopus tropicalis* and *Xenopus laevis* databank entries is given. If the sequences previously have been named within the connexin nomenclature, this is normally the name that is used here. The Ensembl gene identification is given in abbreviated form. For the ortholog of Cx25, the abbreviated form *Xt18051* should be read as ENSXETG00000018051. Other information is given in short form in the title line where needed. If our interpretation of the most likely connexin sequence differs from the prediction found in the databases, the database sequence is given first, and immediately followed by our suggested sequence, which is the one used in our analyses. The conserved domains are marked in **yellow** (unmodified sequence) or **green** (modified sequence). Alternating black/blue sequences are exon predictions copied from Ensembl. Other colors are explained at the relevant sequences. XtDB, *Xenopus tropicalis* database v3.0 (<http://genome.jgi-psf.org/Xentr3/Xentr3.home.html>).

```
>Xt18051 Ensembl prediction. No modification
ATGAGTTGGTCATTTTTGCGTGATATCTTGAGCGGAGTGAATAAATATTCACAGGCATT
GGAAGAATCTGGCTTTCCGTTGTTTCATTTTCCGACTTTTGGTCTATGTCTGTTGCTGCA
GAAGCTGTGTGGAAGATGAACAGAAAGAATTTGAGTGCAACATCAGACAGCCCGGGTGC
GAAATGTCTGCTTTGATCAGTTCTTCCCATCTCCAGGTGCGCCTTTGGGCTCTACAG
CTGATTATGGTGTCAACACCCTCGCTTCTTGTGTTCTTCATGTGGCCTATCGGGAAAGC
AGGGAGAAGAGGCACAACAAGAAGCTGTATGCCAACACTGCAGACATGGATGGGGGATTA
TGGTGATACATATGTCTAAGTCTTCTCTTAAGACTGTGTTTGAATTTGGATTTCTTGTA
CTTTTCTACAACTGTATGGTGGTTTCAAGTGTGCCACGACTGGTGAAATGTGATATGGAT
CCCTGCCCAATGTTGTGGACTGCTATATCTCCAAACCTACAGAGAAGATGATTTTTCTC
TATTTTGTGGTCAACACCTCAGGACTTTGCATTTTGTCTAATGTTAGCGAGTTATTTTAC
CTTATTTTCAAGATACTGTACAAAATTCTACCTGAAAAAGCAGGCAACAAAATGGCACAA
GAAACATGTGATTGCAAGAATAAGCAACATAACCTAATACCCATTACATCTGAGCAGGCG
GGCACTGATGTGCTGATGGTTAA
```

```
>XtCx30 Xt26112 Ensembl prediction. No modification
ATGGATTGGGGAACGCTCTATGAGGTTATTGGCGGAGTGAACAGACATTCTACCAGCATT
GGCAAGATCTGGCTTTCCGTTCTCTTCATCTTCCGCATCATGATCCTTGTGTGGCTGCC
GAGAGCGTGTGGGGGATGAGCAGTCCGATTTCACCTGCAACACCTTACAGCCCGGATGC
AAGAATGTCTGCTATGACCATCATTTTCCAGTTTCTCACATCAGGTTGTGGTGCCTGCAG
CTTATATTCGTTGCCACCCTGCCCTTTTGGTAGCCATGCACGTGGCCTATCTAAACAT
GAAGAGAAGAGGCAAGTTGCAGAGAAGGGAACCTTTGGGGGAAGGACTTGAAGAGTTG
AAGCTAAAGAAGGTCTGAGACTACAGGGGAACCTATGGTGGACCCTACACAACAAGCATCTTA
TTCAGAGTCTCTTTGAAGCTGCCTTCATGTACCTATTCTACTTTCTGTATTCTGGCTTC
CAAATGCAGAGGCTCGTCCAGTGTAGCAACTGGCCGTGCCCAACGTTGTCGACTGTTTC
ATTTCCAGGCCTACAGAAAAGACTGTGTTCACTATATTCATGATTATCGTGTCTGGTATC
TGATGGTCTGTAATGTGATGAGGCTTTGTTACCTAATCGTGCAGGCTTCCCTAAGGAGG
TCAAAAAAACATTCACTCAGGCATTCTAACCACATGTCTGGCAAGGAGGAGAAGCAGAAC
ATCATGAATGAGCAAAAAGAGCAATAG
```

```
>Xtscaf223 According to XtDB 134000033
ATGAAC TGGG GTGTATGAAGCACTTATTACAGGTGTAAATAAGTACTCTACAGAGTTT
GGACGTCTTTGGCTTTCTATTGTCTTTATTTTTCGTTTCTGGTGTACGCACTCACAGCA
GGTCGGGTTTGGGGCGATGACCAGAAAGAATTTGTCTGTAACACCCGTCAGCCAGGATGT
ACCAATGTTTGGGCTTCAAGACTACTTCCCTGTATCCCATGTCCGCTCTGGGCTCTTCAG
TTGATTTTGTAAATGTCCATCCCTCATAGTTGTTATGCATGTGGCTTACAGGGAGAAC
CGAGAACAAAAGAGAAGGGAATAATTAGGAAAGGACTGTGAAAAAATTACATGGATACT
ACAAAGAAGAGAGGAGGACTGTGGTGGACCACCTCATCAGTCTTTTTTTCAAGGCCGCA
GTGGACTCAGTATTTATCAGGTTTATCGGCTCTATGAAAATTTCTTTCTGTCTCGT
TTAGTGAAATGCTCTCTTAATCCTTGCCCTAACATTGTGGATTGCTATATCTAGGCCT
TCAGAGAAAAACATCTTCACTCTCTTCATGATCATCACCTCTGCTGTTTGCATTCTCTTG
AACCTAGTTGAGGCTACCTATCTGATTGGCAAGAAGTGCAAGGAAACAATACATAAAGGG
AATAATGACTCTTCAAGAGAGAATTTGAAAAGTGCCATAAAGACCAAGAAACAAA
ACTTGCAAGAGCAAAATTAAGGTTGGAATACATTGTAAGACCTCAAATATCACTAAAATA
AAAGTAGACAATATTTCAAAGTAA
```

```
>Xt16480 Ensembl prediction. Sequence extended by 30 nt in 3'-end to reach a stop codon
(marked in blue).
ATGAAC TGGG CAATCTTTCAAGGTCTCCTCAGTGGTGTGAGCACTTACTCTACTGCCTTT
GGCCGCATATGGCTTTCCGTTGTATTCAATTTTCAAGGCTTATGGTCTATGTGGTTGCAGCT
GAAAGGGTTTGGGGTGATGAACAGGGAGAGTTTGACTGCAATACTAAACAACCAAGGATGC
ACCAATGTCTGCTATGATTATTATTTCCCTGTCTCACATATTCGACTTTGGGCTCTCCAG
CTCATCTGGTCACTGTCCATCCCTTCTGTCTATAATGCACGTTGCTTACCGAGAGGAA
```

CGGGAGAGAAAAACACAGAGAAAAACAGGGAGAGAATTGCGGCCGACTTTATCTAGACGTA
GGAAAAAGAGAGGAGGCTGTGGTGGACATATCTATTAAGCCTTCTAGCTAAAGCATCA
GTGGATTCAAGTTTTCTCTATGTCTTTTCATCGGATCTACAACAACATTTCTCTCCAAAT
TTGGTTAAATGCTCCATAAAGCCTTGTCCTCAACTGTTGACTGTTTTATATCAAGGCC
ACAGAAAAAACCTATTTACCCTATTTATGTTGCCACTACAATACTGTCATTTTTTTA
AATCTTTGTGAGATGATATCTAATTGCCAAAAGGTGTCGAGAGTGCCTGGGACACAGG
AACATAAAATCCAAAACCTCGCTCCAACAATGTCTATGCATACATTGAGAAGCAAGAGATT
TTCCATGCCAACACAAGATAAATGAAAAAATGTCTCGGCTAAATCCATCCATTAGA
>Xt16480 Our suggested modification

ATGAACTGGGCAATCTTTCAAGGTCTCCTCAGTGGTGTGAGCACTTACTCTACTGCCTTT
GGCCGCATATGGCTTTCCGTGTTGATTCAATTTTCAGGCTTATGGTCTATGTGGTTGCAGCT
GAAAGGGTTTGGGGTGATGAACAGGGAGAGTTTGAAGCAATACTAAACAACCAGGATGC
ACCAATGTCTGCTATGATTATTATTTCCCTGTCTCACATATTCGACTTTGGGGCTCTCCAG
CTCATCTCTGGTCACTGTCCATCCCTTCTTGTCTAATGCACGTTGCTTACCGAGAGGAA
CGGGAGAGAAAAACACAGGAGAAAAACAGGGAGAGAATTGCGGCCGACTTTATCTAGACGTA
GGAAAAAGAGAGGAGGCTGTGGTGGACAATCTATTAAGCCTTCTAGCTAAAGCATCA
GTGGATTCAAGTTTTCTCTATGTCTTTTCATCGGATCTACAACAACATTTCTCTCCAAAT
TTGGTTAAATGCTCCATAAAGCCTTGTCCTCAACTGTTGACTGTTTTATATCAAGGCC
ACAGAAAAAACCTATTTACCCTATTTATGTTGCCACTACAATACTGTCATTTTTTTA
AATCTTTGTGAGATGATATCTAATTGCCAAAAGGTGTCGAGAGTGCCTGGGACACAGG
AACATAAAATCCAAAACCTCGCTCCAACAATGTCTATGCATACATTGAGAAGCAAGAGATT
TTCCATGCCAACACAAGATAAATGAAAAAATGTCTCGGCTAAATCCATCCATTAGA **TTG**
ACACCGCCACAGACTGAATTAATAA

>XtCx31 Xt16481 Ensembl prediction. No modification

ATGGATTGGAAGACATTTCAAGGTATCCTGAGTGGTGTGAATAAGTACTCAACAGCATT
GGTTCGATTTGGTTTCTCGGTGGTTTTTGTGTTCCGAGTACTTGTATGTGGTTGCAGCA
GAGAAAGTTTGGGGCGATGAACAAAAAGATTTTGAAGCAATACTCGGCAGCCAGGGTGC
ACCAATGTCTGCTATGACCACTACTTTCCCATATCACACATCAGGCTCTGGGGCTCTCCAG
CTTATATTTGTTACCTGCCCGTCTCTATTGGTTATAATGCATGTTTCTATAGAGAGGAC
CGGGAGAAAAAGAACCGCTTAAAAAATGGAGACAACCTGCTCAAGCTCTACTCCAACACA
GGAAAAAGCATGGTGGCCTTTGGTGGACCACCTCTAAGTCTTTTCTTAAGACTGGT
ATTGAAATTTGTTTCTTTACATATTGCATAAAATGTGGGATAGCTTTGATCTGCCTCCC
CTAGTGAATGTTCAAACGAACCTGCCCTAATGTAGTAGACTGCTACATAGCCCGTCCC
ACAGAGAAAAAGGGTGTTCACTATTTTATGTTGCTGCTTCTGCTCTGTGCATCATCTA
AACATCTGCGAAATATTCTACCTTATATTTAAAAGGATACTGAATTGCATGAACAAATAC
AAGAAAGTCAAGCAAGCACTCTGTGACCTACAGCAAAGCTTCCACCTGCCAGTGCACATC
CATCTAGACACCATTTGAAAGAAACCATGAACAAGCACTGGAGTCTATACGGGCATCT
GCACCTAATCTGACAACAATATAA

>Xt24391 Ensembl prediction in 2 exons (black/blue letters). Two predicted transcripts, of which ENSXETT0000052744 is used here. Sequence extended by 6 nt in 5'-end to reach initiation codon (marked in green). Inclusion of 93 nt into the reading frame (predicted by Ensembl as intron; marked in blue). Extension of 3'-end by 72 nt to reach stop codon (marked in blue).

GAGTGGGGCTTTCTGAGCTCATTGCTGGACGCAAGTGCAGGAGCATTCTCCCATGGTGGGA
CGTTTCTGGCTGGTGGTTATGCTCATATTCGCGATCCTTATCTGGCCACGGTCCGAAGC
GACATGTTTGAAGACGAGCAAGAGGAGTTTGTGCAATACAATGCAGCCGGTTGCCGA
CAGGTTTGTCTATGACCGTGCTTTCCCATCTCCCACTATCGATTCTGGGTTTTTTCACATT
GTGCTGTTATCAGCCCCGGCTGTACTCTTTGTCAATTTACTCCATGCATCAAAATTCGAAG
ATGGGGAGAGACGCCAAGAGAGGAGGAGCTGCGAGAGCCGGACAGCAGGCCATCGTCAGG
TCAAAAACAAGGCCAGCAGCGTGGGCGCCACATACGCACCTTTTATATCGTCAATGTG
GTTATGAGGATTTTGGCGAGGTTGGATTCTAGTGGGACAGTGGCTTCTCTATGGGTTT
CGGGTGGACCCCCAATATCTGTGCCAACGTGAACCTGCACGCATACTGTGGATTGCTTT
GTATCAAGGCCCAAGGAAAGACAATCTTCTCCAGTTTACTTTGTGGTGGGCGTCATC
TCCGCTTTTCTAGCGTGTGGAACCTTCTGCACATTCTGCTAAAGGGAAGGTAGGGCA
GCGGATCAAAACAGGGGACAGGGAATCAGAAAGACCGGACCGGCGACCGCTCTCCATC
CCGTACTCCCAAAGCACAACTACAAGCTC

>Xt24391 Modified, partly according to XtdB 670000005

ATGGGGGAGTGGGGCTTTCTGAGCTCATTGCTGGACGCAAGTGCAGGAGCATTCTCCATG
GTGGGACGTTTTCTGGCTGGTGGTTATGCTCATATTCGCGATCCTTATCTGGCCACGGT
GGAAGCGACATGTTGGAAGACGAGCAAGAGGAGTTTGTGCAATACAATGCAGCCGGGT
TGCCGACAGGTTTGTATGACCGTGCTTTCCCATCTCCCACTATCGATTCTGGGTTTTT
CACATTGTGCTTTATCAGCCCCGGCTGTACTCTTTGTCAATTTACTCCATGCATCAAAAT
TCCAAGATGGGGAGAGAGCGGAAGAGGAGGAGCTGCGAGAGCCGGACAGCAGGCCATC
GTCAGGTCAAAAACAAGGCCAGACGAGCTGGGCGCCACATACGCACCTTTTATATCGTC
AATGTGGTTATGAGGATTTTGGCCGAGGTTGGATTCTAGTCGGACAGTGGCTTCTCTAT
GGGTTTGGGTGGACCCCAATATCTGTGCCAACGTGAACCTGCACGCATACTGTGGAT
TGCTTTGTATCAAGGCCCAAGGAAAGACAATCTTCTCCAGTTTACTTTGTGGTGGGCG
GTCATCTCCGCTTTTCTAGCGTGTGGAACCTTCTGCACATTCTGCTAAAGGGAAGGTGT
AGGGCAGCGGGTGGAGGGGGACGACCTAGCCCCCGCCAGCCTATGAAAGGGACAAC**TGG**
TCCAACAGGGAGCAGAGAGGACAAATCATTCTTGTCCAAATCAAAACACGGGACAAC
AGGGAAATCAGAAAGCCGACGGGACCGCCTTCCATCCCGTACTCCAAAGCACAAACC
TACAAGCTCAAGGTGGTCCAGCTCATCAATAGTAGCAATCTTCCAGGTACTGGTC
AAAGGAGATTGACTGTGTAG

>XtCx32 Xt26283 Ensembl prediction. No modification

ATGAATTGGGACAGGATTATATGCCATCCTGAGTGGCGTAAACCGCCACTCTACTTCAATT

GGACGCATATGGCTCTCAGTGGTCTTCATCTTTTCGCATCATGGTCCTCGTGGTGGCTGCA
 GAAAGTGTATGGGGGGATGAGAAGTCTGCGTTTACGTGCAACACACAACAGCCTGGTTGC
 AACAGTGTATGCTATGATCACTTCTTCCCATCTCCCATATCCGTCTGTGGGCCCTTCAG
 CTCATTATTGTGTCCACCCCTGCCCTTCTTGTGGCCATGCATGTGGCTCACCTACAGCAC
 CAAGAGAAGAAGGAACACTAGTATGACCGGCCATGTGAAGGACCAGGAGCTGGCAGAAAGTG
 AAGAAACACAAAGTCAAGATCTCCGGCACCTTGTGGTGGACC **TACATCATTAGTGTTC**
TTCAGAATCATATTTGAGGCAGCATTTATGTACCTCTTCTACCTCATCTACCCCGTTAC
TCCATGATCCGTCTTGTTAAGTGCATGCCTATCCCTGCCCAACACTGTTGATTGTTTT
GTTTCTCGTCCACGGAGAAGACCATATTCACAGTCTTTATGCTCGTTGCCTCTGGATTCT
TGCATCTTTCTGAATGTTGCTGAGATAGTCTTCCTCATTGCC CAGGCCGTGCTCCCGCAA
 GTCCATCGCAGACATTCTGACAGCATCAGCAAGGAGCATCAACAGAATGAAATGAACCTTA
 CTGATAACGGGAGGTAGTCTCATGAAACGGTCCCTGGGGGGCAGGAGAAAGGAGATCAC
 GTTCTACTTCCTAG

>Xt12912 Ensembl prediction. No modification

ATGGGAGAATGGACTATCTTAGAAAGACTTTTGGAGGCTGCTGTGCAGCAGCACAGTACT
 ATGATAGGAAGGATTCTGCTGACAGTGGTTGTCTATCTTCCGGATACCTTATTGTGGCCATT
 GTTGGAGAGACTGTTTATGATGATGAGCAAACTATGTTTGTGTGCAACACTCTTCAGCCA
 GGATGTAATCAGGACATTCACAGCATCAGCAAGGAGGCTTTTCCAATCTCACATATAAGATACTGGGTG
TTCCAATCATCATGGTCTGTACACCAAGTCTCTGCTTTATTACTTACTCCGTCCACCAG
 TCTGCCAAGCAGAGGGAGAGGAGGTACTCTACAGTCTTCTTAACGGTAGACAGAGACCCA
 GACACTGTAAACGAGAGGACAGCAAGAAGATTAAGAACACCTTGGTTAATGGTGTCTG
 CAAAACACAGAAAACCTCACCAGGAGGCCGAGCCTGATTGCTTAGAGATCAAGGAAATC
 CCAAACCATCCATTGCAACCAAAATCCAAGATGAGGCGCCAAGAAGGAATCTCAAGA
 TTTTATATCATCCAGGTGGTCTTCAGAAATGCTCTGGAGATTGGTTTCTAGTGGGACAA
TATTTCTGTATGGGTTAATGTCCCGTCCATGTACGAATGTGACAGATATCCTTGCATC
AAAGAGGTGGAGTGTATGTGTCTAGGCCAACTGAGAAGACGGTATTCTGGTTTTTATG
TTTGCGGTAGTGGCCTCTGTGTGTGCTCAATTTGGCAGAGCTGAACCACTTTGGGCTGG
 CGAAAGATCAAGATGGCAGTCAGGGGAGTACAAGCAAAAAGGAAATCCATATATGAGATC
 AGGAATAAGGATTTGCCAGAATGGGGGTGCCCAACTTTGGCAGGACTCAGTCCAGTGAC
 TCAGCCTATGTGTGA

>Xt02781 Ensembl prediction. 36-like gene. Sequence extended by 67 nt in 5'-direction to reach the 5'-end of exon 2. Exon 1 (red letters) found approx 17.7 kb upstream.

GAAAGTAGCAGACAATGTTCTAGTGCAATACACTTCAACCTGGTTGTAACCAAGCGTGC
 TATGATAGAGCATTTCCCATTTCTACATTCGCTACTGGGTATTTAGATCATTTTGGTT
 TGCACACCCAGTCTCTGCTTACCTACTCTGTGCACCAAGTCAGCCAAGCAGAAAGAC
 CGAAGATATTCCTTTCTATATCCCTTATTGGAGAAGGAACCTCCAAGAGAGAACAAAAGA
 ATTAATAATATTAAACGGTATTCTAGTGCAAAATCAGGACATTTCTCCAAGATGAGCCG
 GACTGCTTGGAGGTCAAGGAAATTCAAACACTCCTAAGAAAGCCGTCAAGAGTACAAAA
 GTCAAGCGTCAGGAGGGAATTTCTCGATTCTACATCATTACAGGTCTTTTTTCAGGAATGCT
 CTAGAGATTGGGTTTTGGCTGGACAATATTTCTGTATGGCTTCAGTGTGCCTCCAATC
 TTTGAGTGAACCGATACCCTTGTGTGGAAGGAGGTAGAGTGCTATGTTTCAAGGCCACT
 GAAAAGACTGTCTTCTAGTCTTTATGTTTGTGTGTCAGTGGCATCTGTGTTCTCTTAAAC
 TTAGCAGAACTAAACCACTTGGGTGGAGAAAGATCAAGACAGCTATGAGAGGGGTTCAA
 GCACGAAGAAAGTCAGTCTCCGATGTACGCAAAAAGGAGCCATCAACACTGGCCAGGTT
 CCTACATTAGGTAGGACACAGTCCAGCGAGTCGGCTTATGTTTGA

>Xt02781 Our suggested modification

ATGGGTGAGTGGACCATCTTGGAGCCCTGCTGGAGGGTGGCTGTCAGCAGCAGCTCTACA
ATGATAGGAAGGATCCTATTGACCGTAGTGGTTATATTCGGGATTCTAATTGTGGCTATC
GTTGGAGAGACAGTCTAT GAAGATGAGCAGACAATGTTATGTGCAATACACTTCAACCT
 GGTGTGAACCAAGCGTGCATATGATAGAGCATTTCCCATTTCTCACATTCGCTACTGGGTA
 TTTAGATCATTTTGGTTTGCACACCCAGTCTCTGCTTACCTACTCTGTGCACCAG
 TCAGCCAAGCAGAAAGACCGAAGATATTCCTTTCTATATCCCTTATTGGAGAAGGAACCTC
 CCAAGAGAGAACAAAAGAAATTAACGGTATTCTAGTGCAAAATCAGGACATT
 TCCTCCAAGATGAGCCGACTGCTTGGAGGTCAAGGAAATTCAAACACTCCTAAGAAA
 GCCGTCAAGAGTACAAAAGTCAAGCGTCAGGAGGGAATTTCTCGATTCT **TACATCATTAG**
GTCTTTTTCAGGAATGCTCTAGAGATTGGGTTTTTGGCTGGACAATATTTCTGTATGGC
TTAGTGTGCCTCCAATCTTGTAGTGTAAACCGATACCTTGTGTGAAGGAGGTAGAGTGC
TATGTTTTCAAGGCCCACTGAAAAGACTGTCTTCTAGTCTTTATGTTTGTCTGTCAGTGGC
ATCTGTGTTCTCTTAACTTAGCAGAACTAAACCACTTGGGTGG AGAAAGATCAAGACA
 GCTATGAGAGGGGTTCAAGCACGAAGAAAGTCAGTCTCCGATGTACGCAAAAAGGAGCCA
 TCAACACTGGCCAGGTTCTACATTAGGTAGGACACAGTCCAGCGAGTCGGCTTATGTT
 TGA

>XtCx38 Xt16484 Ensembl prediction. No modification

ATGGCAGGATGGGAGTTGTTAAAGCTCTTATTGGATGATGTTCAAGAACATTCACACTT
 ATTGGGAAAGTCTGGCTGACTGTACTTTTTATCTTCAGAATATTTATTCTGAGTGTGTGCT
 GGTGAGTCAGTTTGGACTGATGAGCAGTCAGATTTCTGTAACACTCAGCAGCCTGGC
 TGTACCAATGTCTGTTATGATCAGGCCCTTCTCTATTTACATGTTCTGCTACTGGGTGCTC
 CAATTTCTCTTTGTTAGCACACCTACCTTGATCTACTTGGGTCATATGGTTTACCTGTCT
 AAGAAGGAGGAGAAGGAAAGACAGAAGGAGAATGAGGCAAGCGAAGGACAGACAGAAGTA
 CGTTCTTTTACTTCTAAGAAGATAAGGATTCAAGGGCCCTTAATGTGTACCTACACCACC
 AGCGTAATCTTCAAAAGCCTTTTTGAAGCTGGGTTCTTCTCGGCCAGTGGTACTTATAT
 GGCTTTGTGATGTCTCAATATATGATGTGAACGGATCCCATGCAAAACCAAGTTGAG
 TGTTTTGTCTCAAGGCCTATGGAAAAACAATATTTATCATATTTATGCTGGTCTGTTTCT
 CTCATTTCTCTAATGTTAAATTTGATGGAGCTTATTCACTTAAGCTGT **AAGTGTTCAC**

CATAGTGTAAAGAAGGAGCAACTTGTCTCAAACAGGGGTTCCCTTGATGGGTTAGGA
AACAGACTACCACCACAGGAATACATCAATCCCCCTCCAGCAACCAAGACATTGATCTG
CCAGCTTATAACAAAATCCCTGGGGGGCAGAATTGGTCAACTATCCAGATAGAACAGCAA
GCGAATGGTCTTTTGAACCCAAGTGTCAGTGTGACTGTTGGTCCCAAAGTGCTATTTCT
GTTGTGGTTTCTGGAGCACCAGGTATAGTCCCCAATTTGGATGCTATCAAAAGTAATCAC
CAGACCTCTTCCAAGCAGCATGTTTAA

>XtCx41 Xt26789 Ensembl prediction. No modification

ATGGGAGACTGGGAGTTTCTTGAAAACTATTGGATCAGGTGCAAGAGCATTCCACATCT
ATCGGGAGAGTCTGGCTTATGGTGCTTTTCATCTTCCGAATCTTGATCTTGGGCTGGCT
GGGGAGTCTGTTTGGGGAGATGAACAGTCTGATTTTATCTGTAATACAGAGCAGCCCGGA
TGCACCAATGTTTGCTATGACAAAGCCTTCCCCATATCTCATGTTGCTACTGGGTACTT
CAGTTTTCTATTTGTCACTACCCACCATTTGTTCTACTTGGGACATGTAATCTATCTTTCA
AGGAGGGAGGAAAAAGTTGAAACAAAAGGAAAGTGAAGTGAAGGCTTTGGATGACAAAGAG
CAAGTTGAACAGGCTATAGCTATTATAGAAAAAGAAAAAATGAACTGTATATCCAAGAA
GATGGGACAGTCAAGATCAAAGGAGCCCTAATGTGCACCTATCTTACTAGTGTTATCTTC
AAAAGCCTATTTGAAAGCTGGATTCTTCTTGGCCAATGGTACCTTTATGGGTTTGTGATG
ACTCCAATTTATGTCTGTGAAAGAGTACCGTGTCCCCACAAAGTAGACTGCTTTGTATCT
CGACCCATGGAAAAAATATTTTCTTATTGTTTTCATGTTAGTGGTGTCTCTCATATCTTTA
TTTCTAAATGTAAGTGAAGTATTTTCTAGTCTGTAAAGAGCATGATCGATACCTTGAAG
AAATATTACCCGTATGTCCCTCCAAACGATACCCCAAGAAATGAGGACACATATCCAGAA
AAGACATCAGAAACAGCCACAGCACCATTCCAAGACAAATCATATATTTACCTTCCCATG
AATGATAATATCTATACCCACAGTACAAGATGCCAAATGAGCAGAATTGGGCAAACTTT
AATACAGAGCACCAGCTAGCTATAAGTGGTAATAACAGTCAACACTAGGCCATTACAGC
CTCAGTGCCTTTCGTGCCTGTGCCCTCCAAAACTCATTCAACAATGGAAAGCCTAGCACC
AGAGCCAGTAGTCTGCTTCCAAGAAGCAATATGTTTGA

>XtCx43 Xt25411 Ensembl prediction. No modification

ATGGGTGACTGGAGTGCCCTTAGGAAGACTTCTTGACAAAGTTCAAGCCTATTGCACTGCA
GGAGGGGAGGTGTGGCTGTCTGTCTTATTTATATTTTGAATCCTTTTATTAGGGACAGCA
GTGGAGTCAGCCTGGGGAGATGAACAATCAGCATTTTGTCTGCAATACTCAACAACCAAGGT
TGTGAAAAATGTTTGCTATGATAAGTCATTCCCGATCTCTCATGTGCGTTTCTGGGTACTT
CAAATCATATTTGTCTCAACGCCCTACCCCTCCTTTATTTAGCCCATGTGTTTTACTTGATG
AGAAAAAGAGAAGCTAAACAGGAAAGAAGAAGAACTTAAAAATGGTTCAGAAATGAAGGT
GGTAATGTAGACATGCACCTTAAACAAATTGAAATAAAGAAATTCAAATATGGCCTTGAA
GAGCATGGAAAAGTTAAGATGCGAGGTGGACTTCTTCGCACTTACATCATCAGCATTTTG
TTTAAATCAGTCTTTGAAAGTTGGCTTTATTTTCAATCAATGGTACATGTATGGGTTTAGC
TTAAGTGCCATCTACACATGCAAAAGAGATCCTTGCCACATCAAGTGGACTGTTTCCTT
TCTCGACCCACAGAGAAGACCATTTTTATCTGTTTCATGTTAATAGTTTCTATAGTTTCT
CTTGCCCTAAATATCATCGAGCTGTTCTATGTACCTACAAAAGTATCAAAGATGGCATC
AAGGGGAAAAAGACCCATATTTGCCACAAATGATTCTGTGGTTCAGGAAAGGAATGT
GGCTCCCTTAAATATGCTTATTTAATGGCTGTTCTTCTCAACTGCTCCCATGTCTCCA
CCTGGTTACAACTTGTACTGGAGAGAAGAAATCCTTCTTCTGTGCAATTATAACAAG
CAAGCAAGTGAACAAAATGGGCCAATTACAGTGCAGAACAAAACAGGATGGGGCAGGCT
GGAAGCACAATTTCTAACTCATGCACAGCCTTTTGATTTTTCGGATGAACATCAGAAT
ACTAAGAAAAATGGCACCAGGACATGAGATGCAGCCCTGACTATTTGGACCAAAGACCA
TCTAGTAGAGCAAGCAGTCATGCAAGTAGCAGACCTCGACCTGATGATTTAGAGATCTAA

>Xt09451 Ensembl prediction. Two predicted transcripts; ENSXETT00000020836 is used here.
Predicted as 3 exons (black/blue/black letters), with an 3.6 kb intron between exon 1 and exon
2. Exon 2 (blue letters) and 3 (black letters) are removed (marked gray in upper sequence).
The sequence is extended by 307 nt into the intron (marked in blue) to reach a stop codon.

ATGGGGGACTGGTCTATCTTGGGGCGTCTCTTGACAGAAGTCCAGAACCATTCCACTGTG
ATTGGAAGATATGGCTGACAATGCTTCTTATTTTTCGCACTCTTTTGGAACGTTGGTG
GGAGATGCAATGTATGGAGACGAACAGTCCAAATTTACCTGCAACACTCTTCAGCCTGGT
TGCAATATGTGTGCTATGACAGTTTTGTCTCCTGTTTCACACCTCAGGTTTGGATATTT
CAGATTGCTGCTGGTTTCAACACCTTCCATTTTTATATCATCTATGCTCTGCATAAAATT
GCCAAGGATGAGAAGACTGAACTGGAAAAGGCCATGTCATGGAGCTGATTGAGAGCCTA
TCTGCTGAAAAACACTGTCTGAGAGTAATGTAACCCAGACTGAAACAACAAATAGAT
GGTAAAGATGCTTCAATGGAAAAGAACCGGATTTAAGCCAAAGTTAAGAAAAGATTTCAAA
GGGTCCATCCCAGTCTTGACAAAATGCACATCATCTATATTGTCACGTGGTACTCAGG
TCTCTGATGGAGGCTGGGTTCTGGTGGGCCAGTATTATCTTTTGGTTTGAAGTGCCT
CACCTTTTCCGATGTACATACCTACCTTGTCCAACATACTGACTGCTTTGTCTCAAG
GCCACAGAAAAGACTGTTTTCTTAAACTTCATGTTTGGAGTTGGACTTGGCTGCTTCATT
CTGAACCTTGTGGAAGTGCATTACTTAGGTTGGTTGAAATTAATACTAAGG
GAGTGGGCACTTCAGCTCCCCCAACATCCTTCAGCCCCCAACATCCTTCAGCCCCC
CAACATCCTCCAGCCCCCAACATCCTCCCCAATATCCTTCAGCTCCCCCAACATCCT
TCAGCCCCCAACATCCTTCAGCTCCCCCAACATCCTCCCCAACATCCTTCAGTTTCCC
CCAACATCCTTCAGCTCCCCCAACATCCTCCCCAACATCCTTCAGCTCCCCCAACATCCT
TCAGCTCCCCCAACATCCTTCAGTTTCCCCCAACATCCTTCAGCTCCCCCAACATCCTTC
AGCTCCCCCAACATCCTTCAGCTTCCC

>Xt09451 Our suggested modification

ATGGGGGACTGGTCTATCTTGGGGCGTCTCTTGACAGAAGTCCAGAACCATTCCACTGTG
ATTGGAAGATATGGCTGACAATGCTTCTTATTTTTCGCACTCTTTTGGAACGTTGGTG
GGAGATGCAATGTATGGAGACGAACAGTCCAAATTTACCTGCAACACTCTTCAGCCTGGT
TGCAATATGTGTGCTATGACAGTTTTGTCTCCTGTTTCACACCTCAGGTTTGGATATTT
CAGATTGCTGCTGGTTTCAACACCTTCCATTTTTATATCATCTATGCTCTGCATAAAATT

GCCAAGGATGAGAAGACTGAACTGGAAAAGGCCTATGTCATGGAGCTGATTGAGAGCCTA
TCTGCTGAAAAACACCTGTCTGAGAGTAATGTAACCCAGATACTGAACAACAAATAGAT
GGTAAAGATGCTTCATGGAAAGAAACCGGATTTAAGCCAAAGTTAAGAAAAGATTTCAA
GGGTCCATCCCAGTCTTGACAAAATGCACATCATCTATATTGTCACAGTGGTACTCAGG
TCTCTGATGGAGGCTGGGTTCCCTGGTGGGCCAGTATTATCTTTTTGGTTTTGAAGTGCCT
CACCTTTTCCGATGTCTACCTACCCTTGTCCTCAACTACAAGTACTGCTTTGTCTCAAGG
GCCACAGAAAAGACTGTTTTCCCTAAACTTCATGTTTGGAGTTGGACTTGGCTGCTTCATT
CTGAACCTTGTGGAACCTGCATTACTTAGGTTGGTTCTATGTGTTTCGCATTTTATCCTTA
GCCTGCTCTGCCTGCTGTGATTTTAAGAGTATCAAGAAGCCAAAGGTTCTACACCCAGAG
CAGAACAGCTTACTCATACCTTAAAGGACACCATACAAGAAAGGTTTGGCTAAAGAGT
TCTTCCAGTTTATCCCAGGAGAAGATCAGTTGTCTCCCTCTCATTGAGCTGCCATTTCC
TTTACACAGGATGACAACAGTGATACTACCTCCAGGAAAGGCCAGACAGTATGCATAGT
AGCCTCGATAAAATATATAAAATCAAAAATCTTGGCTTTAA

>Xt16303 Ensembl prediction. Predicted as 2 exons (black/blue letters). We have included the predicted intron into the reading frame (149 nt; blue), and further extended the 5'-end by 181 nt to reach a stop codon (the 3'-most blue mark).

ATGGGCGACTGGAGCTTCTTGCCAGTTCTTGAAGAGGTGCAGAAACATTCCACGGTG
GTGGGCAAGGTCTGGCTGACCATCTGTTTCATCTTCCGGATGTTGGTGTGGGCACGGCG
GCGGAATCCTCCTGGGGGACGAACAGTCGGACTTTACGTGCGACACGAACACGCCCGGT
TGGGAGAACGTCTGCTAGCACAAGCGTTCCCATCTCCATATCCGATACTGGGTGCTC
CAGATCATCTTCGTCTCCACCCCTCCCTGCTGTACATGGGACACGCCATCCACACCGTC
CGCGTCCAGGAGAAGAAGAACGCCGGAAGAGGAGAAGCGAAGCAGAACCTGAGCGGG
GAGCCGTCTACCATCAGAAGGAATATGCGGTGGAGAAAGAGGAGATCGTTTACACCGAC
GAGTTGAGCGGCAGGATCCGGCTTCAGGGGACGTTGCTGAACACCTACATATGCAGCATC
CTCATCCGACCCCTGATGGAAGTTGCTTTATCGTGGGACAGTATATGCTCTATGGAATA
TTCTTGGAACCCCTGTACACGTGCCGGAGGTGCGCCTGCCCGCAGCCCGTCAACTGCTTT
GTCTCTAGGCCGACGGAAGAAGACGTTTTTATTGTATTTATGATGGCTGTGGGGGTGCTC
TCTCTCCTTCTCAGCCTCGTAGAACTCTACTACCTGGCGTGGAAAAAGATCAAGCACAAA
TTTGCCGGTGCCGTCAGAACACGGCCAACTTTGCAACGGAAAGGT

>Xt16303 Our suggested modification

ATGGGCGACTGGAGCTTCTTGCCAGTTCTTGAAGAGGTGCAGAAACATTCCACGGTG
GTGGGCAAGGTCTGGCTGACCATCTGTTTCATCTTCCGGATGTTGGTGTGGGCACGGCG
GCGGAATCCTCCTGGGGGACGAACAGTCGGACTTTACGTGCGACACGAACACGCCCGGT
TGGGAGAACGTCTGCTAGCACAAGCGTTCCCATCTCCATATCCGATACTGGGTGCTC
CAGATCATCTTCGTCTCCACCCCTCCCTGCTGTACATGGGACACGCCATCCACACCGTC
CGCGTCCAGGAGAAGAAGAACGCCGGAAGAGGAGAAGCGAAGCAGAACCTGAGCGGG
GAGCCGTCTACCATCAGAAGGAATATGCGGTGGAGAAAGAGGAGATCGTTTACACCGAC
GAGTTGAGCGGCAGGATCCGGCTTCAGGGGACGTTGCTGAACACCTACATATGCAGCATC
CTCATCCGACCCCTGATGGAAGTTGCTTTATCGTGGGACAGTATATGCTCTATGGAATA
TTCTTGGAACCCCTGTACACGTGCCGGAGGTGCGCCTGCCCGCAGCCCGTCAACTGCTTT
GTCTCTAGGCCGACGGAAGAAGACGTTTTTATTGTATTTATGATGGCTGTGGGGGTGCTC
TCTCTCCTTCTCAGCCTCGTAGAACTCTACTACCTGGCGTGGAAAAAGATCAAGCACAAA
TTTGCCGGTGCCGCTACGCACAACCCGAGCGCCCTCAGGGTCGCAGCCGTAGAGCCGGGG
CACCGTTCCCAAAGCTGCACTCCGCCTCCGGACTTTAACCAGTGTTTAAGCAGCAATAAG
GGTAAGTTTGAGAGCCCTTTCAGCAATACGTTGGCTTCCAGCAGAACACGGCCAACTTT
GCAACGGAAAGGTCCAGGGCGAAGAAGACATTCGAGACGGGCAATTTATCCACATAAGC
TATGCACAAAATGCAACCGGTATCCAACCACTCCACTCCGCACCATACACCGAACGGCTAC
TGCCAGAGCCCCGAAGCCTGAGCAAGACCAGCCAAGCCAGCAGCAAAAGCCCGTTGAGAT
GACTTGTCTGTGTGA

>Xt26500 Ensembl prediction. No modification

ATGGAGTATTTAGACTCAGTGGGATATTTAATCATCACTTTTAACTACAATGTTACTGCA
ATTGGGAAGGTCTGGCTCACATTAATGCTTTTCTGAGGATGGCTATGGTGTATTTGCC
GGATACCCCTTTATCAAGACGAGCAGGAGCGTTTTATCTGCAACACCTTACAGCCCGGC
TGCTCCAATGTCTGCTACGACATCTTCTCCCTGTGTCCACATGAGATTTTGGCTTTTC
CAGACCATGATAGTCTTTCTACCTACGCTATGTTCTGCTGTTTACATTTTGCACCGAGTG
GTTGGCTACATGGCACCACAAATGATCACTTCCGGAAAGCAACACCCCGACTTATCAT
GGCGTGGGCTTACAGATGGAGATACCTGATTTTCTTGCGGTATATAGTCCAATTACTT
CTGAGAACTCTTCTAGAGGCTGGTTTCTGCACGGGGCAGTATTTTCTCTTTGGGATATTA
GTTCCAAAGAGATTGAGTGTCTCAGGCCCTGCACCAGCAGTGTGATTGTACATC
TCTAGACCCACTGAAAAGTCTCTCATGATGATCTTTATCTGGGGGACAGGAGTATATCT
CTCATACTTGGGCTGGTGGACCTCATTTACGTAATCCATCGGCGAAGGCGCTCCGAGCGC
CTAAAGAGAGAACTGCTTCTGCTGGAAAACGACTCCCTAAAAGGTGGCTGCTGTTCTGAT
TCACCTCCAAATTCAACTTGCACGGCGCTCCGGAGCCAATGGCTTGTACGCTGATTCT
GCTTCCAATGATCTTGAAAAATAGAGGAGAAAATTAAGGGCGACTCCCATTTCTTGCCA
TCGTATGAAGGGGAAACGGCTTCTTTTTCAGACAGAAAGTTCTACCATGATGTGGGAAAG
TCCAACATGAACACCAACAGCAATAAATCATGTTCTTGGCAAGTCAGCCTGACTTCTGCC
GACGGCAAACCCACCAAGAGTGATCACTTGGTCACCCCAAGGCAGCTCAAACCAAG
CAGGAGAAATGCAAGCAGCTCTAATCTGGGAATACTCAGCCCCCTGAAAAATCCCAAGCG
TCCAATCCAAGAAATCTGAGTGGGTTTAA

>Xt03391 Ensembl prediction. No modification

ATGAGCTGGAGCTTCTTGACGCGCTCCTAGAAGAAATCCATAACCACTCAACTTTTGTG
GGCAAGATTTGGTTGACTGTTCTGATTGCTTCCGCAATTGTGTTGACTGTTGTGGGAGGA
GAGTCCATCTACTACGATGAGCAAGCAAGTTTGTGTGCAACACTGGGCAACCGGGCTGC
GAGAACGTGTGTTATAATGCGTTTGTCTCCCTCTCCACGTTCTGCTTGGGTTTTCCAA

ATCATCTTAGTGTCTACACCATCCCTCTTTACCTCGGATACGCCATCCACAAAATAGCA
 AGGACGGAGGACCACATGGAGCAAGAAAGAGGACAAAGCACAGAATGTTCCCGATGCGC
 TGGAAAGCAGCAACCAAGGCATAGGGAATGCTGACGATGACAAATGAAGAGGACCCCATGATG
 TACCCCGAAACTGACATTGTAAGCGAGAGGAGCAAGCGTGAATGTATGCACAGCAGGGTA
 AAACACGACGGGCGACGGCGCATCCATGCTGATGGACTGATGAGGATTACGTTCTTCAG
 CTACTTTTCAGAACATTTTTGAAGTGGGATTTCTGGCTGGACAGTTCTTCTGTATGGG
 TTTGAAGTCACCCCAATATATGTGTGTGATACAAAGCCTTGCCCTCACAATGTGGATTGC
 TTCATATCTCGACCCACTGAAAAGACCATATTCTTGCTTATTATGTATGGAGTAAGCGGC
 CTTTGCTTGCTCCTCAATGTCTGGGAAATGCTCCACCTCGGGTGGGGCGATACTTGAC
 GCGATTAAACAGTAGGAAGAAGGAGTCTGATGATACCTCTGCTTACAATTACCTTTTCAG
 TGGAAACCCCATCAGCTCCCCCGGGTACAATATTGTTGTGAAACCTGATCAAATACAG
 TTCACAGATCTCTCAAATGCCAAATGGCTTACAAACAGAACAGGGCCAACATCGCTCAG
 GAGCAACAATATGGCAGCAGTGAGGACAATATCCCCGCCAGTCTGGAGCACTTGCAACGG
 GAAATCAGGGTTGCTCAGGAGAGGCTACAAATGGCCATCCAAGCCTACAACAGCCAAAGC
 CATCTGCCCTCCTCCAAAGAAAAGAGATCAAGAAGGGCTCCAACAAGAGCAGCGTCAGC
 AGCCGCTCTGGGGATGGAAGATGTCTGTGTGGATTTAA

>Xt06175 Ensembl prediction. Sequence extended by 51 nt to reach a stop codon (marked in blue). Modified according to XtdB 451000007.

ATGAGCTGGAGCTTTTTAACCCGCTTTTTAGAAGAAATTCACAATCACTCCACCTTTGTG
 GGGAAAGTGTGGCTCACGGTTCTGATCATCTTCAGGATAGTCCTGACAGCAGTGGGAGGG
 GAGTCTATCTATTCGGATGAACAGAGCAAGTTCATGTGCAACACCAAGCAGCCGGGCTGC
 GACAATGTCTGCTGACGATGCTTTTGGCCCTCTTTCCACGTCCTGCTTTTGGGTCTTCCAG
 ATCATCATGATATCTGCCCTTCTGTAATGTATCTCGGCTACGCTATACACAGGATCGCC
 AGGTCGGCAGAGAAGATCGCCGGCACCATTTCATGTAGGAAGAAAAAGTCGTTTTCCAAA
 TGGCGGCCAGGCCAAAATGCTGAGGACCCGGAAGCCGAAGAAGAAGAACCTATGATATAC
 GATGGCCCATCTGAGCCAGAGAAGGTGGAAGCAAGAAAAAGGAAACGAAAAAGCACGAC
 GGTCGGAGGCGCATCAAAGAGGAGGGGTTGATGAAAATTTATGTTTTCCAGTTGGTATCA
 AGGGCCATTTTTGAAGTGAGTTTTTGGTAGGCCAATACTTCTATATGGGTTTCGAGTC
 AAGGCCTATTTGAGTGACGACTAACCCGTGCCCCCATAAGGTCGACTGCTTTGTTTCA
 CGACCAACGGAGAAGACGGTTTTCTTCTTAATAATGTATGTGGTGAGTTGCCTTTGCTTG
 TTCCTTAATGTCTGTGAGATGTTGCATCTCGGCGTTGGAACCTGCGGGACGCTATCCGC
 AACAGGAGAATCCTAGAAACAAGCAGGCCCTTACAATTATTCTTATCCGAAGAATC
 ACTTGTCTCCAGAATACAACATGGTGGTCAAGTCGGACAAATCTACCAAAATCCAGTA
 GGCTTAATGGCTCATGACCAAGAACTTGCCCAACGTTGCTCAGGAACAAAATGCCCCAGT
 GACAACGCCCTTCTGACCTGTCTGCTTTGCATAAACATTTGAGCATTGCTCAGGAACAA
 TTAGATATAGCCTTTCAGAACTATAATACTCACGTCAACGGCCAGCCGTCAGGACGAGC
 AGTCCGGCTTCTGGTGGGACGGTGGTGAGCAGAACCAGGCCAACACCGCTCACGAAAAAG
 CAAGGAGCCAAGCCCAAATGCGGCTCA

>Xt06175 Modified according to XtdB 451000007.

ATGAGCTGGAGCTTTTTAACCCGCTTTTTAGAAGAAATTCACAATCACTCCACCTTTGTG
 GGGAAAGTGTGGCTCACGGTTCTGATCATCTTCAGGATAGTCCTGACAGCAGTGGGAGGG
 GAGTCTATCTATTCGGATGAACAGAGCAAGTTCATGTGCAACACCAAGCAGCCGGGCTGC
 GACAATGTCTGCTACGATGCTTTTGGCCCTCTTTCCACGTCCTGCTTTTGGGTCTTCCAG
 ATCATCATGATATCTGCCCTTCTGTAATGTATCTCGGCTACGCTATACACAGGATCGCC
 AGGTCGGCAGAGAAGATCGCCGGCACCATTTCATGTAGGAAGAAAAAGTCGTTTTCCAAA
 TGGCGGCCAGGCCAAAATGCTGAGGACCCGGAAGCCGAAGAAGAAGAACCTATGATATAC
 GATGGCCCATCTGAGCCAGAGAAGGTGGAAGCAAGAAAAAGGAAACGAAAAAGCACGAC
 GGTCGGAGGCGCATCAAAGAGGAGGGGTTGATGAAAATTTATGTTTTCCAGTTGGTATCA
 AGGGCCATTTTTGAAGTGAGTTTTTGGTAGGCCAATACTTCTATATGGGTTTCGAGTC
 AAGGCCTATTTCCGAGTGACGACTAACCCGTGCCCCCATAAGGTCGACTGCTTTGTTTCA
 CGACCAACGGAGAAGACGGTTTTCTTCTTAATAATGTATGTGGTGAGTTGCCTTTGCTTG
 TTCCTTAATGTCTGTGAGATGTTGCATCTCGGCGTTGGAACCTGCGGGACGCTATCCGC
 AACAGGAGAATCCTAGAAACAAGCAGGCCCTTACAATTATTCTTATCCGAAGAATC
 ACTTGTCTCCAGAATACAACATGGTGGTCAAGTCGGACAAATCTACCAAAATCCAGTA
 GGCTTAATGGCTCATGACCAAGAACTTGCCCAACGTTGCTCAGGAACAAAATGCCCCAGT
 GACAACGCCCTTCTGACCTGTCTGCTTTGCATAAACATTTGAGCATTGCTCAGGAACAA
 TTAGATATAGCCTTTCAGAACTATAATACTCACGTCAACGGCCAGCCGTCAGGACGAGC
 AGTCCGGCTTCTGGTGGGACGGTGGTGAGCAGAACCAGGCCAACACCGCTCACGAAAAAG
 CAAGGAGCCAAGCCCAAATGCGGCTCA

ACCTCCGATGGATATAA

>Xt14395 Ensembl prediction. Sequence extended 96 nt to reach a stop codon (marked in blue).

ATGAGCTGGAGTTTCTTAACGCTCTCCTCGAGGAGATCAACAACCACTCGACGTTTCGTC
 GGGAAAGTCTGGCTGACGGTGCTCATATATTCGGATCGTCTGACCGGGTCGGCGGG
 GAGTCCATATACTACGATGAACAGAGTAAATTCACGTGCAACACGACAGCAACCCGGTTGC
 GAGAACGTGTGCTACGACGCTTTTGGCCCCCTGTCGACGTTTCGGTTCTGGGTCTTCCAG
 ATCATCTCTCATCACACCCGTCATTATGTACCTGGGGTTTCGCCATGACCCGATTGCG
 CGCCAGCCCGAGATGCAGATAAGGCGCTCGGAGAAAACCAAAAGCAAAAAGCGCGCTCCC
 ATAATTCATCGGGGGGCCATGAGAGACTACGAAGAGGCTGAAGATAACCAAGAGGAAGAT
 CCCATGATTTGTGAAGAAGAGGAACCCGAGAAAGACAGCGAGAAGGGAGACAAAAAGAAG
 CACGATGGGCGCCGGGATCAAAACAAGATGGCCTGATGAAGGTTACGTGTTGCAAGTTG
 CTCTTCCGTTCCGTTGTTGAAGTCGGCTTCTGATGGGCCAGTACATACTGTATGGGTTT
 GAGGTCAATCCCTTTTTCTGTGTCTCCCGGAAACCGTGCCCTCACACCGTCGACTGCTTC
 GTGTCACGCCCCACCGAGAAGACCATCTTCTGCTGATCATGTACGCCGTCAGTGCGCTC
 TGCCCTTCTCTCAACTGTGCGAGCTCTTCCACTTGGGCATAGGGGGCATCCGAGATGCC
 CTCCGGCAGAGAAGAAGAAGAACTTCAGGAATCCCGTAAAAAGACCCCAAGTCCCCACCA

AATTACCACTCTGTCTGAAGAAAGGCAGGCTGCCCAACGGGAAACCCGTTTTCCAGGC
AACGGAGTGTCCGAAGGCTTCGAGTTGCCGACACACGAGCTGGACAGGTTACGGCAGCAC
CTCAAGTTGGCTCAGGAGCATCTCGACTTGGCTTTCCACCTCAACCCAAACGGGCGATAAT
ACCCACGCGTCCCGGAGCAGCAGCCCCGAGTCCAACAGCATTGCCGCCGAACAGAATCGC
CTAAACTTGGCACAAGAGAAGGGAGTGGGAAACCGGGAAAAATCTGGTAAG

>Xt14395 Our suggested modification

ATGAGCTGGAGTTTCTTAACGCGTCTCCTCGAGGAGATCAACAACCACTCGACGTTTCGTC
GGGAAGGTCTGGCTGACGGTGCTCATATTTCCGGATCGTCTGACCAGCGGTCGGCGGG
GAGTCCATATACTACGATGAACAGAGTAAATTCACGTGCAACACGCAGCAACCCGTTGC
GAGAACGTGTGCTACGACGCTTTGCGCCCTGTGCGACGTTCCGTTCTGGGTCTTCCAG
ATCATCTCATCAACACCCCGTCCATTATGTACCTGGGGTTCCGCATGCACCGATTGCG
CGCCAGCCCGAGATGCAGATAAGGCGCTCGGAGAAAACCAAAAGCAAAAAGCGCGCTCCC
ATAATTCATCGGGGGGCCATGAGAGACTACGAAGAGGCTGAAGATAACCAAGAGGAAGAT
CCCATGATTTGTGAAGAAGAGGAACCCGAGAAAGACAGCGAGAAGGGAGACAAAAGAAG
CAGCATGGCGCGCCGGGATCAAAACAAGATGGCCTGATGAAGGTCACGTGTTGCAAGTTG
CTCTTCCGTTCCGTTGTTGAAGTCGGCTTCTGATGGGCCAGTACATACTGTATGGGTTG
GAGGTCAATTCCTTTTTCTGTGTCTCCCGGAAACCGTCCCTCACACCGTCGACTGCTTC
GTGTACGCCCCACCGAGAAGACCATTCTCTGCTGATCATGTACGCCGTCAAGTGCCTC
TGCTCTTCTCTCAACCTCTCGAGGCTTCCACTTGGGCATAGGGGGCATCCGAGATGCC
CTCCGGCAGAAGAAGAAAGAACTTCAGGAATCCCGTAAAAAGACCCCAAGTGGCCCAACA
AATTACCACTCTGTCTGAAGAAAGGCAGGCTGCCCAACGGGAAACCCGTTTTCCAGGC
AACGGAGTGTCCGAAGGCTTCGAGTTGCCGACACACGAGCTGGACAGGTTACGGCAGCAC
CTCAAGTTGGCTCAGGAGCATCTCGACTTGGCTTTCCACCTCAACCCAAACGGGCGATAAT
ACCCACGCGTCCCGGAGCAGCAGCCCCGAGTCCAACAGCATTGCCGCCGAACAGAATCGC
CTAAACTTGGCACAAGAGAAGGGAGTGGGAAACCGGGAAAAATCTGGTAAGTCAGTTGGA
TTTTGTGGGTATCTATGACTGCTTGGAGGAGGGGGGGGTATTTGTGTCAGCCAAGGGG
TGCTGGTTGTCCGGCTGCCTTCTTAA

>Xt02644 Ensembl prediction. No modification

ATGGGCGACTGGAGTTTCTGGGGAGACTATTGGAAAAATGCTCAAGAACACTCAACAGTC
ATTGGCAAGGTTTGGTTGACAGTGCTCTTTATCTTCCGGATCCTGGTACTGGGAGCTGCA
GCAGAGGAGGTTTGGGGGATGAACAGTCTGACTTCACCTGCAATACGCAGCAACCAAGGT
TGTGAGAACGTCTGCTATGACAAAGCATTTCCTATTTCTCACATTAGGTTCTGGGTCCTC
CAGATTATTTTGTGCTACTCTACCTTATCTACCTGGGCCATGCTCTGCACATTGTG
CGAATGGAAGAAAAGAGAGAAAGAGGAGGATTAAGAAAGTGCAATTAAGGTGCC
AGTGAAAAGGAACATTTAATAAAGGGTGGTGAAAAGAGGAAAAGCCACCTATTAGAGAC
GAGAGGGGAAAAATCAGGATAGGAGGTGCCCTTCTTCGCACCTATGTCTTTAATATCATC
TTCAAGACTTTGTTGAGATTGGCTTTATTGTAGGACAGTATTTCTTATATGGTTTTAG
CTAAACCCCTATACAGGTGACGCCGTTGGCCTTGTCTTAACACCGTGGACTGTTTTATT
TCACGGCCAACGGAAGACCATTTTTATCATATTTATGCTTGTAGTGGCTTGTGTGCT
CTTTTGTGAATATGTTAGAGATATATCATCTGGGGTGAATAAACTCAAGCAAGGCATG
ACAAACAGATACATCCCTGATCCTTCTTGCAACAAATCAGAACCTCAAGCCCTCAAAC
GCCCATCTAGCATGTCTTCCACCATACTATCTGATGTGGCAGCTGCATCTCCAGTC
TTGCATCATGCATATGCCAGGTTCTCTGCTCCAGACTTCAAAATAACATCTGTGCCGGAG
GAAGCTCTAGAGTCTCCTTCTCTGATGAGCCATCATGAACATCTCATAGCTGCATCTGAG
CAAACTGGACCAATCTAGCAGTGGAACGTGAGAGAATGTCTAGGTCAAGCAGTGGCAGT
CTTTGCAAGTCTACTACCTTGCCTTAGACAGTCTGAGAAATGATGGAGGATGTGAGGAG
ATGCTAACAACAGAAGCCAGTGGAAGCAGTGCAAGGGCAGAAGACAGACCTGTGACCACC
AGGGTGGAAATGCACAGCCGCTGTCTGATGATAGACGAAGAAGACTCAGCAGGTCTAGC
AGGTCCAGCAGTAGCAGAGCAAGGTGAGATGACTTAGCTGTCTAA

>Xt26282 Ensembl prediction. No modification

ATGGGAGACTGGAACTTGTCTGGGAAAGCTTCTTGAGAGTGCCAGGAACATTCTACGGTG
GTGGGCAAGTCTGGCTGACTGCTCTTCATCTTCCGAATACTTGTCTTGGGAACAGCC
GCTGAAAAAGTCTGGGGGACGAGCAGTCTGGTTTTACCTGTGACACTAAACAGCCCGGT
TGTCAAAACGTTTGTACGATCACACGTTCCCATTTTCGCACATTGATTCTGGGTGCTG
CAGATCATTTTCTGTGCTCACTCCAACCTCTCATCTACCTTGGCCACATATTGCATCTTGTG
CGCTTGGAGAGAAGCGAAGCAGAAGGAAAAGTTGCGGCAAGCAGCAGAGACCCACAG
GATGACAAGCAGTGTCTCTGGTGCCAAACAAGCCGGCCAAACCCCTGTCAAGGATGAC
TTGGGCAAGGTGCGCTGCGAGGGGCCCTGCTGCGCACATACGCTCTTCAACGTCATTTTC
AAGACCATGTTTGAAGTAGGATTATTGTGGCGCAATATTTCTTTATGGATTCCAGCTG
CAGCCTCTCTACACCTGCAGCCGCTGGCCTTGCCTTAACACAGTTAACTGTTACATTTCT
CGACCTACTGAGAAAACCATATTTATTTATTTTATGTTGGCCGTTGCATGCTTGTCTCTA
CTGCTTAACCTGATTGAGGTTTACTATTTGGGATTACAAAAGTGCAAGGCAAGGACTGTCC
CCATCCCATATGAACAGTAACCAAGTTACCAACTAAGACTAGCAGTACAACCTCTACC
CACATTATCCCCCACTATCCCCATTATTCTCAACTGAGAAGGGATCTACATTCAGTCCA
ACTCTGGCATTTCCTCTGTCTCCAAGCAAGACCACGTCTAATGTAATGGACAGCAGCCAA
GCTGCGCGCAACAAAACCGTGATAACCAAGGCCGTGGAAACGCAATGGGATTCTGATCTG
AACGTTGCATCTGAAAGCAGCAGTGAGTTGGACCAATTTGAGAAGCAAGGCTCCCCAGC
AGAAACAGCAGGCACAGCAGCAACCGCTCTCGGCTGGAGTCTTGCAGTATAA

>Xt16305 Ensembl prediction. 5 predicted transcripts, of which ENSXETT00000035580 (below) is the longest. Prediction of 4 exons (black/blue/black/blue letters). The introns on either side of exon 2 (in blue letters) are predicted as 5 and 4 nt, respectively. The 5 nt intron is included into the reading frame without any changes (marked in green). The 4 nt intron (TGGG) gives a frame shift, and one G is removed, giving the sequence TGG (also marked in green). Furthermore, also to keep the reading frame, two single nt are removed (marked in pink in

Ensembl sequence, and by underlining the adjacent nt in our modified sequence), and three single Ns are added (marked in pink in the modified sequence).

```

ATGGGAGACTGGAGTTTCTTGGGTAATATTTTGGAAAGAAGTCAATGAGCACTCGACCGTG
ATTGGTCGAGTTTGGCTCACTGTTCTCTTCATTTTTCGGATACTATTCTGGGAACGGCT
GCTGAGTTTGTGGGGCGATGAACAATCTGATTTTGTGTGCAACACCCAGCAGCCCCGGT
GTAAAATGTCTGTACGACGAAGCTTTTCCTCCACATTAGGCTGTGGGTTCTCCAGATT
ATCTTTTCCACGCCATCCTCGGTCTACGTGGGGCACGCGGTGCACACGTTTCGCATGAAG
GAGAACTGGAAGGAGAGAGAAGAGGCTGAAATGAGCAGGCAGCAGGAGATGAACGAGGAA
CGATTACCCCTTGCCCTGATCAAGGGAGCATCAGGACCACGAAGGAGACAAGTACCAAA
GGAACCAAGAAGTTTCGCTTGAAGGAACCCCTCCTCAGAACCTACATCTGCCACATTATA
TTCAAAACTCTTTTTGAAGTCGGCTTTTGTGGTCGGCCAATACTTTCTCTACGGCTTCTGG
ATCAAACCCCTTGATCGTTGCAGTAGATGGCCCTGCCCAAACACGGTGGACTGCTTCGTT
TCCAGGCCGACGGAGAAAACGATTTTTCATTACGTTTCATGTTAGCCGTGGCTGCGGTCTCT
CTCTTTTGAACGTGGTGGAGATTAGCCACTTGGGATGGAAAAAGATCCGACTTGCTTTT
CGCAGACCCGACCAACATCCCGACCACTGCCAGGAGAGGTTTCCAGAAAATCATTGCAT
TCTATAGCAATTTCTTCTATTCCAAAGCTAAAGGCTATAAGCTGCTCGACGAAGAGAAA
GTAGTCTCGCGTTTCTACCCGATGACCGAAGTCGGATTAGAAGCCAGCCCGCTTCCGGCG
CACTTCAACGGCTATGACAAAGGCAGCACAGGGCTCCTAGAAGACATCTCAAAGGGTTAT
GATGAAACCTGCCCTGACCTCCCGAGCTGCCCTCTGAATTCATTCCCGACTCCAGATCG
CTGATGCTTTTGAACAGGCTTCTTCTGAACGTGCTGCCTCAGTGAGGTCCCCTTCGGTG
CGTGGCCCATCTGACCGAGCCCTCTGAACGTGCACCTTCTGAAAGGGCACCCACCCGT
GCTGCTTCTGAGCATGCAGGTTCAAGGACACACCCCTCTGTGCGAACACCTTCCAACAGT
GCAACTCCCGAACCTGCCCTCCCGAGCTGCCCTCTGAATTCATTCCCGACTCCAGATCG
CTGAGCAGGTTGAGCAAGCTAGCAGCAGGGCCCGCTCAGACGACTTAACGGTCTGA

```

>Xt16305 Modified according to Xl BC046682:35 (one nucleotide removed between underlined positions)

```

ATGGGAGACTGGAGTTTCTTGGGTAATATTTTGGAAAGAAGTCAATGAGCACTCGACCGTG
ATTGGTCGAGTTTGGCTCACTGTTCTCTTCATTTTTCGGATACTATTCTGGGAACGGCT
GCTGAGTTTGTGGGGCGATGAACAATCTGATTTTGTGTGCAACACCCAGCAGCCCCGGN
TGTAAAAATGTCTGTNACGACGAAGCTTTTCCTCTCTTCCACATTAGGCTGTGGGTTCTC
CAGATTATCTTTTGGTCCACGCCATCCCTGGTCTACGTGGGGCACGCGGTGCACCACGTT
CGCATGAAGGAGAACTGGAAGGAGAGAGAAGAGGCTGAAATGAGCAGGCAGCAGGAGATG
AACGAGGAACGATTACCCCTTGCCCTGATCAAGGGAGCATCAGGACCACGAAGGAGACA
AGTACCAAGGAACCAAGAAGTTTCGCTTGAAGGAACCCCTCCTCAGAACCTACATCTGC
CACATTATATTCAAACTCTTTTGAAGTCGGCTTTGTGGTCGGCCAATACTTTCTCTAC
GGCTTCTGGATCAAACCCCTTGATCGTTGCAGTAGATGGCCCTGCCCAAACACGGTGGAC
TGCTTCGTTTCCAGGCCGACGGAGAAAACGATTTTTCATTACGTTTCATGTTAGCCGTGGCT
GCGGTCTCTCTCTTTTGAACGTGGTGGAGATTAGCCACTTGGGATGGAAAAAGATCCGA
CTTGCTTTTTCGACAGCCCGACCAACATCCCGACCACTGCCAGGAGAGGTTTCCAGAAA
TCATTGCATTCTATAGCAATTTCTTCTATTCCAAAGCTAAAGGCTATAAGCTGCTCGAC
GAAGAGAAAGTAGTCTCGCGTTTCTACCCGATGACCGAAGTCGGATTAGAAGCCAGCCCG
CTTCCGGCGCACTTCAACGGCTATGACAAAGGCAGCACAGGGCTCCTAGAAGACATCTCA
AAGGGTTATGATGAAGCCCTGCCATCCTACCGTCAAGCGGAAGAGCTAGATGGTGTTAA
GTGCAAGCTTCTGATGCTTTTGAACAGGCTTCTTCTGAACGTGCTGCCTCAGTGAGGTCC
CCTTCGGTGCCTGCCCATCTGACCGAGCCCTCTGAACGTGCACCTTCTGAAAGGGCA
CCCACCCGTGCTGCTTCTGAGCATGCAGGTTCAAGGACACACCCCTCTGTGCGAACACCT
TCCAACAGTGCACCTCCGAACGAGTACCCTCCCGAGCTGCCCTCTGAATTCATTCCCGAC
TCCAGATCCCTTGAGCAGGTTGAGCAAGCTAGCAGCAGGGCCCGCTCAGACGACTTAACG
GTCTGA

```

>Xt05268 Ensembl prediction. 3 predicted transcripts, of which ENSXETT00000011516 (below) is the longest. 2 exons predicted (black/blue letters). We regard the 60 nucleotide predicted intron as a part of the reading frame (marked in blue).

```

ATGGGAGACTGGAAATTTCTTGGGGTCACTAGAAAGAAGTTCACATCCATTCCACAATT
GTTGAAAAAATTTGGCTTACAATTCTGTTTATATTCGCATGCTTGTATTGGGAGTGGCA
GCTGAAGAGGTTTGGAACTGATGAGCAGTCGAGTTTGTATGTAACACAGACCAACCTGGA
TGCAAAAAATGTTTGTATGACCAAGGCTTTTCTATCTCACTTATCAGATATTGGGTTCTG
CAAGTAATATTGTATCTCACCTTCTTTGGTTTACATGGGCCATGCTATTTACAGGCTG
AGAGCATTGGAGAAAGAACGACATAGGAAAAAGCTCAGCTTAGAGGGGAGCTTGAAGGA
GTTGAGTATGATCTGAGTGAAGACCGCAGAAAGTTGGAGCGTGAACCTCAGACACTTAGAG
CAAAAAAGGCTGAATAAGGCTCCATTACGGGGTCTCTTCTCTCACCTATGTAATACAT
ATTTTTACAAGATCTGCTGTGAGATTGGTTTTATGATTGGACAGCATTGCTTTATGGA
GTTTCAGTGAATCCTTTTACAAGTGTGAAAGAGACCCATGTCCCAATGTTGTTGATTGT
TTTGATACAAGGCTACAGAAAAAATGATTTTATGCTTTTTATGCAAGTGCATAGCAGCT
GTCTCTTTGTTCTAAATATTTTGGAAATTGCACATCTTGGGGTAAAAAGAATAAAAAAA
GCATATATGATGAGGTATAAATGAAAGAAGAGTTTGATGACTTTTATGTAATTAACCC
AAGAAAACTCCACAGTAAGCCACACTTGCATAGGTACGTCTCTAGTCCACAAAAGACT
CTTCTTCTGCACATGAAAATATACATTGTTGATAGAAAAACAGAGTACCCCACTGTG
TATACCGTTTTAAATCCTCCATGTGGGTACCATCCCATAGAAGATGGTCAGAATGATATT
TGTAATACTTTTCATGGAGGATGAGCAAGAAACAAAATCAGCCAATGGTATCAATGCCACA
AACACAGTTGAAAGTCACTATCATGACAACCTTGACAATAATAATGGGAGGATAAACAAA
ACATTAGATGCAGATGAAAAATGCTCCTCTTATACAGACAACAACAACTGATGACCATCT
GATGACACTGTGGATTACTCCTTACAAAATATATCCTGTCCAATGCCACTAACGTACTC
CGGAAATCCAGAAAGAGTCAGCGCTCCATGGAATGGTCTAAAGCAGCAGAGAGAGGAGGC
TTTGTTCAAGATCTTGGATGCAGGGGGCGCAGTAGCTTCAGTGCAATTTATCATGGGGA
CTCTCAAGTCTGATCTGAAACGAGTTAGCAGGCCATCCACACCAGAAAGTAATGGGAGAA
ATGAGCCCAAAATCCAACAGGATAAAACATGTGAAGGTTCTGTGACCTTCTCTCCATCA

```

CGACGGATGTCCTGGAAGTAATGCCAGCAGCCGGCGTGCCCCAACTGATCTACAGATC
TAA

>Xt05268 Modified according to XtDB 442000006

ATGGGAGACTGGAATTTCTTGGGGTCATCCTAGAAGAAGTTCACATCCATTCCACAATT
GTTGGAAAAATTTGGCTTACAATTCTGTTTATATTCGCATGCTTGTATTGGGAGTGGA
GCTGAAGAGGTTTGAATGATGAGCAGTCGAGTTTGTATGAACACAGACCAACCTGGA
TGCAAAAATGTTTGTATGACCAAGGCTTTTCTATCTCACTTATCAGATATTGGGTTCTG
CAAGTAATATTTGTATCCTCACCTTCTTTGGTTTACATGGGCCATGCTATTTACAGGCTG
AGAGCATTGGAGAAAGAACGACATAGGAAAAAGCTCAGCTTAGAGGGGAGCTTGAAGGA
GTTGAGTATGATCTGAGTGAAGACCGCAGAAGTTGGAGCGTGAACTCAGACACTTAGAG
CAAAAAAGGCTGAATAAGGCTCCATTACGGGGTTCTCTTCTCTTCACTATGTAATACAT
ATTTTTACAAGATCTGCTGTCGAGATTGGTTTTATGATTGGACAGCATTTGCTTTATGGA
GTTTCAGCTGAATCCTCTTTACAAGTGTGAAAGAGACCCATGTCCCAATGTTGTTGATTGT
TTTGTATCAAGGCCTACAGAAAAACTGTATTTATGCTTTTTATGCAGTGCATAGCAGCT
GTCTCTTTGTTTCCTAAATATTTTGGAAATGACATCTTGGGGTTAAAAGAATAAAAAAA
GCATATATGATGAGGTATAAATGAAAGAAGAGTTTGTGACTTTTATGTAATTAAACCC
AAGAAAACTCCACAGTAAGCCACACTTGCATAGGTACGTCTCTAGTCCACAAAAGACT
CTTCCTTCTGCACCTAGCAATTATACATTGTTGATAGAAAAACAGAGTGACCCCACTGTG
TATACCGTTTTAAATCCTCCATGTGGGTACCATCCCATAGAAGATGGTCAGAATGATATT
TGTAATACTTTTATGAGGATGAGCAAGAAACAAAATCAGCCAATGGTATCAATGCCACA
AACACAGTTGAAAGTCACTATCATGACAACCTTGACAATAATAATGGGAGGATAAACAAA
ACATTAGATGCAGATGTTAAGACTTTGCATTGCCATGGAGAAAGAAGAACTGGTTGTATA
AGAAGTACTTTATGTGATAAATGCTCTCTTATACAGACAACAACAACTGATGACCATCT
GATGACACTGTGGATTACTCCTTACAAAATATATCCTGTCCAATGCCCACTAACGTACTC
CGGAAATCCAGAAGAGTCAGCGTCCATGGAATGGTTCTAAAGCAGCAGAGAGAGGAGGC
TTTGTTCAGATTCTGGATGCAGGGGGCGCAGTAGCTTCAGTGCAAAATTTATCATGGGGA
CTCTCCAAGTCTGATCTGAAAGAGTTAGCAGGCCATCCACACCAGAAGTAATGGGAGAA
ATGAGCCCAAAATCCAAACAGGATAAAACATGTGAAGGTTCTGTGACCTTCTCCATCA
CGACGGATGTCCTGGAAGTAATGCCAGCAGCCGGCGTGCCCCAACTGATCTACAGATC
TAA

>Xt19125 Ensembl prediction. Two predicted transcripts, of which ENSXETT00000041461 (below) is the longest. Predicted as 2 exons (black/blue letters). Insertion of the 129 nt predicted intron as a part of the reading frame (marked in blue). Further extended the 3'-end by 114 nt (also marked in blue) to reach a stop codon.

ATGGGGGACTGGAATTTGCTGGGAAGCATACTGGAGGAGGTGCACATCCATTCCACCATT
GTCGGAAGATCTGGCTGACAATTCTTTTTGTATTCCGGATGCTTGTCTTGGGGTGCT
GCCGAAGATGCTGGGATGACGAACAAGATGAATTTATTTGTAATACTGAGCAACCAGGC
TGCAAGGAATGTTTGTATGACCAAGCTTTCCCAATCTCACTTATCAGGTACTGGGTACTG
CAAATTATATTTGTCTCTTCTCCATCCCTTGTTTACATGGGCCATGCACTGTACAGACTG
AGGGCTCTGGAAGGAGAGGCAAAAAGAAGAAAGCCCACTTAGAGCAGAGCTGGAGGAA
CTTGAAGTGGTTGTTGATGAGCAGAGAAACTGGAGAGGGAACCTAGGAAATTAGAAGAA
CAAAGGAAAGTAAACAAGGCTCCATTGAGGGGGTCCCTATTACGCACATATGTTCTGCAC
ATACTGACAAGGTGCGTAGTAGAAGTTGCTTTTATGATAGGGCAGTACTTGCTGTATGGA
TTTGACCTAGACCCCTTGATATAATGTTACGGTCCCCTTGCCCTAATATAGTGGACTGC
TTTGTCTCAAGACCAACAGAAAAGACTATTTTTATGGTATTTATGCACAGTATAGCAGCA
GTGTCAATTGTTCCCTAAATGTTTGAATACTTCCATTTAGGCATTAGAAAAATAAAGCAG
GGGCTTTCTGGTAAATCCAATGCTGAGACAACAGAGGATGAAATAAGCAGCTGCAAGTCC
AAGAAGAACTCTATGCAGCCTGTTTGTATCATGAGCAATTCCTCCCAAGTAAGATCATC
CCCTTGCTTCAAGTGGCTACAAACTGCTGCCTGATCAACAAATGGATCCGGTAGGACTA
CCAGCTTACTTGCCCTCCAGGACTTTAAGGAGGCACAAATGATGAGTGATCAAAAT
CATTATGGGAAAGTTCTTGCCCTGACAAGCATAGGATACCCACAATTCAGCCTACCCCT
GCCCTCCAGAAGCGCAATCTGTCTCCAGCAGTGAAGATTCCCAAGAACTTAAACAC
AACACCTGTTTGGGCTCAAAAAGTTGCCTCAAAAGCCAGCATTCCCTGGATCAGCCAGG
CACAGTTCTGTGACCCACTAAGGTGATGCTGCACTCCAGCCACATGGAGCTTCACTCT
GCACACGAAAGTACAGCAGGTTGGCAGCTGCAAGGACTTTGCTGAGGATCGAGTGAT
TCTGCAGACAGTGGACATAGAAAAGTGAAGTTCTACATCCAGGGGATTATCCGAGAGCAGA
CTGATCAGTGACACTGAGAGTTTGGACTCTAGAAATGGCTCTGGCTCTGAATCTAGAAGG
AGAGACGATAGTCAAGCATCACTCCACCTCCACCATCTGGTCGAGGATGTCAATGGCA
AGTAAAGCA

>Xt19125 Our suggested modification. Position corresponding to potential intron start in other Cx62 orthologs is underlined.

ATGGGGGACTGGAATTTGCTGGGAAGCATACTGGAGGAGGTGCACATCCATTCCACCATT
GTCGGAAGATCTGGCTGACAATTCTTTTTGTATTCCGGATGCTTGTCTTGGGGTGCT
GCCGAAGATGCTGGGATGACGAACAAGATGAATTTATTTGTAATACTGAGCAACCAGGC
TGCAAGGAATGTTTGTATGACCAAGCTTTCCCAATCTCACTTATCAGGTACTGGGTACTG
CAAATTATATTTGTCTCTTCTCCATCCCTTGTTTACATGGGCCATGCACTGTACAGACTG
AGGGCTCTGGAAGGAGAGGCAAAAAGAAGAAAGCCCACTTAGAGCAGAGCTGGAGGAA
CTTGAAGTGGTTGTTGATGAGCAGAGAAACTGGAGAGGGAACCTAGGAAATTAGAAGAA
CAAAGGAAAGTAAACAAGGCTCCATTGAGGGGGTCCCTATTACGCACAATGTTCTGCAC
ATACTGACAAGGTGCGTAGTAGAAGTTGCTTTTATGATAGGGCAGTACTTGCTGTATGGA
TTTGACCTAGACCCCTTGATATAATGTTACAGGTCCCCTTGCCCTAATATAGTGGACTGC
TTTGTCTCAAGACCAACAGAAAAGACTATTTTTATGGTATTTATGCACAGTATAGCAGCA
GTGTCAATTGTTCCCTAAATGTTTTAGAAATCTTCCATTTAGGCATTAGAAAAATAAAGCAG
GGGCTTTCTGGTAAATCCAATGCTGAGACAACAGAGGATGAAATAAGCAGCTGCAAGTCC
AAGAAGAACTCTATGCAGCCTGTTTGTATCATGAGCAATTCCTCCCAAGTAAGATCATC
CCCTTGCTTCAAGTGGCTACAAACTGCTGCCTGATCAACAAATGGATCCGGTAGGACTA

CCAGCTTACTTGCCCCATCCCAGGACTTTAAGGAGGCACAAATGATGAGTGATCAAAAT
CATTATGGGAAAGTTCTTGCCCCTGACAAGCATAGGAAACTGCAGACCATAGTTGCACT
GAGCATCAGCAGCAGGGAAGCCATTGCAAGCAAATCCTCCACTCTGCAGAGCCCCACAGG
ATACCCAATCACACTGACCCTACCATGGGCACATGACACACAGGATACCCACAATTCAG
CCTACCCCTGCCCTCCAGAAGCGCAATCTGTCTCCAGCAGTGAAGATTCCCAAAGAAA
CTTAAACACAACACCTGTTTGGGCTCAAAAAGTTGCCTCAAAAGCCAGCATTCCCTGGAT
CAGCCCAGGCACAGTTCTGTGCACCCACTAAGGTCATGTCTGCACTCCAGCCACATGGAG
CTTCACTCTGCACTACGAAAGTACAGCAGGGTTGGCAGCTGCAAGGACTTTGCTGAGGAT
CGCAGTGATTCTGCAGACAGTGGACATAGAAAAGTGAGTTCTACATCCAGGGGATTATCC
GAGAGCAGACTGATCAGTGACACTGAGAGTTTGGACTCTAGAAATGGCTCTGGCTCTGAA
TCTAGAAGGAGAGACGATAGTCCAAGCATCACTCCACCTCCACCATCTGGTCGCAGGATG
TCAATGGCAAGTAAAGCAGCCAGCCCTGGGCTAATACTGGCCAGTGGTTCCCACTCTCAG
TCTAGGCACCCTCATTATCATTATTACATATTACAAACAGTAAACAAAAATCCCTTAATA
AAAAGGTCATAG

Suppl. Figure 3 *Connexin* sequences from zebrafish (*Danio rerio*, Dr). The sequences are found in the Ensembl database or GenBank. The sequences were updated as of July 1 2006. If the sequences previously have been named within the connexin nomenclature, this is normally the name that is used here. In Suppl. Table 6, information on zebrafish databank entries is given. The Ensembl gene identification is given in abbreviated form. For the ortholog of Cx25, the abbreviated form *Dr59215* should be read as ENSDARG00000059215. Other information is given in short form in the title line where needed. If our interpretation of the most likely connexin sequence differs from the prediction found in Ensembl (or GenBank), the database sequence is given first, and immediately followed by our suggested sequence, which is the one used in our analyses. The conserved domains are marked in yellow (unmodified sequence) or green (modified sequence). Alternating black/blue sequences are exon predictions copied from Ensembl. Other colors are explained for the relevant sequences. Please note that both the Ensembl gene identification number and the Ensembl gene prediction might change on future updates of the zebrafish genome database.

```
>Dr59215 Ensembl prediction. No modification.
ATGAACCTGGGTTTCTCTGGAGAACGTGTTGAGCGGGGTGAACCGCTACTCCACCGTGGTG
GGCCGGGTCTGGCTCTCCATCCTCTTCCGTCATCTGGTGTTCGTGGCGGCCGCC
GAGCAGGTGTGGAAGGACGAGTTCAAGGACTTCGTCTGCAACACGACGAGCCGGGTGC
GAGCAGGTGTGCTTCGACCATTCTTCCCATCTCTCAGGTGCGTCTGTGGGCGCTGCAG
CTCATCATGGTGTCCACGCCGCTGCTGCTGGTGGCTCTGCACGTGGCCCTACCGAGAGCAC
CGCGAGCGCAACACAAGCGCAGGCTCTACCAGGACAAGGGCAGCATTGACGGCGGCTG
CTGTTACAGTACATCACCAGCCTGGTCTCAAGACGTCTTTCGAGGTGGGACGCTGCTG
GCCTTCTACCTGCTGTACAGCGGTTTCCACGTGCCGCGGCTGCTGCGTGCGCCGAGAGC
CCCTGCCCAACACGCTGGACTGCTACATCGCCAGAGCCACCGAGAAGAAGATCTTCTC
TACATCATGGGCTGCACCTCCATCCTGTGCATCGCGCTCAACCTGCTGGAGATGGGCTAC
ATCGTGTCCAAGCAGTGTGGAAGAGCTTCAGCAAGAGATACACTCCGGTGCGGGATGGG
GCCACTCGCCCGCTCCACCTTTACTCTCGGCCACCGCAACCTCATGCACAGCCAAG
GAGGAGGCGATCAGTCTGCCCGCGGGACAGGAGACAGCCTGA
```

```
>DrCx33.8 Dr42707 Ensembl prediction. No modification.
ATGAGTTGGGGAGCACTTTATGCTCAGCTGGGAGGAGTGAATAAACACTCCACCAGCTTG
GGGAAGATCTGGCTGTCTGCTCTTCCCTCTTCCGTCATTGTCATCTGGTTCATAGCAGCA
GAGACGGTCTGGGGAGACGAACAGTCAGACTTCACCTGCAACACACAAGCCTGGTTGC
AAAAACGTCTGCTATGACCACTTCTTCCAGTCTCGCACATACGTTTCTGGTGTCTGCAG
CTCATCTTTGTGTCCACACCGGCTTTACTGGTAGCTATGCATGTGGCATATCGCAAGCGC
AACATGAAAAAGAAAAGCACTTTTAGCCAAGCGTGGAGGTAATGTTAAAGAGATGACCTG
GAGAGCTTGAAGAACCGGCTCTACCCATCACTGGGCCACTGTGGTGGACCTACACATCC
AGCCTGTTCTTACAGACTTCTTTTCGAGGCCGGATTTCATGTATGCTCTCTATTACGTCTAT
GATGGCTTTTCAGATGGCAGCGCTTGTGAAGTGTGAGCAATGGCCTTGTCCCAATAAAGTT
GACTGTCTTTCAGAGCTCAAGGCCGACAGAGAAGACGGTCTTCAACCTCTTTATGGTGGGATCT
TCTGCTATTTGCATTGTGCTCAATGTGGCTGAAGTGGCCTATCTGATTGTCAAAGCATTG
CTCAGGTGCTCAGCCAGAGCCAAAGGGAGGCGCTCATTGTACACCAAGAGAAAATGTCC
ACAGAAAAGGCGCACCTACAGAAATGAAAAAACGCAAGGTTGCTGTATCAGCTTCGGAC
TCATCGAGCAATAAGACTGTTTAA
```

```
>DrCx34.4 Dr59220 Ensembl prediction. No modification.
ATGAATTGGGCTTTTCTTCAGGGTCTCCTGAGCGGGGTCAACAAGTACTCCACAGCGTTC
GGCCGTGTCTGGCTCTCGATAGTCTTCTTTTTCAGAGTCATGGTTTTTGTAGTCGCGGCT
GAAAAAGTGTGGGGCGACGAGCAGAAAGACTTTGCGTGTAAACACCGCCAGCCGGGATGC
CATAATGTATGCTATGACCACTTCTTCCCGTGTCCACATCCGCTCTGGGCTCTGCAA
CTCATCTTCGTCACTTGTCCGTCACTTATGGTGGTTTTACATGTGGCATATCGTGATGAA
CGTGAGCGGAAAAACCGTCTCAAATATGGTGAAGGATGTAACGTTTGTACGACAACACC
GGAAAGAAACGTGGTGGTCTTTTGGTGGACGTACGTGCTCTCGCTGGTTTTTAAATGGGA
GTGGATGCGACTTTTGTGTATCTGCTGTACCACATCTACGAGGGCTACGATTTTCCAGTT
CTGGTGAATGTTCTGAAGCTCCATGCCCAAACATTGTGGACTGCTTCATCTCGCGGCC
ACAGAGAAGCGAATCTTCAACATCTTTATGGTGGTGACCACTGCTGGTGTGCATCTGCTG
TCTCTCTTCGAGATCTCTATCTGCTGGTGGGCAACGCTGCTTTGAATGCATCAATAGGGTG
CAGAGCTCAGACATGTGAACAGAGAGATCCATGGCTAATATGACAACTTGAATGCT
CATTTAGAGTCAAAACAACAAAAAAGTGGCAAGCGAAGACCAGCCGGCACCAGCATACAGT
GTGGTCATGTACGCAAAAAGAAAACAGCTTGGAAAAATCTTCAATCCGAGTTGGACC
TTGAGAGATGACAAATGTTCAACTTCACATTCACTTGGAGGTGAATGCAACGTGACTGA
```

```
>DrCx35.4 Dr42866 Ensembl prediction. No modification.
ATGGACTGGAAGACTTTTCAAGCCCTGCTCAGCGGGGTGAACAAACTCCACTGCATTG
GGCCGGATATGGCTCTCAGTGGTTTTTGTGTTCAAGGGTCATGGTTTTATGTCGTAGCGGCA
GAAAGAGTTTGGGGTGATGAGCAGAAAGACTTTGACTGCAACACCAAGCAGCCGGGCTGC
GCAACAGTCTGCTATGATTTCTACTTCCCATTTCCACATAAGACTATGGGCTCTGCAG
CTCATCTTCGTACAGTGTCCATCACTAATGGTGGTCATGCACGTGAAATACCGTGAGGAA
```

CGTGAACGCAAAGCCAAAGCAAACTCTACGCCAACACGGGAAAGAAGCACGGTGGACTG
 TGGTGGACG**TATCTGATCAGCCTTTTTGCTAAGACTGGCATTGAGATCACCTTCCTGTAC**
ATCCTCCACCACTCTACGACAGCTTCTACCTGCCAAGGCTGGTGAAGTGTGATGTCCAG
CCATGTCCCAATGTTGTGGACTGTTACATTGGCCGGCCACAGAGAAAAGAGTCTTCACT
TATTTTCATGGTGGGAGCGCTCAGCGCTCTGCATAGTGCTCAGTGCTGCGAGATCATCTAT
CTGATCGCCAAACGCATCAGCCGCTGCGCTAACAAATACAAGCAGCATGACAAAAGAAGC
 ACACCGATTAATCAGCGATTCGAGATGAGGACAGCAACTGCATATTCTTTGACAGAG
 CTGGAGAGCAAGCCGAGTATAAACCAGAGACTAAAGAAGACTTTAGGTCTGAATATAAA
 CCTGGCCCTAAACGCAGTTTAAATCTGAAGTCAAGCCCACTTTTAAGCCAGCCTACAGG
 TTGAGTGTGGACATGAGAGCGTCTGCTCCAAATCTCTCAGCACCATGTACAAAATACAG
 TCTGGTATCATCTAA

>DrCx28.6 Dr03925 Ensembl prediction. No modification.

ATGAACTGGTCGGGATTGCAAGTCCCTTTCTGAGCGGGGTCAATCAATATTCGACCGTGTTT
GGTCGAGTGTGGCTACCGTGGTGTGTTTGTGTTTCGCGTCTGGTGTGTTGATGTGACAGCT
CAACGCGTTTGGGGTGACGAAAACCTTAGTGTGCAACACCAGGCAACCCGGCTGTGCCAAC
GTCTGCACCGACACCATTTCCCATCTCTCACACCCATCTGTGGGCATTACAGCTCATC
TTCTGCATATGTCGCTCGCTCATGGTCTATAGCCACGTCAAACTCAGAGAAGACAAAAAC
 AAGAAGTACACAGACGCTCATAGGGGAGAGCATTTATACGCCAACCTGGAAAAAAGCGT
 GGTGGTCTTTGGTGGACT**TACCTGCTGAGCTTGCTCATTAAAGTCATTGTGACGCTGGT**
TTTCTGTATATCTTCATTAACCTGTATACGGCTTTGATCTTCCTCGCCTTGTCAAAGTGC
TCGCTGGATCCTGTCCAAATACAGTGGACTGTTTCATCTCTCGTCCCACAGAGAAGAAG
ATCTTCACTCTTCACTGATTTCCAGTGTGGTCTGCTATCTTCAATGTGCATCTGTGAA
ATGGCTTATCTCATTGGAAAGCGAGTATCCAATAACTGTATGATGGTAAAGGACCAGCA
 CAAAGATCCAAAACACATGATCCAACCTTTCAAGCAGACAGAACTTAATTTACAAAAA
 ATAAAAGAGAAAAAATAGACAATACAGCTCTTTAA

>DrCx30.9 Dr28833 Ensembl prediction. No modification.

ATGAACTGGTGTCCCTAGAAGTCTGCTTGGCGGGGTAGCCAATACTCCACTGTGTTT
GGCCGTGTCTATCTCTCCGTGGTGTTCATCTTCCGAATCCTGGTGTGTTGTTGTTGCTGTCTGTC
CAGCAAGTCTGGAACGACGAGCAAGAAAGACTTCATCTGCAACACGGCCAGCCAGGCTGC
ACCAATGTCTGCTACGACCAAGTCTTCCCATATCCACATCCGTCTATGGGCTCTCCAG
CTTATTTTCGTCACTTGTCCATCTCTCATGGTGGTCTGCTCAGCTCAAATATCGACAAATG
 AAGAATGTGAAGTACAACACTGCCCGCAATGGTGAAAACATGTATGCAAAACCAGGAAGA
 AAGCGTGGAGGCGCTGTGGTACACCT**TATATCCTCAGCCTGTTATTCAAAGCCGGCTTTGAT**
GCAGCATTTCTGTATATTTGTACTACCTTTATAAATTCGACATGCCAAATGTTACCAA
TGCACTGCGGAACCTTGTCCAAATACAGTGGACTGCTATATCTCCCGTCCCACAGAGAAG
AAAATCTTCACTCTGTTTATGGTGGTCTCCTCCTCTGTGTGCATCTTCATGTGTATCTGT
GAGATGGTGTATCTGATTACAAAGAAAGCTGGCAAAATTTCTGCACAAAAGAGTGAGGAG
 AATAGAAACTACATAAGCAGAGGGTGGATCCGACGGCCATGTCCAACCAGAACTCTCAAC
 AATCTTAAATGGCACAGGCTGCAGAAGAACTAAAAAATCTGAGATGATCCATTCGGAG
 TTGAACAAATTATATGACGCTGA

>DrCx27.5 Dr35533 Ensembl prediction in 2 exons (black/blue letters). Exon 1 and 14 nt of exon 2 removed (marked gray) to reach initiation codon consistent with other connexins. The sequence is then in accordance with NM_131811.

ATGCCACTAACACCACCTGCGGAGCCTGATCTGGAACCAAAATGAACTGGGCGTCATTT
TATGCCGTGTATCAGCGGCGTGAACCGACATTCCACCGGCATTGGGCGGATTTGGCTGTCT
GTCTCTTTCATTTCCGGATCCTGGTCTGGTGGTGGCGGCGGAGAGCGTGTGGGGCGAC
GAGAAAGCGCATTTTCATCTGCAACACCCAACAGCCGGGATGCAACAGCGTGTGCTACGAC
CACTTCTTCCGATCTCCACATCCGACTGTGGGCCCTGCAGCTCATCATGCTCTCCACC
CCGCGCCTGCTGCTCGCATGCACATTGCACACCGTGGGACATCGACAAAAAGTTGTAT
CGCCAGGCTGGCCGACACAGCCGAAAGACTTGGAGGCGATAAAGAACCAGAAGATGAAG
ATTACCGGCGCCCTCTGGTGGACATATATGATCAGCCTGCTGTTCCGTGTGTTTCGAG
TCCGCTTTATGTATCTGTTTACATGATTTACCGGGCTATAAGATGTTCCGGCTGGTG
AAGTGTGACTCGTATCGGTGCCAAACATTGTGCACTGTTTCGTGTCCAGGCCGACAGAG
AAAACAGTCTTCACTATATTTATGCTGGCGGTGTCCGGCTGTGATCCTGCTCAACATC
GCCGAAATCGTCTTTCTGTGGCGAGAGCAACAGTCGACATCTCAATAACTCCAAAGAT
TCCGCTGTGGGAGCCTGGATCTCCAAAAACTCTGCTCCTTCTAG

>DrCx27.5 Dr35533 Modified according to NM_131811.

ATGAACTGGGCGTCATTTTATGCCGTGATCAGCGGCGTGAACCGACATTCCACCGGCATT
GGGCGGATTTGGCTGTCTGTCTCTTTCATTTTCCGGATCCTGGTCTGGTGGTGGCGGCG
GAGAGCGTGTGGGGCGACGAGAAAGCGCATTTTCATCTGCAACACCCAACAGCCGGGATGC
AACAGCGTGTGCTACGACCACTTCTTCCGATCTCCACATCCGACTGTGGGCCCTGCAG
CTCATCATGGTCTCCACCCCGCCTGCTGGTGGCATGCACATTGCACACCGTGGGCAC
ATCGACAAAAAGTTGATCGCCAGGCTGGCGCACAGCCGAAAGACTTGGAGGCGATA
AAGAACCAGAAGATGAAGATTACCGGCGCCCTCTGGTGGACA**TATATGATCAGCCTGCTG**
TTCCGTGTGTTTTCGAGTCCGCCCTTTATGTATCTGTTTACATGATTTACCCGGGCTAT
AAGATGTTCCGGCTGGTGAAGTGTGACTCGTATCCGTGCCCAACATTGTGCACTGTTTC
GTGTCCAGGCCGACAGAGAAAACAGTCTTCACTATATTTATGCTGGCGGTGTCCGGCGCTC
TGTATCCTGCTCAACATCGCCGAAATCGTCTTTCTGTGGCGAGAGCAACCAAGTCGACAT
 CTAATAACTCCAAAGATTCCGCTGTGGGAGCCTGGATCTCCAAAAACTCTGCTCCTTC
 TAG

>Dr58064 Ensembl prediction. 108 nt added at 3-end to reach stop codon (marked blue).

ATGAATTGGGCATCCTTTTATGCTGTGATCAGTGGTGTGAACAGGCATTCAACAGGCATT
GGACGCATCTGGCTGTACGTCATCTTTCATCTTCCGTATCTTGGTGTAGTAGTGGCAGCC

GAGAGCGTGTGGGGAGATGAAAAGTCAGGCTTTACCTGCAATACTCAGCAACCCGGCTGC
AACAGCGTGTGTTATGACCAGTCTTTCCAATCTCACACATCCGCCTTTGGATTTTGCAG
CTCATTCTGGTGTCCACACCAAGCCCTACTGGTCACTATGCATGTTGCACATCGGCGACAC
GTTGAGAAAAAGATCCTCAAGATATCTGGTCAGGGAACGAAAAGGACTTCGAGAGCATT
AAAACCCGAAAAGTTCAAATTTGTTGGTGCATATGGTGGACTTACATGATAAGTATCATA
TTTCGCATAATTTTGAAGTGGTTTTCTGTACATTTTCTACTTAATCTATCCAGATATC
ACTATGGTTCGTCTTGACAAATGTGACTCATATCCATGTCCAAATACAGTAGACTGTTTT
GTGTCTCGTCTACAGAGAAGACCATTTTTACTGTCTTTATGCTGGTGGTGTCTGGACTT
TGTGTCTTGCTAAATATCACAGAGGTTATGTATTTAATAAATCGGGCATGTATCAAATAT
TTTCAAGGAGCAGTACATCAAATAAAGGACCTTGGCTCACTCATAAACTGGGAACCTAT
>Dr58064 Our suggested modification.

ATGAATTGGGCATCCTTTTATGCTGTGATCAGTGGTGTGAACAGGCATTCAACAGGCATT
GGACGCATCTGGCTGTCAAGTCATCTTCACTCTTCCGTATCTTGGTGCTAGTAGTGGCAGCC
GAGAGCGTGTGGGGAGATGAAAAGTCAGGCTTTACCTGCAATACTCAGCAACCCGGCTGC
AACAGCGTGTGTTATGACCAAGTTCTTTCCAATCTCACACATCCGCCTTTGGATTTTGCAG
CTCATTCTGGTGTCCACACCAAGCCCTACTGGTCACTATGCATGTTGCACATCGGCGACAC
GTTGAGAAAAAGATCCTCAAGATATCTGGTCAGGGAACGAAAAGGACTTCGAGAGCATT
AAAACCCGAAAAGTTCAAATTTGTTGGTGCATATGGTGGACTTACATGATAAGTATCATA
TTTCGCATAATTTTGAAGTGGTTTTCTGTACATTTTCTACTTAATCTATCCAGATATC
ACTATGGTTCGTCTTGACAAATGTGACTCATATCCATGTCCAAATACAGTAGACTGTTTT
GTGTCTCGTCTACAGAGAAGACCATTTTTACTGTCTTTATGCTGGTGGTGTCTGGACTT
TGTGTCTTGCTAAATATCACAGAGGTTATGTATTTAATAAATCGGGCATGTATCAAATAT
TTTCAAGGAGCAGTACATCAAATAAAGGACCTTGGCTCACTCATAAACTGGGAACCTAT
AAGCAGAATGAAATAAATAATTTGATATCAGAGCATTCAATTTAAACCTAGATTTAATGTT
GGGCGGAAACCTCCAGTGTGAAAAATGAGCGCTGCTCAGCTTTCTAG

>DrCx35 Dr58009 Ensembl prediction. No modification.

ATGGGGGAATGGACAATTCTCGAGCGTCTCCTGGAGGCGGCTGTCCAACAGCACTCTACT
ATGATTGGGAGGATCCTGCTAAGTGTGTGGTGTATCTCCGGATTCTAATTGTGGCGATT
GTTGGAGAGACCGTGTACGACGACGAGCAGTCAATGTTGTGTGTAACTCTGCAGCCA
GGCTGTAAACCAAGCTTGCATGACAAAGCGTTCCCTATATCTCACATCAGATACTGGGTT
TTCCAGATCATCATGGTTTGCACACCCAGTCTCTGCTTCATCATACTCTGTGCATCAG
TCGGCCAAACAGAAAGGAGCGGAGGTATTCTACCATCTACCTGTCCCTCGACAAAGACCCC
GATACGATGAGGCGAGACGACAGCAAAAAGATCAAAAACACCATTTGTGAACGGAGTACTT
CAAAACGCGGAGACTCCACCAAGAGTCCGAGCCTGACTGTCTAGAGGTCAAAGAGATC
CCCAATTCAGCCATGAGAACTACCAATCTAAAATGAGAAGACAAGAGGGCATCTCCAGG
TTTTACATCATTCAAGGTGGTGTTCAGAAACGCACTGGAATTTGGCTTCTGGTGGGCCAG
TATTTCTTGACGGATTCAACGTGCCCGCCGCTGTACGAGTGCGACCGCTATCCTTGCATC
AAAGACGTCGCAATGCTACGATCAAGACCTACGGAGAAAACCGGTGTTCCTCGTCTTCATG
TTTGACGTCAAGTGGGATTTCGCTGGTGTCAACCTGGCTGAACCTCAATCACCTGGGCTGG
AGGAAAATTAAACAGCGGTGAGGGGGGTTACGGCCAGGAGGAAGTCCATCTATGAGATC
AGAAACAAGGACTTACCGCGGATGAGCATGCCGAATTTGGAAGAACCAGTCCAGTGAC
TCTGCCTACGTTTAA

>Dr35340 Ensembl prediction. 36-like sequence. 71 nt corresponding to exon 1 found approx 22
kb upstream (marked green). To fit this to the general pattern of splicing sites, the two
first nt of the Ensembl prediction are regarded as the two last nt of the intron, and
therefore removed (marked gray in upper sequence).

AGGATCCTGCTGACAGTAGTGGTGATATTCCGGATCCTGATCGTGGGTATAGTGGGAGAG
AAGGTGTATGAGGACGAACAAATTATGTTTATATGTAACACACTGCAACCGGGTTTGAAC
CAGGCTGTCTACGATAAAGCCTTCCCAATCTCCATATCCGTTACTGGGTTTTCCAGAT
ATCCTGGTTTGCACGCCAGCCTCTGCTTTATCACATACTCTGTGCACAGTCTGCCAAG
CACAAAGACCAGCGTTACACACTTCTGCATGGCCCTTACATCGACCACGGTCATGGGCCG
AGCCGCAAGCTCCGCAACATCAACGGCATCCTGGTGCACCCGGAAGCAAGACGACCGC
GAATGTCTGGATCTGAAAGACATTCCCAACATCCCGCGGGGGTGACATACTCCAAAAGT
GCCAAGATCCGCGCAGCAGGAAGGCATCTCCCGCTTCTATGTCAATCAAGTGGTGTCCGG
AACGTTCTGGAATCGGCTTTCTAGCCGGCCAGTACTTCTCTACGGATTCAATGTTCCC
GCCATGTTTGAGTGCGACCGCTACCCCTTGCCTGAAGGAGGTTGAATGTTACGTGTGCGGT
CCCACAGAGAAGACAGTTTTCTGGTGTTCATGTTTCGAGTTAGCGGAATCTGCGTGGTG
CTAAATTTGGCTGAGCTCAACCACTTGGCTGGCGCAAGATTAAAGACGGCCATCCGTGGC
GTCCAAGCTCGTAGAAAGTCTATTTGTGAGATCCGAAAAAAGGATGTGTCTCACTTGTCC
TCCGTGCCCAACTTGGGCGCACCCAGTCTAGCGAATCAGCTTATGTCTGA
>Dr35340 Our suggested modification.

ATGGGAGAAATGGACTATATTGGAGAGGTTGCTGGAGGCTGCGGTACAGCAGCACTCTACT
ATGATCGGCAGGATCCTGCTGACAGTAGTGGTGATATTCCGGATCCTGATCGTGGGTATA
GTGGGAGAGAAGGTGTATGAGGACGAACAAATTATGTTTATATGTAACACACTGCAACCG
GGTTGCAACCAAGCCTGCTACGATAAAGCCTTCCCAATCTCCATATCCGTTACTGGGTT
TTCCAGATCATCCTGGTTTGCACGCCAGCCTTGTCTTTATCACATACTCTGTGCACAG
TCTGCCAAGCACAAAGACAGCGTTACACACTTCTGCATGGCCCTTACATCGACCACGGT
CATGGGCCGAGCCGCAAGCTCCGCAACATCAACGGCATCCTGGTGCACCCGGAAGCAAA
GACGACCGCAATGTCTGGATCTGAAAGACATTCCCAACATCCCGCGGGGGTGACATAC
TCCAAAAGTGCCAAGATCCGCCAGCAGGAAGGCATCTCCCGCTTCTATGTCAATCAAGT
GTGTTCCGGAACGTTCTGGAATCGGCTTTCTAGCCGGCCAGTACTTCTCTACGGATTC
AATGTTCCCGCATGTTTGTAGTGCGACCGCTACCCCTTGCCTGAAGGAGGTTGAATGTTAC
GTGTCGCGTCCCACAGAGAAGACAGTTTTCTGGTGTTCATGTTTCGAGTTAGCGGAATC
TGGTGGTGTCTAAATTTGGCTGAGCTCAACCACTTGGCTGGCGCAAGATTAAAGACGGCC
ATCCGTGGCGTCCAAGCTCGTAGAAAGTCTATTTGTGAGATCCGAAAAAAGGATGTGTCT

CACTTGTCCTCCGTGCCCAACTTGGGCCGCACCCAGTCTAGCGAATCAGCTTATGTCTGA

>Dr35765ic Ensembl prediction. 36-like sequence with no predicted exon 1. A reasonable exon 1 is found approx. 60 kb away, but it is on the opposite strand. An assembly error can not be excluded, but the sequence is not included here.

NNGATTTTGCTGACAGTTGTGGTGATATTCGGTATCCTAATCGTGGCCATTGTTGGTGAA
ACCGTCTATGAGGATGAGCAGACCATGTTTATCTGCAACACCCCTCCAACCGGGCTGCAAC
CAGGCCTGCTACGACAAAGCCTTCCCCATTTCCACATCCGATACTGGGTCTTCCAGATT
ATCCTTGTCTGTACACCAAGCCTGTGCTTCATCACCTATTTCAGTACACAGTCAGCCAAA
CAGCGTGACCGCCGCTACTCCTTCTCTACCAATAATGGAAAAGGACTACAGTCGTGAG
GGTACGCGGAAACTACGCAACATTAAACGGCATTGTTGGTGCAGCACTCTGAAAGTGCGGT
GGAAAGGATGAACCTGATTGCTTAGAGGTAAAGGAGATCCCAATGCTCCAAGAGGCCTC
TTGCATGGCAAGAGTTCAAAGGTACGCCGACAAGAGGGGATTTCTCGTTCTACATAATC
CAGGTAGTGTTCCGCAATGCACTGGAGATAGGTTTCTTAGCAGGGCAGTACTTCTATAT
GGCTTCAGCGTTCTGCGCATCTTCGAGTGTGACCGATACCCCTGCGCTGAAGGAAGTGGAG
TGCTACGTATCCCGACCCACTGAGAAGACGGTCTTCTAGTGTTCATGTTTGCAGTGAGT
GGCATCTGCGTTGTGCTGAACCTCGCTGAACCTCAACCCTCGGCTGGAGAAAAATCAAA
GCTGCCATTTCGAGGTGTCCAAGCCCGCAGGAAGTCCATCTGTGAGATCCGCAAAAGGAC
ATGGCCCACTGTCAACCGCCAAACCTGGGCAGGACCCAGTCCAGCGAGTCTGCTTAT
GTTTGA

>Dr17927 Ensembl prediction. Predicted as 2 exons (black/blue letters). 92 nt removed at 5'-end to achieve an initiation codon more in accordance with other connexins (marked gray). The 48 nucleotide intron (marked blue) is regarded as a part of the reading frame. Sequence extended 252 nt in 3'-end to reach stop codon (marked blue).

ATGAAAACGGCCGCTCCTCGTGGCAGCAGAGCTAAAGTGGAGCTATTTAGGTGCTGTAC
TTCAATCGGCCGGCAGCAGACATGACAGAATGGACGCTGCTGAAGCGGCTGCTGGAC
GCCGTGCACCACTCCACCATGATCGGACGCCCTCTGGCTCACAATAATGGTGATTTTC
AGATTGCTGATCGTCGCAAGTGGCGACCGAAGACGTCTACACTGATGAGCAGGAGATGTT
GTGTGCAATACTCTCAACCGGGATGTCCGAACGTCTGCTATGATGCATTGCGCCAATA
TCGCAACCAAGTCTTTGGGTCTTCCAGATCATCACGGTCTCCAGCCGCTCGCTTTGTTTT
ATTATCTACACCTGGCACAACCTGTCCAAACAACCCGAAGGCTCCATTAAATCGCATAAA
CACATAAATCCAAGCCTTGAAGGAGTCACCAACAGAAACCATCGCAAGCATCGAAAACC
TCTTCGGGAGTCCCTTCGAAGTATTACATCTTCCACGTTTGTCTTCTACCATTTCTGGAA
GTGGCCTTTGTAGTGGCCAGTGGCTGCTTTTTCGGCTTCGCGCTCCAGCTCATTTTCGTT
TGCACGTCTTCCCTTGCATGCAAGCGTGCAGTGTACGTTTCTCGTCCACGGAGAAA
ACCGTTTTTCTTATCTTTATGTTCTGCGTTGGAGTTTTCTGCATCTTCTGAACCTCTTA
GAGCTCAATCATTTGGCTTGAAGATGATCAAGAGG

>Dr17927 Our suggested modification.

ATGACAGAATGGACGCTGCTGAAGCGGCTGCTGGACGCCGTGCACCACTCCACCATG
ATCGGACGCCCTCTGGCTCACAATAATGGTGATTTTCAGATTGCTGATCGTCGCAAGTGGCG
ACCGAAGACGTCTACACTGATGAGCAGGAGATGTTCTGTGCAATACTCTCAACCGGGA
TGTCCGAACGTCTGCTATGATGCACTTCGCGCCAATATCGCAACCAAGTCTTTTGGGTCTT
CAGATCATCACGGTCTCCACGCCGTGCTTTGTTTTATTATCTACACCTGGCACAACCTG
TCCAAACAACCCGAAGGTGAGCAGATAAAGGAAGCGCTTGAGAGGAGCTGCGACTCGGAG
AGTTGCTCCATTAAATCGCATAAACACATAAATCCAAGCCTTGAAGGAGTCACCAACCG
AAACCATCGCAAGCATGAAACCTCTTCGGGAGTCTTTTGAAGTATACATCTTCCAC
GTTTGTCTTTGCTACCATTTGGAAGTGGCTTTGTAGTGGCCAGTGGCTGCTTTTTCGGC
TTCCGCGTCCAGCTCATTTCTGTTGACGTCTTCCCTTGCATGCAAGCGTGCAGTGC
TACGTTTCTCGTCCACGGAGAAACCGTTTTTCTTATCTTTATGTTCTGCGTTGGAATT
TTCTGCATCTTCTTGAACCTTTAGAGCTCAATCATTTGGCTTGAAGATGATCAAGAGG
TCTGTGCTGGTCAAGGACGGATCCTGGAATGGATACGGCGCTATAAACCAAGACTCGCAG
TCGATCGCTTCTTTGACGTTTCGAGATGTTACCACTACATCACTACCGACTCTTGAT
CTAGTTGTGGATCACCACCTGACTGGACATGCGCTGCAAACTGCTCGACAAAGAGGAC
AATAGAGGAACACAGAGTAAACCAAGCAACAGAAAAGCAAAACAGAGGAGCACTGAG
GTTTGGATATAA

>DrCx28.9 Dr41797 Ensembl prediction. No modification.

ATGGGAGAATGGGGATTTCTCTCAAGCTGCTGGACAAAGTGCAGTCTCACTCCACAGTG
GTTGGGAAGGTGTGGCTACGGTCCTGTTTGTCTTCAGGATCATGTTTCTCGGGGATGGT
GCTGAAAAGGTGTGGAGCGACGAACAATCAAAAATGATCTGCAACACGAAACAGCCTGGT
TGACAGAACGTATGCTACGATCACACCTTTCCCATCTCCCATATTTCGTTCTGGGTTCTT
CAAATCATCTTCGTGTCCACGCCAACACTTCTATACTTCGGCCACGTCCTGCATGTCTC
CACAAAGAAAAGAACTGCGACACGAGATCGAATCCCATGCTGAAAAACAAGGCCTCAAA
CAGCCGAAATATATAGCAGATTACGGCAAGTCAATCAAGGGCCAATTATTGGGTAGT
TACCTATCCAGCCTGTTTGTGAAGATCTTGTAGAGGCCGCGTTATCGTTGGCCAGTAT
TATATTTACGGTTTCATAATGATCCCGAAGATCGAATGCTCCAGTCTCCTTGCCCTCAT
ACAGTTGAGTGTACATGCTCCGTTCCACAGAGAAGACCATCTTCATCATCTTCATGCTG
GTGGTGGCGTGCATCTCTGCTTCTGAACGTGGTTGAGATGTTCTACCTGATATGCCGC
AGGTCAAAGAGACACCGCGCCGCAAGATGACTTCATTTCATAAGGTTTAAACGGATCC
AAGGTGTACATTTCAAGAACTTCAAAGTCTAGCAAACTTAA

>Dr41792 Ensembl prediction. No modification.

ATGGGCGACTGGGGATTCTCTCAAACTTTTGGACAAAGTGCAGTCTCACTCGACCAAGC
ATTGGAAGGTTTGGCTGACAGTTCTGCTGATCTTCAGAATAATGGTTCTAGGTGCCGGA
CTGGATAAAGTCTGGGGAGACGAACAGTCCAGAATGGTCTGCAACATCAACACTCTGGT
TGCTGAAACGCCGTGTACGACCACATCTTCCCATATCTCACATGCGATTCTGGGTGCTC

CAAATCATCTTCGTGGCCACTCCGAATCTGGTCTACCTCTTTTATGTTCTGCATGTCATC
 CATAGAGAAAACAACTGAGGCAGCGTTTAGAAAATCAGGCAGAGAAGCACGGTGTCAAG
 CTACCGAAATACACAGACGGCAATGGGAAGGTTTATTATAAAGGGAACCTTCTCGGTTGT
 TATATGTTTAGCCTCATTGTGACTATTTTGTGGAGGCTGGCTTTCTTGTAGGCCAGTAT
 TTTTAAATTGGCCTTTTGTATGCCCATGCAGCTTGACTGTAATGTAGAGCCATGTCCTAGT
 GTTGGTCTGCATTGTTTTACGTCCCGTCCAACCTGAAAAGAGCATCTTCATTGTGTTTCATG
 CTCATTGTGGCTTGGCTGTCTTTAGCTCTGAATATTGGAGAGATTTTTATCTGATTGGT
 CGCAGGAATGTGTATAAGCAAGGACTCGTTCGAATGCTGTGGATGAGATGCACAAATTG
 AGCCCCACTGAAACGTTTTGCTGA

>DrCx32.2 Not predicted in present (July 1 2006, Zv6) genebuild of Ensembl. In previous genebuild predicted as Dr21696. According to AY340236.

ATGGGAGACTGGGGGTTTCTCTCAGCCTTACTGGACAAAGTACAGTCTCACTCCACTGTC
 ATCGGGAAGATATGGATGAGCGTCTATTTCATCTTCCGCATCTTGGTGTGGGAGCAGGA
 GCCGAGAAATGTTTGGGGCGACGAAAGATCCAACCTTAGTGTGCAACACCAACACCCCTGGC
 TGCATAACCTGTGCTACGACTGGCAGTTCCCATTTCCGCACATCCGCTTCTGGGTCATG
 CAAATCATCTTTCATTTCCACTCCAACCTTAGTGTATCTGGGGCACGTGGTGCACATCATC
 CACCAGGAGAACAAACAGAGAGAACTTCTCAAAAGCAATCCCATGGCAAAGTCGCCGAAA
 TACACTGACGAAACGGAAAGGTCGAAATTAAGGAAGTATGTTGGGTAGCTACTTGCAG
 CAACTGTTTCAATTAAGATCATTTTAGAGGTGGCCTTCATCGTCGGACAGTATTATCTGTTT
 GGATTCATCATTGACCACAAGTTCATCTGTGAAAGGTCACCCTGTATGAGGGCTGAGTGT
 TTCGTGTCCAGACCCACGGAGAAAGCATCTTCATTATCTTCATGCTGGTGGTGGCTTGC
 GTGTCTCTGGCCTTAAATGTTCTGGAGATATTTATTTGCTTTGTAGGAGGATCAGTCGG
 AGAAGTAAGAAGTGTAGACAAGCAATGTATAATGGTGAATCTCGTTATCCGGGACATTTT
 ACAACAGAACTCGAGTCTATGAATGGGATGAGGCATAATGAGTTTAAATGTGGCCTTTCAG
 ACAAAGTGGAGTCAAAGAAAAGGCAGTCTGGACGCAGCCAAACCTGAGGCTTAA

>DrCx32.3 Dr41787 Ensembl prediction. No modification.

ATGGGAGACTGGGGATTTCTCTCATCGTTATTAGACAAAGTACAGTCTCACTCCACCGTT
 GTTGGCAAAATATGGATGAGCGTGCTTTTCATCTTTCGGATCCTGGTGTGGGAGCAGCG
 GCGGAGAGCGTTTGGGGTGACGAACAATCAAGTTTGGTTTGCACACCTTGCACCTGGT
 TGTGAAAACGTGTGCTACGACTGGCAGTTCCCATCTCACACATCCGCTTCTGGGTCCTG
 CAGATCATATTTGTCTCCACTCCGACTTTGGTGTACCTCGGCCATGCGGTGCAGGTCATT
 CACAATGAGAACAACTAGGGGAGAAGAAAAAATCCTTGGTGATGGCCACATGTTGAAG
 GAACCAATACACCTACCGACGCCAAGGCCACGTCAAGATTAAAGGAAACCTCTCGGCAGC
 TATCTAACGCAGTTGTTTTTAAATCATCCTTGAAATCGCGTTTCATTGTTGGACAGTAT
 TATTTATACGGCTTTATTTATGGTCGCCAAGTTTACATGCTCCCGTTCCCTTGGCCTTAC
 ACTGTTGAATGTTTATGTCCCGTCCCACCGAAAGACCATCTTCATTATATTTATGCTA
 GCGGTGGCTCGCTTACTGTTACTGAATGTATAGAGGTGTTTTACCTGCTGTTTACC
 AGAGTGGGATGTGCGAAGAGACGATCACATACTGTTACTACGGCTAAAACCCGGCCAGT
 TTGTCTTCTCCTCGGCAGATGAACTCTGAAGACGCTCTGAAGCAAAACAACTCAACAAG
 CAGTTTGAGAGCGGACAGAGCCTTGGAGGAAGCCTGGATGGGGCGAAAGACATGCAA
 TTGATGGAAGATCACTAG

>Dr53083 Ensembl prediction. No modification.

ATGGGCGAGTGGGATTTCTCGGACGACTGCTGGATAGAGTCCAGACACACTCCACCGTA
 GTGGGAAAAATCTGGCTACCGTCTGTTTGTCTTCAAGATTCTAGTCCTTGGAGCGGCG
 GCTGAGAGAGTGTGGGCGATGAGCAGTCCGACTTCATCTGCAACACAGAGCAACCGGGA
 TCGGAAAACGTGTGCTACGACCGCGCGTTCCCAATCTCGCACGTCGCTACTGGGTGCTT
 CAGATCATCTCTGTGTCCACGCCAACTCTGGCCTACCTGGGCCACGTCTCCACGTGCATA
 CACGCCGAAAAGAAAGTCGAGAGATGATGAAGAAAGAGCTTCAGAATGAGCAATCAAC
 CTCTTCTCAAGAAAGGCTACAAGATTCCCAAGTACAGCCGGGAAAACGGGAAGGTCAAC
 ATCCGTGGACGCTCTCTGAGAAGCTACATTCTGAGTTTGTCTTGTAAAGATACTACTGGAG
 GTGGGTTTCATTTTGGGCCAATACTATCTTTATGGCTTCACTCTTAGGGCTCAATATGTC
 TGCAGCTATTTCCCGTCTCTCAAGGTGGACTGTTTTTTGTGAGGGCCACTGAGAAA
 ACCATCTGTTTCTGTTTATGCTGGTGGTGGCTTGCAATTTCTTCTCTTAAATGTGATT
 GAGATCTTATATCTTTGCGCTAAGAAGATCAGCGAGTGTCTCAGCCGCAAAAAGGACTAC
 ACCATCACTCCGGTGACCCCTGTGGTGAGCAAAAAGAACTTTAAAAATACAGATCAGGTG
 ATACAGAATTGGATGAACCGCAGTTGGAGCTTCAGAGGAGGGAACCTGGCAATGAGGCG
 ACCAAGAGCTTGGCTTCAGAGGGTCGACGTGACATGCAAGAGGTTTATATCTGA

>Dr09593 Ensembl prediction. Predicted as 2 exons (black/blue letters). 6 nt added at 5'-end to reach initiation codon (marked green). The predicted 57 nt intron is included as a part of the reading frame (marked blue). Sequence extended 309 nt in 3'-direction to reach stop codon; 3 first nt as a part of the second conserved domain (marked green), the remaining corresponding to the cytoplasmic C-terminal tail (marked blue).

AGAGCTGACTGGGGGTTTCTGGAGCACCTGCTAGAGGAAGGCCAGGAGTACTCGACGGGC
 GTAGGACGCGTGTGGCTGACCGTCTCTTCTCTTCCGCATGCTTGTGCTTGGACGGCT
 GTGAATCTGCATGGGACGACGAGCAGTCTGACTTCGTCTGCAACACCAACAGCCCGGC
 TGTGAGTCCGTCTGCTATGACAAGGCCTTCCCATATCCCACTTCCGTTACTTTGTCTT
 CAGGTCATCTTCGTCTCAACTCCGACCATTTTCTACTTCGGCTACGTGGCCTTGAGGGCT
 AGAAATGAGAAGAGGCCAGAGGAAAAGCTGGAGGAAGATGGCAGAAGGCATAGACATCGA
 AAGACCAACGCCTGTATATTAGAAAGAACCTCCAAGCTTAAAGGGAAGTCTGTGCGCT
 TACGCGGCGAGCATAATAGTGAAAGTTCTCATGAAAGTTGGCTTCATCCTGGGTTTATGG
 ATACTTTATGGGTTTGTGATCGAGGCCAAGTACGTGTGTGAGAGGCTTCCCTGTCTCAC
 ACGGTTGACTGTTTGTCTGCTCAGCACCACAGAAAAAACATCTTACCATATACACAAA
 GTCATCGCCGTAGTGTGATTCTCTCAACGTTGTGGAGCTTTTCCATCTGCTT

>Dr09593 Our suggested modification.

ATGTCAGAGCTGACTGGGGTTTCTGGAGCACCTGCTAGAGGAAGGCCAGGAGTACTCG
ACGGGCGTAGGACGCGTGTGGCTGACCGTCTCTTCTCTTCCGCTATGCTTGTGCTTGGC
ACGGCTGTGGAATCTGCATGGGACGACGAGCTGACTTCTGCTGCAACACCAACAG
CCCGGCTGTGAGTCCGCTGCTATGACAAGGCCTTCCCATATCCCACTTCCGTTACTTT
GTCTTTCAGGTCTCTTCTGCTCAACTCCGACCATTTTCTACTTCGGCTACGTGGCCTTG
AGGGCTAGAAATGAGAAGAGGCCAGAGGAAAAGCTGGAGGAAGATGGCAGAAGGCATAGA
CATCGAAAGACCAACGCCTGTATATTAGAAATAATAAAGAGGAAGACGAAGATGGGAGT
GAAACTGAAAAGAACCGCAAAGCGCTGGAACCTCCAAAGCTTAAAGGGAACTGCTGTGC
GCTTACGCGGCGAGCATAATAGTGAAAGTTCTCATCGAAGTTGGCTTCATCCTGGGTTTA
TGGATACTTTATGGGTTTGTGATCGAGGCCAAGTACGTGTGTGAGAGGCTTCCCTGTCTT
CACACGGTTGACTGTTTCTGCTCACGACCAACAGAAAAAACATCTTACCATATACACA
CAAGTCATCGCCGTAGTGTGATTCTCTCAACGTTGTGGAGCTTTTCCATCTGCTTCAG
TTGGTGATAACACGTCGACTAGAGAAGAAATATCAGGCAGAGGTGCAGATTTCATAATAGA
GTTAGAACAGCACCACTCAACGTTCAACAACCATCATTTGAGGAAAGGAACCATCTCTTT
CTTCTGTAGCACATGGTGGATACCCTACCGAAGGGTTGGATTGGGAAAAAGAGACCCT
TCTTGGCAGAAACATGCTTCCAAGCTACTCAAACGCTTATGGAATATGAAGCCTGCA
ATAAACAAAAACAATCTTCTTAAAGACCTCAAGGCTGATGACAAACCGAGACATTAT
GTTTGA

>DrCx43 Dr41799 Ensembl prediction. No modification.

ATGGGTGACTGGAGTGCGTTGGGAAGGCTTCTTGACAAGGTGCAGGCCACTCCACGGCC
GGAGGGAAGGTCTGGCTCTCTGTGCTTCTCATCTTCCGGATCCTTGTCTGGGAACAGCA
GTGGAATCGGCCCTGGGGTACGAGCAGTCAGCTTCAAGTGCAATACCCAGCAGCCTGGT
TGGGAGATGTCTGCTATGACAATCGTTCCCATCTCGCACGTGCGCTTCTGGGTGCTT
CAGATCATCTTCTGTCCACGCCGACGCTCCTGTACCTGGCGCATGTCTTCTACCTGATG
CGAAAGGAGAGAGAACTCAACCGTAAAGAGAGGAGCTGAAGGCCGTGCAGAACGACGCG
GGCGACGTTGAGTCCATCTCAAGAAAATCGAGCTCAAGAAAGTTAAGCATGGCCTAGAG
GAGCACGGCAAGGTGAAGATGAAGGGTAGCCTGCTGCGCACCACATCTTCAGCATCATT
TTCAAGTCCATCTGTGAGGTGGTCTTCTGGTCATCCAATGGTACCTCTACGGCTTCAGC
CTCTCTGCGGTGTACACATGCGAACGACGCGCTTGGCCCTCATAGGGTGGACTGTTTCTT
TCTCGGCCACCGAGAAGACCATCTTCATCATCTTCATGCTAGTGGTTTCGCTCTTCTCG
CTTTTGCTCAACATCATCGAGCTCTTCTACGTGCTCTTCAACGAATCAAGGACCGCGTC
AAAAGCCGACAAAACACAGATTTCCCACTGGCACTTTGAGCCCCACGCCGAAGGAATG
TCTACGACCAATACGCTACTACAAATGGTTGCTCCTACCAACTGCACGCTCTCACCA
ATGTCACTCCAGGCTACAACTGGCCACCGGCGAAAGGACCAACTCTTGCCGCAATTAC
AACAAAGCAGGCTAATGAGCAGAAATTGGGCCAATACAGCACAGAACAGAAATCGTTGGGC
CAGAAATGGCAGCACCATCTCAAATTCACATGCACAAGCCTTCGACTACCTGATGATACA
CATGAGCACAAAGAACTGACGCCAGGGCATGAGTTGACGCAATTGGCGTTGATAGATGCA
CGGCCGTGACGCGTGCCAGCAGCCGATGAGCAGTCGAGCGAGGCTGATGACCTGGAC
GTCTAG

>Dr07288 Ensembl prediction. No modification.

ATGGGTGACTGGAGCGCACTGGGGAACCTTCTTGACAAGGTCCAGGCGTACTCCACTGCT
GGAGGCAAGGTCTGGCTCTCCGCTCTTCATCTTCCGGATCCTGGTGTGGGGACGGCG
GTGGAGTCCGCTGGGGAGACGAGCAGTCGGCCTTCAAATGCAACACGCTGCAACCTGGA
TGTGAGAACGTGTGCTATGATAAGTCTTCCCATCTCCACGTCGCTTCTGGGTGCTG
CAGATTATATTTGTGTCCATGCCGACCTCTTATATCTCAGCCATGTGGTGTCTTATG
AACAAAGAGGAGAACTGAATAAAAAAGAGGACAACTACGAGACATCCAAAGCAAAGGC
GGAGATGTGGAGCTGCTCTGCGCAAAATCGAAACGAGGAAAGTTCAAGTACGGATTGGAG
GATCACGGGAAGATCAAGATGAGGGGAGGGATTTTACAGCTATATAGTGAGCATCGTG
TTGAAAGTCCGTATTTGAAATTGTCTTCTTTAATACAGTGGCATCTTTACGGATTCAAG
CTGTGCGCTGTTTATACGTGCGAGAAGTTCCCTTGTCCGCATAAGGTGGACTGTTTCTG
TCCCGTCCACAGAGAAGACAGTTTTCATCATCTTCATGCTGGTCTGCTGCTGGTCTCT
CTGGCTCTCAACGTATTTGAGTTTTTTATGTGATTTTAAAGAAATGAAAGACCAAT
AGGGAAGTCTGAGAAGAAATTTGACAGTGCTTGAATATCAAGCCCTGTCCGAGGAATCTG
TCCGGCTATGAGTATTACATGACTGCTCGGCCCGTCCCAATCTAGGCTACAATCTA
GACACTGTCGATAAATCAACTCCTCTGATAATTACGACAAGCAGGCTAATGAGCAGAAC
TGGACTAATTACAGCACAGAACCAGTTGGGTACAGCCAGCGCTTTCCTTACCCG
GAGAAAGTGAAGTCTAGGGAAGGATCTTCTGCTGCTAAAAGAGCTTGAACCTCGACCCAG
AGTCGAGCGAGCAGTCGAGCCAGGCCGATGATCTTGACATCTAG

>Dr54744 Ensembl prediction in 2 exons (black/blue letters). Sequence extended 6 nt in 5'-direction to reach initiation codon (marked green). Predicted intron of 111 nt regarded as part of reading frame. Sequence extended 321 nt in 3'-direction to reach stop codon, with the 6 first nt as a part of the second conserved domain (marked green), and the remaining corresponding to the cytoplasmic C-terminal tail (marked blue).
GATTGGTCAATTCTGGGCGCTTTCTAACCAGAGGTTCAAGATCATTCAACAGTCATCGGC
AAAATCTGGCTGACGGTGCTACTGATCTCCGCATACTGCTGGTACCCTGGTGGGAGAT
GCAGTGTACAGCGACGAGCAGTCCAAATTCACCTGCAACACCTTCAGCCGGGCTGCAAC
AACGCTGCTATGATACATTCGCCCCGCTCACACTTACGCTTCTGGGTCTTCCAGATC
GTTCTGGTCTCAACGCCCTCCATCTTCTACATCATCTATGTGCTGCACAAAATACCAAA
GATGAGAAGATGGAGACGGAGAGGATCCACGCAGAGGAAGATTCGGTGAGCTCGGCAAA
GATCCAACCACTTTCCAGTCAGGTTCTACTATTTATATTGTTACGTCCTGATCCGC
TCCGCTCTGGAGATCACTTTCTGTGCGTCAGTATTACTTGTGGATTGAGGTGCCT
CATTTGTTCCGCTGCCAAACGTACCTTGGCCAAACAGGACTGACTGTTTGTGTCTCGA
GCCACCAGAAAGACTATTTCTTAATTTTATGTTTCAGCATCAGCTTGGGTTGTTTCTC

CTGAACATTGTGGAGCTTCACTACCTG

>Dr54744 Our suggested modification.

ATGGGTGATTGGTCAATTCTGGGCCGCTTTCTAACCGAGGTTCAAGATCATTCAACAGTC
ATCGGCAAAATCTGGTGACGGTGCTACTGATCTTCCGCATACTGCTGGTCACCTGGTG
GGAGATGCAGTGTACAGCGACGAGCAGTCCAAATTACCTGCAACACCTTTCAGCCGGGC
TGCAACAACGCTGCTATGATACATTGCCCCCGTCTCACACTTACGCTTCTGGGTCTTC
CAGATCGTTCTGGGTCTCAACGCCCTCCATCTTCTACATCATCTATGTGCTGCACAAAATC
ACCAAAGATGAGAAGATGGAGACGGAGAGGATCCACGCAGAGGCCAGTCACCCTAGTCGA
ATACAAGGAGATGGTTCCAGACTGGCCTACGGAGCCCAGGGAGAAGAATGGGGTGGTCAG
GACGAGGGAAGCGTTGAGCAAAGTCTCCTGCAGGAAGATTTCCGTGAGCTCGGCAAAGAT
CCAACCACTTTCCAGTCAGGTTCTACTCATTATATTGTTACGTCCTGATCCGCTCC
GTCCTGGAGATCACCTTTCTTGTGCGTCACTATTACTTGTGTTGGATTGAGGTGCCTCAT
TTGTTCCGCTGCCAAACGTACCTTGGCCAAACGAGCTGACTGTTTTGTGCTCGAGCC
ACCGAAAAGACTATTTTCTTAATTTTATGTTACGATCAGCTTGGGTGTTTTCTCTCTG
AACATTGTTGGAGTCTCAGACCTGGTTGGGTCTACATTTTCCGTATGCTCTGCGCCGCC
TGCTTCTTGTGCTGCAAGTCAGAGAGGGATTGTATGCTCAACGAAACCCACTGTTGCTT
CGCCTCAGACACTCGATCAGAGCAGGCTGGTCTGCACTCTTCAACAACCACTCTGTCT
CAGGAGAAGACCGGGACGGCTTGTCTTACATGGTCTGTCTCTTTTGTGAGACCGAC
TCGACTCTTGAAGCTCTCAAGAGGAACCTGAGGAGAGGGAACGCATGAGGTCAAA
TTGGCCAACATGGTTAGATTTACTGGTAAAAAGTCTTGGCTGTGA

>DrCx45.6 Dr40065 Ensembl prediction. No modification.

ATGGGGGACTGGAGTCTCTTGGGTAATTTCTCGAAGAAGTGCAGGAGCACTCTACGTCG
GTTGGGAAGGTGTGGTTAACCGTGCTGTTTATCTTCCGTATCCTGGTGTAGGCACAGCG
GCAGAGTCATCATGGGGCGACGAGCAGTCCGACTTTATGTGTGATACTCTACAACCTGGT
TGCAACATGTCTGTTATGACCGAGCTTTCCCATTTGCTCATATCCGCTACTGGGTGCTG
CAGATCGTCTTGGTATCAGACCTCCCTCATCTACATGGGCCATGCCATGCACACCGTC
CGCATGGAGGAGAAGAGAAAACAAAAGGAGCAGGAGGAGAAGGCAGAGCGGGAAAAGGA
GAGAAGGAGTATCTGGAACATAAAGAGAAATTCGAAAAACAAAGACAAAATCCACCTG
AAGGGGGCACTGCTGCAGACAATGTTCTGAGCATTGTGATCCGCCTGGTCTGGAAGTG
ACCTTCATTGTGATTACATGATGTACGGGATCTTCTGGATGCTCTGTATCCATGT
TCAATGCTTCCCTGCCCAACCTGTGAAGTGTACATGTCCGCTCAACTGAGAAAAAT
GTCTTTATTGTGTTTATGCTGGTGGTTGCGAGCTGTTTCGCTCCTCCTCAGCGTCATAGAG
TTATATCACCTTGGATGGAAACAGTGCAAAAAATGCCTTAGGAAACATGCTGACAAGCAT
GCCAATGACAAAATTCAAAACGTCAAAGCTGTTTCTGCAATTGAACCGATCAGGACAAGC
ATTCCAATGGATCTGGCTGAGAATCCAGCCTCGTCTTTCTCAAACCTGCACTCCACCT
CCAGATTTCAACCACTGCTTAAGATCAAACCAAGGTCCAACATCTCCTCCACATCTTCAT
TCTCACCATCTTCATACATCCACCAACCTGCCAGCCCTTCAACCAACATCTGGCACAC
CAGCAGAACTGCTGATCAGGCGGAGCGGATCACCACAGCCATGATGGCTGGAG
CCAGCCGTGGACTTCTGCGATGCAATTATGGGAGTCTGAGGCTCGGGTTGGAAGTGAA
ATGACACCCAGTACCCCTTCCACACCATCTCCATCCAGGTTCTTCAGAGACAAGCGC
CGGCTTAGCAAGACAGTGGTACTAGTAGTAACCGACTCAGACCAAGTATCTGGCCGTG
TAG

>DrCx41.8 Dr36688 Ensembl prediction. No modification.

ATGGCCGACTGGAGTCTGCTGGGGAACCTTTCTAGAAGAAGTCCAGGAGCATTCAACCTCG
GTGGGAAAGGTGTGGCTCACTATTCTTTTCTATCTTCCGATCTTGGTACTAGGCACGGCT
GCTGAATCCTCGTGGGGCGACGAGCAGGAAGACTTACCTGCGACACAGAGCAGCCCGGC
TGCGAGAACGTTTGTACGACCGAGCCTTCTATAGCGCATATACGCTTCTGGGTGCTC
CAGATCGTGTCTGCTACACCTTCTCTGATCTACATGGGGACGCAATGCATATCGTC
CGCCGAGAGAGAAGAAGAGAAAGAGTGGATGATGAAGGAGCGCAGAGATGGAGAA
AAGTACCCAGAAAGTACAAAGAACAAAGGAGGACGAAGGTGGAGGTAGGAGGTACGATTG
AAGGGTGCCTTGTGCAACAATACGTCCTCAGCATCTCATCCGCACTGTGATGGAAGTG
ATCTTCTATCATATCCAGTACCTGATCTACGGAGTCTTCTTAGTGCACTCTATGTGTGT
AAAGCCCTCCGTGTCCACATCCGGTCAACTGCTACATCTCAGACCAACAGAGAAGAAC
GTGTTCAATTGTCTTCTACATAGCAGTAGCAGCGGTGTCACTGTGCTTAGTATCGTGGAA
CTGTATCATTTGGCATGGAAGCAGTTGAGGAAGTATGTGCACGGATACAAGGCTTCCAAA
CAACGACCAAAACAGCCGTCCACCATGCTGCACTTTACCAAATCCGTCCACCCCAAAC
CGAGCCTGCACCCACCTCCAGATTTCAACCAATGCTTGACCTCGCCGCATCTTCTCCT
ACTTTACAGACACACTCGCTTTTACATCCGACCTGCCCTCCATTTACGACCGACTGGCG
CACCAGCAGAACTCTGCAACATGGTCACTGAAAGGCACAGAGGACAAGACTACTTAGGG
GTCAACTTCTTGAGCTTCTCAGACACCTACAGAGACTCCCAACTCTGTGCTCACCT
TCATTCTAAGCAGTGATTTTGGGACAAGCAAGGTTTAGTAAGAGCAGCGGACACAGC
AGCCGCATGAGACCGGACGACCTTGCCTATAG

>Dr45438 Ensembl prediction. Cx40.1-like sequence predicted as 2 exons (still lacking exon 1).
The predicted 66 nucleotide exon is regarded as a part of the reading frame (marked blue). 456
nt added at 3'-end to reach stop codon (marked blue).

TTAGGGAAAGCTTGGCTCGTCTGGTGGTATTTCTAAGGATCCTGGTGTGCTGTTTGTGCT
GGTTATCCTCTCTACCAGGATGAGCAGGAACGATTTGTGTGTAAACACATTACGCCCCGT
TGTGCCAATGTGTGCTATGACATGTTTGCCCTCTGTCCCTCTTCCGCTTCTGGCTTGTGA
CAACTGACCACTGTGCTCCCTCCCTACATAATGTTTCATCATCTATGTGATCCAAAAATC
CATCAAGAATCATTCAGGAAAGCATCTCTTTGTAAGATGGTCAAGGCTGAGAAGGGA
AGGGTGCAGTACTTACGGGAGCCTACATCTTGATCTTCTGCTTCGGATAATGGTAGAA
GCTGGATTTGAGCTGCCATTATTACTTGTTCGGCTTTCACATCCCCAGACGCTTTATG
TGTACGAGGACACCTGTGCTCCCAACAAATGGTGGACTGCTACATCTCTAGACCCACTGAGAAA
ACCGTCATGCTGAAGTTCATGTTAGGGGACGCGCTTGTCCCTGCTATTAAACATCTGC

GACCTGATTTGTGCAATCAAGCGCTCTGTGAGGCCAAAAACAAACGAAAGATGCTAGTG
 CAAAAGATGTATGCAGAGGAGCAGTACTACGTGTACAGGAATGGAAATCAAGGTGTGGAC
 GCTAGCAGCCCTCCAAATCAAGATGTG

>Dr45438ic Our suggested modification.

TTAGGGAAAGCTTGGCTCGTCTCGTGGTATTTCTAAGGATCCTGGTGTGCTGTTTGTCT
 GGTATCTCTCTACAGGATGAGCAGGAACGATTTGTGTGAACACCATTACGCCCGGT
 TGTGCCAATGTGTGCTATGACATGTTTGCCCTCTGTCCCTCTCCGCTTCTGGCTTGT
 CAAGTACCACTCTGTGCTCCCTACATAATGTTTCATCATCTATGTGATCCACAAAGTG
 AGTTCTGGCTTAGCTACCGATACCGGAACCTCCGAGTCCATAAAAGCGGACTCCATCTAC
 AAGATCCATCAAGAACATTCAGGAAGCATCTCTTTGTAAGATGGTCATGAAGGCTGAG
 AAGGGAAGGGTGCAGTACTTCACGGGAGCCATCATCTTGCATCTTCTGCTTCGGATAATG
 GTAGAAGCTGGATTTGGAGCTGCCATTATTACTTGTTCGGCTTTCACATCCCAGACGC
 TTTATGTGTGTCAGCAGGCAACCCTGCACAAACATGTTGGACTGCTACATCTCTAGACCCACT
 GAGAAAACCGTCATGCTGAATTCATGTTAGGGGCGAGCCGCTTGTCCCTGCTATTAAAC
 ATCTGCCAGCTGATTTGTGCAATCAAGCGCTCTGTGAGGCCAAAAACAAACGAAAGATG
 CTAGTGCAAAAGATGTATGCAGAGGAGCAGTACTACGTGTACAGGAATGGAAATCAAGGT
 GTGGACGCTAGCAGCCCTCCAAATCAAGATGTGATGAGTCCAGGAGTGTTCGCAAAAGA
 GGGACAGAACTCAAGTGGCGATGAAGCTGCTTCTGTGCTTTTGGATGATGACCCTCCG
 CCATCCTTACCTCAAGAAAGGAAAGCCATCAATTTACAGGAATGCCTGGTGCAGAAAGCAAT
 GATGACAGCAGCAGTTACCAACCCACCAAGAAGGGGGGATGGTAAGAGAGGGCAGTGAA
 GTGGCCCTGTGCCCGCAGTGAGCCCTTGGGAACCCCTAGATCCATCCGAGTAAGCAACAGT
 AGTCGGCTGAAACCTCCACCACCTCTCGAAGGGATAAACTTGCCGTGCAAGGTGCAATT
 GATGGCTCTGGAGCGACAGCAATTGTGTACAGAGAAGAGTAGGGCAATATACTCTGGTAGAA
 ATGACCACTGGTGAAGACATAACAACCTTGCAGTGGAGATGGGAAAGAGAAAAGTCAGAG
 TGGGTTTGA

>Dr38313 Ensembl prediction. Predicted as 2 exons (black/blue letters). Sequence extended 27
 nt in 5'-direction to reach initiation codon (marked green). Predicted intron of 120 nt
 regarded as a part of the reading frame (marked blue). Sequence extended 540 nt in 3'-
 direction to reach stop codon. Modelled according to XM_679922.

CTGGAGGAGATCCAGCACCATTCCACGTCGGTGGGGAACCTCTGGCTCACCACGCTAGTT
 GTGTTTCGCATCGTGTGACCGCGGTGGGCGGCGAGTCTATATACTACGACGAGCAGAGC
 AAGTTCATCTGCAACTCTGCACAGCCGGATGTGAGAACGTCGTACGATGCCCTTCGCG
 CCGCTGTCCCACGTGCGCTTTTGGGTCTTTAGATCATCTTTCTCCCTGCCCTCTCTG
 CTGTACATGGGCTATGCGCGCAATAAGATCTCACACAGACATGAGGATGGGAGGAAGAG
 CGAGAAGATGACCCCATGATCTACGAAGTGCCTGAGATAGACACCACACGTCGGGAATTA
 GTGCCACCGCAGCTAAACCCAAAGTGCCTCACGATGGGCGTAGACGATTCAAATGAT
 GGCCTGATGCGAGTTTACGTGCTACAGCTGTTGACACGATTTGTTCTGGAGGCCATTTT
 CTTGACGAGCATCTGTTGTATGGCTTCCGCGTGGAACTGTTTCTGTGTGACGGAT
 GTTCCCTGCCCGCACCAGGCTGGACTGCTTCATCTACGCGCCACTGAAAAACCATCTTC
 CTTAGAATCATGTATGGCGTCAGCTGCCGTGTGCTGTGCTAAACCTGTGGGAAATGATT
 CATCTCGGCGTGGGAACCATCAGCGATGTCTTGCAGCAACGAAACGCGGCAGCTAGCGAT
 GATGAGTATCAGCTCGGCTGCTGGCA

>Dr38313 Modified according to XM_679922.

ATGAGCTGGAGCTTCTGACGCGTCTGCTGGAGGAGATCCAGCACCATTCCACGTCGGTG
 GGGAAACTCTGGCTCACCACGCTAGTTGTGTTTCGCATCGTGTGACCGCGGTGGGCGGC
 GAGTCTATATACTACGACGAGCAGAGCAAGTTCATCTGCAACTCTGCACAGCCGGATGT
 GAGAACGCTGTGCTACGATGCCCTTCCGCGCGCTGTCCACGTCGCTTTTGGGTCTTTCAG
 ATCATCTTTCTCCCTGCCCTCTCTGTGTACATGGGCTATGCCGCAATAAGATCTCA
 CACAGAAGAGATTACGGGCGGCTCGGGGGCCGGGCTCCTTCCACAGGAGATTTCGGCA
 GGGGGCGGATACATCAACGCGAGGCCGAGAAAGATGTACTTTGGGGCAGCGCAGCATCGG
 CCAGGAATCATGAGGATGGGAGGAGAGCGAGAAAGATGACCCCATGATCTACGAAGTGCCCT
 GAGATAGACACCACACGTCGGGAATTAGTGCCACCGCGACCTAAACCCAAAGTGCCTCAC
 GATGGGCGTAGACGATTCAAATGATGGCCTGATGCGAGTTTACGTGCTACAGCTGTTG
 ACACGATTTGTTCTGGAGGCCATTTTCTTGCAGGACAGTATCTGTTGTATGGCTTCCGC
 GTGGAACTGTGTTGTCACGGATGTTCCCTGCCCGCAGGGTGGACTGCTTCATC
 TCACGGCCCACTGAAAAACCATCTTCTTAGAATCATGTATGGCGTCAGCTGCCGTGTGC
 CTGCTGCTAAACCTGTGGGAAATGATTATCTCGGCGTGGGAACCATCAGCGATGTCTTG
 CGCAACGAAACGCGGAGCTAGCGATGATGAGTATCAGCTCGGCTGCTGGCATCCGGC
 GCGTTTTCCGTTGGGAGTCGCGGGCTTCACTTAGTGAAAGGAGAACCTGTAGGTGGAGTC
 GCGGTTGGTGTTCGAGAAGCGGATTACGTTGGTTATCTTTCTCTGGAACACCCCATCT
 GCACCACCTGGGTATAATATCGTGGTGAAACCTGAAACGATGCCCTACACAGACTTGAGC
 AATGCTAAAATGGCGTGAAGCAGAACCGGCCAACATCGCCAGGAGGAGCAGCAGCAA
 TACGGCTCCAACGAGGACAACCTCCCATCTGCAGGTGAAACACGCGCCGCCACCTATTAAC
 AAAGATGTGATACAGTTGGAGGCGGCCATTACGGCTTATACCTTGACGACCATGCTAGC
 AATAACCACGATGAACCTGGACGACACTAATGATATTGATGAGAAGCCTCAGAGCAATATC
 ACCACAGCGCCACAGAAGGAGCGAAAGCAACGGTCCAAGCATGGGAAATCCGGGAGCGCT
 GGGAGCAGCAGCAGCAGCAAAATCAGGGGAGGGAAGCCATCTGTCTGGATCTGA

>DrCx47.1 Dr06961 Ensembl prediction. No modification.

ATGAGCTGGAGCTTCTCACTCGACTCTTGGAGGAAATCCACAACCACTCCACATTTGTG
 GGGAAAGTCTGGCTGACGGTCTTGATCATCTTCCGGATCGTTTGCACCGAGTCGGGGGC
 GAGTCGATCTACTCGGATGAGCAGACAAAGTTCACCTGCAACACAAAGCAGCCCGGCTGT
 GACAACGCTGTGCTACGACGCTTTCGACCGCTCTACACGTCGCTTCTGGGTCTTCCAG
 ATAATCATGATTTCCACACCTCCGTATGTATCTGGGATAGCCATCCATAGATCGCC
 AAAACCTCAGAGGAGGAACGACACAAGAACAGATTTACCAGAAGAGGAGGACACAGT
 CGCTGGAGAAACGACACCATCTAGAGGACGCTTTAGAGGAGGAAGATGAGGACGCGGAG

CCAATGATCTACGAAGAAGATGCACGAGAGATCAAAGCAGAGACTGTCCGAGATCCCCTA
 AAACACGATGGCCGCCGAGGATCATGCAAGAAGGTTTAAATGAGGATG**TATGTTCTTCAA**
CTTTTATCCCGCCGCATCTTCGAGGTGGGATTCTCACGGGTCAAGTATCTCCTCTACGGC
TTCGCGTCAACCCCTTGTACGTCTGCAACAAGATCCCATGCCACACAGGGTGGACTGC
TTTGTTTCAAGACCCACCGAGAAGACCATCTTTTTGCTCATCATGTATGTGGTGAGCTGT
CTATGCTGCTGCTCAATGTTTGGCAGAGATGTTTCACTTGGGGATCGGTGCCTTTCGAGAC
 ACTCTTCGCAACGTCGAAACCAGAATCAGCGACCTTCTATGGCTACCCTTACTCCAGG
 AATATTTCCAGTTCTCCGCCAGGATACAACCTAGTTGTTAAATCCGACAAACCCGGTCGC
 ATTCCTAACAGCATCGTCTGCCTGATCAGAACATGGATAGAGAGATCGCAGAACAACAC
 TGCACAAGTCTGATGAGAACATCCCCTGACCTAGCAACCTTGACCAACCATTTACGA
 GTAGCTCAGGAGCAGCTTGACATGGCTTTTCAGACATACAACACAAAAACCACTCATATT
 TCAAGAGCCAGCAGCCCCGTTTCTGGTGGCACAACGACAGAGCAGAACCGCATCAACATG
 GCTCAGGAGAAGCAGGGCGCTCGGCCAAAGCAAGCACCGAGAGAGCTGGGACACTAGTA
 AAAAATGGAAAACTTCGGTGTGGATTTAA

>DrCx44.2 Dr34207 Ensembl prediction. 6 nt added at 3'-end to reach stop codon (marked blue).

ATGAGTTGGAGCTTCTTGACACGTTTGCTTGATGAAATCTCCAACCATCTACTTTTGTG
 GGAAAGATCTGGCTCACTCTCCTCATTATCTTCCGAATCGTCTGACAGTGGTGGGCGGT
 GAGACCATCTATCAGGATGAACAGAGCAAGTTTGTATGCAATACACAACAGCCTGGTTGT
 GAGAACGTGTGCTATGATGCTTTTGCCCTTTATCGCATGTTAGATTTTGGGTTTTTCAA
 ATTATTGTGATAACCACTCCATCCATTATGTATCTCGGCTTTGCCATGCACAAAATTGCT
 CGAATGGCCGATGATGAATATCGACCACGCAACGCAAAATGCTGTCTATGGTTCATCGA
 GGTATGAGCCGTGACTGATGTTGACGATGGTTGACGAGATGAGTGAAGAAGTTCCCATGATCCCA
 GAAGAGATTGAGCCCTCAGAGAAAAACAACAAATCAGCAGCTTCAACCAAGACTACCGCT
 GCTTCTGATGCTGCCGTGAAACATGATGGTGCAGCGCCGATCAAGAGAGATGGTCTCATG
 AAGGTGTACGTGTTACAGTTGATCTCTCGTGTTCCTTTGAGATAGCTTTCTCTTTGGC
 CAATATATTCTTTATGGTTTATGGTCTCTCCGTCTACATTTGCACCCGAAGCCCTTGC
 CCACACACTGTGGATTGCTTTGTCTCACGTCCCACTGAAAAAACCATCTTCTGGTCATC
 ATGTATGTTGTGACGTGACTCTGTCTGGCACTGACTGTATTGGAATCCTGCATCTGGGA
 ATTGGTGGCTTGAAGGACTCTTTCGTAATCGAGCAATCGGAGACTCCCTGTTTCATAGG
 CCATCCAGCTCCACCATCTGTACCGCCTTCCAGTGCTCCACCTGGATACCAAGCTGTC
 CTAAAAAGTACTCTCAGGCAAGCTGAAGGCTGAGTTTCTAGCAGACTCGGGACGGGAT
 TCAATGGGTGGTGACAATACTACTCGTGATCTAGACCGTCTGCGGAGGCATCTGAAAAAT
 GCACAGCAACACCTGGACCAGGCCTACCACACTGAGGAAGTAGGGGCTTCACACAACAGC
 GGGCCTGACTCTAAAAGCATCGCTGCTGAGCAAAACCGACTCAACCAAGCGCAGGAAGGC
 TTTGGCAGCACTGAGGAGAAA

>DrCx44.2 Our suggested modification.

ATGAGTTGGAGCTTCTTGACACGTTTGCTTGATGAAATCTCCAACCATCTACTTTTGTG
 GGAAAGATCTGGCTCACTCTCCTCATTATCTTCCGAATCGTCTGACAGTGGTGGGCGGT
 GAGACCATCTATCAGGATGAACAGAGCAAGTTTGTATGCAATACACAACAGCCTGGTTGT
 GAGAACGTGTGCTATGATGCTTTTGCCCTTTATCGCATGTTAGATTTTGGGTTTTTCAA
 ATTATTGTGATAACCACTCCATCCATTATGTATCTCGGCTTTGCCATGCACAAAATTGCT
 CGAATGGCCGATGATGAATATCGACCACGCAACGCAAAATGCTGTCTATGGTTCATCGA
 GGTATGAGCCGTGACTATGACATGGTTGACGAGATGAGTGAAGAAGTTCCCATGATCCCA
 GAAGAGATTGAGCCCTCAGAGAAAAACAACAAATCAGCAGCTTCAACCAAGACTACCGCT
 GCTTCTGATGCTGCCGTGAAACATGATGGTGCAGCGCCGATCAAGAGAGATGGTCTCATG
 AAGGTG**TACGTGTTACAGTTGATCTCTCGTGTTCCTTTGAGATAGCTTTCTCTTTGGC**
CAATATATTCTTTATGGTTTATGGTCTCTCCGTCTACATTTGCACCCGAAGCCCTTGC
CCACACACTGTGGATTGCTTTGTCTCACGTCCCACTGAAAAAACCATCTTCTGGTCATC
ATGTATGTTGTGACGTGACTCTGTCTGGCACTGACTGTATTGGAATCCTGCATCTGGGA
ATTGGTGGCTTGAAGGACTCTTTCGTAATCGAGCAATCGGAGACTCCCTGTTTCATAGG
 CCATCCAGCTCCACCATCTGTACCGCCTTCCAGTGCTCCACCTGGATACCAAGCTGTC
 CTAAAAAGTACTCTCAGGCAAGCTGAAGGCTGAGTTTCTAGCAGACTCGGGACGGGAT
 TCAATGGGTGGTGACAATACTACTCGTGATCTAGACCGTCTGCGGAGGCATCTGAAAAAT
 GCACAGCAACACCTGGACCAGGCCTACCACACTGAGGAAGTAGGGGCTTCACACAACAGC
 GGGCCTGACTCTAAAAGCATCGCTGCTGAGCAAAACCGACTCAACCAAGCGCAGGAAGGC
 TTTGGCAGCACTGAGGAGAAA**GGTTAG**

>DrCx43.4 Dr07099 Ensembl prediction. 3'-end (marked in gray) replaced by sequence taken from NM_131069 (marked blue) to achieve stop codon.

ATGAGCTGGAGTTTTCTTACGCGTTGTTGGATGAAATCTCCAACCACTCCACCTTCGTG
 GGCAAGATATGGCTCACGTTATTCATCATCTTCCGATTGTTTGAAGTGTGGGGGGA
 GAATCGATATACTACGATGAACAGAGCAAAATTTGTGTGTAATACCCAGCAACCTGGTTGT
 GAGAACGTTTGTACGATGCATTTGCACCACTCTCTCATGTCCGGTTCTGGGTTTTCCAG
 ATCATTTTGTATCACACCCCCACTATCATGTACTTGGGATTTGCTATGCACAAGATCGCT
 CCGTCAAATGATGTGGAGTACAGGCCAGTCAACAGGAAACGCATGCCAATGATCAACCGC
 GGAGCCAACCGGGATTATGAGGAGGCCGAAGACAACGGTGAGGAAGATCCTATGATTATG
 GAAGAGATCGTGCTGAGGAGAAAAAGGCTCCAGAGAAGTCTGCTGTTAAACATGACGGC
 CGGCGGAGAATAAAGCGAGATGGGCTCATGAAGGTGTACATCTGCAGCTTCTGTGAGG
 ATTATTTTCGAGGTGGGCTTTCTCTTTGGCCAGTATATCCTGTATGTTTCGAGGTCGCC
 CCGTCATACGTGTGCACTCGCAGTCCCTGCCCGCACACCGTAGACTGCTTTGTGTACGT
 CCGACAGAGAAAAACCATCTTCTGCTGATTATGTATGCCGTGAGCTGTCTCTGCTTGTCT
 CTTACGGTGTGGAGATACTTCATTTGGGCCTCAGCGGAATTCGTGATGCTTTTCGACGA
 CGTGACGCCCATCAAAGTGTTCAGCGCCACGTGCCCCCATATGCAGACAGGTGCCCACT
 GCCCCGCCAGGGTACCACACTGCCCTGAAAAAGACAAGCTGTCTTTGGGAATGAAACCG
 GAGTATAACTTGGACTCCGGTCGGGAGTCTTTGGTGACGAGTCGTATCGCGAGACATT
 GACCGCTGCGCAGGCACCTGAAACTGGCTCAGCAACATTTAGATTTGGCCTATCAGAAT

GGCGAGAGCAGTCTTCACGCAGCAGCAGCCAGAGTCCAACGGCACTGCTGTGAGCAG
AACAGACTTAACCTTTGCTCAGGAGAAGCAGGGGAGCAAATGTGAAAAAGGTAAGAATGGT
>DrCx43.4 Modified according to NM_131069.

ATGAGCTGGAGTTTTCTTACGCGGTTGTTGGATGAAATCTCCAACCACTCCACCTTCGTG
GGCAAGATATGGCTCACGTTATTCATCATCTTCCGCATTGTTTTGACTGTTGTGGGGGGA
GAATCGATATACTACGATGAACAGAGCAAAATTTGTGTGTAATACCCAGCAACCTGGTTGT
GAGAACGTTTTGCTACGATGCATTTGCACCACTCTCTCATGTCCGGTTCTGGGTTTTCCAG
ATCATTGTTGATCACACCCCCACTATCATGTACTTGGGATTTGCTATGCACAAGATCGCT
CGGTCAAATGATGTGGAGTACAGGCCAGTCAACAGGAAACGCATGCCAATGATCAACCGC
GGAGCCAACCGGATTATGAGGAGGCCGAAGACAACGGTGAGGAAGATCCTATGATTATG
GAAGAGATCGTGCCTGAGAAAGAAAAGGCTCCAGAGAAGTCTGCTGTTAAACATGACGGC
CGGCGGAGAATAAAGCGAGATGGGCTCATGAAGGTGTACATCCTGCAGCTTCTGTGAGG
ATTATTTTCGAGGTGGGCTTTCTCTTTGGCCAGTATATCCTGTATGGTTTCGAGGTCGCC
CCGTCATACGTGTGCACTCGCAGTCCCTGCCCGCACACCGTAGACTGCTTTGTGTACGT
CCGACAGAGAAAAACCATTTTCTGCTGATTATGTATGCCGTGAGCTGTCTCTGCTTGTCT
CTTACGGTGCTGGAGATACTTCATTTGGGCTCAGCGGAATTCGTGATGCTTTTCGACGA
CGTGACGCCCATCAAAGTGTTCAGCGCCACGTGCCCCCATATGCAGACAGGTGCCCACT
GCCCCGCCAGGGTACCACACTGCCCTGAAAAAGACAAGCTGTCTTTGGGAATGAAACCT
GAGTATAACTTGAAGTCTGAGTCCGGTGGGAGTCTTTTGGTGACGAGTCTGTCGAGACATT
GACCGCTGCGCAGGCACCTGAAACTGGCTCAGCAACATTTAGATTTGGCCTATCAGAAT
GGCGAGAGCAGTCTTCACGCAGCAGCAGCCAGAGTCCAACGGCACTGCTGTGAGCAG
AACGACTTAACCTTTGCTCAGGAGAAGCAGGGGAGCAAATGTGAAAAAGGTATCCATGCT
TGA

>DrCx48.5 Dr21889 Ensembl prediction. No modification.

ATGGGTGACTGGAGCTTTCTTGGGCGGCTCTTGGAAAAATGCGCAGGAACACTCGACAGTG
ATCGGCAAAAGTCTGGCTGACGGTACTCTTCATTTTAGGATTCTGGTGTGGGAGCGGCA
GCTGAGGAGGTCTGGGCGATGAACAGTCGGACTTCACCTGCAACACTCAGCAGCCTGGC
TGTGAGAACGTCTGCTACGATGAGGCCCTTCCCATCTCCACATCCGCTTCTGGGTGCTC
CAGATCATCTTCGTGTCCACGCCGACGCTCATCTACCTGGGCCACGCTCTGCACATCGTT
CGTATGGAGGAGAAGCGGAAAGAGCGTGAGGAGGAGTTGCGAAAGGCCAGCCGGCTCCAG
GAGGAGAAAGAACTCCTGTATAGAAATGGAGGGGGAGGGGAGCCTGGTGGACGGGGTGGG
GGCGGCAAAAAGGAAAAGCCGCAATCAGAGACGAGCATGGCAAAATCCGCATTAGAGGT
GCCTTGTGTGCGCACCTACGTGTTCAACATCATTTTCAAGACCTGTTTGAAGTGGGGTTC
ATTTTAGGTCAAGTTTCTCTATGGTTTCCAGTTGCGGCCCTGTATAAGTGTGCGCGG
TGGCCTTGCCCCAACACGGTGGACTGCTTCATTTCCCGGCTACAGAAAAGACCATCTTC
ATCATATTTATGCTTGTGGTGGCTTGCCTGTCCCTTTTGTGAATTTGTTAGAAATCTAT
CACCTCGGATGGAAGAAGGTCAAACAGGGCATGACCAATGAGTTTGCCCCGACCGTGAA
TCGCTGCGTGAAGCGGAGTGAAGCTGAGCCGAGTCCCCAGAACTGCGCTCCAACCTC
AGCTACCCGCCAGACTACACGGAAGTGGCGGTGGCGGGTGGCGGTTCTCCAGCCTGTG
TCAGCGCCCTCCACCGCAGAGTTAAGATGGACCTTTGCGCGAGGAGCTTGAGGAGTCC
TCACCTTTCTACATCAGCAACAACAACACAGGCTAGCTGCTGAGCAGAACTGGGCC
AACCTGGCCAGGAGCAGACTCGGGAGATGAACGCCACCTCCCTGCTCTCTCTC
TCATCTTGCTCCTCTGATAACGTACGGCAATCCAAAGATGCCGCTCAGCTTGCCAGCACC
CCCTCCTCTGCTGGTGGTGGTTAAGCACTGGGCCGAGGAGGGGACGTCAACACCACG
GTGGAGATGCACGAGCCGCCGTCATTTCACTGACGCTGACGACTGAGCAGGGCTAGT
AAAGCCAGCAGTGTGAGAGCGAGGCCCAATGATCTGGCGGTGTAG

>DrCx39.9 Dr04082 Ensembl prediction as pseudogene in 2 exons (with a 5 nt intron) in present
genebuild (July 1 2006, Zv6). Previously predicted as a potentially functional gene. The
predicted 5 nt intron included as a part of the reading frame (blue), and the sequence was
further extended 7 nt into the Ensembl predicted exon 2 to reach the stop codon (marked pink).
The remaining of exon 2 removed (marked gray in upper sequence).

ATGGGAGACTTTAGCTCTCTTGGGAAGCTTTTAGAAAGCGCCAGGAACATTCTACAGTG
GTGGGCAAAAGTCTGGCTACCGTCTATTCTATCTTCCGTATCTTGGTGCTCAGTGCTGCC
GCAGAGAAGGTTTGGGGTACGAGCAGTCTGGCTTACCTGTGACACCAACAACTGGT
TGTGAGAAATGTGTAGATGTAACCTTCCCATCTCTACATCCGATTCTGGGTGCTC
CAGATCATCTTCGTCTCAACCCCAACTGATCTATCTGGGCCATATCTCCACCTGGTG
CGCATGGAGCAAAAAGTAAAAAGCAATGAAACAAGCGGAGCAGACAAAACGCACTTCTG
GGCCACAAACCAAAAGGTCCCATACGTGATGAACAGGGGAAGATCTGTCTGAAGGGAGTC
TTGCTGCGCACGTATGTCTTCAACATCATCTTCAAAACACTGTTTGAGGTGGGCTTCATT
GTGGCACAGTATTTCTCTATGATTGAGCTCAAGCCTCTGTACACCTGCAGCAGGTGG
CCTTGCCCAACACTGTCAACTGCTACATCTCGAGACCGACAGAAAAAACCATCTTCATT
ATTTTCATGCTGGCCGTAGCCTGTGTCTCTGTCTGCTCAACCTGGTGGAGATGTATCAC
CTGGGCTTTACCAAGTGACAGACGGGTCTCCGTTACCGGCGTGCTCACTCAGTCTGCGAC
ACTGAGTCTAAAGTGCTAGTGAGGACGTTGTTGTCTTTTGTCAAAATTACCTTAT
TTTCTGTCTCAGCACTCTCCGCTTCTTCCACAGAGCCGCACTTCAACCTCTCA
GAACCTGATGGGACCTTTCCGGTTCATAACAGCCGCTCTGTTTACAAGCAAAACCGAGAG
AACATGGCTGTAGAGCGCAACGGCAACCTGACACTGCTGATGTTAAGATCAGCAAAACA
GTGAGCTCTGTACCCGGATCACCATCAAGCCAGCAGCGCAGGCCAGCCATTTCAGTCCG
TATAGCAGCAACAGACAGGATGGATGACAATCTGAGCAGAAGGAAACGAGAGTCTT
GACTCCATTACTGAATGTGGGATCTCCGTTCTGAACAATAAGAGAGTCCCTCTCCCG
CTCACAGGCAACATCGCAACGATGAGGTTCAATTATCAGGAGTGTCAATTAG

>DrCx39.9 Dr04082 Our suggested modification (identical to prediction in previous Ensembl
genebuild, Zv5).

ATGGGAGACTTTAGCTCTCTTGGGAAGCTTTTAGAAAGCGCCAGGAACATTCTACAGTG

GTGGGCAAAGTCTGGCTCACCCTCTATTCATCTTCCGTATCTTGGTGCTCAGTGCTGCC
 GCAGAGAAAGTTTGGGGTGACGAGCAGTCTGGCTTACCTGTGACACCAAAACACCTGGT
 TGTGAGAAATGTGTGTACGATGTAACCTTCCCATCTCTCACATCCGATTCTGGGTGCTC
 CAGATCATCTTTCGTCTCAACCCCAACACTGATCTATCTGGGCCATATCCTCCACCTGGTG
 CGCATGGAGCAAAAACATAAAAGCAATGAAACAAGCGGAGCAGACAAAACAGCACTTCTG
 GGCCACAAACCAAAAGGTCCCATACGTGATGAACAGGGGAAGATCTGTCTGAAGGGAGTC
 TTGCTGCCGACGTATGTCTTCAACATCATCTTCAAAACACTGTTTGAGGTGGGCTTCATT
 GTGGCACAGTATTTCTCTATGGATTTGAGCTCAAGCCTCTGTACACCTGCAGCAGGTGG
 CCTTGCCCAAACACTGTCAACTGCTACATCTCGAGACCGACAGAAAAACCATCTTCATT
 ATTTTCATGCTGGCCGTAGCTGTGTCTCTGCTGCTCAACCTGGTGGAGATGTATCAC
 CTGGGCTTTACCAAGTGACAGAGGGTCTCCGTTACCGGCGTGCTCACTCAGTCTGCGAC
 ACTGAGTCTAAAGTGCCTAGTGAGGACGTTGTTGTTCTTTTGTGCAAAATTACCTTAT
 TTTCTGCTCACGCACCTCCTCCGGCTTCTTTCCACAGAGCCGCACTTCAACCTCTCA
 GAACCTGATGGGACCTTCCGGTTCATAACAGCCGCTCTGTTTACAAGCAAAACCGAGAG
 AACATGGCTGTAGAGCGCAACGGCAAAACCTGACACTGCTGATGTTAAGATCAGCAAAACA
 GTGAGCTCTGTACCCGGATCACCATCAAGCCAGCAGCGCAGGCCAGCCATTGAGTCGC
 TATAGCAGCAACAAGACCAGGATGGATGACCTCAAATCTGA

>DrCx44.1 Dr15076 Ensembl prediction. No modification.

ATGGGTGACTGGAGTTTCTTGGGCAACATCCTCGAGGAAGTAAATGAGCATTGACGCGTA
 ATCGGTAGGGTGTGGCTCACGGTCTCTTCATCTTCCGAATCCTCATCTTGGGCACAGCC
 GCAGAGTTTGTGTGGGGCGACGAGCAATCGGATTATGTGTGAACACGCGAGCAGCCGGGT
 TGGGAGAACCTTGTCTACGACGAGGGCTTCCCATCTCGCACATTGCGCTCTGGGTGCTC
 CAGATCATCTTTGTTTCCACACCCTCATTAGTGTACGTGGGCCACGCCGTGCACCATGTA
 CACATGGAGGAGAAAGCTAAGGAACGGGAGGAGCTGAGCTCAACCGGCAGCAAGAGAAC
 GAGGAGAGGGTGCCTGCTGGCGCTGATCAGGGAAGCGTCCGTACGGCCAAGGAGACGAGC
 ACAAGGGGACGAGCAAGAAAGTTTCGTCTGAGGGCACTCTCTGAGGACCTATATCTGCCAC
 ATTATCTTCAAGACTCTCTTGGAGATCGGCTTTGTGGTGGGTCAATACTACTTGTACGGC
 TTCCGAATCTTGGCACTCTACAAGTGCAGCCGTTGGCCGTGCCCAAACACGGTTGACTGC
 TTCGTCTCAAGACCAACCGAGAAACCGTTTTCATCATCTTTCATGTTAGCTGTGGCCTGC
 GTCTCACTGTTCCCTCAATTTTGTGGAAATCAGCCATTTAGGCTTAAAAAGATCCACTTT
 GTGTTTCGTAAGCCGGTGCAGCCGCAAGTGCAGGGACAGGAGCAGCCGAAAGGCATTG
 CCTTCCATAGCTGCCTCATCGATCCAGAAAGCCAAAGGTTACAAGTTGTTGGAGGAGGAC
 AAATCCACGTGCGCACTTCTTCTCTGACTGAGGTTGGGGGATGGAGGCTGGACGCCTG
 CGGGCTTCATATGAGCGCAATTTGAGGAGAAATCTGACGAGGCCATGGCACCTAAGAAAGAC
 ATGTCTAAGATGTATGACGAAACGCTGCCCTCTTACGCCAGACGACCGTATTGGACCG
 AGTGCATCGTCAGGAATCTGCGCAGGGATGAAGATGAGGACGAGTTGGCTGTGGAGGCA
 GACATGGAGGCCAGCGAGACGATAGAAGATACAGACCGCTCAGCAGCCTGAGCAAGGCC
 AGCAGTCGCGCAAGGTCAGATGACTTGACGGTATAA

>Dr56505 Ensembl prediction. No modification.

ATGGGAGACTGGAGCTTCTTGGGTAATATTTTGGAGGAAGTAAATGAACATTCAACTGTG
 ATTGGTCGTGTTTGGCTACAGTACTATTCATTTTCCGAATTTTAAATCTGGGCACGGCT
 GCTGAGTTTGTCTGGGGAGACGAGCAGTCCGATTACGTGTGAACACTCAGCAGCCAGGA
 TGGGAGAAATGTGTGCTATGATGAGGCCCTTCCAATCTCTCACATTGCGCTATGGGTGTTA
 CAGATCATCTTTGTATCCACACCTTCACTTGTATACGTGGGCCATGCTGTCCATTACATC
 CACATGGAGGAGAAACGAAAGGAGCGGGAAGAAGCTGAGGTGAGCCACGCAAGCAAGT
 TGGAGGAGCGCCAGGCAATGGATCAAGGAAGCGTTGCACTGCCAAGGAGACGACGAGC
 AAGGGGAGCAAGAAGTTTGAAGTGAAGGAACCTGCTGTGCACCTACATCTGCCATATC
 ATCTTCAAAGCTCTGTTTGAATAGGCTTTGTAGTGGGACAATATTTCTTTATGGGTTT
 CGCATCCTGCTTTGTACAAGTGCAGTCTGTTGGCCATGCCCTAACACAGTGGACTGCTAC
 GTCTCCCGGCCACCGAAAAGACCATCTTCATCATTTTTCATGCTTGCAGTGGCTTGTGTT
 TCACTGTTCTCAACTTTGTGGAGATTAGCCACCTCGGCTTGAAGAAGATCCGCTTTGTG
 TTCCACCGACCACTCCAGCACAGTTAGAGTCGCTTGGACCTCCTGAGAGGAGTTTACCG
 TTTCTTCTAACCACCCCTGTCCAAAAAGCCAAAGGCTACAGGCGCCTCGAAGAGGAGAAG
 AAAGACGAGGTGGCTCATATCTATCCACTAGCTGAGGTTGGGATGGAAGAGGGACAGTTT
 TTCTTACCTCAGCTGGAGAAGGAGCAGAAAAGCAGTCAAGGAGCAATTCTGCCAACAGCG
 CCACCTGTAGAGGAGACAATTATATGCGATGAGACTCAACCTCCTTCTCAGGTGACA
 GAGACATTACCAGAGCTCCCACTGAAGAGCCGCTAGGGAAGGAGATGAGGAGATGATA
 GATGCTCCACTTGTGTCTCTGCTTCTAGATTTAGAACCACAGAGGAAACAAGATCTTCAAGTCGT
 CTCAGCAAAGGTAGCAGCAGAGCCAGGTGAGATGATCTAACTATATGA

>DrCx55.5 Dr04403 Ensembl prediction. No modification.

ATGGGAGACTGGAACTTTCTTGGTGGGATTTTGGAAAGAGTGATATCCACTTACCATG
 GTGGGAAAAATCTGGCTACAATCCTCTTCATATTTTCGATGCTGGTGCTTGGTGTTGGCA
 GCGGAGGATGTGTGGAATGATGAACAGGCTGACTTTATCTGCAACACCGAGCAGCCTGGT
 TGCCGCAATGTGTGCTATGATAAAGCTTTCCCATTTCCCTAATCCGTTACTGGGTTCTG
 CAGGTCATATTTGTGCTCTCTCCCTCACTGGTGTACATGGGCCATGCCCTCTACCGCCTC
 CGTGCTCTTGAGAAGGAGCGTCAGCGCAAAAAGATGGCACTGAGACGAGAACTTGAGGGT
 GTGGATGTGGAGATGGCGGAGGTACGGCGCAAAATAGAACGTGAGCTTCGGCAGATCGAC
 CAGGGGAAATTAACAAGGCTCCATTGAGGGGGTCTCTTCTGCGCACAACGTTGGCTCAC
 ATAGTCACCCGCTCAGCTGTAGAAGTATTCTTCATGACAGGACAATATGTTCTGTATGGG
 TTTCAACTGAATACACTGTACAAATGTGAACGGGAGCCTTGTTCCCAATGCAGTGGATTGC
 TTTGTATCTCGACCCACTGAGAAAAGCGTCTTCATGGTGTTCATGCAATGCATAGCTGGC
 ATTTTCATTGTTCTCAACATTCTGGAGATCCTGCATTTGGGGTACAGAACTTAAAAAG
 GTCATCTGAACACTGATGACAGCTGAGGGATGATCCCAATGACAGCTACTATCCCAAC
 AAAGTGAAGAAAGATTCTGTTGTGCATCAGACATGCATTGGCACCTCCACTGGCCGAAG

GCCACCATTGCTTCTGCACCCAGTGGATACAACCTTCATCTTGATCGACCACCTGATGGA
 GCTGCCTATCCTCTTTGATTAACCCATCCTCTGCTTTCTTGCTGTTTCAAGGGTGATTTA
 CCAGCTAAAAACGGTGGCTGATGAACCAAGTACTTGCAGAACAGTCCCACAGAGCACAAAC
 AGCAATTCAAACAATACCAGCAGTGAAGTCTGCACTACCCACCTTGCAACTCTGCACTCCA
 CCCAAGCAAGACGAAGGGGAAGATTCTGTCCAACTTTACCACTGCACAGAAAGGGCAG
 GAGTCTAAGTTATCAGAGTCATCGAGCCACACCAGAGAATCATCTCATGCCTCATCCAGC
 ATGGTAAAGAAACCTTGGAGGGTAGTGCTCCCTGGAATTGCTCAACAGTCGTAGAAGGT
 AATGGCTCAGACTCAGATTCCCTGGAAGGCTCCAAAGCTCGTTGCTTATTCTGCTGTG
 CGGGCACGTACCTCATCCAGGTCTGATACCAAGCTGAGCAGGCCACCTCCCCTGATTCA
 GTCGAAGAATCGAGCTCTGAGTCACGGCATAGTCCACGAGCGTCACCAAGCCATCGTGCC
 TCATTGGCCAGCAGTTCCAGCAGCAGACGAGCAGCTCCACAGACTTACAAATTTAA

>DrCx52.9 Dr39845 Ensembl prediction as pseudogene in present genebuild (July 1 2006, Zv6).
 Predicted as 2 exons (black/blue letters) with a 2 nucleotide intron (GA) in-between. Modelled
 according to NM_207093, a G is added to the 2 nt (GAG; marked in blue). Sequence extended 3 nt
 in 3'-direction to reach stop codon (marked in blue).

ATGGGGGACTGGAACCTTCTGGGGGGGATTCTGGAGGAGGTGCACATCCACTCCACTATG
 GTGGGAAAGATCTGGCTACCATCCTCTTCATCTTCCGCATGCTGGTGTGGGCGTTGCG
 GCGGAGGACGTGTGGAACGATGAGCAGTCCGACTTCATCTGCAACACTGAGCAGCCCCGGC
 TGCCGCAACGTCTGCTATGACCAGGCCCTCCAGTGTCCCTGATCCGCTACTGGGTGCTG
 CAGGTTATCTTCGTGTCTCACCTCGCTGGTCTACATGGGCCATGCCATCTACAAGCTG
 CGCGCCCTGGAGAAGCAGCGCCACTGCCAGCGTGTGACCCTGCGGCGGGAGCTGGAGACG
 GTGGAGCCGCGAGCTGACGAGACGCGGCGGCATCGAACGGGAGCTGCGGCAGCTGGAG
 CAGGGCAAGCTGAACAAGGCTCCGCTGCGGGGCTCGCTGCTGCGCACATATCTGGCCAC
 GTGCTGACCCGCTGCGTGTGGAGGTCTGCTTCATGATGTGCCAGTATCTGCTATATGGA
 CACCGGCTGCAGCCGCTCTACAAGTGCAGCCGCGAGCCGTGCCCAAACGTGGTGGACTGC
 TTTGTGCTCGGCCACGAGAAAAGCCTGTTTCATGGTGTTCATGCAAGGCATCGCGGCG
 GTGTGCTGTTCTGAGCTGCTGGAGCTCCTGCATCTGGCTATAAGAAGCTGAAAAAG
 GGGCTGCGGGACCACTGCCCTGCCCTCAGAGACGGGCCAGCGGACTCCAGTGCCCCCAAT
 AGGAACCTCTGTGGGACAGCAGGCTCGCAAGGCCACCATGCCCATGGCACTCAGTGGGTAC
 ACCGCACTGCTGGAGAAACAGGGCAATGGCCCCACGTACCCCTCAATCATACACCCGCTG
 TCTGCCTTTGTGCCGATCCAGGGCCCGTCAGCACCAGACATGGAAAACCGGGATGCCCTG
 CGCAGCCCGCTGGAACACAACAGCTCCAACAACACCAGCAGCGGGTCCCGGTCTCCATTG
 GTGCAGGAGCGTGGTCCAGCTGCGTCCCCCAGCGACTACCACAGCTCCCCCAGACAGAC
 TCCTTGCCCTGATAGCCACACTCCCCCTGACGAGCCCTGCCCCAGATGGATTTGTTT
 CCCTTGAGGGATGCAAGCTCTTGTCCAGCTGTCTGCACAAGCAGCGGCGGGTCAGTCCG
 CCCTGGAACCTGCAAGCACTGTGATGGAGAGCACCAGGCTCGGACAGCGAGTCCAGCAGTGGT
 GGGCTCAGTCAGGGAAGGAGGACTCGCTGTGGCCGGTCCGTGTCCCGCTCAGACCTGCGG
 CTCAGCCCGGACTCACAGCCTCACTGTGCTGCAGAAAGCCCTCACTCTACCCACGCGG
 CAGCGGCCAGCGATAGGCAGCAGCAGACGGGCAGACCTGCAGATC

>DrCx52.9 Modified according to NM_207093

ATGGGGGACTGGAACCTTCTGGGGGGGATTCTGGAGGAGGTGCACATCCACTCCACTATG
 GTGGGAAAGATCTGGCTACCATCCTCTTCATCTTCCGCATGCTGGTGTGGGCGTTGCG
 GCGGAGGACGTGTGGAACGATGAGCAGTCCGACTTCATCTGCAACACTGAGCAGCCCCGGC
 TGCCGCAACGTCTGCTATGACCAGGCCCTCCAGTGTCCCTGATCCGCTACTGGGTGCTG
 CAGGTTATCTTCGTGTCTCACCTCGCTGGTCTACATGGGCCATGCCATCTACAAGCTG
 CGCGCCCTGGAGAAGCAGCGCCACTGCCAGCGTGTGACCCTGCGGCGGGAGCTGGAGACG
 GTGGAGCCGAGCTGATGGAGACGCGGCGGCATCGAACGGGAGCTGCGGCAGCTGGAG
 CAGGGCAAGCTGAACAAGGCTCCGCTGCGGGGCTCGCTGCTGCGCACATATCTGGCCAC
 GTGCTGACCCGCTGCGTGTGGAGGTCTGCTTCATGATGTGCCAGTATCTGCTATATGGA
 CACCGGCTGCAGCCGCTCTACAAGTGCAGCCGCGAGCCGTGCCCAAACGTGGTGGACTGC
 TTTGTGCTCGGCCACGAGAAAAGCCTGTTTCATGGTGTTCATGCAAGGCATCGCGGCG
 GTGTGCTGTTCTGAGCTGCTGGAGCTCCTGCATCTGGCTATAAGAAGCTGAAAAAG
 GGGCTGCGGGACCACTGCCCTGCCCTCAGAGACGGGCCAGCGGACTCCAGTGCCCCCAAT
 AGGAACCTCTGTGGGACAGCAGGCTCGCAAGGCCACCATGCCCATGGCACTCAGTGGGTAC
 ACCGCACTGCTGGAGAAACAGGGCAATGGCCCCACGTACCCCTCAATCATACACCCGCTG
 TCTGCCTTTGTGCCGATCCAGGGCCCGTCAGCACCAGACATGGAAAACCGGGATGCCCTG
 CGCAGCCCGCTGGAACACAACAGCTCCAACAACACCAGCAGCGGGTCCCGGTCTCCATTG
 GTGCAGGAGCGTGGTCCAGCTGCGTCCCCCAGCGACTACCACAGCTCCCCCAGACAGAC
 TCCTTGCCCTGATAGCCACACTCCCCCTGACGAGCCCTGCCCCAGATGGATTTGTTT
 CCCTTGAGGGATGCAAGCTCTTGTCCAGCTGTCTGCACAAGCAGCGGCGGGTCAGTCCG
 CCCTGGAACCTGCAAGCACTGTGATGGAGAGCACCAGGCTCGGACAGCGGAGTCCAGCAGT
 GGTGGGCTCAGTCAGGGAAGGAGGACTCGCTGTGGCCGGTCCGTGTCCCGCTCAGACCTG
 CGGCTCACGCGGAGCTCACAGCCTCACTGTGCTGCAGAAAGCCCTCACTCTACCCACG
 CGGCAGCGGCCAGCGATAGGCAGCAGCAGACGGGCAGACCTGCAGATCTGA

>DrCx52.6 Dr34930 Ensembl prediction. No modification. Sequence corresponding to potential
 intron start is underlined.

ATGGGAGATTGGAACCTTGCTTGGGAGTATCTTAGAGGAGGTTACATTCACTCCACCATT
 GTGGGCAAAATCTGGCTTACCATTCTTTTCATTTCCGCATGCTGCTACTTGGGGTTGCT
 GCGGAAGACGATGGGATGATGAGCAAGTGAGTTTGTGTGAACACAGAAGAGCCTGGA
 TGCAAGAATGTCTGTTATGACCAAGCGTTCCCGATATCCCTCATCCGATACTGGGTCTTG
 CAGATCATCTTTGTCTCTTCGCCATCTTTGGTGTACATGGGACACGCACTCTACCGGCTC
 CGAGCTCTGGAGAAGAGCGGCCACAAGAAGAAAGTCCAGCTGAAGGTGGAGCTGGAAGAG
 AGTGAGGCTTTGGAAGAACACAAAGAAATTGAAAAAGAGCTTCGGAAGCTGGAGGAGCAG
 AAGAAAGTGAGGAAGGCCCTCTGAGGGGTTCTTGTGCGCACATATGTGTTCCATATC
 TTAACCAAGTCACTGGTGGAGGTGGGATTTCATAGTGGGCACTACATGCTCTACGGTATC

GGACTGACGCCGCTGTACAAGTGTGAGCGTGATCCTTGCCCCAACAGTGTAGACTGCTTT
 GTTTCTCGGCCGACAGAGAAGAACATCTTCATGATTTTCATGCTGGTCATATCAGGAGTG
 TCCTTGTTCTCAATCTCCTGGAGATTTTCCACCTTGGTGTGAAAAAGATCAAGCAGACC
 ATATATGGATCCATGTACAGCGACGACGACAGCATTTCAGGTCAAAGAAAACTCCATG
 GTCCAACAAGTGTGCTTCCTTACAAATTCCTCACCACAAAAACAGTTGCATTTGACACAC
 ACTTCCCTTGGCATGGCACCTGATGGACAGATGGTACCTTTGCCTATCTATATGCAGACA
 GCTGGTCATGTAGTGTCCAACATTAAACCAAATGGATCTGTGCAGCCTCTGAGACAGGAC
 CGTCTTCCAAGCCAACCCGAAATTCAGTTCTTCAACAACCTGGGAATCAAGAACGAAGG
 TCCATTCCAGACAACCGTCTGCAATCCTGCAGCAGTGAGGATTCTGGCCCCAAAGGTTCC
 GAACCACCGAAATATTCTCAGCAGCCCCGAGCTTCAATTCAGAGCCAGTCATATAGAAATA
 CCAGCAGCATTAAGGAAACACAGCCGTGTTAGTCAGTGTAAAGACTTCAGTGAGGAGAGT
 GATTCCGGTGGAAAGTGGGAACATCCTACTGCCAGGAAAGCCAGCTTCATGTCAAGGGGA
 CTTTCAGAGAGCCCGTCTGAGAGTGCAGCCTCAAGAGTGGATCAGACACAGAGGCCAAC
 CGTATCACTCAAGGGGAGAGTCCCGCTATGACACCACCCCTGCAGCAGGACGTAGAATG
 TCAATGGTAAGAAAAATTCTGA

>Dr57792 Ensembl prediction. Sequence extended by 6 nt in 3'-direction to reach stop codon (marked blue).

ATGGGGGACTGGAACCTGTTGGGGAGCATTTTAGAGGAGGTTACATTCACTCCACCATC
 GTGGGTAAATCTGGCTGACCATCCTGTTCAATTTCCGGATGTTGGTTCTTGGTGTTCG
 GCAGAGGACGTTTGGGTGGACGACGAGCGAGTTCTGCTGCAACACGACGACGCTGGA
 TGCAAGAACGCTGTCTACGACAGGCATTCCCAATATCGCTCATTGCTTCTGGGTATTG
 CAGATCATTTTTGTTTCTCGCTCGCTTGGTGTACATGGGACATGCTGTACCGCTA
 AGGTCTTTAGAGAAAGAACGGCACAGGAGGAAAATCCAGCTGCGAGCAGAGCTGGAAGAG
 ACCGAGCCCTCTTGGAGGAGCAGAGAAAGTTGGAGAAGGAACAGGAGGTTGGAAGAG
 CAGAAGAAGATGAAGAAGGCTCCTCTAAGGGGCTCCTTGCCTCGTACATATATTATCAT
 ATCCTCACCAGATCAGTGGTGGAAAGTGGCGTTTCAATGTCGGGCAGTATATCTTATATGGC
 ATTGGACTGGATCCTTTGTACAAGTGTGAGAGGGTGCCTTGCCGAACAGCGTGGACTGT
 TATGTTTCCAGGCCAACAGAGAAAACCATCTTCATGGTTTTCATGATCGTCATCGCAGGA
 GTTTCGTTGTTCTGAACCTTTTGGAAATATCCCACTTGGGGGTGAGAAAGATTAAACAG
 ACTCTGAGCGGACTGCAGTTTGTGCGAGGAGGACAGTCTTTCGCAACCCCAAGCATTGACA
 ATTCAGCAGCTCTGCGTGATGACGGAAGTATTCGCTCACAAAACCCACAATTGAAAACG
 TTTATCCCGCAGGGACAAATGGACAAACATCTGTTCAAGCTCTTCGAGCAATGATATTCTG
 CGGCACAACAGTTTAGCAGCGTCCACCTTACCGGTGTCCTGCATTACCCAGCAGCCTCGC
 CAAATGCGCCAGCCAGCCAGGGAATGATTCATGAAGTGCAGTCCAGGGGTCACTGAGG
 CTCTGGAGGACCAAGGAAAACAGCATCCAGATAGCAGTAAGTCTCAGAAAGGGACATT
 AGGCCTTTTAACTCGGGCCATCCGGGTTCTGAGGGCCATACCGAGATACCAGCCTGCCTT
 CGCAATGCTCTGCACAGGCCAGCCGCTTGGCGGACCTAGCAGATGATGCTATGGAGTCC
 TCCGAGAGCGACTTCTGTCCACCCAACAGGAAAGCCAGTTTCATGGTTCCGATGCCCTCT
 GAAAGCATGTCCGGCAGTCCGTCTGTCCCTCCACCAGGAGTTCAGAGTCTGAGCTGGGA
 TCCCTTAACGACCTGCCATGAACCCACCACAGGGGGAGGACGACGGATGTCTATGGCA
 AGTAGATGG

>Dr57792 Our suggested modification. Position corresponding to potential intron start in other Cx62 orthologs is underlined.

ATGGGGGACTGGAACCTGTTGGGGAGCATTTTAGAGGAGGTTACATTCACTCCACCATC
 GTGGGTAAATCTGGCTGACCATCCTGTTCAATTTCCGGATGTTGGTTCTTGGTGTTCG
 GCAGAGGACGTTTGGGTGGACGACGAGCGAGTTCTGCTGCAACACGACGACGCTGGA
 TGCAAGAACGCTGTCTACGACAGGCATTCCCAATATCGCTCATTGCTTCTGGGTATTG
 CAGATCATTTTTGTTTCTCGCTCGCTTGGTGTACATGGGACATGCTCTGTACAGCTA
 AGGTCTTTAGAGAAAGAACGGCACAGGAGGAAAATCCAGCTGCGAGCAGAGCTGGAAGAG
 ACCGAGCCCTCTTGGAGGACACAGAAAGTTGGAGAAGGAACAGGAGGTTGGAAGAG
 CAGAAGAAGATGAAGAAGGCTCCTCTAAGGGGCTCCTTGCCTCGTACATATATTATTCAT
ATCCTCACCAGATCAGTGGTGGAAAGTGGCGTTTCAATGTCGGGCAGTATATCTTATATGGC
ATTGGACTGGATCCTTTGTACAAGTGTGAGAGGGTGCCTTGCCGAACAGCGTGGACTGT
TATGTTTCCAGGCCAACAGAGAAAACCATCTTCATGGTTTTCATGATCGTCATCGCAGGA
GTTTCGTTGTTCTGAACCTTTTGGAAATATCCCACTTGGGGGTGAGAAAGATTAAACAG
 ACTCTGAGCGGACTGCAGTTTGTGCGAGGAGGACAGTCTTTCGCAACCCCAAGCATTGACA
 ATTCAGCAGCTCTGCGTGATGACGGAAGTATTCGCTCACAAAACCCACAATTGAAAACG
 TTTATCCCGCAGGGACAAATGGACAAACATCTGTTCAAGCTCTTCGAGCAATGATATTCTG
 CGGCACAACAGTTTAGCAGCGTCCACCTTACCGGTGTCCTGCATTACCCAGCAGCCTCGC
 CAAATGCGCCAGCCAGCCAGGGAATGATTCATGAAGTGCAGTCCAGGGGTCACTGAGG
 CTCTGGAGGACCAAGGAAAACAGCATCCAGATAGCAGTAAGTCTCAGAAAGGGACATT
 AGGCCTTTTAACTCGGGCCATCCGGGTTCTGAGGGCCATACCGAGATACCAGCCTGCCTT
 CGCAATGCTCTGCACAGGCCAGCCGCTTGGCGGACCTAGCAGATGATGCTATGGAGTCC
 TCCGAGAGCGACTTCTGTCCACCCAACAGGAAAGCCAGTTTCATGGTTCCGATGCCCTCT
 GAAAGCATGTCCGGCAGTCCGTCTGTCCCTCCACCAGGAGTTCAGAGTCTGAGCTGGGA
 TCCCTTAACGACCTGCCATGAACCCACCACAGGGGGAGGACGACGGATGTCTATGGCA
 AGTAGATGGAAATGA

Suppl. Figure 4 *Connexin* sequences from pufferfish (*Takifugu rubripes*, or more commonly *Fugu rubripes*, Fr). The sequences are found in the Ensembl database and were updated as of July 1 2006. Although predicted as connexins, none of the *Fugu* sequences have formally been called by a connexin name, and we here use the abbreviated form of the Ensembl gene identification. For the ortholog of Cx25, the abbreviated form *Fr153654* should be read as NEWSINFRUG00000153654. After a recent update in *Fugu* assembly and genebuild (December 2005, Ensembl v36 and later), a number of genes changed their Ensembl gene identification numbers. The old gene identification is given in parenthesis in the title line. Other information is given in short form in the title line where needed. If our interpretation of the most likely connexin sequence differs from the prediction found in Ensembl, the database sequence is given first, and immediately followed by our suggested sequence, which is the one used in our analyses. The conserved domains are marked in yellow (unmodified sequence) or green (modified sequence). Alternating black/blue sequences are exon predictions copied from Ensembl. Other colors are explained for the relevant sequences. The following gene predictions in Ensembl are not included here:

Fr163054: identical to Fr122421

Fr150812: nearly identical to Fr130034

Please note that both the Ensembl gene identification number and the Ensembl gene prediction might change on future updates of the *Fugu* genome database.

>Fr153654 Ensembl prediction. Extension by 147 nucleotids in 3'-direction to reach stop codon (marked in blue).

```
ATGAACTGGGGCTTTCTGGAGAACATCCTCAGCGGAGTGAACAAATACTCCACGGTGATC
GGGCGCATCTGGCTCTCCGTCGCTCTTCTTTCAGAACTCTGGTGTATGTCGCAGCAGCC
GAGCAAGTGTGGAAGGACGAGATGAAGGAGTTTGTGTGCAACACCCGTCAGCCTGGCTGC
GAGACTGCCTGCTTCAACCATTCTTCCCCATCTCGCAGGTGCGCCTCTGGGCCATGCAG
CTCATCCTGGTATCAACCCCATCCCTGCTGGTGGCCCTGCATGTGGCCTACAGGGAGCAC
CGTGAGGCCAAGCACAAGAAACAACGTACAAGGACAAAGCAACATTGACGGAGGATTG
TTTTAACTTACATCGCCAGTCTGGTTTTTAAGACTGCTTTTGAAGTCTGCTCCCTGCTC
ATCTTCTACTTTGTTTACAACGGTTTCGAGCTCCCTGTGTGTGCTCCGCTGCAACCAGAGT
CCCTGTCCAAACACAGTGGACTGTTTCAATTGGCAAAGCCACCGAAAAGAAGATTTTCTC
TACATCATGGCCTGCACCTTCTGTACTTTGCATCTTTCTCAATTCGGTGGAGCTCCTCTAC
ATTATATGGAACAATTAAGTCAAATGCGTCATCCGGCATCCGGCATCACGTT
```

>Fr153654 Our suggested modification.

```
ATGAACTGGGGCTTTCTGGAGAACATCCTCAGCGGAGTGAACAAATACTCCACGGTGATC
GGGCGCATCTGGCTCTCCGTCGCTCTTCTTTCAGAACTCTGGTGTATGTCGCAGCAGCC
GAGCAAGTGTGGAAGGACGAGATGAAGGAGTTTGTGTGCAACACCCGTCAGCCTGGCTGC
GAGACTGCCTGCTTCAACCATTCTTCCCCATCTCGCAGGTGCGCCTCTGGGCCATGCAG
CTCATCCTGGTATCAACCCCATCCCTGCTGGTGGCCCTGCATGTGGCCTACAGGGAGCAC
CGTGAGGCCAAGCACAAGAAACAACGTACAAGGACAAAGCAACATTGACGGAGGATTG
TTTTAACTTACATCGCCAGTCTGGTTTTTAAGACTGCTTTTGAAGTCTGCTCCCTGCTC
ATCTTCTACTTTGTTTACAACGGTTTCGAGCTCCCTGTGTGTGCTCCGCTGCAACCAGAGT
CCCTGTCCAAACACAGTGGACTGTTTCAATTGGCAAAGCCACCGAAAAGAAGATTTTCTC
TACATCATGGCCTGCACCTTCTGTACTTTGCATCTTTCTCAATTCGGTGGAGCTCCTCTAC
ATTATATGGAACAATTAAGTCAAATGCGTCATCCGGCATCCGGCATCACGTTCTGTGGAGAAGAA
CCGTCCTCTCGCTACCACTCACAAGGGTCCAACATCAACAGATATGTCTCTGTTGAGCCT
GGTGTCTTAACGAAGGTGGCCCATAAAGGCTAAAGCTGAAAACCTCCCAAGTTCATTCT
GGGACTCATTAA
```

>Frscf38 (previous id: Fr137409) not predicted in present genebuild

```
ATGTCTTGGGCCACGCTTTACAGTCAAGCTGGGTGGTGTCAACAAACACTCCACCAGCCTG
GGAAAGATCTGGCTTTCTGTCTCTTCTATCTTCCGCGTCACCATTTCTGGTTCTGGCCGCT
GAGAAAGTCTGGGGCAGCAACAGTCCGACTTTAAATGCAACACGACGACGAGGTTGC
AAAAATGTCTGCTACGATCATTTCTTTCCCGTTTCGCACATCCGCTGTGGTGCCTGCAG
CTGATCTTTGTGTCCACCCCGGCCCTTCTGGTGGCCATGTATGTGGCCTACAGAAAACGT
GGAGAACAGAGAACCGTTATGGCCTCCGGAGGCGATGAGAAGGTGAAGGAGACCGACCTG
CAGATACTGAGGACGAAGCGCCTGCACATCACGGGCCCTCTGTGGTGGACCTACACCTGC
AGCTTGTCTTCAAGATTGCTGTTTGAAGGTGGCTTCATGTACGCTCTGTACTTTATCTAC
GATGGCTTCCAGATGCCGCGACTGGTCAAGTGCAGCAGTGGCCTTGCCCAACAAGGTC
GACTGCTTCATCTCCAGGCCAACAGAGAAAACCGTCTTCAACATCTTCAATGGTGGTCTCG
TCGGCCATTGTATGGTTCTCAATGTTGCTGAGCTCTTCTACCTTTTTGCCAAGGCCCTC
ATGCGGTTATCAGCCAGGTCAAAGCAGCGTAAGCGGAGATATACCAGCGAATCAAACCTC
AACCAGGACACGCTTCTGGACAACAGGAGGAATGAAACCTTTGTAG
```

>Frscf64 (previous id: Fr124726; this gene id now predicts another gene that is a Cx46 ortholog) not predicted in present genebuild

```
ATGTCTTGGACCACTCTGTACGCTCAGCTGGCTGGAGTAAACCGTCACTCCACCAGCTTG
```

GGTAAAGTCTGGCTCTCTGTGCTCTTTATTTTTCGAGTTATGGTTTTGGTCGTGGCGGCC
 GAGAGTGTGGGGGAGATGAACAGTCTGACTTCACCTGTAACACCCTACAGCCTGGTTGT
 GAGAATGTCTGCTATGATCAGTTCTTCCCGTCTCCACATCCGGCTCTGGTGTCTTCAG
 CTTGTCTTTGTCTCCACTCCAGCTCTCCTGGTTGCGATGTACGTGGCCTACCGGAACCAC
 GGCGACAGAAAAAGCTCTCCAGGTGTTCACTTTCAACAAAGGTGAGGAGGAAGAGTTG
 GAGAGCCTCAGGAACAGGAGACTGCCTATATCTGGTGCCCTCTGGTGGACATACGCCCTTC
 AGCCTTCTGTTTCAGGCTCTTGTGTTGAAGGAGGATTTCATGTACGCCCTTGATGTGATTTAT
 GATGGCTTCCGGATGCCGCGCCTGGTGCAGTGCAGCAGTGGCCATGCCCAAACCTAGTG
 GACTGTTTCATCTCACGGCCAACAGAGAAAAACAGTCTTCACCATTTTCATGGCCACCTCA
 TCCTCCATCTGCATGCTCCTTAACATGGCAGAGCTTGCAATCTTGTGGCCAAGGGAGTC
 ACGAGGTAG

>Fr129456e1 (previous id: Fr129456) Ensembl prediction in 2 exons (black/blue letters). We suggest that exon 2 belongs to another connexin (which we named Fr129456e2 and is found at the top of next page). Exon 2 is therefore removed from this sequence (marked gray in upper sequence). Sequence extended by 304 nt in 3'-direction to reach stop codon (marked green for part of sequence that is considered as part of second conserved domain; marked blue for remaining sequence).

ATGAAGTGGGCATTCTCCAGGGCTCTCAGCGGGGTGAACAAGTACTCCACCGCCTTT
 GGCCGAGTGTGGCTCTCCATTGTCTTCTCTTCAAGGTCATGGTGTCTGGTGGCGGCT
 GAGAAGGTGTGGGGCGACGAGCAGAAAGACTTCAAATGCAACACGGCTCAGCCCGGCTGC
 CACAACGTCTGTACGACCACTTCTTCCCGTTTCCCACGTCCGGCTGTGGGCGCTGCAG
 CTCATCTTCTGTCACCTGCGCTCTCTCTGGTGGTGTATGCACGTGCGCTACAGGGAGGAC
 AGGGAGCGGAAAAACAGGCTTAAATATGGCGACGACTGCCGCGTCTCTACCAGAACACC
 GGAAGAAGCGCGGAGGCTGTGGTGGACCTACGTCTCTCACGTTGGTCTTCAAATCGGC
 GTCGACGCCACCTTTGTCTACCTTCTCTACCACATCTACGAGGGTTACGACTTCCCCTCG
 CTCATCAAAGCCACGACGCGGAGGAAAGCTGTACAACAACACGGGCAAGAAGCAC
 GGCGGCCCTGTGGTGGACGTATATGTGAGCCTGTTCTGTGAAGACGGGCATCGAGGTCGCC
 TTCTCTACATCTCCACACGCTACGACAGTTTCTACCTGCCGAGGCTGGTCAAGTGC
 GAGGTGTGCGCCTGCCCAACAGGTGGACTGCTACATCGGCCACCCACCGAGAAGAAG
 GTCTTACCTACTTCATGGTTGGAGCCTCGGCCCTCTGCATCATCTCAACATTTGCGAG
 ATCATTTACCTC

>Fr129456e1 (previous id: Fr129456) not predicted in present genebuild.

ATGAAGTGGGCATTCTCCAGGGCTCTCAGCGGGGTGAACAAGTACTCCACCGCCTTT
 GGCCGAGTGTGGCTCTCCATTGTCTTCTCTTCAAGGTCATGGTGTCTGGTGGCGGCT
 GAGAAGGTGTGGGGCGACGAGCAGAAAGACTTCAAATGCAACACGGCTCAGCCCGGCTGC
 CACAACGTCTGTACGACCACTTCTTCCCGTTTCCCACGTCCGGCTGTGGGCGCTGCAG
 CTCATCTTCTGTCACCTGCCGCTCTCTCTGGTGGTGTATGCACGTGCGCTACAGGGAGGAC
 AGGGAGCGGAAAAACAGGCTTAAATATGGCGACGACTGCCGCGTCTCTACCAGAACACC
 GGAAGAAGCGCGGAGGCTGTGGTGGACCTACGTCCTCACGTTGGTCTTCAAATCGGC
 GTCGACGCCACCTTTGTCTACCTTCTCTACCACATCTACGAGGGTTACGACTTCCCCTCG
 CTCATCAAAGTCCAGCAGAAGCCTGCCGAACACGGTGGACTGCTTCATCGCGCGGCC
 ACCGAGAAGCGGATCTTCACTCTTCACTCTTCTGTTGGTGGTACACAGCCTGGTCTGCATCTTCTC
 TCCATCATCGAAATCCTCTACTTGGTGGGCAACGCTGCCGTGAGTGTGTTGACGGCCGGT
 CACCACACTCACCACCCCATGACCAACAACATCTCGAGCGGAAGCCACCTGATGGAGTCC
 AGCACTCTAAAGAGGGTTCCAAAGTGACCCCCGAAACGCCGGCACCTTCGTACAGCTCT
 GCCATATCCTGA

>Fr158780 Ensembl prediction in two exons (black/blue letters). Exon 2 extended 178 nt in 3'-direction to reach stop codon (marked blue).

ATGAAGTGGGCCTTCTCGAGGGCTCTCAGCGGGGTGAACAAGTACTCCACAGCGTTC
 GGCCGTATCTGGCTCGCCATCGTTTTTCATCTTCAAGGCTCCTGGTCTTCTGGTGGCCTGT
 GAGAAGGTCTGGGGCGACGAGCAGAAAGACTTTGACTGCAACACCCTGCAGCCCGGGTGT
 CACAACGTCTGTTACGACTACTTCTCCCGTCTTTACACCCGACTCTGGTCCCTGCAG
 CTGATCTTCTGTCACCTGCCGCTCCCTTCTGGTACGCTTACGCTGCTTACAGGAAGGAT
 CGTGAACGTGTAACATCGGCTGAAGCAGGAGAGAAAAACAGCCCCCTCTGTGACAAACA
 GGAAGAAGCGAGGAGGTCTTTGGTGGACCTACTTCTCAGCCTGCTGTTTAAGATAACG
 GTGGACGTGGTGTCTCTCTCTGTTCTACATCTACGAGGCCACCTTCTTCCACCG
 CTGGTGAATGCGAGGAAGACCCGTGTCCCAACGTGGTGGACTGCTACATTGCCAGGCCG
 ACGGAGAAGAAAAATATTACCATCTTTCATGGTGGTACACAGCTTCGTGTGCATCTGCCTC
 ACGGTTTGGCAGGTTTTCTACCTGTGCGGGAAGAGGATCTGGGAGTGCAGCAGGGGCGG

>Fr158780 Our suggested modification of exon 2.

ATGAAGTGGGCCTTCTCGAGGGCTCTCAGCGGGGTGAACAAGTACTCCACAGCGTTC
 GGCCGTATCTGGCTCGCCATCGTTTTTCATCTTCAAGGCTCCTGGTCTTCTGGTGGCCTGT
 GAGAAGGTCTGGGGCGACGAGCAGAAAGACTTTGACTGCAACACCCTGCAGCCCGGGTGT
 CACAACGTCTGTTACGACTACTTCTCCCGTCTTTACACCCGACTCTGGTCCCTGCAG
 CTGATCTTCTGTCACCTGCCGCTCCCTTCTGGTACGCTTACGCTGCTTACAGGAAGGAT
 CGTGAACGTGTAACATCGGCTGAAGCAGGAGAGAAAAACAGCCCCCTCTGTGACAAACA
 GGAAGAAGCGAGGAGGTCTTTGGTGGACCTACTTCTTCAAGCCTGCTGTTTAAGATAACG
 GTGGACGTGGTGTGTTACTCTCTGTTGTTCTACATCTACGAGGCCACCTTCTTCCACCG
 CTGGTGAATGCGAGGAAGACCCGTGTCCCAACGTGGTGGACTGCTACATTGCCAGGCCG
 ACGGAGAAGAAAAATATTACCATCTTTCATGGTGGTACACAGCTTCGTGTGCATCTGCCTC
 ACGGTTTGGCAGGTTTTCTACCTGTGCGGGAAGAGGATCTGGGAGTGCAGCAGGGGCGG
 TGCCACCTCTGACAGAGAGGACTCTTCTGGTGGAGGTTCTCTGAGCGCAGGAACGCT
 GTGAACAAAGGCTCGGTGGCGGCTGAGGCTGCGGCGCTCGACAGAGACGGAGAAGCCTTC
 AGTCCCGCCCCAGCGCTACGCCATCGCCGCTCTCGTCAAGTCTGATGTGTGACGATTGA

>Fr129456e2 (previous id: Fr157125) Ensembl prediction in 2 exons (black/blue letters). Exon 1 is considered to belong to another connexin (see previous page), and is therefore removed (marked gray). Exon 2 extended 317 nt in 5'-direction to reach initiation codon (this includes the complete first conserved domain in green, and a part of the intracellular loop in blue). Sequence extended 387 nt in 3'-direction to reach stop codon (6 nt included in the second conserved domain, in green; remaining corresponds to cytoplasmic C-terminal tail in blue).

```
ATGAACCTGGGCATTCTCCAGGGCCTCTCAGCGGGGTGAACAAGTACTCCACCGCCTTT
GGCCGAGTGTGGCTCTCCATTGTCTTCTCCTTCAGGGTCATGGTGTTCGTGGTGGCGGCT
GAGAAGGTGTGGGGCGACGAGCAGAAAGACTTCAAATGCAACACGGCTCAGCCCGGCTGC
CACAACTGTGTACGACCACTTCTTCCCGTTTCCACGTCCGGCTGTGGGCGCTGCAG
CTCATCTTCGTACCTGCCCCTCTCTCTGGTGGTGATGCACGTCCGCTACAGGGAGGAC
AGGGAGCGGAAAAACAGGCTTAAATATGGCGACGACTGCCGCCGTCTTACCAGAACACC
GGGAAGAAGCGCGGAGGCTGTGGTGGACCTACGTCTCAGTTGGTCTTCAAATCGGC
GTGCAGCGCCACCTTTGTCTACCTTCTTACCACATCTACGAGGGTTACGACTTCCCCTCG
CTCATCAAAGGCCAAGCAGCGGACGGGAAGAAGCTGTACAACAACACGGGCAAGAAGCA
GGCGGCCGTGTGGTGGACGTATATGCTGAGCCTGTTCTGTAAGACGGGCATCGAGGTCGCC
TTCTCTACATCTCCACACGTCTACGACAGTTTCTACCTGCCGAGGCTGGTCAAGTGC
GAGGTGTGCGCCTGCCCAACAGGTGGACTGCTACATCGGCCACCCACCGAGAAGAAG
GTCTTACCTACTTTCATGGTTGGAGCCTCGGCCCTCTGCATCATCTCAACATTGCGAG
ATCATTTACCTC
```

>Fr129456e2 (previous id: Fr157125) Our suggested modification.

```
ATGGACTGGAAGACCTTCCAAGCCCTCTCAGTGGGGTGAATAAATACTCCACGGCGTTT
GGGAGGGTCTGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
GAGCGCGTGTGGGGCGACGAGCAGAAAGGACTTTGACTGCAACACCAAGCAGCCGGGCTGC
GCTAACGTCTGTACGACTTCTTCTTCCCATCTCCACATCCGCTGTGGGCCCTGACAG
CTCATCTTCGTACGTCGCCCTCTTCTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CGCGAGCGCAAAATTCAGGGCCAAGCAGCGGCGACGGGAAGAAGCTGTACAACAACACGGGC
AAGAAGCAGCGGCGCTGTGGTGGACGTATATGCTGAGCCTGTTCTGTAAGACGGGCATC
GAGGTGCGCTTCTCTACATCTCCACACGTCTACGACAGTTTCTACCTGCCGAGGCTG
GTCAAAGTGCAGGTTGTGCGCCTGCCCAACAGGTGGACTGCTACATCGGCCACCCACCC
GAGAAGAAGGTCTTACCTACTTTCATGGTTGGAGCCTCGGCCCTCTGCATCATCTCAAC
ATTTGCGAGATCATTTACCTCATCGCCAAGCGCGTGTGCGGTGCGCCAAACAGGTCAAG
AGGCACCATCGCAACAGAGCCCCGTGCCCTCCGAGAACTACAGCGACGACCCCTTCAAC
AACTGCAACGTGACGATGGCGAAGCGGAGCTGAAGGACAACCCCGTCTCTCAGGACC
GCGTGAAGTCTACATCTACATTAAGCTGGACAGTCTTGGATGAACGACAAGATCCGGGCTCT
GCCCCAATCTGTCTGCATGGCCGGTTGCGGGGCCAGTGTAGGCATAAACGGGCCGACA
GTGGTTGGACTGGAACCCAGCAGGAAACCCCTGAACGACAGCAGGTGCTGGACTTCAGG
TCGTTGTGTAAGAACCTGGAAGCGGAGTCATCCCGGCTGGACTCTTAA
```

>Fr158780e3 (previous id: Fr137254) Ensembl prediction. 19 nt removed at 5'-end to reach initiation codon in accordance with other connexins (marked gray). Sequence extended by 75 nt in 3'-direction to reach stop codon (marked blue).

```
AACCCCCCGCAGAGCCAAATGAGTTGGAAGTTTCTTGAGGGTCTCTCAGCGGAGTCAA
CAAGTACTCCACTGGCTTCGGACGCATCTGGCTGTGCGTGGTCTTCGTCTTCCGCGTGCT
GGTCTTCTGTCGTGGCTGCCGAGCGGGTCTGGAGCGACGACAGGGACACTTTGAGTGCAA
CACCCGTGAGCCAGGCTGCACCAACATCTGCTATGACTACTTCTTCCCATCTCCACAT
TCGCTGTGGGCGTCCAGCTGATCTTACCTACCTGCCCTCCTTACATGGTGGTGTGCA
CGTGGCTTACAGGGAAGAACGGGAACGCAAGTACAAAGCCAAGCAGCGGAGGACGCCG
GCTGTACGACAACCCAGGCCAGAACGACGGCGGTCTGTGGTGGACGACTTGTGAGCCT
CTTCAACCAAGACCACTTCGAGATGCTGTTCTCTACCTGCTCAACTACATCTACGACAG
CTTCAAACTGCCAGGAAGTCCAGTGTGACGCGAGTCCCTGCCCAACTGGTGGACTG
CTACATATCCCGGCCACTGAGAAGACGGTTTTTACCTACTTTCATGGTGGGTGCTCGGT
GCTGTGCGTGGTGTCAACATCTGTGAGATCTATCTGATTGCTGCTGTGGTGGTGA
TCGGAAGTATCGGGGAAGCAACCGT
```

>Fr158780e3 (previous id: Fr137254) Our suggested modification.

```
ATGGATTGGAAGTTTCTTGAGGGTCTCTCAGCGGAGTCAACAAGTACTCCACTGGCTTC
GGACGCATCTGGCTGTGCGTGGTCTTCTGCTTCCGCGTGTGGTCTTCTGCTGTGGCTGCC
GAGCGGGTCTGGAGCGACGACAGGGACACTTTGAGTGAACACCCGTGAGCCAGGCTGC
ACCAACATCTGCTATGACTACTTCTTCCCATCTCCACATTCGCTGTGGGCGCTCCAG
CTGATCTTTCATCACCTGCCCCCTCTTCTATGGTGGTGGTGGTGGTGGTGGTGGTGGT
CGGGAACGCAAGTACAAAGCCAAGCAGCGGAGGACGCGCGGCTGTACGACAACCCAGGC
CAGAAGCAGCGCGGTCTGTGGTGGACGTACTTGTCTGAGCCTCTTCAACCAAGACCACTTC
GAGATGCTGTTCTCTACCTGCTCAACTACATCTACGACAGCTTCAAAGTCCCGAGGAAA
GTCCAGTGTGACGCGAGTCCCTGCCCAACCTGGTGGACTGCTACATATCCCGGCCACT
GAGAAGACGGTTTTTACCTACTTTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
ATCTGTGAGATCTCTATCTGATTGCTGCTGTGTGGTGAATCGGAAGTATCGGGGAAGC
AACCGTGGCTCTCTAGGAAGGTCACGCGCGGCCAGCGCGGGCTGCGATGGTTGCAAG
TCCTCTCTTGTGCATGATTAG
```

>Fr143712 Ensembl prediction in 3 exons (black/blue/black letters). Last exon extended by 99 nt in 3'-direction to reach stop codon (marked in blue).

```
ATGAACCTGGTGGGGCTGGAGAGCCTTCTCAGTGGAGTCAACAATACTCCACGGCCTTT
GGGAGAATCTGGCTGTCCATGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CAGAGGGTCTGGGGTGTGAGAGCAAGGATTTCTGTGTGCAACACTCGACAGCCGGGTTGT
ACCAACATCTGCTACGACCACATCTTCCCATCTCCACATCCGCTCTTGGGCTCTGCAG
CTGATCTTCTGTCAGCGAGTCCCTGCCCAACCTGGTGGACTGCTACATATCCCGGCCACT
GAGAAGACGGTTTTTACCTACTTTCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
ATCTGTGAGATCTCTATCTGATTGCTGCTGTGTGGTGAATCGGAAGTATCGGGGAAGC
AACCGTGGCTCTCTAGGAAGGTCACGCGCGGCCAGCGCGGGCTGCGATGGTTGCAAG
TCCTCTCTTGTGCATGATTAG
```

AAGAGAGGGGGGCTGTGGTGGACCTATCTGCTGAGTCTGATCCTCAAAGCTGGATTGAC
 GCCTCGTTTCTTTATATCTGTACAAGATATATGATGGTTATGACTTGCCAGGTTGTCG
 AAATGTTTCGCTGGATCCGTGTCCCAACACGGTCGACTGCTTCATCAGTCGCCCCACAGAG
 AAAAAGATCTTCATGTTGTTTCATGGTCGTGTCAGTGCGCTTTGCATTTTCATGTGCCTC
 TGGCAAAATGCTCTATCTTGTGGGAAGCGCATCGCCAACTGGTAAAGATCCGCCACCAG
 AACGAACAGATCCTATTTGCTGAGCAGCAGCAACTACCGACATGGTCCCACCCAGA
 >Fr143712 Our suggested modification.

ATGAACTGGTTCGGGGCTGGAGAGCCTTCTCAGTGGAGTCAACAAATACTCCACGGCCTTT
 GGGAGAATCTGGCTGTCCATGGTGTGTTGTTCCGTGTGCTGGTGTGTTGGTGGCAGCA
 CAGAGGGTCTGGGGTGATGAGAGCAAGGATTTTCGTGTGCAACACTCGACAGCCGGGTTGT
 ACCAACATCTGCTACGACCACATCTTCCCATCTCCACATCCGTCTCTGGGCTCTGCAG
 CTGATCTTCGTACCTGCCCGTCTTGATAGTGTGATGGCTCACGTCAAATTCGTGAAGGG
 AAGGATGCCAAATACGTGGAGCAGCACCACGGCTCTCACCTATACAGCAACCCCGGCAAG
 AAGAGAGGGGGGCTGTGGTGGACCTATCTGCTGAGTCTGATCCTCAAAGCTGGATTGAC
 GCCTCGTTTCTTTATATCTGTACAAGATATATGATGGTTATGACTTGCCAGGTTGTCG
 AAATGTTTCGCTGGATCCGTGTCCCAACACGGTCGACTGCTTCATCAGTCGCCCCACAGAG
 AAAAAGATCTTCATGTTGTTTCATGGTCGTGTCAGTGCGCTTTGCATTTTCATGTGCCTC
 TGGCAAAATGCTCTATCTTGTGGGAAGCGCATCGCCAACTGGTAAAGATCCGCCACCAG
 AACGAACAGATCCTATTTGCTGAGCAGCAGCAACTACCGACATGGTCCCACCCAGA
 TCCAGTATCACAAAGACTGACCCAACTGACGGACAGTCAGCTCAGTTTAAACAGAAAGGAG
 AAGGTCAGAGAAGGTACTGTGACGACCACACTGTAA

>Fr125568 Ensembl prediction in two exons (black/blue letters). Last exon extended by 120 nt in 3'-direction to reach stop codon (marked in blue).

ATGAACTGGTCTGCACTGGAGGCCCTGATCAGCGGGGTCAACAAGTACTCCACCGTGTT
 GGACGCGTCTGGCTGTCCATGGTCTTCGTTTTCCGAGTGATGGTGTGTTGGTGGCGGCT
 CAGCGGGTGTGGGGCAGCAGCAGCAAGGACTTTGTCTGCAACACGGCCAGCCGGGCTGC
 AACACGCTGTGCTACGACAGCATCTTCCCATCTCACACATCCGCTGTGGGCCCTGCAG
 CTCATTTTCGTCACTTGGCCGTGCTGATGGTGGTGGGCCACGTCAAGTATCGGGAGAAG
 AAAGACTCCCACTACACCACCTCGCACCACGGGAAACACCTGTACGCCAATCCTGGAAAG
 AAGCGTGGAGGGTGTGGTGGACCTACCTGCGGAGTCTGATTTTCAAGGCCGGCTTTGAC
 GCTGGTTTCTGTACATCCTCTATCACGTCTACGACGGTTATGACATGCCCGGCTCTCT
 AAGTGCTCCCTGGAGCCGTGCCCAACACAGTGGACTGCTTCATCTCGCGGCCACTGAG
 AAGAAGATCTTACCCTCTTCATGGTGATCTCTTCTGCCGTCTGCATCCTGATGTGCCTC
 TGTGAGATGATCTACCTCATCTGCAAGCGGTTACAAACTCATTAAAGCGAAGGAACGAG
 GTGGAGAGAAGGTTGTTCCGCGAGAGTCATGAGATGGCCCT

>Fr125568 Our suggested modification.

ATGAACTGGTCTGCACTGGAGGCCCTGATCAGCGGGGTCAACAAGTACTCCACCGTGTT
 GGACGCGTCTGGCTGTCCATGGTCTTCGTTTTCCGAGTGATGGTGTGTTGGTGGCGGCT
 CAGCGGGTGTGGGGCAGCAGCAGCAAGGACTTTGTCTGCAACACGGCCAGCCGGGCTGC
 AACACGCTGTGCTACGACAGCATCTTCCCATCTCACACATCCGCTGTGGGCCCTGCAG
 CTCATTTTCGTCACTTGGCCGTGCTGATGGTGGTGGGCCACGTCAAGTATCGGGAGAAG
 AAAGACTCCCACTACACCACCTCGCACCACGGGAAACACCTGTACGCCAATCCTGGAAAG
 AAGCGTGGAGGGTGTGGTGGACCTACCTGCGGAGTCTGATTTTCAAGGCCGGCTTTGAC
 GCTGGTTTCTGTACATCCTCTATCACGTCTACGACGGTTATGACATGCCCGGCTCTCT
 AAGTGCTCCCTGGAGCCGTGCCCAACACAGTGGACTGCTTCATCTCGCGGCCACTGAG
 AAGAAGATCTTACCCTCTTCATGGTGATCTCTTCTGCCGTCTGCATCCTGATGTGCCTC
 TGTGAGATGATCTACCTCATCTGCAAGCGGTTACAAACTCATTAAAGCGAAGGAACGAG
 GTGGAGAGAAGGTTGTTCCGCGAGAGTCATGAGATGGCCCTCTGGCAGCACCAGGTCC
 GAGCTGAGGTCCAAATCATCGATCAGGGTGGATCCAAACCGCTCTGTCCAGGACCTCACC
 GAAGAGAAGCGGCCACCTGAGAAACAAAGATCGCAGCATAG

>Fr133739 Ensembl prediction. Sequence extended by 147 nt in 3'-direction to reach stop codon (3 nt considered as part of second conserved domain, marked green; remaining marked blue).

ATGGGGGAATGGGGCTTCTCGGTGGACTCTTCGACAGCCTCCAGGCTCACTCGCCCATG
 CTCGGCCGCTTCTGGCTCTGCTCATGCTCATCTTTCCGATAGTGATCCTCGGAACCTGTG
 GCCAGCGACCTGTTTGAAGACGAACAGGAGGAATTTGCCTGCAACACCCCTCCAGCCGGGC
 TGCAAAACAGGTGTGTTACGACATGGCCTTTCCCATCTCGCAGTACAGATTCTGGGTGTTT
 CACATCGTCTCATCGCCACGCCATCGCTGCTTTTTCTGGTTTACACCATGCATCACCAC
 AATAAGAAGAACTCCAAATTCATCAGAGGTACAATGAAGACATCCGTTTAAAGGAGGCTT
 TACATCGTCAACGTGGTGTTCGATCCTGGCAGAAAGTTGGGTTTCTCGTGGGTCAAGTGG
 CTGCTCTATGGCTTCAAGGTGGAGGCCAGTTCCCTGCAAGCCGCTTCCCTGCCCTAC
 ACCGTGGACTGCTTACCTCCCGCCCGGCGGAGAAAACCGTCTTCTCTGCTTCTACTTT
 GTCGTGGGGCGATTGCAAGCCCTTTTACGTGTGCTGAGCTCTTCCACAGCTCC

>Fr133739 Our suggested modification.

ATGGGGGAATGGGGCTTCTCGGTGGACTCTTCGACAGCCTCCAGGCTCACTCGCCCATG
 CTCGGCCGCTTCTGGCTCTGCTCATGCTCATCTTTCCGATAGTGATCCTCGGAACCTGTG
 GCCAGCGACCTGTTTGAAGACGAACAGGAGGAATTTGCCTGCAACACCCCTCCAGCCGGGC
 TGCAAAACAGGTGTGTTACGACATGGCCTTTCCCATCTCGCAGTACAGATTCTGGGTGTTT
 CACATCGTCTCATCGCCACGCCATCGCTGCTTTTTCTGGTTTACACCATGCATCACCAC
 AATAAGAAGAACTCCAAATTCATCAGAGGTACAATGAAGACATCCGTTTAAAGGAGGCTT
 TACATCGTCAACGTGGTGTTCGATCCTGGCAGAAAGTTGGGTTTCTCGTGGGTCAAGTGG
 CTGCTCTATGGCTTCAAGGTGGAGGCCAGTTCCCTGCAAGCCGCTTCCCTGCCCTAC
 ACCGTGGACTGCTTACCTCCCGCCCGGCGGAGAAAACCGTCTTCTCTGCTTCTACTTT
 GTCGTGGGGCGATTGCAAGCCCTTTTACGTGTGCTGAGCTCTTCCACAGCTCCATAAAG
 TGGTCTGCTGACGACGGAGGTAGGAAGGCAGAAACAGAACTTGCCTGTATCAGCGAG
 AACCTCTTCAACTTCAAGCAGGAAGAAGAGACAGCAAAAGGAGAAGCGGCAGATGAAGAAT

CGCACGCACCCAACAGCGTGA

>Fr144150 Ensembl prediction. No modification.

ATGAACTGGGGAACCTTTTACGCCCTGATCAGCGGCGTAAACAGGCACTCGACCGGCATC
GGGAGGGTTTGGCTCTCCGTCATCTTCGCTCTCCGAATCCTGGTGTGGTGGTGGCTGCT
GAGAGCGTTTGGGGAGACGAGAAGTCGGGCTTACCTGCAACACCCAGCAGCCTGGCTGC
AACAGCGTCTGCTACGACCAAGTTCTTCCCATCTCGCACATCCGCCTTTGGGCTCTGCAG
CTGATCCTGGTCTCCACCCCGGCCCTGCTGGTGGCCATGCACGTAGCCACAGACGGCAC
ATCGACAAGAAGATCCTGAAGAGGGCCGGCCGGGGCACCCCAAGACCTGGAGCAGATC
AAGAACCAGAGGTTCCAGATCACCAGGCTCTGTGGTGGACGTACATGATCAGCATCATC
TTCAGGATCGTCTTTGAGGTGGCTTTTCTCTACATCTTCTACCTGATCTATCCAGGTTTC
AAAATGGTGCCTTTGGTCAAGTGCAGCTCGTACCCCTGCCCAACACCGTGGACTGTTTT
GTGTCCAGACCCACAGAGAAGACCATATTTACCGTGTTCATGCTGGGGGTCTCGGGGGTG
TGTGTGCTTCTGAACCTTGGCTGAGATGGTCTACCTCATCGGC CGGGCTGCAGGCAGTGC
ATCAGAGGCTCGGAGAGCAACCTCAAAGTCCCTGGATCAGTCAAAAATTTGCTCTTAC
AGGCAAAATGAGATTACGAACCTGATATTGGACCATCCCTCAGGTCAAAGTTCGGCGTG
ACCAAAAAGAAGCCAGCTGA

>Fr157320 Ensembl prediction. Sequence extended by 9 nt in 3'-direction to reach stop codon (marked blue).

ATGAACTGGGCATCATTTTACGCCGTATCAGCGGTGTGAACAGACACTCCACGGGCATC
GGCCGCATCTGGCTTTCTGTGCTCTTTATTTTCCGCATCCTGGTCTGGTGGTGTGCTGCG
GAGAGCGTGTGGGGAGACGAGAAGTCGGGCTTACCTGCAACACCCAGCAGCCGGGCTGC
AACAGCGTCTGCTACGATCACTTCTTCCGATCTCCACATCCGCCTCTGGGCACTCCAG
CTCATCTGGTCTCCACCCCTGCCCTGCTGGTGGCTATGCATGTGGCTCATCGCCGCCAC
ATCGACAAGAGGCTCTACAACTGTCAAGGGCGGACCAACCCCAAGATCTGGAGCAGATT
AAGACCCAGAAAATGAAAATCAGAGGCGCGCTGTGGTGGACGTACGTACATCAGCCTGCTC
TTTCGCGTTATCTTCGAGGTGACCTTTATGTACCTATTTTACATGATCTACCCCGTTAC
AAGATGATCCGGCTGGTGAAGTGTGACTCGTACCCCTGTCCCAACACAGTGGACTGCTTT
GTCTCCAGGCCACAGAGAAGACGGTTTTACCGTCTTCATGCTGGCTGTGTCAAGGGTCT
TGATTCTGCTCAACATTGACAGAGGTGGTGTCTTGGTGGGGAAGGCCCTCGCGTAAACAT
TTACACCATGCTGGAGACTCAGCCATGGGGGCTTGGATCCAACAAAAGCTCTGC

>Fr157320 Our suggested modification.

ATGAACTGGGCATCATTTTACGCCGTATCAGCGGTGTGAACAGACACTCCACGGGCATC
GGCCGCATCTGGCTTTCTGTGCTCTTTATTTTCCGCATCCTGGTCTGGTGGTGTGCTGCG
GAGAGCGTGTGGGGAGACGAGAAGTCGGGCTTACCTGCAACACCCAGCAGCCGGGCTGC
AACAGCGTCTGCTACGATCACTTCTTCCGATCTCCACATCCGCCTCTGGGCACTCCAG
CTCATCTGGTCTCCACCCCTGCCCTGCTGGTGGCTATGCATGTGGCTCATCGCCGCCAC
ATCGACAAGAGGCTCTACAACTGTCAAGGGCGGACCAACCCCAAGATCTGGAGCAGATT
AAGACCCAGAAAATGAAAATCAGAGGCGCGCTGTGGTGGACGTACGTACATCAGCCTGCTC
TTTCGCGTTATCTTCGAGGTGACCTTTATGTACCTATTTTACATGATCTACCCCGTTAC
AAGATGATCCGGCTGGTGAAGTGTGACTCGTACCCCTGTCCCAACACAGTGGACTGCTTT
GTCTCCAGGCCACAGAGAAGACGGTTTTACCGTCTTCATGCTGGCTGTGTCAAGGGTCT
TGATTCTGCTCAACATTGACAGAGGTGGTGTCTTGGTGGGGAAGGCCCTCGCGTAAACAT
TTACACCATGCTGGAGACTCAGCCATGGGGGCTTGGATCCAACAAAAGCTCTGCTTCCTC
TAG

>Fr137912 Ensembl prediction. No modification.

ATGGGGGAATGGACTATACTAGAGAGGCTCCTGGAGGCTGCTGTCCAGCAGCACTCTACT
ATGATAGGAAGGATCTTACTAACAGTGGTGGTATCTTCCGGATTCTAATCGTGGCGATA
GTTGGAGAGACTGTCTATGATGATGAGCAGACCATGTTTGTGTTGTAACACCTTACAGCCG
GGCTGCAACCAAGCGTGCTACGACAAAGCATTTCCCATTTACACATTAGATATTGGGTT
TTTCAGATTATCATGGTGTGCACGCCGAGCCTTTGTTTCATCAGTACTCGGTGCACCAG
TCGGCCAGCAGAGGAGCGGCCCTACTCCACAGTCTATCTGACACTAGATAAGGATCAA
GATTCACTCAAACGAGATGAGAGCAAAAAGATAAAGAACACCATCGTCAACGGAGTCTTT
CAGAACACGGGAGAACTCCACAAAGAGCCGAACCGGACTGTTTGAAGTGAAAGAGATC
CCCAATTCCGGCCATGAGAACTGCAAAAGTCCAAAATGAGGCGCCAGGAAGGCATCTCCAGG
TTTACATCATCCAGGTGGTTTTTCAGAAACGCGTTGGAAATCGGCTTTTGGTGGGCCAG
TACTTTCTGTACGGATTCAACGTCCTCGGTGACGAGTGCACCGCTACCCCTGCATA
AAAGACGTGAGTGCTACGTCTCAAGACCCACGGAGAGACAGTGTTCCTGGTGTTCATG
TTCGCCGTGACGGCTTTTTCGCTGGTGTGAACCTGGCGGAGCTCAATCACCTGGGCTGG
AGGAAAATCAAGACGGCGTGCAGGGGCTGCAGGCTCGGCGGAAGTCCATCTATGAGATC
AGAAACAAGGACTTCCGAGGATGAGCGTGCCCAATTTCCGGCGCACTCAGTCCAGTGAC
TCTGCGTACGTGTAG

>Fr164573 (previous id: Fr150812) Ensembl prediction. No modification.

ATGGGGGAATGGACCATCTTGGAGCGTCTGCTGGAGGCGGCTGTCCAGCAGCACTCCACT
ATGATTGGAAGGATCCTGCTGACAGTGGTGGTATCTTCCGCATCCTAATAGTCGCGATA
GTGGGTGAGAAGGTGTACGAGGACGAGCAGATCATGTTTCATCTGCAATACCATGCAGCCC
GGCTGCAACACAGGCTGCTACGACAAGGCCCTCCCATCTCACACATCCGCTACTGGGTCT
TTTCAGATCATCTTGGTGTGCACGCCGAGCCTGTGCTTCATCAGTATTCCGTGCACCAG
TCTGCCAAGACGCGACCGAAGCTACTCCCTCCTGCATCCGTACATGATCACCATGGC
CATGGTCAACACGGTCCGCATCAGACCATCACGCTCGCAAAATCCACTCGCGTTACATC
AACGGTATCCTGGTGCATCCTGAGGGCAGTAAAGAAGACCACGACTGCCTGGAGGTCAAG
GAAATCCCAATTGGACCCCGGGACTGCCTCCAACACACAAGAGCGCAAGGTTCCGCGG
CAGGAAGGTATTTCCCGTTTCTACGTATCAGGTGGTGTTCGTAAATGCGCTGGAGATA
GGCTTCTTGGCAGGCCAGTACTTCTGTATGGCTTCAACGTTCCAGGGATGTTTGAAGTGC

GATCGCTACCCCTGCGTGAAGGAAGTCGAGTGTACGTATCTCGTCCCACAGAGAAGACT
GTGTTTTTGGTCTTTATGTTCCGCGGTACGCGGCATATGTGTGCTGCTCAACCTGGCTGAG
CTCAACCACATCGGCTGGAGGAAGATAAAGACGGCCATCCGAGGGGTGACGGCTCGGAGG
AAGTCCATCTGCGAAGTCGCAAGAAGGACGTGTCTCACCTGTCTCAGGCCCAACCTG
GGCAGGACCCAGTCCAGCGAGTCCGCTACGTCTGA

>Fr130034 Ensembl prediction. No modification.

ATGGGAGAATGGACCATCTAGAGCGCTCCTGGAGGCTGCAGTGCAGCAGCATTCCACT
ATGATTGGGAGGATCCTGCTGACAGTGGTGGTATCTTCCGGATCCTGATCGTGGCCATC
GTCGGGGAACCGGTGTACGAGGACGAGCAGACCATGTTTATCTGTAAACCCCTGCAGCCG
GGCTGCAACCAAGGCTGTACGACAAAGCGTTCCCATCTCCACATCCGCTACTGGGTG
TTCCAGATCATCTGGTGTGACCCCCAGTCTGCTTATCACTTACTCAGTCCACCAG
TCAGCCAAGCAGAAAGACGCTCGGTACTCCTTCTCTATCCCATCATGGAGAGGGACTAC
GGGGGAGGGACGGCAGCGGAAGCTCCGCAACATCAATGGGATTCTGGTTCAACACGGC
GGCGATGGTGGAGGAGGGAAGGAAGCAAGACTGTCTGGAGGTGAAGGAGATCCCAAC
GCCCCGCGGGGCTCACTCATGGCAAGAGCTCAAGGTCCGGCGCCAAGAAGGGATCTCC
CGCTTCTACGTCAATCAAGTGGTTTTCCGGAACGCTCTGGAGATCGGATTCTTGGCCGGC
CAATACTTCTCTACGGCTTCAGCGTGCCTGGGATTTTGAATGTGACCGCTACCCGTGT
CTGAAGGAGGTGGAGTGTGCTGCTGCCGGCCACCGAGAAAACCGTGTCTTGGTGTTC
ATGTTCCGCGGTGAGCGGCATCTGCGTGGTGTCAACCTGGCGGAGCTCAACCATCTGGGG
TGGCGCAAGATAAAGGCCGCCATCAGGGGGGTCCAGGCCCGCAGGAAGTCCATCTGCGAA
ATCCGGAAGAAGGACATGGCTCATCTCTCCAGCCCCCAACCTGGGACGCACGCAGTCT
AGTGAGTCGGCTACGTGTGA

>Fr123845 Ensembl prediction in 2 exons (black/blue letters). The 69 nucleotide intron is regarded as a part of the reading frame (marked blue). Sequence extended by 312 nt in 3'-direction to reach stop codon.

ATGACTGAATGGACGCTGCTCAAACGCCTCCTGGACGCGCTCCACCAGCACTCCACCATG
ATTGGCCGTCTGTGGCTGACCGTTATGGTCATCTTCAGGCTGCTGGTTGTCGCCGTGGCG
ACCGAGGACGTGTACGCCGACGAGCAGGAGATGTTTGTGTGAACACCTTGCAGCCGGGA
TGCTCCACCGTCTGCTACGACGCTTTCGCTCCCATCTCGCAGCCACGCTTCTGGGTGTT
CACATCATCAGCGTCTCCACGCCATCGCTCTGCTTATCGTCTACACGTGGCACAACCTG
TCCAAGTCCCCCAGCGCTCTGGGGCGGCGCGCGCGCTTAAGGAGGGAGGGCAGCAC
CTGGGCGCACAGCTGGCGGACATCTTAGAGGGCGGAGGCTGGTGACCTCCCGTCACGTG
CCGGCAGGAAGTTCTGAGGGCTTGGCGGTCTCTGGAGGAGTCTTCCAAATGTTACATC
TTTCACGTGTGTTACGAGCGCTCTGGAGGTGGGCTTTGCTCTGGCCAGTGGAAAGCTG
TTGGGTTTGCAGGTTCCGGTCTCTTCTGTGTGAGCTCCTCGCCCTGCAACACGCCGTG
GACTGTACGTCTCCAGGCCACAGAGAAGACCATATTCCTGATTTTCATGTTTTGTGTT
GGTCTTTTCTGCATCTTCTCAATCTGCTGGAACCTCAATCACCTGGGCTGGAAGAAGATC
CGGCAGGCGGTGCAGTGAAGGAGGAG

>Fr123845 Our suggested modification.

ATGACTGAATGGACGCTGCTCAAACGCCTCCTGGACGCGCTCCACCAGCACTCCACCATG
ATTGGCCGTCTGTGGCTGACCGTTATGGTCATCTTCAGGCTGCTGGTTGTCGCCGTGGCG
ACCGAGGACGTGTACGCCGACGAGCAGGAGATGTTTGTGTGAACACCTTGCAGCCGGGA
TGCTCCACCGTCTGCTACGACGCTTTCGCTCCCATCTCGCAGCCACGCTTCTGGGTGTT
CACATCATCAGCGTCTCCACGCCATCGCTCTGCTTATCGTCTACACGTGGCACAACCTG
TCCAAGTCCCCCAGCGCTCTGGGGCGGCGCGCGCTTAAGGAGGGAGGGCAGGTG
GGCGGCGGACAGGGGGTCTCTCCAGCTGCAGCTCGGACAGCTGCTCCGTCTCTCC
CATCAGCACCTGGGCCACAGCCTGGCGGACATCTTAGAGGGCGGAGGCTGGTGACCTCC
CGTCACGTCGCCGACAGGAAGTTCTGAGGGCTTGGCGGTCTCTGGAGGAGTCTGTCCAAA
TGTTACATCTTTACGTGTGTTTACGAGCCGCTCTGGAGGTGGGCTTTGCTCTGGCCAG
TGGAAAGTGTTCGGTTTGCAGGTTCCGGTCTCTTCTGTGTGAGCTCCTCGCCCTGCAAC
CAGCCCGTGGACTGCTACGTCTCCAGGCCACAGAGAAGACCATATTCCTGATTTTCATG
TTTTGTGTTGGTCTTTCTGCATCTTCTCAATCTGCTGGAACCTCAATCACCTGGGCTGG
AAGAAGATCCGGCAGGCGGTGCAGCTGAAGGAGGAGCCGTCTGGCCAGGCTGCACAGGT
GGACGACAGGGCTATGAAGCCCTTCTCCAGTCACTCTTACCCAAAGTCTCTTAAAGG
CTGAAGGACATCAGCTCCACCCCTCTGCCACCCCTGGATGTGGTGATGGCTACCGCCCC
GAGTGGAGCTGCGTGTGAACCTGTGGCAAGAAGAGAGAGTTCCAGAAGGTCCAGGAGATC
CGCTTAGAGGTGTCCAAGAGAGCCGACCCACAGAGGACAGAGGCCACCTGTGAAGAAC
GCGGAGACCAAGAGGCTCCAAGCAGAGCTGCACCGAAGTCTGGATTAA

>Fr158218 Ensembl prediction. Sequence extended by 9 nt in 5'-direction to reach initiation codon (marked green). Sequence extended by 178 nt in 3'-end to reach stop codon (marked blue).

TGGGGATTCTTATCGTCTTGTGGACAAAGTCCAGTCCACTCCACGGTCATCGGAAAG
ATCTGGATGAGCGTCTTCTTCTGTTTCAAGATCATGGTTTTGGGCGCCGGCGCCGAGAGC
GTTTGGGCGACGAGCAGTCTGGGTTTCTATCTGCAACACTCAGCAACCCGGTTGCGAGAAC
GTCTGTACGACTGGACCTTCCCAATTTTCGACATTCGCTTCTGGGTCCTCCAGATCATC
TTTGTGTCACGCCGACGCTGGTGTACCTGGGCGACGCCATGCACATCATCCAAGGAG
AACAAGCTGAGGGAGAAGCTGCTGAGCCCCGGCGGGCCCGCTTGTCTAAGGTGCCAAG
TACACTGACGAAAAAGGGAAGGTGAAGATCAAAGGAAACCTGCTGGGGAGCTACCTGACC
CAGCTCGTGTCAAGATCTCATCGAGCGGCCCTCATCGTGGGCCAGTACTACCTGTAC
GGCTTTCATAGGTGCCAATGTTCCCTTGTTCAGGGAGCCCTGCCCTTACCCTGGAG
TGCTACATGTCCCGTCCACCGAGAAGACCATCTTCATCATCTTCATGTTGGTGGTGGGC
TGGCTCTCCCTGCTCTCAACGTGGTCGAGGTGCTTACCTCCTGTGCACCAGGCTCAAA
TGTGCTCCAG

>Fr158218 Our suggested modification.

ATGGGAGACTGGGGATTCTTATCGTCTTGTGGACAAAGTCCAGTCCCACTCCACGGTC

ATCGGAAAGATCTGGATGAGCGTCTCTTCTGTTTCAGGATCATGGTTTTGGGCGCCGGC
 GCCGAGAGCGTTTGGGGCGACGAGCAGTCGGGTTTCATCTGCAACACTCAGCAACCCGGT
 TGGGAGAACGTCTGCTACGACTGGACCTTCCCAATTTGCGACATTTCGTTCTGGGTCTCTC
 CAGATCATCTTTGTGTCCACGCCGACGCTGGTGTACCTGGGCCACGCCATGCACATCATC
 CACAAGGAGAACAAGCTGAGGGAGAAGCTGCTGAGCCCCGGCGGGCCCCGCTTGCTAAG
 GTGCCCCAAGTACACTGACGAAAAAGGGAAGGTGAAGATCAAAGGAAACCTGCTGGGGAGC
 TACCTGACCCAGCTCGTGTTCGAAGATCCTCATCGAGGCGGCCCTTCATCGTGGGCCAGTAC
 TACCTGTACGGCTTCATCATGGTGCCAAATGTTCCCTTGTTCAGGGAGCCCTGCCCTTC
 ACCGTGGAGTGCTACATGTCCCGTCCCAACGAGAAGACCATCTTCATCATCTTCATGTTG
 GTGGTGGGCTGCGTCTCCCTGCTCTCAACGTGGTCGAGGTGCTCTACCTCCTGTGCCACC
 AGGCTCAAATGTGCTCCAGGTCCCGCGCACAGAAGCTCACGTGCGCGGAACATCCCGCC
 ACCCTGCCCGCTCCCAATGGCCGACGGTGGACGATGCGCTCATGCAGAACAAAGATAAAC
 CTGGAGAAAGAACGCGGTGAGAGCATCGGCGGGAACCTGGATGGCGCCAAGGAGGAGACG
 CAGCTGCTGCGCCATTAA

>Frscf483 (previous id: Fr159440) not predicted in present genebuild. Sequence extended by 111 nt in 3'-direction to reach stop codon (marked blue).

ATGGGCGAGTGGGATTTGTTGGGCCGCTGTTGGATAAAGTGCAGAGTCACTCCACAGTT
 CTGGGCAAGGTTTGGCTCAGGTGCTTTTGTCTTCCGCATCCTGGTGCTGCAGACCGCC
 GCCGACAAGGTGTGGGGTGATGAGCAGTCTGACTTTGTCTGCAACACTCAGCAGCCGGGC
 TGTGAGAACGTCTGCTACGACCTCGCGTTCCCATCTCTCACGTGCGCTTCTGGTTTCTT
 CAGATTATTGCCATAGCGACGCCGAAGCTGCTCTACCTCGGCCACGTCCTCCACGTGATC
 CACATTGAGAAGAAGGAGAAGGAAAAAGATGAAGAAACAGGCCGAGTTGGATGCTCAGGCG
 TGTCTGTTCTCAGGACCTACAAAGTTCCCAAGTACATCAAAAGCTCTGGCAAGATCAGC
 ATCCGCGGCCGCTCCTCCGAGTTACACCTTCCACCTGCTGGCCAAGATCCTCTGGAA
 GTTGTCTTCATCGCCGGCCAGTACTTCTCTTTGGCTTCACCTGGACTCTCGCTACGTC
 TGCCAGCGCCACCCCTGCCCCCAAGGTGGACTGCTTCTGTCCAGGCCTACGGAGAAG
 TCGGTATCATCTGGTTCATGCTGGTGGCGGCAGTCTGCTCCCTGGCCCTCAGCTGGTG
 GAGCTGTTCTACCTGTGCGTGAAAGCGACGAAGGAGTGATGGCGAGGAGGCAGGACTAC
 ACGGTGACGCCGTGACGCCCGGTTTCGGGAAGGAAAGCTTCAAATCTCCGATGAG
 ATGATCCAAAATTGTATCAACCTGGAG

>Frscf483 (previous id: Fr159440)

ATGGGCGAGTGGGATTTGTTGGGCCGCTGTTGGATAAAGTGCAGAGTCACTCCACAGTT
 CTGGGCAAGGTTTGGCTCACCCTGCTTTTGTCTTCCGCATCCTGGTGCTGCAGACCGCC
 GCCGACAAGGTGTGGGGTGATGAGCAGTCTGACTTTGTCTGCAACACTCAGCAGCCGGGC
 TGTGAGAACGTCTGCTACGACCTCGCGTTCCCATCTCTCACGTGCGCTTCTGGTTTCTT
 CAGATTATTGCCATAGCGACGCCGAAGCTGCTCTACCTCGGCCACGTCCTCCACGTGATC
 CACATTGAGAAGAAGGAGAAGGAAAAAGATGAAGAAACAGGCCGAGTTGGATGCTCAGGCG
 TGTCTGTTCTCAGGACCTACAAAGTTCCCAAGTACATCAAAAGCTCTGGCAAGATCAGC
 ATCCGCGGCCGCTCCTCCGAGTTACACCTTCCACCTGCTGGCCAAGATCCTCTGGAA
 GTTGTCTTCATCGCCGGCCAGTACTTCTCTTTGGCTTCACCTGGACTCTCGCTACGTC
 TGCCAGCGCCACCCCTGCCCCCAAGGTGGACTGCTTCTGTCCAGGCCTACGGAGAAG
 TCGGTATCATCTGGTTCATGCTGGTGGCGGCAGTCTGCTCCCTGGCCCTCAGCTGGTG
 GAGCTGTTCTACCTGTGCGTGAAAGCGACGAAGGAGTGATGGCGAGGAGGCAGGACTAC
 ACGGTGACGCCGTGACGCCCGGTTTCGGGAAGGAAAGCTTCAAATCTCCGATGAG
 ATGATCCAAAATTGTATCAACCTGGAGCTGGAGCAGCATAAAGAGCAGAGGGGGGGAAG
 AGGATCACCGCGCGGCCAACGAGGTCCCCAGTATCATCTACCTGACAGCAAGAGCAAG
 GGGGAGGTCCGTATCTGA

>Fr159647 (previous id: Fr125574) Ensembl prediction in 2 exons (black/blue letters). Sequence extended 9 nt in 5'-direction to reach initiation codon (marked green). 60 nucleotide intron regarded as part of reading frame (marked blue). Exon 2 extended 342 nt in 3'-direction to reach stop codon (marked blue).

GATTGGTCTTCTCGGAGAACCTGCTGGAGGAGGGCCAGGAGTACTCTACGGGCATCGGC
 CGGTCTGGCTCAGGTGCTCTTCTGTTTCGATGCTCGTGTGGGGCATCGGCAGAG
 TCCGCTGGGATCGAGGACGCAAGCCAAATTTTCATCTGCAACACGCATCAGCCGGCTGCACC
 AACGTGTGCTACGACAAAGCCTTCCCATCTCCCACTTCCGCTACTTTGTCTCCAGATC
 ATCTTTGTTTCCAGCGGACCATCTTCTACTTCGGATACGTCGCTTTCGGGTACGGAAG
 ATCAATAAAGACGTGGAGGGCAGCTCTGATGAAGAGGAGAAGAACTAGAGGAAGTGAGG
 AAAAGCAGAAAAGCTGATAAGGAACTTCTGAGGCACCTAAGCTGAAAGGCAGACTGCTG
 TGTGCGTACACCCTCAGCATCCTCTTAAAGTCTCTCTAGAAGGTGGCTTCATGACAGGC
 CTGTATTTCTGTACAACGGCTTCTACATCGCAGCAAAGTTTCGAGTGTCAAAGGAACCTT
 TGTCACACAGGCTGGACTGCTCTGCTCGCGGCCACGGAGAAGACCATCTTTGTGTTA
 TACACTCAGGTATCTCTGCGCTCTCCCTGCTCCTCAACCTGGTGGAGCTCCTCCACCTT
 CTGAGCTAGCCGTGCTCACCGGCTGGAGAAAAGCCACGGTCACCGC

>Fr159647 (previous id: Fr125574) Our suggested modification.

ATGTCAGAGGTGATTGGTCTTCTCGGAGAACCTGCTGGAGGAGGGCCAGGAGTACTCT
 ACGGGCATCGGGCGCGCTTGGCTCACGGTGCTTCTCTGTTTCGATGCTCGTGTGGGG
 GCATCGGCAGAGTCCGCTGGGATGACGAGCAAGCCAAATTTTCATCTGCAACACGCATCAG
 CCGGGCTGCACCAACGTGTGCTACGACAAAGCCTTCCCATCTCCCACTTCCGCTACTTT
 GTCTCCAGATCATCTTTGTTTCCACGCCGACCATCTTCTACTTCGGATACGTCGCTTTG
 CGGGTCAGGAAGATCAATAAAGACGTGGAGGGCAGCTCTGATGAAGGTGAGAGAGGAGGG
 ATGGCGAAGGAGACGGACAGTAACCTGCAACGAAACGCAGCTCAGAGGAGAAGAACTA
 GAGGAAGTGAGGAAAAGCAGAAAAGCTGATAAGGAACTTCTGAGGCACCTAAGCTGAAA
 GGCAGACTGCTGTGTGCGTACACCTCAGCATCCTCTTAAAGTCTCTCTAGAAGGTGGC
 TTCATGACAGGCTGTATTTCTGTTACAACGGCTTCTACATCGCAGCAAAGTTTCGAGTGT
 CAAAGGAACCTTGTCCCAACACGGTGGACTGCTTCTGCTCGCGGCCACGGAGAAGACC

ATCTTTGTGTTATACACTCAGGTCATCTCTGGCGTCTCCCTGCTCCTCAACCTGGTGGAG
 CTCCTCCACCTTCTGCAGCTAGCCGTCGCTACCGGCTGGAGAAAAGCCACGGTCACCGC
 GGCGTCTACCTGCTCCCGCTGAGCAGGCGACCGTGGAGGCCGCGCGAATCCAAATGGAG
 GTGTTACAGTCCAGTAAAGCAGGGAGCAACGGCGACCTTCCAACCCGGCATGAGGTGGGG
 CGTCACACCAATCCCTGCGAGAGTTCCGGGGAAACCAGGGATCGAGGTGAACAGGGGACCC
 GGAGAGCCTGGGGACGACCTCCTCCCTAGTTATGTGACTTGCCTTGAAGCCACGAGGGCT
 ATGCTTTTCACCCAGAGTCCATTATAAGAAGAACACAGTCCAGAGTCCGAAAAGCACCAAG
 GCAGCTCAGAAAGGACATTCAAACAGAAACATTATGTATGA

>Fr130494 Ensembl prediction in 3 exons (black/blue/black letters). The middle exon has a 5'-intron of 40 nt and a 3'-intron of 2 nt. Both are included into the reading frame as parts of the first conserved domain (marked in green). Sequence extended by 42 nt in 3'-direction to reach stop codon (marked blue).

ATGGGTGACTGGAGTGCTCTGGGTGCGTTGTTGGACAAGGTTCAAGCCTACTCCACTGCT
 GGAGGGAAGGTGTGGCTCTCTGCTCTTCATATCCGGATCCTTGTCTGGGCACTGCT
 GTGGAATCAGCGTGGGGAGATGAGCAGTCTGCCCTCAAATGCAACACCCAGCAGCCTGGT
 TGCAGAAATGTCTGCTATGACAAATCATCTTCGTGTCAACACCTACCTCCTGTACTTGG
 CTCATGTTTTACCTGAATAGGAAAGAACAGAAATTCAGCAAGATTGAAGAGGTGCTCAAA
 GCTGTACAAAATGACGGAGGCGACGTTGATGTCCCACTGAAGAAGATTGAGATGAAAAAG
 CTTAAATATGGCATCGAGGAACATGGGAAAGTGAAGATGAAAGGAGCCCTGCTGAGAACT
 TACATAGTCAGCATCTTCTTCAAGTCACCTTTTGAAGTGGGCTTCTGCTGATTACAGTGG
 TACATCTATGGCTTCAGTTTGTCTGCTGTCTACACCTGTGAGAGGTCCCCATGTCCACAC
 AGGGTGGAGTGTCTTCTTCCCGTCCCACTGAGAAGACGGTCTTCATTATTTTTCATGTTG
 GTGGTCTCGCTCGTATCCCTGGTACTTAATATTATTGAGCTCTTCTATGTGCTTTTCAAG
 AATATCAAAGATCGTGTGAAGGGCAAACAGCAGCCACGCTCTACCCAGCGCTGGCACC
 CTCAGCCCTATGCCCAAAGAGCTGTCCACTACCAAGTATGCCTACTACAACGGTTGTTCC
 TCGCCAACTGCACCATTTTCAACCATGTGCGCCGCAAGTTATAAGATGGCCACAGGGGAG
 CGGGGGGCCGGATCGTCCGTAATTATAAAGCACGCTAGCGAGCAGAATTGGGCCAAC
 TACTCCACAGAGCAGAAGCGACTCGGACAGAATTGGAGGAGGAAGCACAATTTCAAATTCC
 CACGCCAAAGCGTTTGACTTCCCGGATGATACCAAGAGCACAAGAAATGTCCTCATTTG
 GCAAGTCATGAGTGCACCGTTAGCGCTGATGGACGCTCGTCTTGCAGCCGGGCAAGC
 AGC

>Fr130494 Our suggested modification.

ATGGGTGACTGGAGTGCTCTGGGTGCGTTGTTGGACAAGGTTCAAGCCTACTCCACTGCT
 GGAGGGAAGGTGTGGCTCTCTGCTCTTCATATCCGGATCCTTGTCTGGGCACTGCT
 GTGGAATCAGCGTGGGGAGATGAGCAGTCTGCCCTCAAATGCAACACCCAGCAGCCTGGT
 TGCAGAAATGTCTGCTATGACAAATGCTTCCCTATCTCCCATGTTGCTTCTGGGTCTCTA
 CAGATCATCTTCGTGTCAACACCTACCTCCTGTACTTGGCTCATGTTTCTACCTGAAT
 AGGAAAGAACAGAAATTCAGCAAGATTGAAGAGGTGCTCAAGCTGTACAAAATGACGGA
 GGCGACGTTGATGTCCCACTGAAGAAGATTGAGATGAAAAAGCTTAAATATGGCATCGAG
 GAACATGGGAAAGTGAAGATGAAAGGAGCCCTGCTGAGAACTTACATAGTCAGCATCTTC
 TTCAAGTCACCTTTTGAAGTGGGCTTCTGCTGATTACAGTGGTACATCTATGGCTTCAGT
 TTGCTGTGCTGCTACACCTGTGAGAGGTCCCATGTCCACACAGGGTGGACTGTTTCTTG
 TCCCGTCCCACTGAGAAGACGGTCTTCATTATTTTTCATGTTGGTGGTCTCGCTCGTATCC
 CTGGTACTTAATATTATTGAGCTCTTCTATGTGCTTTTCAAGAATATCAAAGATCGTGTG
 AAGGGCAAACAGCAGCCACGCTCTACCCAGCGCTGGCACCCTCAGCCCTATGCCAAA
 GAGCTGTCCACTACCAAGTATGCCTACTACAACGGTTGTTCTCGCCTGCAACCACTT
 TCACCAATGTGCGCCGCAAGTTATAAGATGGCCACAGGGGAGCGGGGGCGGATCGTGC
 CGTAATTATAAAGCACGCTAGCGAGCAGAATTGGGCCAACTACTCCACAGAGCAGAAG
 CGACTCGGACAGAATGGAGGAGGAAGCACAATTTCAAATTTCCACGCCAAAGCGTTTGAC
 TTCCCGGATGATACCAAGAGCACAAGAAAATGTCCTATTGGCAGCTCATGAGCTGCAA
 CCGTTAGCGCTGATGGACGCTCGTCTTGCAGCCGGGCAAGCAGCAGATTGAGCAGTCTGA
 GCACGGCCAGATGACCTAGATGTTTGA

>Fr142334 Ensembl prediction in 2 exons (black/blue letters). Sequence extended 9 nt in 5'-direction to reach initiation codon (marked green). Intron of 102 nt included as a part of reading frame (marked blue). Sequence extended 324 nt in 3'-direction to reach stop codon, 3 first nt considered part of second conserved domain (green), remaining corresponds to cytoplasmic C-terminal tail (marked blue).

TGGTCCATTCTTGGCCGCTTCTTAACAGAAAGTTCAAGATCATTCCACAGTCATTGGCAAG
 ATATGGCTGACAAATGCTGCTCATCTTCCGCATCTTGTGGTGGCGTTGGTGGGCGACGCG
 GTGTACAGTGACGAGCAGTCTAAGTTTACCTGCAACACCTACAGCCTGGATGCAACAAC
 GTCTGCTACGACACCTTTGCTCTGTGTACACTTGCCTTCTGGGCTTTTTCAGATTGTC
 CTGCTCTCCACACCTTCGATTTTCTACATCGTCTACGCTTGTGCAAAAGATCACCAGAAT
 GAAAAATTATAACACCACCTATAACAACGAAGAGTGGAGCTCTCAGGAGGATGAGTGTGAG
 GAGAGGAGCCAGCTGAACGAGGAAATGAAAGAGGTGCGAAAGGACCCGACCCAGCTCTCC
 AGTCAAGTGTGCTCATCTACATCATCCATGTTCTCTGCGCTCCATCATGGAGCTCATC
 TTCTGTATCGGACAAATATACCTCTTTGGATTTGAAGTGCCGCTCTTTTCCGCTGTGAC
 ACCTACCCGTGTCAAACCAACCGACTGCTTTGTGTCCGAGCCACAGAGAAGACCATC
 TTCTGAACTTCATGTTACGCGTCAGTCTGGGATGCTTCATCTGAACATTGTGGAGCTG
 CATTATCTCGGC

>Fr142334 Our suggested modification.

ATGGGAGAGCTGGTCCATTCTTGGCCGCTTCTTAACAGAAAGTTCAAGATCATTCCACAGTC
 ATTGGCAAGATATGGCTGACAAATGCTGCTCATCTTCCGCATCTTGTGGTGGCGTTGGT
 GGCGACGCGGTGTACAGTGACGAGCAGTCTAAGTTTACCTGCAACACCTACAGCCTGGA
 TGAACAACGCTGTGCTACGACACCTTTGCTCCTGTGTACACTTGCCTTCTGGGTCTTT
 CAGATTGCTCTCGTCTCCACACCTTCGATTTTCTACATCGTCTACGCTTGTCAAAGATC

ACCAAGAATGAAAA**GT**TAGAGGTGAAGAAGGTGGTCTGTGGTACCCAGGTCTCCACACCG
CTCAGAGGGGAGAAGGATCCGGGGGAGATAAAGAGGCAATGCTGGAGGGGGTAGTTAT
AACACCACCTATAACAACGAAGAGTGGAGCTCTCAGGAGGATGAGTGTGAGGAGAGGAGC
CAGCTGAACGAGGAAATGAAAGAGGTGCGAAAGGACCCGACCCAGCTCTCCAGTCAAGTG
TTGCTCATCT**TACATCATCCATGTTCTCCTGCGCTCCATCATGGAGCTCATCTTCTGATC**
GGACAATATTACCTCTTTGGATTGAAGTGCCGCATCTTTCCGCTGTGACACCTACCG
TGTCCAAACCAACCGAGCTGCTTTGTGTCCCGAGCCACAGAGAAGACCATCTTCTGAAC
TTCATGTT**CAGCGTCAGTCTGGGATGCTTCATCCTGAACATTGTGGAGCTGCATTATCTC**
GGCTGGATTTATATCTTCAGGGTGTGTTCTCTGCATGCTGCACGTGCTGCAAGTCAGAC
AGAGACCCCGTTCCGCGAGGTGGAGTTGTATTCCGACAACAACCCGCTGCTGCTGGAGCTC
AAGCATTCACTGCGGGGAGGGTCGTGCTGCAGGCCACCTCTGCTGTGTCACGGGACAAG
AGCAGCAGCGTCCCAATCAAGCCCGGCCATCTCTTTGAAACAGACTCTACGCTGGAG
TGCACGTCAAAGAGGAACCTAGATGAGAAGGAACGCGCGAAAAACAAGACTACATAAAATG
GGAAGAGGCAAAAAGTCATGGCTGTAA

>Fr125574 (previous id: Fr137519) Ensembl prediction. No modification.

ATGGGTGACTGGAGCCTCTGGGAACTTTCTAGAAGAGGTCCAGGAACACTCCACCTCG
GTGGGAAAGGTCTGGCTACCGTCTTGTTCATCTTCCGGATCTTGGTGTGGGCACGGCC
GCCGAGTCTCTGGGGGAGGGTCGTGCTGCAGGCCACCTCTGCTGTGTCACGGGACAAG
TGCACCAACGTCTGCTACGACAGCGCTTCCCCATCGCCATATCCGCTACTGGGTGCTG
CAGATTGTGTTTGTCTCCACGCCGTCCCTCATCTACATGGGTACGCCATGCACACCGTG
CGCCGAGAGGAGAAACAGCGCAGGAGGGAGCAAGAGGAGAGGGAAGCGAGGGCGGAAAGT
GGAGGAAGCCTGGAGGCTGAGGAATTCTTCCAACAGAAGGAGAGGAAAGAGCCCGGCG
TCGGGCGGGACCAAGCCGGTTCAACTGAGAGGAGCCCTGCTGCAACG**TACATACTCAGC**
ATCATGATCCGCACGGTGATGGAGGTGACGTTTATTGTGGTGCAGTACCTGATGTACGGG
GTGTTCTCTAATGCGTTGTACCTGTGCAAGGCCGTGGCCCTGTCCAAACCCGGTCAACTGC
TACATGTCAGGAGCACAGAGAAGAACGTCTTATCGTCTTATGCTGTTGGTGGCGGGT
GTGTCTCTGCTGCTCTCCGTGTTGGAGCTCTACCACCTCGGCTGGAGGAGCGTCAGAAGA
CACCTACGCAATAAGATGAGCGAAAAGAGCAACACAGAACTGTGACAGTGGCTGTGTCC
ACGGCCTTGGAGCCCAACAGTCCACCACAGCCTTCGCTTTCCTGCACCCACCCCAAGAT
TTCAGCCAGTGCCTTGCAGCCTCTGGATCCATGAATGCCATAACGTCCATGGCCGCTCAC
CCCTTCAACAACAGGATGGCGCTGCAGCAGAACTCGGTCAACCTGGCCACCGAACAGCAT
CACAGCTGCGACAACCTGGAGGACGAGTCAGACTTCTGAGGATCAGATACGACCAACTA
CCCATGGAGCTGCCCAAAGCTGCTCGCCATCCCCCTCTGCACTCCAGCTACACGAGG
GACAAACGGCGCTGAGCAAGACCAGCGGGAGCAGCAGACCTCGTCCCGATGATCTT
GCGGTGTAG

>Fr129429 Ensembl prediction in 7 exons (alternating black/blue letters). The 159 nt intron between the predicted exons 3 and 4 included as a part of the reading frame (marked in blue in lower sequence). The Ensembl predicted exon 5 is a part of the second conserved domain. To make a better fit to the orthologs (and all other connexins), the last 13 nt of exon 5 is removed (marked gray in upper sequence). This creates a new splicing site between the predicted exons 5 and 6 (marked by underlining last and first nt of the corresponding exons in lower sequence). The first part of exon 6 (marked in bold blue letters in upper sequence) appears to contain several potential errors. We exchanged this sequence with the corresponding sequence from the previous releases of Ensembl (v32-35) (marked in red letters in lower sequence).

ATGGCAGACTGGAGCTTACTGGGAACTTCTGGAGGAGGTGCAGGAGCATTCCACCTCT
GTTGGAAG**GTGTGGTTGACCATCTCTTTATCTTCCGGATCCTTGTGCTGGGGACCGCC**
GCCGAGTCCTCATGGGAGACGAGCAAGAAGATTTCAACTGTGACACAGAACAGCCAGGC
TGCGAGAACGTTGTATGACCGAGCCTTCCCAATAGCACATATACGATACTGGGTGCTG
CAGATTGTGTTTGTGTCCACGCCAGCCTGATCTACATGGGTACGCCATGCACAGGGT
CGCTTGCGTTTGAGGGGAGCGCTACTACAGACCTATGTGCTGAGTATACTGATACGAAGC
GTCATGGAGGTGGTGTCTCACTCTCCAGTATTTAATGTACGGGATCTTCTTAATCTT
CTGTATGCTCGCAAGGTAAAAACAGAAA**GCTTGCCGTGCCCTCAGCGGGAAGTGAAGT**
CTCCAGGCCAACAGAGAAAATGTCTTTATTGTGTTTATGCTGCGCGTTTCCGGTGTCTT
CTGGTCTCTCAGCGTCTAGAGCTGCAACACCTGGCGTGGAGGCACTGCTGTAGAAAGACG
GCGGCGGCTAACAAAGGCTCGCTAGGCCGACAGGTCTCTGTGCCCTCCACCACAGTCC
ATCCCACTCCAGAATTAGCCAGTGCATGATGGGTCCACACACTTCTCTCTTTCG
TTCCCCAACCAACACCTGGCACACCAACAGAACTCAGAGAACATGGCCACCGAGAAGCAC
AAAATAGCCGCCCGCTAGAGGAGGCCACCTCCTGCAGATGGGCTGCTACTCACACGGA
TGGCAGAGACCAACCCAGTCAGATCCAGGACGACCCCTACCTGAGGAACGACAACAGT
CGCTACGGTCCCGGCAGCAGGGAGATCAGCTGTTACAGATCCAAAATGGAGGCTCCGAC
AGGCTCTTGCTTTGCCCAAGCGGCGCTACAATCAGAAAGACAAGCGGAGATTACAGAGA
ACCAGCGGCACAGCAGCCGAACAAGAGCGGACGACCTGTCCGTTTAA

>Fr129429 Our suggested modification.

ATGGCAGACTGGAGCTTACTGGGAACTTCTGGAGGAGGTGCAGGAGCATTCCACCTCT
GTTGGAAGGTGTTGGTTGACCATCTCTTTATCTTCCGGATCCTTGTGCTGGGGACCGCC
GCCGAGTCCTCATGGGAGACGAGCAAGAAGATTTCAACTGTGACACAGAACAGCCAGGC
TGCGAGAACGTTGTATGACCGAGCCTTCCCAATAGCACATATACGATACTGGGTGCTG
CAGATTGTGTTTGTGTCCACGCCAGCCTGATCTACATGGGTACGCCATGCACAGG**GT**
CGCAGGGAGGAGAAGAGGAGGAGCGGGAGGAGGAGCGGGAGGGGAGAGGGGAGAG
GAGGACCCGGGCGGCGGACGAGGAAATGACAGTGGGGAAGAAGATGAGAAAGGTGGG
AGAGAAGTGGAGAAGCACGGAGAGAAAGAGGGTGGAGGTCGCTTGCCTTGGGGGAGCG
CTACTACAGACC**TATGTGCTGAGTATACTGATACGAAGCGTCATGGAGGTGGTGTCTCTC**
ACTCTCCAGTATTTAATGATACGGGATCTTCTTAATCTCTGTATGCTGCAAGGCTGG
CCGTGCCCTCATCTGTGAAGTGTATGTCTCAGGCCAACAGAGAAAATGTCTTTATT

>Fr161792 (previous id: Fr139615) Ensembl prediction. Sequence extended 705 nt in 3'-direction to reach stop codon; first 144 nt considered as part of second conserved domain (marked green), remaining corresponds to cytoplasmic C-terminal tail (marked blue).

```
>Fr161792 (previous id: Fr139615) Our suggested modification.
```

>Fr150243 Ensembl prediction in 2 exons (black/blue letters). 63 nucleotide intron considered as part of reading frame (marked blue). Sequence extended 390 nt in 3'-direction to reach stop codon; first 12 nt as part of second conserved domain (marked green), remaining corresponds to cytoplasmic C-terminal tail (marked blue).

```
>Fr150243ic Our suggested modification.
```

43

GACAGCAGTGACTCTCAAGAACGGCGAGCCTGGGTGTGA

>Fr145344 Ensembl prediction in 2 exons (black/blue letters). The predicted 60 nucleotide intron is regarded as a part of the reading frame (marked blue). Sequence extended 204 nt in 3'-direction to reach stop codon (marked blue).

ATGAGCTGGAGCTTCCTCAGCGGGCTGCTGGAGGAGATCCACAACCACTCCACCTTCGTG
GGGAAGCTGTGGCTCACCCTGCTCATCGTCTTCCGCATCGTTCTACTGCCGTTGGGGGA
GAGTCCATCTACTACGATGAGCAGAGCAAGTTCTGTGCAACTCGGGACAGCCGGGCTGC
GAGAACGTTTGTACGACGCCTTTGCCCTCTGTCTCAGCTCCGCTTCTGGGTATTCCAG
ATTATCCTGGTGGCGATGCCCTCTCTCATGTACATGGGCTACGCCATCAACAAGATCGCT
AGATTAGATGAAGCCAAAGGAGGTGGAACCTCCACTGCTGTTAGAACGGGAGGGGGGGC
TACACGCACAGGAAGCCAGGAAAATCTGCTTTGGAGCGCGGCAGCACCAGGGGTATCGAG
GAGACCGAGGAGGACAGGAGGACGATCCCATGATCTACGAGGTACCGGAGATCGAGCCC
CCCAAGAGGCGGAGGGATCCGCTGCAGCCCGCTCCGAGACCCAAAGTCCGGCAGCATGGA
CGCAAGCGCATCAGAGACGAGGGGCTGATGCGGGTTTACGTTTGCAGCTGGTGACCCGT
ACGGTGTCTGGAAGCCTGCTTCTCGCCGGCCAGTATTTACTGTACGGGTTCCGTGTGATG
CCCGTGTTCGTGTGCTCGGGGAAACCGTCCCCCACAACGTTGACTGCTTCGTCTCACGA
CCCACAGAGAAGACCATCTTCTGCGCATCATGTACGGGGTCACAGTCCTTTGCCTCATT
CTCAACATTTGGGAGATGCTTCATTTAGGGATCGGCTCCATATACGACATCCTCCGCCGG
CGGCGCAGCCACCCAGGATGATGAGTACCAGCTGGGCTTGTGGGAGGGGTCGGCGGT
GACGGGGCTGCCGATTATGTGCGCTACCCTTTCTCGTGAACACGCCGTCGGCTCCGCCT
GGCTACAACATTGTGGTAAAGCCGAGCAGATGCCCTACACAGACCTCAGCAACACCAAG
ATGGCGTGCAAGCAAAACCGGGCAAACATTGCCAAGAAGAGCAACAGCAGTTTGGTAGT
AACGAAGACAACCTTCCCAACGGAGGAGAAGCCGCGTGGCTTTGAACAAGACATGATC
CAGCAGGCTCAGCAGCAGCTGGAGGCGGCCATCCAGGCCCTACAGCCAGCAGCAC

>Fr145344 Our suggested modification.

ATGAGCTGGAGCTTCCTCAGCGGGCTGCTGGAGGAGATCCACAACCACTCCACCTTCGTG
GGGAAGCTGTGGCTCACCCTGCTCATCGTCTTCCGCATCGTTCTACTGCCGTTGGGGGA
GAGTCCATCTACTACGATGAGCAGAGCAAGTTCTGTGCAACTCGGGACAGCCGGGCTGC
GAGAACGTTTGTACGACGCCTTTGCCCTCTGTCTCAGCTCCGCTTCTGGGTATTCCAG
ATTATCCTGGTGGCGATGCCCTCTCTCATGTACATGGGCTACGCCATCAACAAGATCGCT
AGATTAGATGAAGCCAAAGGAGGTGGAACCTCCACTGCTGTTAGAACGGGAGGGGGGGC
TACACGCACAGGAAGCCAGGAAAATCTGCTTTGGAGCGCGGCAGCACCAGGGGTATCGAG
GAGACCGAGGAGGACAGGAGGACGATCCCATGATCTACGAGGTACCGGAGATCGAGCCC
CCCAAGAGGCGGAGGGATCCGCTGCAGCCCGCTCCGAGACCCAAAGTCCGGCAGCATGGA
CGCAAGCGCATCAGAGACGAGGGGCTGATGCGGGTTTACGTTTGCAGCTGGTGACCCGT
ACGGTGTCTGGAAGCCTGCTTCTCGCCGGCCAGTATTTACTGTACGGGTTCCGTGTGATG
CCCGTGTTCGTGTGCTCGGGGAAACCGTCCCCCACAACGTTGACTGCTTCGTCTCACGA
CCCACAGAGAAGACCATCTTCTGCGCATCATGTACGGGGTCACAGTCCTTTGCCTCATT
CTCAACATTTGGGAGATGCTTCATTTAGGGATCGGCTCCATATACGACATCCTCCGCCGG
CGGCGCAGCCACCCAGGATGATGAGTACCAGCTGGGCTTGTGGGTACCACTGGAGGCT
GTAGAGGGGGCCGTAGGGGGTACAGCCCTGAGGCGGGCTCTGAAGGAGGGGTCGGCGGT
GACGGGGCTGCCGATTATGTGCGCTACCCTTTCTCGTGAACACGCCGTCGGCTCCGCCT
GGCTACAACATTGTGGTAAAGCCGAGCAGATGCCCTACACAGACCTCAGCAACACCAAG
ATGGCGTGCAAGCAAAACCGGGCAAACATTGCCAAGAAGAGCAACAGCAGTTTGGTAGT
AACGAAGACAACCTTCCCAACGGAGGAGAAGCCGCGTGGCTTTGAACAAGACATGATC
CAGCAGGCTCAGCAGCAGCTGGAGGCGGCCATCCAGGCCCTACAGCCAGCAGCACAGGCT
GAGGTGCAGCTCGGGGAGAACCAGGACGACAAACCCAGAGTAACATCATTGAGGCTCAA
CCGCGAGCTGCAGCCTCAGCCCCATAAGGAGCGCAACACAGATTCAAGCACGGCAAGGA
GGCAGCAGTGCAGGAGGACGAGCAGCAACAGCAGCAGCAAAATCGGGAGAGGGGAAG
CCCTCCGTGTGGATTAA

>Fr160540 (previous id: Fr148221) Ensembl prediction in 2 exons (black/blue letters). The predicted 60 nucleotide intron regarded as a part of the reading frame (marked blue). Sequence extended by 447 nt to reach stop codon (marked blue).

ATGAGTTGGAGTTTCTGACACGCCTGTTGGAAGAAATTCACAACCATTTCTACGTTTGTG
GGCAAGATATGGCTGACTGTCTTATTGTCTTCCGCATCGTGTGACGGCTGTTGGCGGG
GAGTCCATCTACTACGATGAGCAGAGCAAGTTTGTGTGCAACTCGGGCCAGCCGGGCTGC
GAGAATGTCTGCTACGATGCCTTCGCTCCACTGTCACACGTCCGCTTTTGGGTGTTCCAA
ATCATTCTGGTGGCCACCCATCGCTCATGTATCTGGGATACGCTGTCAACAAAATTGCT
CGTGCCGAAGAGCGGGCAGACAGGAGGCAGCACAGAGGCATCGAAGAAGCTGAGGATGAC
CAAGAGGAAGACCTATGATCTACGAACAGCAGACATTGGCAGTGAAGACGCAAAAGGG
AGCGTCACTAAGGGAAGAAAGATAAGGTCAAGGTGCGCCACGACGGACGCCAGCGTATTA
GAGGATGGCTTGTGCGGATTATGTCTTTCAGCTCTTGGCCCGCTCCCTGCTGGAGGTG
GCTTTCTGTGTGGGCAATCACCCTGTATGGATTGCTGTTCCCGCCACCTATGCTCTGC
TCTCAGCTGCCATGCCCCCAGCGTGGACTGCTTTGTGTCTCGGCCACTGAAAAAACC
ATCTTCTCTCATTATGTACACAGTCTCCCTGCTCTGTCTGATGCTGAATATCTGGGAG
ATGCTTCACTAGGCATCGGCACCATCTGCGAG

>Fr160540 (previous id: Fr148221) Our suggested modification.

ATGAGTTGGAGTTTCTGACACGCCTGTTGGAAGAAATTCACAACCATTTCTACGTTTGTG
GGCAAGATATGGCTGACTGTCTTATTGTCTTCCGCATCGTGTGACGGCTGTTGGCGGG
GAGTCCATCTACTACGATGAGCAGAGCAAGTTTGTGTGCAACTCGGGCCAGCCGGGCTGC
GAGAATGTCTGCTACGATGCCTTCGCTCCACTGTCACACGTCCGCTTTTGGGTGTTCCAA
ATCATTCTGGTGGCCACCCATCGCTCATGTATCTGGGATACGCTGTCAACAAAATTGCT
CGTGCCGAAGAGCGGGCAGGTGGGAAGGGGCGCAGGCTATTGCGAGAGGAACTCAAG
AGGAAGCTGTATCTGGCAGACAGGAGGCAGCACAGAGGCATCGAAGAAGCTGAGGATGAC

CAAGAGGAAGACCCCTATGATCTACGAAACAGCAGACATTGGCAGTGAAGACGCAAAAGGG
 AGCGTCACTAAGGGGAAAGATAAGGTCAAGGTGCGCCACGACGGACGCCAGCGTATTTAA
 GAGGATGGCTTGATGCGGATTATGTCTTTCAGCTCTTTGGCCCGCTCCCTGGAGGTTG
 GCTTTCTCTGTGTGGGAGTACACCTGTATGGATTGCTGTTCCCCCACCTATGTCTGC
 TCTCAGCTGCCATGCCCCACAGCGTGGACTGCTTTGTGTCTCGGCCCACTGAAAAAAC
 ATCTTCTCTCTATTATGTACACAGTCTCCCTGCTCTGTCTGATGCTGAATATCTGGGAG
 ATGCTTACCTAGGCATCGGCACCATCTGCGAGATCATCCGTTCCCACTCCCTGAT
 GAGGAGCTGTACGGACTGACACAATCAAAAGGAGCCACGCTGACGCTGGATTGAGCCGA
 GAGGAGTACAGCAGCTACGCTTTCTCTTGAATGCCCATCAGCTCCGCTGGGTACAAC
 ATCGCAATCAACCCCTCTGGTAACAGCAGGACACCGTGATCAACCCCTGCCGTGCTACT
 GATCTCACCACGCGAAGATGGCGTGCCGGCAAAATCACGCAAACTCGCTCACGAGGAG
 CAGCAACAGTACAGCAATAACGACGAAACCTATGCAAGCCGGGATGGGCGATGACCAC
 ATGCGTGCCAATCACTCTCAGAACAGACTGGAGGCGGACAGCGCAGCTCACAGCCAGCTG
 GAGGGTCAAAGCAACAAGGCTCATCGTGACCGCAACAGCGACAGGCCCTCAAACACAG
 TCCAGCAAGGATGACACGACCGGGGAGCACCAGCACCAGCAATACCAGCAATATGGC
 GTCATCAAAGTTGAGAGTGATATGA

>Fr164495 (previous id: Fr136128) Ensembl prediction in 2 exons (black/blue letters). 6 nt added in 5'-direction to reach initiation codon (marked green). The predicted 96 nt intron is regarded as a part of the reading frame (marked blue). Sequence extended by 447 nt in 3'-direction to reach stop codon (marked blue).

TGGAGCTTCTCACACGCTCTGCTGGAAGAGATCCACAATCATTCCACATTTGTGGGGAAA
 GTGTGGCTCACTGTGCTCATTATCTTCCGCTTGTGCTCACGGCAGTTGGAGGCGAATCC
 ATCTACTCGGATGAGCAGACAAAGTTACCTGCAACACAAAGCAGCCGGTTGTGACAAC
 GTATGCTACGATGCCTTTGCCCTCTCTCGCATGTCCGTTTCTGGGTCTTCCAGATCATC
 ATGATCTCCACTCCTTCCATCATGTACATGGGCTATGCCATTACAAAGATTGCTCGGAAG
 GAGGACGAAGATGACGACGCTGAGCCCATGATCTATGAGGATACACTGGAGGTGCAAGAT
 GCCAAACCAAGCCGGGACAGCAGCAGGCGCAAGACCCACCAAAATACGACGGCTCGA
 AAAATCATGCAGGAAGGCCTGATGAGGATCTACGTCCTCAGCTGATGTCAGAGCTGTT
 TTTGAAATTGCCTTCTTGTCTGGACAGTACCTTCTCTATGGTTTTCTGTTAGTCCATCC
 TATGTATGCAACAGGATCCCCTGCCACACAGAGTGGACTGTTTATCTCGAGACCCACA
 GAAAAAACAATCTTCTCCTTATCATGTACGTGGTGGCTGCTTTGTCTTGTCTAAAC
 ATCTGCGAGATGCTTACCTGGGAATGGGAACATTTGGGACACCCCT

>Fr164495 (previous id: Fr136128) Our suggested modification.

ATGAGCTGGAGCTTCTCACACGCTCTGCTGGAAGAGATCCACAATCATTCCACATTTGTG
 GGGAAAGTGTGGCTCACTGTGCTCATTATCTTCCGCTTGTGCTCACGGCAGTTGGAGGC
 GAATCCATCTACTCGGATGAGCAGACAAAGTTACCTGCAACACAAAGCAGCCGGGTTGT
 GACAACGTATGCTACGATGCCTTTGCCCTCTCTCGCATGTCCGTTTCTGGGTCTTCCAG
 ATCATCATGATCTCCACTCCTTCCATCATGTACATGGGCTATGCCATTACAAAGATTGCT
 CGGATACAGATGAGGAGCGCAGGAAACTCCACAGGCTTCGCAAAAAGCCTCCCCCGCAT
 TCCAGATGGAGAGAGAGCCATCACCTGCAGGGCGTCTTAGAGGAGACGAAGATGACGAC
 GCTGAGCCCATGATCTATGAGGATACACTGGAGGTGCAAGATGCCAAACCAGAACCCGGG
 AACAGCAGGGGCAAGACCCACCAAAATACGACGGCCGTCGAAAAATCATGCAGGAAGGC
 CTGATGAGGATCTACGCTCTTACGCTGATGTCAAGAGCTGTTTTGAAATTGCCTTCTT
 GCTGGACAGTACCTTCTCTATGGTTTTCTGTTAGTCCATCCTATGTATGCAACAGGATC
 CCTTGCCACACAGAGTGGACTGTTTCTCTCGAGACCCACAGAAAAACAATCTTCTCTC
 CTATCATGTACGTGGTGGCTGCTTTGTCTTGTGCTAAACATCTGCGAGATGCTTCA
 CTGGGAATGGGAACATTTGGGACACCCCTTCGATGAAGAGGAGCAGGGGCGAGACAGTCA
 TCCTACGGCTACCTTTTTCTCGCAATATTACAGCTTCCCTCCAGGGTACAACCTCGTA
 ATGAAGACAGACAAACCCAGCAGGATTCCCAACAGCCTCATTGCCCATGGGCGAAGCGTA
 GCCAACGTGGCTCAGGAGCATCAGTGATCAGCCCGGACGAGAACATCCCTCTGATCTC
 GCAAGCCTACACCGGCACCTAAGAGTTGCTCAAGAGCAACTTGATATGGCATTTACAGAC
 TACCAGACCAACAAACCCAGCAAACTCCAGAACAGCAGTCCAGTGTCTGGAGGCACC
 ATCGCAGAGCAGAACAGAGTCAACGCAGTTCAAGAGAAGCAGGGCGCAAGGCCGAAATCA
 GCCACAGAGAAGGCTGCAACTATTGTAAAAAATGGAAGAGCTCCGTTTGGATCTAG

>Fr122421 Ensembl prediction. Sequence extended by 48 nt in 3'-direction to reach stop codon (marked blue).

ATGAGCTGGAGTTTCTGACGCTTTGCTGGACGAGATCTCAACCACTCCACCTTCGTG
 GGGAAAGTCTGGCTGACCGTCTGATCATCTTCCGCTCGTGCTGACCGCGTTGGCGGC
 GAAACCATCTACTACGATGAGCAAAAGTAAATTTGTTTGAACACGACGACCGTGGATGT
 GAGAACGTCTGCTACGATGCCTTCGCCCCGCTCTCACAGTACGATTCTGGATCTTTCAG
 GTGATTTTGATAACCAACCCCAACATAATGTACCTGGGCTTCGCCATGCACAAGATCGCA
 CGCATGGATGACAGCGAGTACCGCATCGTCCGGAACCCAAAAAGAGGATGCCATCGTT
 AGCCGCGGAGCTGTTAGGACTATGAGGAGCGAGAGGACAATGGAGAGGAAGACCCATG
 ATTGCCGAGGAGATTGAACAAGAAAAGCCTGACAAAACGGAGAAGGATCCTAACTGGTAT
 CTGTGTATCCAGGCACAGAGAAGAAGCACGACGGCCGGCGGCAATCCAGCGCGACGGC
 CTAATGAAGGTCTACGTGTGCCAGCTGCTGTGGCGCTCATCCTTTGAGGTTGCGTTTCTT
 TTGGCCAGTACATCCTCTACGTTTGTAGGTTTTGAGGTTTTTCCGTCCTTTGTGTGACCCGCTCA
 CCATGCCCCACACCGTGGACTGCTTCTGTGTCGCGCCACAGAGAAGACCATCTTCTCTG
 CTGGTATGTACGTCTGTCTTCTCTGCTGCTCTCACCCTTTTGAATGATCCAT
 TTGGGGATAGGAGGTGTCGCGACACCTTTGGAGACGGGCCACTCTCAACCCACGCGCC
 CCTCGTCCGTCCACACGACGACATACCCACAGCCCGCCAGGATACCAAGCCACTATG
 AAGAAGGAGAACTGAAAGGACAGCTGAGGGACTCGCCGATAGGGGACTCTGGGCGGGAG
 AGCTTCGGTGATGAGGGTCCGTATCCAGGGAAGTGGAGCGCTTGAGGAGGCACCTGAAG
 CTGGCCCGACCAACCTGGATTTGGCCTACCAGGTGAGGAAGGAAACCCCTCACGGAGC

AGCAGCCCCGAGGTGAACACGGCTGCACAGACGGCTGCCGAGCAGAACCGACTCAACTTT
 GCCCAGGAGAAGCAGGGAGAAACAACCGAAAAA
 >Fr122421 Our suggested modification.
 ATGAGCTGGAGTTTCTGACGCGTTTGGTGGACGAGATCTCCAACCACTCCACCTTCGTG
 GGGAGATCTGGCTGACCGTCTCTGATCATCTTCCGCATCGTGCTGACCGCCGTTGGCGGC
 GAAACCATCTACTACGATGAGCAAAATAATTTGTTTGCAACACGACGAGCCTGGATGT
 GAGAACGCTGCTACGATGCCTTCGCCCGCTCTCACACGTACGATTCTGGATCTTTCAG
 GTGATTTTGATAACCAACCCCCACCATAATGTACCTGGGCTTCGCCATGCACAGATCGCA
 CGCATGGATGACAGCGAGTACCGCATCGTCCGGAACCCAAAAAGAGGATGCCATCGTT
 AGCCGCGGAGCTGTTAGGGACTATGAGGAGGACAGAGGACAATGGAGAGGAAGACCCCATG
 ATTGCCGAGGAGATTGAAAGAGAAAAGCCTGACAAAACGGAGAAGGATCCTAACTGGTAT
 CTGTGTATCCAGGCACAGAGAAGAAGCAGACGCGCCGGCGGCAATCCAGCGCGACGGC
 CTAATGAAGGTCACGTGTGCCAGCTGTGTGGCGCTCATCTTTGAGGTTGCGTTTCCTT
 TTCGGCCAGTACATCTCTACGGTTTTGAGGTTTTTCCGTCCTTTGTGTGCACCCGCTCA
 CCATGCCCCACACCGTGGACTCTTCTGTGTGCGCCCCACAGAGAAGACCATCTTCTG
 CTGGTCATGTACGTCGTGCTTTCTCTGCTGCTCCTCACCGTTTTTGAATGATCCAT
 TTGGGGATAGGAGGTGTCCAGCACCTTTTCGAGACGGGCCACTCTCAACCCACGCGCC
 CCTCGTCCGTCCACCACACGACATACCCACAGCCCCGCCAGGATACACGCCACTATG
 AAGAAGGAGAACTGAAAGGACAGCTGAGGGACTCGCCGATAGGGACTCTGGCGGGAG
 AGCTTCGGTGATGAGGGTCCGTCATCCAGGAACTGGAGCGCTTGAGGAGGCACCTGAAG
 CTGGCCAGCAACACCTGGATTTGGCCTACAGGTCGAGGAAGGAAACCCCTACGGAGC
 AGCAGCCCCGAGGTGAACACGGCTGCACAGACGGTGCCGAGCAGAACCGACTCAACTTT
 GCCCAGGAGAAGCAGGGAGAAACAACCGAAAAAAGGTAACCGGCGAGAACTGGCTGAAA
 TGTGCTATAATGTTATCTGA

>Fr162962 (previous id: Fr121008) Ensembl prediction in 2 exons (black/blue letters). Sequence extended by 6 nt in 5'-direction to reach initiation codon (marked green). Predicted intron of 54 nt regarded as part of reading frame (marked blue). Sequence extended by 294 nt in 3'-direction to reach stop codon (marked blue).

TGGAGTTTCTCACGCGCTGTTGGACGAGATCTCCAACCACTCGACCTTCGTGGGCAAA
 ATCTGGCTCACCTCCTCATCGTCTTCCGCATCGTGCTGACGGCCGTCGGGGGCGAGTCT
 ATATACTATGATGAACAGAGTAAGTTTGTGTGCAACACAAACAGCCTGTTGCGAGAAC
 GTGTGCTACGACGCTTCGCGCCGCTGTGCGACATTGCTTCTGGGTTTTCCAGGTGATT
 ATGATCACCGCCCCACCATCATGTACCTCGGCTTCGCCATGCACAAAATCGCCCGGATG
 AACGACGACGACTATCGACCCCGCAGCAGGAAGAAGATGCCATCGTCAGCCGGGAGCC
 AACCGGGACTACGAGGAAGCGGAGGACAACGGCGAGGAGGACCCGATGCGTGCAAGAAA
 CACGACGGGCGGCGTCGGATCAAGCGCGATGGCCTGATGAAGGTCTACGTGTTCCAGCTG
 CTATCGCGGGCCATCTTGAAGTCTCCTTCTGTTGGACAGTACATCCTCTACGGGCTG
 GAAGTCGACCCCTCGTACGTTTGACGCGCTCCCTTGCCCGCACAGGTGGACTGTTTT
 GTTTCCGTCACGAGGAAAACCATCTTCTGCTCATCATGTATGCGGTGAGCGGTCTT
 TGCTTGCTCTTACCCTGCTGGAGATCATCCACCTCGGCATCAGCGGTCTTCGGGACTGC
 TTCTGCGCCCCCGGCTCGCCCTCCACCCCGCGCCACTCGGCTCTCGCCAGCCAGAGG
 TCCTCCATCTCCCGCAGCCGTCGCTCCGCGGGGCTACCACAGCGCTCTGAAAAAGGAC
 CTTTCGGGAAAGATG

>Fr162962 (previous id: Fr121008) Our suggested modification.

ATGAGCTGGAGTTTCTCACGCGCTGTTGGACGAGATCTCCAACCACTCGACCTTCGTG
 GGCAAAATCTGGCTCACCTCCTCATCGTCTTCCGCATCGTGCTGACGGCCGTCGGGGGC
 GAGTCTATATACTATGATGAACAGAGTAAGTTTGTGTGCAACACAAACAGCCTGTTGCG
 GAGAACGCTGTGCTACGACGCTTCGCGCCGCTGTGCGACATTGCTTCTGGGTTTTCCAG
 GTGATTATGATCACCGCCCCACCATCATGTACCTCGGCTTCGCCATGCACAAAATCGCC
 CGGATGAACGACGACGACTATCGACCCCGCAGCAGGAAGAAGATGCCATCGTCAGCCGG
 GGAGCAACCGGACTACGAGGAAGCGGAGGACAACGGCGAGGAGGACCCGATGATCTCTG
 GAAGAGATCGAACCGGAAAGGAGAAGGAGAAGGAGACCGGAGAGCCGATGATGCAAGAAA
 CACGACGGGCGGCGTCGGATCAAGCGCGATGGCCTGATGAAGGTCTACGTGTTCCAGCTG
 CTATCGCGGGCCATCTTGAAGTCTCCTTCTGTTGGACAGTACATCCTCTACGGGCTG
 GAAGTCGACCCCTCGTACGTTTTCGACGCGCTCCCTTGCCCGCACAGGTGGACTGTTTT
 GTTTCCGTCACGAGAAAACCATCTTCTGCTCATCATGTATGCGGTGAGCGGTCTT
 TGCTTGCTCTTACCCTGCTGGAGATCATCCACCTCGGCATCAGCGGTCTTCGGGACTGC
 TTCTGCGCCCCCGGCTCGCCCTCCACCCCGCGCCACTCGGCTCTCGCCAGCCAGAGG
 TCCTCCATCTCCCGCAGCCGTCGCTCCGCGGGGCTACCACAGCGCTCTGAAAAAGGAC
 CTTTCGGGAAAGATGGGCTTTAGGGACAACCTGGGAGACTCCGGCCGGGAGTCTTTTGGT
 GACGAGACTTCATCGCGGGAACGGAGAGGCTGCGTAAACACCTGAAACTGGCGCAGCAG
 CACCTGGACATGGCCTACCAGAACGGGGAAGCAGCCCGTCCGCGACGACGACCCCGAG
 TCCAACGGCACGGCGGTGGAGCAGAACCAGGCTGAACTTTGCTCAGGAGAAGCAGAGTGAC
 AAAGGTCAGACCCTAATTTTATTTGTTGGCGCATCTTTGCCAGGGTTTTTGATGGAATTCTA
 GATCTGTAA

>Fr149277 Ensembl prediction in 2 exons (black/blue letters). Predicted intron of 72 nt regarded as part of reading frame (marked blue).

ATGGGCGACTGGAGCTTCTGGGGCGGCTGTTGGAGAACGCTCAGGAGCATTCTACGGTC
 ATCGGCAAGTCTGGCTGACTGCTCTTTCATCTTCAGGATCCTGGTGCTGGGGGCGAGCC
 GCCGAGGAGGTCTGGGGGACGAGCAGTCTGATTTACCTGCAACACCCAGCAGCCCGGT
 TCGGAGAAATGCTGCTACGACGAGGCGCTTCCCATTTTCGCACATCCGCTTCTGGGTGCTG
 CAGATCATCTTGTGTCGACGCCACCTCATCTACCTGGGCCAGTGCTGCACATCGTC
 CGCATGGAGGAGAAGCGGAAGGAGAAGGAGGAGGAGCTCGGAAAGCAACCGGTTACAG
 GAGGAGAAAGAACTCCTTTACAGAAACGGGGGAGACGACGAGGAGGAGGCGCAAGAAGGAG

AAGCCGCCCATCAGGGATGAGCACGGCAAAATCCGCATCAGAGGTGCACTGCTGCGTACC
TATGTGTTCAACATTATATTTCAAACCCCTGTTTGAAGTGGGATTCAATTTTGGGCCAGTAT
TTCTCTACGGCTTCCAGCTGAGGCCCTGTACAAGTGTGGACGTTGGCCCTGCCCAAC
ACTGTAGACTGCTTCATTTCCAGACCTACTGAAAAGACAATTTTTATTATATTTATGCTT
GTGGTGGCTTGCCTGTCTCTTTTGTGAATTTGTTAGAGATCTATCACCTCGGATGGAAG
AAAGTTAAACAGGGCGTGACAAACCAAGTTTGTCCCGACCGCGAATCAATGCGCCGGGTC
AACATTGCAGAGCCCGAGTGTGTTGGCCTTGGCCTCCAGAACTGCCCATCCAGTTACCCC
CCCACTACACTGATGTGACGGCGGGCAGTGGGGCGTTCTTCAGCCCGTGGCGCCGCCG
GCCGTGCCTTCGACCACGGAGTTCAAGACCGACGACCTCCAGCGGGAGCCGCCTCGCCAC
CAGCCCTCCGCCTCTACTACTACATCAGCAACAACAACACAGGCTGGCCACGCAG
CAGAAGTGGGCCAACCTGGCCACTGAGCAGCAGACTCGGGAGATGAAGGCCAC**CAGCAAC**
ATCACCAACACAACCACCACCGTCACCGCGTCAGGTAGCAGTAGCCCGGGCTCGGCCCTCC
AACACGGGCAGCTGGGGCGGAGGAAGGAAGGAGCGGGGGGAAAACGGCGTCTCCACCACC
AGGGTGGAGATGCACGAGCCTCCGGCGACGGCCGGCGTGGACCTCGGCGACTTAGCCGA
GCCAGCAAGACGACGAGCGTCAGAGCGAGGCCAAGCGACCTGGCTGTCTAA
>Fr149277 Our suggested modification.
ATGGGCGACTGGAGCTTTCTGGGGCGGCTGTTGGAGAACGCTCAGGAGCATTCTACGGTC
ATCGGCAAAAGTCTGGCTGACTGTCTCTTCACTTTCAGGATCCTGGTGTGGGGGAGGCC
GCCGAGGAGGCTGGGGGGGACGAGCAGTCTGATTTACCTGCAACACCCAGCAGCCCGGT
TGCGAGAATGTCTGCTACGACGAGGCCCTTCCCCATTTCCGCACATCCGCTTCTGGGTGCTG
CAGATCATCTTCGTGTGACGCCGACCCCTCATCTACCTGGGCCACGTGCTGCACATCGTC
CGCATGGAGGAGAAGCGAAGGAGAAGGAGGAGGAGCTGCGGAAAGCAAAACGGTTACAG
GAGGAGAAAGAACTCCTTTACAGAAACGGGGAGACGAGGAGGAGCGGCAAGAAGGAG
AAGCCGCCCATCAGGGATGAGCACGGCAAAATCCGCATCAGAGGTGCACTGCTGCGTACC
TATGTGTTCAACATTATATTTCAAACCCCTGTTTGAAGTGGGATTCAATTTTGGGCCAGTAT
TTCTCTACGGCTTCCAGCTGAGGCCCTGTACAAGTGTGGACGTTGGCCCTGCCCAAC
ACTGTAGACTGCTTCACTTCCAGACCTACTGAAAAGACAATTTTTATTATATTTATGCTT
GTGGTGGCTTGCCTGTCTCTTTTGTGAATTTGTTAGAGATCTATCACCTCGGATGGAAG
AAAGTTAAACAGGGCGTGACAAACCAAGTTTGTCCCGACCGCGAATCAATGCGCCGGGTC
AACATTGCAGAGCCCGAGTGTGTTGGCCTTGGCCTCCAGAACTGCCCATCCAGTTACCCC
CCCACTACACTGATGTGACGGCGGGCAGTGGGGCGTTCTTCAGCCGTGGCGCCGCCG
GCCGTGCCTTCGACCACGGAGTTCAAGACCGACGACCTCCAGCGGGAGCCGCCTCGCCAC
CAGCCCTCCGCCTCTACTACTACATCAGCAACAACAACACAGGCTGGCCACGCAG
CAGAAGTGGGCCAACCTGGCCACTGAGCAGCAGACTCGGGAGATGAAGGCCAC**CTCCTCC**
TCCTCCTCTCCACCAATGCAGAGCAGCAGCCCGTCGATGCGGAGCTGCTCCCTTCC
GCCAGCAGCAACATCACCAACAACCAACACCGTCACCGCGTCAGGTAGCAGTAGCCCG
GGCTCGGCCTCCAACACGGGAGCTGGGGCGGAGGAAGGAAGGAGCGGGGGGAAAACGGC
GTCTCCACCACAGGGTGGAGATGCACGAGCCTCCGGCGACGGCCGGCGTGGACCTCGG
CGACTTAGCCGAGCCAGCAAGAGCAGCAGCGTCAGAGCGAGGCCAAGCGACCTGGCTGTC
TAA

>Fr124726 Ensembl prediction. Sequence extended by 339 nt in 3'-direction to reach stop codon (marked blue).

ATGGGTGACTGGAGCTTTCTAGGGCGGCTGCTGGAGAATGCTCAAGAACTCCACTGTG
ATTGGAAGGTTTGGCTGACTGTCTGTTATCTTCCGCATCTTGGTGTGGGTGCAGCA
GCCGAAGAAGTTTGGGGTGATGAGCAGTCTGATTTACCTGTAAACACGACGAGCCTGGT
TGTGAGAACGTTGCTATGATGAAGCCTTCCCTATCTCCACATTGCTTCTGGGTGCTG
CAGATCATTTTGTCTCCACTCCAACATCATCTACCTGGGCCACGTCTTCAATTGTC
CGCATGGAGGAGAAGAGGAGAGAGAGGGAGGAGACCTCCGAAAGGCAGGACGGCACCAG
GAGGACCATGATCCTCTCTATCATAATGGAGTTAGCAATGGAGGAAGCAGAGGTGGTGGC
AAAAAAGAAAAGCCACCTATTCGTGATGAACACGGGAAGATTTCGATCCGTGGGGCGTTA
CTGAGGACCTACATCTTTAACATCATCTTCAAGACTCTGTTTGAAGTGGGTTTCATCCTG
GGGCACTACTTCTCTATGGCTTCCATCTGAGGCCGCTCTACAAATGTGGCCGCTGGCCC
TGCCCAAACACCGTGGACTGTTTCATCTCCAGGCCACTGAAAAGACGATCTTCATCATC
TTTATGCTGGTGGTGCATGCATCTCGTTGGCCCTCAACCTGTTGGAGATCTACCACCTG
GGATGGAAAGGTCAAGCAGGGAGTCAACCAATGAGTTTGTCCCGATGGCGAGTTGCTG
CCGCAGAGTGCAGACGACACAGAGACATGGAGAAGATCCACGAGCAGACTTCTCCATCG
GCATTTGAATGTTTGTGACATATTCCAGCATGAATATGGCAGGACGCGTAGCTGAAGAA
GGAGGAACCTACAGTCCACCCGAGGCCCTCTCTAGCAGTAATGTCTTACCTACCGGTCTC
AAGATGGACGGCACAGTGTTCACCCAGATGACTTCTGTTGGAGGCACTGCCTCCTTCT
TTTTGCAGCAGTAATGACAAAGTGAAGCCATGGGCGACTAACAGAAGTGGAGCAAACTGG
AGCAACATGGCACTGGAGCTCCGCACT
>Fr124726 Our suggested modification.
ATGGGTGACTGGAGCTTTCTAGGGCGGCTGCTGGAGAATGCTCAAGAACTCCACTGTG
ATTGGAAGGTTTGGCTGACTGTCTGTTATCTTCCGCATCTTGGTGTGGGTGCAGCA
GCCGAAGAAGTTTGGGGTGATGAGCAGTCTGATTTACCTGTAAACACGACGAGCCTGGT
TGTGAGAACGTTGCTATGATGAAGCCTTCCCTATCTCCACATTGCTTCTGGGTGCTG
CAGATCATTTTGTCTCCACTCCAACATCATCTACCTGGGCCACGTCTTCAATTGTC
CGCATGGAGGAGAAGAGGAGAGAGAGGGAGGAGACCTCCGAAAGGCAGGACGGCACCAG
GAGGACCATGATCCTCTCTATCATAATGGAGTTAGCAATGGAGGAAGCAGAGGTGGTGGC
AAAAAAGAAAAGCCACCTATTCGTGATGAACACGGGAAGATTTCGATCCGTGGGGCGTTA
CTGAGGACCTACATCTTTAACATCATCTTCAAGACTCTGTTTGAAGTGGGTTTCATCCTG
GGGCACTACTTCTCTATGGCTTCCATCTGAGGCCGCTCTACAAATGTGGCCGCTGGCCC
TGCCCAAACACCGTGGACTGTTTCATCTCCAGGCCACTGAAAAGACGATCTTCATCATC
TTTATGCTGGTGGTGCATGCATCTCGTTGGCCCTCAACCTGTTGGAGATCTACCACCTG
GGATGGAAAGGTCAAGCAGGGAGTCAACCAATGAGTTTGTCCCGATGGCGAGTTGCTG
CCGCAGAGTGCAGACGACACAGAGACATGGAGAAGATCCACGAGCAGACTTCTCCATCG
GCATTTGAATGTTTGTGACATATTCCAGCATGAATATGGCAGGACGCGTAGCTGAAGAA
GGAGGAACCTACAGTCCACCCGAGGCCCTCTCTAGCAGTAATGTCTTACCTACCGGTCTC
AAGATGGACGGCACAGTGTTCACCCAGATGACTTCTGTTGGAGGCACTGCCTCCTTCT
TTTTGCAGCAGTAATGACAAAGTGAAGCCATGGGCGACTAACAGAAGTGGAGCAAACTGG
AGCAACATGGCACTGGAGCTCCGCACT
ATGGGTGACTGGAGCTTTCTAGGGCGGCTGCTGGAGAATGCTCAAGAACTCCACTGTG
ATTGGAAGGTTTGGCTGACTGTCTGTTATCTTCCGCATCTTGGTGTGGGTGCAGCA
GCCGAAGAAGTTTGGGGTGATGAGCAGTCTGATTTACCTGTAAACACGACGAGCCTGGT
TGTGAGAACGTTGCTATGATGAAGCCTTCCCTATCTCCACATTGCTTCTGGGTGCTG
CAGATCATTTTGTCTCCACTCCAACATCATCTACCTGGGCCACGTCTTCAATTGTC
CGCATGGAGGAGAAGAGGAGAGAGAGGGAGGAGGACCTCCGAAAGGCAGGACGGCACCAG
GAGGACCATGATCCTCTCTATCATAATGGAGTTAGCAATGGAGGAAGCAGAGGTGGTGGC
AAAAAAGAAAAGCCACCTATTCGTGATGAACACGGGAAGATTTCGATCCGTGGGGCGTTA
CTGAGGACCTACATCTTTAACATCATCTTCAAGACTCTGTTTGAAGTGGGTTTCATCCTG
GGGCACTACTTCTCTATGGCTTCCATCTGAGGCCGCTCTACAAATGTGGCCGCTGGCCC
TGCCCAAACACCGTGGACTGTTTCATCTCCAGGCCACTGAAAAGACGATCTTCATCATC
TTTATGCTGGTGGTGCATGCATCTCGTTGGCCCTCAACCTGTTGGAGATCTACCACCTG
GGATGGAAAGGTCAAGCAGGGAGTCAACCAATGAGTTTGTCCCGATGGCGAGTTGCTG
CCGCAGAGTGCAGACGACACAGAGACATGGAGAAGATCCACGAGCAGACTTCTCCATCG

>Fr141787 Ensembl prediction. No modification.

```
>Frscf89 (previous id: Fr144151) not predicted in present genebuild.
```

>Fr160171 (previous id: Frscsf387) Ensembl prediction in 2 exons (black/blue letters). The predicted 96 nt intron is regarded as part of the reading frame (marked blue). Sequence extended by 531 nt in 3'-direction to reach stop codon (marked blue).

```
>Fr160171 (previous id: Frscaf387) Our suggested modification.
```

48

TGCTTCGTGTCTCGTCCCACTGAGAAGACTGTCTTCATCATCTTCATGCTGGCCGTGGCC
 TGTGTCTCTCTCTCTCAACTTTGTGGAGATTAGTCACTTGGGCTGAAGAAGATTTCGC
 TTCGTCTTTTCGCAAGCCCGTGCCTGCGCCCAAGGCGAGGGCTCGGCCCGCTTCCA
 GCACCAAGGAAGAGCTGCTTCCCTAGCTGTGCCCTCCCTGCAGAGAGTGAAAGGTTAC
 CGGCTGCTGGAGGAGGAGAAAGCTCCCAACTCATCTCTACCTCTGGCTGAGGTGGGC
 ATGGAGGCCGCGAGAGGACCCACCTTCCAGGGGCTAGAGGAGAAGTCCAGGAGGTG
 CTACCCATGGAGGACATCTCTAAGGTGTATGACGAGACTCTGCCCTCTACACCCAGACC
 ACTGAGACTGGGGGGTACATTACACGAAGAGGAAGAGGAAGAGGTAGAGGAAGAGCAG
 CCAGCAGAATCAGAGAAGGGAGAGGTGGTGGTGAAGAGGAGGAGGAGGAAAGTGGTGAAC
 GTGGAGGAGGTAGCTGAAGCGGAGGCCACGGATACGATAGAAGACACCAAGCTTGAGC
 CGACTGAGTAAAGCCAGCAGCAGGGCCAGGTCAGACGATCTCACAGTATGA

>Fr164448 (previous id: Frscaf797) Ensembl prediction in 2 exons (black/blue letters). Sequence extended by 6 nt in 5'-direction to reach initiation codon. Predicted 96 nt intron regarded as part of the reading frame, with the first 4 nt taken as a part of the first conserved domain (green), and the remaining 92 nt corresponding to a part of the cytoplasmic loop (marked blue). Sequence extended by 618 nt in 3'-direction to reach stop codon (marked blue).

GACTGGAACCTTCTCGGAGGGATTTTGGAGGAGGTGCACATTCCTCCACCATGGTGGGC
 AAGATCTGGCTCACCTTCTCTTCGTTTTCCGCATGCTGGTCTCGGAGTGGCGGCCGAA
 GACGTGTGGAACGACGACGAGCGCTGACTTCATTTGCAACACCGAGCAGCCGGGATGCAGG
 AACGTTTGCTACGACCTGGCTTTTCCCATCTCCCTCATCCGCTACTGGGTGCTGCAAGTT
 ATCTTCGTGTCCTCCCTCGCTGGTTTACATGGGCCACGCTCTGAAGAGGATTGAACGG
 GAGGTGAAGCAGCTCGACCAAGGCAAGCTGAACAAAGCTCCGCTCCGGGGATCCCTGTTG
 CGCACTTATGTGGCAGATGTGGTCACCCGGTCCGTTGTTGAAGTGGGCTTCATGACGGGC
 CAATATGTCCTTATGGGTTTACCTCTACCACTTTTCAAGTGCAGCGGGATCCTTGT
 CCTAATGCTGTGGCTGTACGTCTCCAGGCCGACGGAGAAAGGCTCTTCATGGTCTTC
 ATGCAATTCATCGCCGAATTTCCCTCTTCTTAACATTTTGGAGATGGCGTATCTTGGC
 TACAAGTGGATTAAACAGGGCATCTTGGATCTTTAC
 >Fr164448 (previous id: Frscaf797) Our suggested modification.
 ATGGGA GACTGGAACCTTCTCGGAGGGATTTTGGAGGAGGTGCACATTCCTCCACCATG
 GTGGGCAAGATCTGGCTCACCTTCTCTTCGTTTTCCGCATGCTGGTCTCGGAGTGGCG
 GCCGAAGACGTGTGGAACGACGACGAGCGCTGACTTCATTTGCAACACCGAGCAGCCGGGA
 TGCAGGAACGTTGCTACGACCTGGCTTTTCCCATCTCCCTCATCCGCTACTGGGTGCTG
 CAAATTATCTTCGTTGCTCTCCCTCGCTGGTTTACATGGGCCACGCTCTGTACAGACTG
 CGGGCCCTGGAGAAAGCAGGCCAGAGAAAGAAAGTCTGCTGAGGAAGGAGCTGGAGTTG
 GTGGACGTGGATCTGGCCGAAGCCAGGAAGAGGATTGAACGGGAGGTGAAGCAGCTCGAC
 CAGGGCAAGCTGAACAAAGCTCCGCTCCGGGGATCCCTGTTGCGCACTTATGTGGCAGAT
 GTGGTCAACCGGTCGTTGTTGAAGTGGGCTTCATGACGGGCCAATATGTCCTTATGGG
 TTTACCTCTACCCACTTTTCAAGTGCAGCGGGATCCTTGTCTAATGCTGTGGACTGT
 TACGTCTCCAGGCCGACGGAGAAAGCGTCTTCATGGTCTTCATGCAATTCATCGCCGCA
 ATTTCCCTCTTCTTAACATTTTGGAGATGGCGTATCTTGGCTACAAAGTGGATTAAACAG
 GGCATCTTGGATCTTACCGCAATTACAGGATGAGCTCGATGATGACTTTATCTCTAAG
 GGGGAAAGGAATCTGTTGCGCAACTCTGCGCCAGTGCAGGCCGGAAGATGACGATTACA
 TTTTCCCAATTGATGGCAACCTAATGCAAGGAGCCGTTGGTCCAGCAGCGGCTTCTCT
 CTCCTGAGTGACCTGTCCAATCAACCACATCTGGGGGTTCCACATGTTTGGCCAAAGC
 CCAAAAGGGCGCAGACGATGCGAAGCTTCTCTCCCTCACAGCCAGGAAAAAGAG
 CGCAGTGACAGTGGCAGCCAGACTATCCAAAGTGCCAAACAGAAACAGCCCCCGCTG
 CCAGTGAGCGTACCGCGAGGCCCTTGGAGGGCTCATTCTTCAAATGCGCCACGGTGCCG
 GAGGGGAAGAGTTCTGACACGGATTCAAACGAGGAGCTGTGCGCTCAGACGTCGCCAAAC
 CAGCGCGTCCGCGATCTCAGCCGTAGCTCCAGCGGGAGTCTCTGCTGCTGCTCCAGCTCG
 GGCTGCGCGCACAGCCCCACGCTGCCCCCTCTTACTGCAAAACATCATCACCGAGCAAA
 AGCAGCAGCAGTCGGGAGCCAGACCTGCAAATTTAA

>Fr163704 (previous id: Fr139587) Ensembl prediction in 2 exons (black/blue letters). Sequence extended 6 nt in 5'-direction to reach initiation codon. Predicted 36 nt intron regarded as part of reading frame (marked blue). Sequence extended by 774 nt in 3'-direction to reach stop codon (blue).

GACTGGAATTTCTTGGAGGAATCTTAGAGGAGGTGCATATCCACTCCACCATGGTGGGC
 AAGATCTGGCTGACCATCTTGTTCATATTTCCGATGTTAGTACTGGGAGTTGCTGCGGAG
 GATGTGTGGAATGACGAACAGTCTGATTTTCATCTGCAACACCGACCGCTGGTTGTCGA
 AATGTCTGTTATGACAGGCTTTCCCATCTCCCTCATTGATGACTGGGTGCTTCAGGTG
 ATTTTCGTGTCCTCCCTCTTGGTCTACATGGGCCATGCCATTTATCAACTGCGAGCT
 CTGGAGAAGGAACGCCACTTGGATGTGGAAATTGGTGGAGGTAAAGGAAGAGATTGAAAAA
 GAGATGAGGCAGCTAGAGCAGGGCAAACTCAACAAGGCACCACTGAGAGGGTCTCTATTG
 TGTACTTATGTGGCCACATTGTGACTCGCTCGTTGGTAGAGGTGAGCTTCATGATGGGT
 CAGTACATCTTGTATGGACACCACTGAAACCTCTTACAAGTGTGAGCGAGAGCCGTGC
 CCAAAATGTGGTGGCTGCTTGTGTCCAGACCCACAGAGAAACAGTTTATGATGTTT
 ATGCAAGCCATTGCTGCTCTCACTCTTCTCAGTCTCTTGAGATTATCCACTGGGA
 TTTAAG

>Fr163704 (previous id: Fr139587) Our suggested modification.
 ATGGGA GACTGGAATTTCTTGGAGGAATCTTAGAGGAGGTGCATATCCACTCCACCATG
 GTGGGCAAGATCTGGCTGACCATCTGTTTCATATTTCCGATGTTAGTACTGGGAGTTGCT
 GCGGAGGATGTGTGGAATGACGAACAGTCTGATTTTCATCTGCAACACCGACCGCTGGT
 TGTGCAAAATGCTGTTATGACAGGCTTTCCCATCTCCCTCATTGATGACTGGGTGCTT
 CAGGTGATTTTCGTGTCCTCCCTCCTTGGTCTACATGGGCCATGCCATTTATCAACTG
 CGAGCTCTGGAGAAGGAACGCCACTGTAAGAAGGTGGCATTACGCCGGGAGATGGAAGCA

GTGGATGTGGAATTGGTGGAGGTAAAGGAAGAGAATTGAAAAAGAGATGAGGCAGCTAGAG
 CAGGGCAAACCTCAACAAGGCACCACTGAGAGGGTCTCTATTGTGTAATATATGTGGCCAC
 ATTGTGACTCGCTCGTTGGTAGAGGTCAGCTTCATGATGGGTCACTACATCTTGTATGGA
 CACCACCTGAAACCTCTTTACAAGTGTGAGCGAGAGCCGTGCCAAATGTGGTGGACTGC
 TTTGTGTCAGAGCCACAGAGAAACAGTTTTTATGATGTTTATGCAAGCCATTGCCTGC
 CTCTCACTCTTTCTCAGTCTTCTTGAGATTATCCACCTGGGATTAAAGAGCTTAAGAAG
 TGTATCTTGAACCTCTTCCACACCTGAAAGATGATCCTGATGAATTTTACATTAGCAAG
 TCAAAAAAGAACTCAGTCGTGCATCAGGTGTGTGCTGGAACATCTGTAGCTGGAAGACA
 ACTATTTCCACAGCGCCATGTGGATACACGTTGCTGATGGAGAAGCAGGGCAATGGGCC
 AACTACTCGCTTCTCAATGCCTCCTCTGCTTTCATTCCAATACAAGGGGACCTGGTGCA
 AAGTCAGATCGGCGTAAGGATGGCAAGGAGGGAATTCCAAGTCTACAGAACAAAACAGT
 AATTCCAACAACACCAGCAGCGACACACATTCTTCTGTGGATAAACATGAGGAGCCA
 GAGGAGCCCTTGGTGACCTCTGAATATCTACGCTCCCTGTTGCCGACGCCACCTCCTGC
 CCAACCTGTGAGGCATTACCAGGAAGTCAGGAGGATCAGCCACCTTGAAGTGTCTCC
 ACTCTACCAAGAGGGAATGGCTCAGACAGTGGGATTCTACCTGGGGGCAACAGCATC
 AAGCAACGCAGCAGCTGTGTTGGGCCCCGTGCAAGGATTCTCTCAAATCAGACACTAAA
 AAGCCTGGCAGACCAAGCCCGGACTCAGCAGGTGAGCTGAGTTCCGCGTCTCGTCAC
 AGCAATGAGAGTAACAGCCCCACAGCTTACCCCCAACCCGACAGGTGTCAGCAGCAAGT
 AGTGCCAGCAGCCGGCGAGCTCCGACTGACCTACAGATATAA

>Fr149617 Ensembl prediction in 2 exons (black/blue letters). Sequence extended 6 nt in 5'-direction to reach initiation codon (marked green). Predicted intron of 72 nt regarded as part of reading frame (marked blue). Sequence extended by 708 nt in 3'-direction to reach a stop codon (marked blue).

GACTGGAACTTATTAGGAAGTATTTAGAAAGTCCACATTCATTCCACCATTTGTGGGC
 AAAATCTGGCTCACCATCCTCTTCATCTTCCGGATGCTTGTGCTTGGGGTTGCCGCCGAG
 GATGTTTGGAGAGCTGCTCAGCATGTAATTTGTTTGAACACGGAGCAACCTGGGTGCAAG
 AACGTCTGCTACGACAGGCTTTCCCGTCTCCCTGATCCGTTATTGGGTCTGCAGATT
 ATTTTGTATCCTCTCCATCACTGGTCTACATGGGACATGCATGTATCGCCTGAGGACC
 CTTGAGAAAGAGCGGCACAGGAGGAAAGTCTGCCTGAAAGCTGAGCTGGAGGTGAGGAAG
 GCCCTCTAAGAGGCTCCTTGCTTCGCACATATGTTCTCCATATCTTAACTAGGTCCGCA
 GTAGAGGTGGGTTTTATCGTAGGACAATGTGCTCTGTACGGCCTTGGACTGTCTCCCTTG
 TACAAATGTGCCAGACTGCCGTGTCCCAACAGCGTCGACTGTTTCGTCTCTCGGCCTACA
 GAAAGAACATTTTATGGTCTTCATGCTAGTCATTGCTGGTGTTCGTTGTTCTCAAC
 ATTCTGGAGATTTTTCATCTGGGTGTGAAGAGGATTAACAAAGTTTGTATGGATATAAA
 TACCGA

>Fr149617 Our suggested modification. Potential intron start is underlined.

ATGGGGGACTGGAACTTATTAGGAAGTATTTAGAAAGTCCACATTCATTCCACCATTT
 GTGGGCAAAATCTGGCTCACCATCCTCTTCATCTTCCGGATGCTTGTGCTTGGGGTTGCG
 GCCGAGGATGTTTGGGACGATGAGCAGAGTGAATTTGTTTGAACACGGAGCAACCTGGG
 TGCAAGAACGTCTGCTACGACAGGCTTTCCCGTCTCCCTGATCCGTTATTGGGTCTCTG
 CAGATTATTTTGTATCCTCTCCATCACTGGTCTACATGGGACATGCATGTATCGCCTG
 AGGACCCCTTGAGAAAGAGCGGCACAGGAGGAAAGTCTGCCTGAAAGCTGAGCTGGAGGT
 ACAGACCCCATCCAGGAGGATCACAAGAGGATCGAGCGAGAATCAGGAACTAGATGAA
 CAGAAGAGAGTGAGGAAGGCCCTCTAAGAGGCTCCTTGTTCGCACATATGTTCTCCAT
 ATCTTAACTAGGTCCGCAAGTAGAGGTGGGTTTTATCGTAGGACAATGTGCTCTGTACGGC
 CTTGGACTGTCTCCCTTGTACAAATGTGCCAGACTGCCGTGTCCCAACAGCGTCGACTGT
 TTCGTCTCTCGGCTACAGAAAGAACATTTTTCATGGTCTTCATGCTAGTCATTGCTGGT
 GTTTCGTTGTTCTCAACATTCTGGAGATTTTTCATCTGGGTGTGAAGAGGATTAACAA
 AGTTTGTATGGATATAAATACCGAGATGACGAGAGCGTGTACCGCTCAAAGAAGAACTCC
 ACGGTACAGCAAGTGTGTTGTTCTCACAATTCGTACCACAGAGGCTGGTGCACTCACA
 CAGATGACTTGTTCGGCTCTGCCTGACACTCACGGGGAGACTTGGCAATAAATGTGTCC
 CACCAGAACCAGGAAGGATCTGGCGCCACCAACCAGCATCCGTACACATTGGCATGCCT
 GCCCAGGGTCTTCATCAGGTGCCACCTGTGAGCAGCAATGCGCTGTAGGTACAAGGAAG
 CCGTCGTATAGCAGCGAAGAATCCAGCGAACCTCACGTGAGGCCACAATATGCAGGACCC
 AGAGCCACCCTCGTCGCCAGCCACATGGAGATCCAGCAGCCCTGAGGAACCCACAGAGG
 AAAATGAGCAGAGTAAGTGTATAAGGACCTTAGCGACATGAGTGACTCGGCAGAGAGC
 GAGCCCCACCCACGGTCCGGAAGTGCAGCTTTATGTCCGGGGTCTGTCCGATGGAAG
 CTGTCTCCCGTCCGACAGCACCGACGCCACAGCGGAAGTGTGCTGAAGCCAGCAT
 CTCACCAAGCCGAGGGTTCAAGTGGTGACCCCCCACCACCGGCCAGCGGAGGAGGATG
 TCCATGGTTAGTAGACAATTTTACAGTCCACAACAACTTCACAAACCTGATTCTGGT
 GTAGATAGTTAG

>Fr161672 Ensembl prediction in 2 exons (black/blue letters). Sequence extended 6 nt in 5'-direction to reach initiation codon (marked green). Predicted intron of 63 nt regarded as part of reading frame (marked blue). Sequence extended 714 nt to reach a stop codon (marked blue).

GACTGGAACTGCTTGGCAGCATCCTAGAAGAGGTTACATACATTCCACCATCGTGGGC
 AAAATATGGCTGAGCTCCTCTTCAATTTCCGATGCTGATATTGGGAGCAGCTGCTGAA
 GATGTGTGGGATGATGAGCTGTCTGAGTTCATCTGTAACACTGACCAACCAGGATGCAAA
 GCTGTCTGCTATGACCGTGCTTCCCTATCTCGCTTATTCGCTTCTGGGTCTGCAGGTT
 ATCTTTGTCTGACCCCTCTTGGTCTATATGGGCCATGCCCTTTATTGCATCCGAGCT
 CTTGAGAAAGAGCGCCACCGCAGGCAGAGAAATTGAGAAGGCTGGATGAACAGAGAAGAA
 GTGAAGAAGGCCCTCTTAAAGGTTCTTTGTTGAGAACTTACATTATCCATATCCTTACT
 CGCTCTGTGGTGGAAATCTGTTTCTTCTTGGCCAGTATTTCTCTACGGTGTCAATTG
 GACCCACTTTATAAGTGTGAGAGGATGCCCTGTCCCAACAGTGTAGACTGTTACATCTCT
 AGGCCACAGAGAAGAGCATTTTTCATGCTTTCATGATTGCCATTGCTGGTATTTCACTT
 TTACTCAACATTTTGAATATCACACCTAGGCATAAGGAAAATTAAAGGGTACTATAT

GGAGAG

>Fr161672 Our suggested modification. Position corresponding to intron start in other Cx62 orthologs is underlined

ATGGGGGACTGGAACCTGCTTGGCAGCATCCTAGAAGAGGTTACATACATTCCACCATC
GTGGGCAAAATATGGCTGACCATACTCTTCATTTCCGCATGCTGATATTGGGAGCAGCT
GCTGAAGATGTGTGGGATGATGAGCTGTCTGAGTTCATCTGTAACACTGACCAACCAGGA
TGCAAAGCTGTCTGCTATGACCGTGCCCTCCCTATCTCGCTTATTCGCTTCTGGGTCCCTG
CAGGTTATCTTTGTCTCTGCACCTCTTTGGTCTATATGGGCCATGCCCTTTATTGCATC
CGAGCTCTTGAGAAAGAGCGCCACCGCAGGCGTATCCAGCTAAAGGAGGAGTTGGATGAG
GCTGAATTAGCAATGGAGGAACAGAGGCGTGCAGAGAGAGAATTGAGAAGGCTGGATGAA
CAGAAGAAAGTGAAGAAGGCCCTCTTAAAGGTTCTTTGTTGAGAACTTACATTATCCAT
ATCCTTACTCGCTCTGTGGTGGAAATCTGTTTCCTTCTTGGCCAGTATTTCTCTACGGT
GTTCAATTGGACCCACTTTATAAGTGTGAGAGGATGCCCTGTCCCAACAGTGTAGACTGT
TACATCTCTAGGCCACAGAGAAGAGCATTTTCATGGTCTTCATGATTGCCATTGCTGGT
ATTTCACTTTTACTCAACATTTTTGAAATATCACACCTAGGCATAAGGAAAATTAAAGGG
ATACTATATGGAGAGCTATACAGAGACGATGACAGTTTGATTTTCAAGTCCAAGAAGAAA
GCCTCCTTACCACAACTTTGTGTCATTAGCAGTGTATCACCTACAATGGGCCTTTGACT
CAAACACTGAAAGTGATTCCAGAGGTAGCCATGAAATTTTCTATTGTAATGCTGGCCTT
AAATCCAGCCAGGACATACAAAGACCCAACCGTAGCCTGCAACCTAAGCCAACCTGGGTGT
GTTGAGAACTTACAAATCAAGCCCCCTAATCAGTGAGGAAATGAACACTTTCAAGGCA
GAAAACAACCCAGAGTGGATCTCTTCTTTTCAATTGGTGGAGGTAGTACAACAACCCAC
GAAGATGATACACGTGATGTAGGACTCCCTCTTTCTAACCATTGCGAAACTCAGTTATTA
CATAACATCCTAAGGTCAAGGGATGCACAAAAAGAGGAGCGAAAGGATTCAGTGATGAAT
GAAGTCCTCATACCAAACCCAGGAAGACGAGCTTTTTGAACAGACCACCATCAGAGAGC
TTGTCTTCTATCAGTAACCTACAAGTCCATCCTTACATACCTCAGAGGAATCTGATGAA
CTGGGCTCATTACAGGGAGACATGCCAATAATGCCACCGGCTGGCCGAAGAATGTCTATG
GCAAGTATAGCGTCATTGGTTTTCTCATGGGTTGAATAGCTCAAACACATTTTTCACATT
TTGGTATGA

Suppl. Figure 5 *Connexin* sequences from green spotted pufferfish (*Tetraodon nigroviridis*, Tn). The sequences are found in the Ensembl database. Although predicted as *connexins*, none of the *Tetraodon* sequences have formally been called by a *connexin* name, and we here use the abbreviated form of the Ensembl gene identification. For the ortholog of Cx25, the abbreviated form Tn35526001 should be read as GSTENG00035526001. Other information is given in short form in the title line where needed. If our interpretation of the most likely connexin sequence differs from the prediction found in Ensembl, the database sequence is given first, and immediately followed by our suggested sequence, which is the one used in our analyses. The conserved domains are marked in **yellow** (unmodified sequence) or **green** (modified sequence). Alternating black/blue sequences are exon predictions copied from Ensembl. Other colors are explained for the relevant sequences. The following predictions in Ensembl were not included in the sequences below:

Tn35460001: essentially identical to Tn35462001 except at 3'-end.

Tn01689001: essentially identical to Tn18746001.

Tn36946001: very similar to Tn13946001, but appears to contain insertion of a G at around pos 135. Runs into a row of unknowns at the border of TM2 and intracellular loop.

Tn26780001, Tn22191001: potential starts of cds (not counted as excluded because percentage of identity to other Tn connexins cannot be decided).

Tn03189001: incomplete, but large parts very similar to Tn01153001.

The sequences were updated as per July 1 2006. Please note that both the Ensembl gene identification number and the Ensembl gene prediction might change on future updates of the *Tetraodon* genome database.

>Tn35526001 Ensembl prediction. Sequence extended by 117 nt in 3'-direction to reach stop codon (marked blue).

ATGAAGCTGGAGCTTTCTGGAGAATCCTCAGTGGGGTGAACAAATACTCCACAGTCATT
GGGCGCGTCTGGCTCTTTGTGGTGTTCCTCTTCAGGATCTTGGGTACGTCGTAGCGGCC
GAACAAGTGTGGAAGGAAGAGACGAAGGAGTTCGTGTGCAACACCCGTCAGCCCGGCTGC
GAGACCACCTGCTTAACCACTTCTTCCCTGTTTCGACGGCGCGTCTCTGGGCCATGCAA
CTCATCCTGGTGTCTACCCCATCTCTGCTAGTGGCCCTGCATGTGGCATATAGGGAACAC
CGAGAGGCCAAGCACAAGAAAAAAGTGTACCAGGACAAGGCGACCATTTACGGGGGATTG
CTTTTCACTTACATCGCCAGTCTCATTTTAAAGACGGCGTTTCGAGGTGGGCACCTGCTG
GTGTTCTACTACGTCTACAACGGTTTTGAGCTTCCGGCGCTGCTCCGCTGCGACCAAGT
CCCTGCCCAAACGTAGTGGACTGTTTCATCGGCAAGGCCACCGAAAAAGAGATTTTCTC
TACATCATGGCCTGCACATCTATCCTCTGCATCGTTTTGAATGTCGTTGAGCTTATTTAC
ATCGTATGGAAGCAAGTCGTTAAATACGTCATCCAGCATTACAAT

>Tn35526001 Our suggested modification.

ATGAAGCTGGAGCTTTCTGGAGAATCCTCAGTGGGGTGAACAAATACTCCACAGTCATT
GGGCGCGTCTGGCTCTTTGTGGTGTTCCTCTTCAGGATCTTGGGTACGTCGTAGCGGCC
GAACAAGTGTGGAAGGAAGAGACGAAGGAGTTCGTGTGCAACACCCGTCAGCCCGGCTGC
GAGACCACCTGCTTAACCACTTCTTCCCTGTTTCGACGGCGCGTCTCTGGGCCATGCAA
CTCATCCTGGTGTCTACCCCATCTCTGCTAGTGGCCCTGCATGTGGCATATAGGGAACAC
CGAGAGGCCAAGCACAAGAAAAAAGTGTACCAGGACAAGGCGACCATTTACGGGGGATTG
CTTTTCACTTACATCGCCAGTCTCATTTTAAAGACGGCGTTTCGAGGTGGGCACCTGCTG
GTGTTCTACTACGTCTACAACGGTTTTGAGCTTCCGGCGCTGCTCCGCTGCGACCAAGT
CCCTGCCCAAACGTAGTGGACTGTTTCATCGGCAAGGCCACCGAAAAAGAGATTTTCTC
TACATCATGGCCTGCACATCTATCCTCTGCATCGTTTTGAATGTCGTTGAGCTTATTTAC
ATCGTATGGAAGCAAGTCGTTAAATACGTCATCCAGCATTACAATCCTGTGGAGAAGAGA
CCTCCCTCTGGCAACAGACACAAGTGTCCAACGTCAATGGATATGCTCTGTTGACCAC
GTTGGAATGAAGAAGACAGCCCAAAAAGACTAAACCATTAG

>Tn29165001 Ensembl prediction in 2 exons (black/blue letters). By extending the first exon with 71 nt into the predicted intron (marked blue), a stop codon is reached. The second exon (blue letters) is discarded.

ATGTCTTGGGCCGCACTTATTAGTAAGCTGGGTGGTGTCCACGAATACTCCACAGCCTG
GGGAAGGTCTGGCTTTCTGTCTCTTTCATCTTCCGCATCGGTATTCTGTTGTGGCCACC
GAGAAAGTCTGGGGAGACGAACAGTCCAGCTTTACGTGCAACACACAGCAGCCGGGCTGC
AAAAACGTCTGCTACGACTGTTCTTCCGGTTTCACACATCCGCTTGTGGTGCCTGCAG
CTGATCTTTGTGTCGACCCCGGCCCTTCTGGTGACCATGTACGTGGCCTACAGAAAACAT
AAAGATGAGAAAAATTGTTTAGACTCCAAGGACAGTGAGACGGGGAAGGAGGAAGAGACG
GGAAAAAGGTTAAAAAGAGGAACAAAAAAGCGTCTGCCATCACAGGTCCTATGTGG
TGGATCTACACGACGAGCTTGTCTTCCAGACTGCTCTTGGAGGGGGGCTCATGTACGCT
CTGTACTTCACTATGATGGCTTCCAGATGCCACGTCTGCTCAAGTGTGAGCAGTGGCCT
TGCCCCAACAAGGTGAGTGTCTCGTCTCCAGGCCGACGGAGAAAACGGTCTTACCATC
TTCATGGTGGCCTCGTCCGTCATTTGTATGTTTCTTAATTTTGCCGAACATCATCTACCTA

ATTGGCAAGGCCCTCTTTAAGAGAAAGTGGACTGACCTCTCCGTGCCACAATGTCTTGGGC
CACACTTTACAGTCAGCTGGGCGGCGTCAACAAACACTCCACCAGCTGGGGAAGGTCTGG
CTGTCTGTGCTCTTCACTTCCGCGTCACTATTCTGGTTCTGGCCGCCGAGAAAGTCTGG
GGAGATGAACAGTCCGACTTTACGTGCAACACACAGCAGCCGGGCTGCAAAAACGTCTGC
TACGATCACTTCTTCCCGGTTTACACATCCGCTTGTGGTGCCGTCAGCTGATCTTGGTG
TCGACCCCGGCCCTTCTGGTGGCCATGTACGTGGCCTACAGAAAACGTGGAGATAAGAGA
ACTGTCCCTGGCCCTCCGGCGGGGATGAGAAGGTGAAGGAGGCAGACCTGCAGACCCTGAAG
ACGAAGCGTCTGCACATCACAGGCCCTCTGTGGTGGACCTACACCAGCAGCTTGTCTTC
AGACTGCTCTTTGAGGGGGGCTTACGTACGCTCTGTACTTCACTATGATGGCTTCCAG
ATGCCACGTCTGGTCAAGTGTGAGCAGTGGCCTTGCCCCAACAAAGGTGGACTGCTTCATC
TCCAGGCCGACGGGAGAAACGGTCTTACCATCTTATGGCGGCCCTCGTCCGCCCTTTGT
ATGGTTCTTAATATTGCTGAACCTCGTCTACCTGATTGTCAAGGCTCTCGTGGGTTATCA
GCCAGGTCCAAGCAGCGGAAACAGAGATACGCTCGGGAGAACTTCCACCGGGATCACATG
CTTCTGGAGAACAAAGAAGATGAGAACATGTTTCTCAGACCCAACAAGCAACAGAACC
ATGTGTTGA

>Tn29165001 Our suggested modification.

ATGTCTTGGGCGGCACCTATTAGTAAGCTGGGTGGTGTCCACGAATACTCCACCAGCCTG
GGGAAGGTCTGGCTTTCTGTCTCTTCTATCTTCCGCATCGGTATTCTGGTTGTGGCCACC
GAGAAAGTCTGGGGAGACGAACAGTCCAGCTTTACGTGCAACACACAGCAGCCGGGCTGC
AAAAACGTCTGCTACGATCACTTCTTCCCGGTTTACACATCCGCTTGTGGTGCCGTCAG
CTGATCTTTGTGTGACCCCGGCCCTTCTGGTGACCATGTACGTGGCCTACAGAAAACAT
AAAGATGAGAAAAATTGTTTAGACTCCAAGGACAGTGAGACGGGGAAGGAGGAAGAGACG
GGAAAAAAGGTTAAAAAAGGAGGAACAAAAAAGCGTCTGCCCATCACAGGTCTCTATGTGG
TGGATCTACACCAGCAGCTTGTCTTCCAGACTGCTCTTTGAGGGGGGCTTCTATGTACGCT
CTGTACTTCTATGATGGCTTCCAGATGCCACGTCTGCTCAAGTGTGAGCAGTGGCCT
TGCCCCAACAAAGGTGGACTGCTTCTGCTCCAGGCCGACGGAGAAAAAGGTTCTTACCATC
TTATGGTGGCCTCGTCCGTCATTGTTGATGGTTCTTAATTTTGCCGAACCTCATCTACCTA
ATTGGCAAGGCCCTCTTTAAGAGAAAGTAGTGGGGGAAAAAATGCACTATAGAATTAAAC
TCCAACCAAGAATACCATGCTTTTGGAGAAAAATTAA

>Tn29164001 Ensembl prediction as lambda crystallin. From this predicted sequence, we use only exon 9 (underlined), which corresponds to the second conserved domain with a few nt on either side (also underlined in lower sequence). Sequence extended 547 nt in 5'-direction to reach initiation codon (green corresponds to first conserved domain; blue corresponds to sequence in cytoplasmic loop). Sequence extended by 68 nt in 3'-direction to reach stop codon (marked blue).

ATGGCCACTATCTCCGTAGTCGGCAGTGGGCTTATAGGCCGTTCTTGGGCAATGGTGTTT
ATAAGTGGAGGTTACAATGTGAAGATCTATGACAACCAACCTGGACAGGCAGCCAAAGCC
ATCACCGAGATCAGGAAGCAGCTGGAGAGCTGGAGGAAGCTGACATGCTGAGAGGGGAA
CTGAAGGCCGCACAGCAATGTGCTCTCCTCAGCAGCCATGATGACCTGGCTCAGGCGCTG
GAAGGAGCCTTCTTTGTCCAGGAGTGTGTGTTGAGCAGCTTGAGGTCAAGCAGAACGTC
TTCCGGGACATAGAGGCCCTGGTGAGGGAAGATGTATCCTGAGCAGCTCCACTTCTTGT
CTGGTACCAAGCAACATCTTCTCAAAGTCCAGAACAGGATCGCTGCCTCGTCTCTCAC
CCGGTAAACCCACCTTACTATGTTAACTGGTGGAGTTGATACCTACCCAGAGACTGCA
GCCGTTGTTATGGACACCACCGAACCTGATGACCAAGGTCGAGCAGGTTCCCGTCTGT
CTGAGGAAGAGGTTGATGGTTTTGCTCTCAACAGGATACAGGCTGCTATCACTGCAAG
TCGTGGAGGCTGGTCCAGGATGGCATCATCTCTGTCCGGACATCGACCTGGTGATGTCA
GAGGGGCTCGGAATGCGTTACGCAATTCATCGGTCCCATGGAAACCATGCACCTCAATGCA
CCTAAAGGTTTAGAGGACTACATGAAGCGCTATGGAGAGGGTATACGGAGGGTGCTGGCT
TCTTTTGGGCCCTTACCAGAACTCTCTGGGGAAGGTGCCAAAAGCATCTTAAAGCCCTC
TGTGGTGGATCTACACAGCAGCTTGTCTTCCAGACTGCTCTTTGAGGGGGGCTTCTATGT
ACGCTCTGTACTTCTATGATGGCTTCCAGATGCCACGTCTGGTCAAGTGTGAGCAGT
GGCCTTGCCCCAACAAAGGTGGACTGCTTCTGCTCCAGGCCAACGGAGAAAACGGTCTTCA
CCATCTTCATGGTGGCCTCGTCCGTCATTTGTGTGGTTCTTATTTTGCCGAACCTTTTT
ACCTTATTGACAAGGCCCTCTTTAAGAGAAAGTGGACTGA

>Tn29164001 Our suggested modification.

ATGTCTTGGGCGGCACCTTTATAAGCGGCTGGGCGGCGTCAACAAACACTCCACCAGCCTG
GGGAAGGTCTGGCTTTCTGTCTCTTCTATCTTCCGCGTCTCTATTCTGGTTCTGGCCGCC
GAGAAAGTCTGGGGAGACGAACAGTCCGACTTTACGTGCAACACACAGCAGCCGGGCTGC
AAAAACGTCTGCTACGATCACTTCTTCCCGGTTTACACATCCGCTTGTGGTGCCGTCAG
CTGATCTTTGTGTGACCCCGGCCCTTCTGGTGACCATGTACGTGGCCTACAGAAAACAA
ACAGATTGGAAAAAAGGTTTAGCCTCCAAGGACAGTGAGACGGGGAAGGAGGAAAAAAGTG
GAGAAGGGCGGTGAGAAGAATGAGGAGGAAATACTCATATTGCTTGAAGAGGGGGAAAGT
AATTCAGTCCCCGAGAACAGTAAGAAAGTGAAGGAGGAAAAATACGACAGAGTTAAAGAAT
GATAAGGACAATGAGAAGAAAAAGGGGAAAAAAGAGGAACAAAAAATGAGCGTCTACCC
ATCACAGGCCCTCTGTGGTGGATCTACACCAGCAGCTTGTCTTCCAGACTGCTCTTTGAG
GGGGGCTTCATGTACGCTCTGTACTTCACTATGATGGCTTCCAGATGCCACGTCTGGTC
AAGTGTGAGCAGTGGCCTTGCCCCAACAAAGGTGGACTGCTTCTGCTCCAGGCCAACGGAG
AAAACGGTCTTACCATCTTCTATGGTGGCCTCGTCCGTCATTTGTGTGGTTCTTATTTTT
GCCGAACCTTTTTTACCTTATTGACAAGGCCCTCTTTAAGAGAAATACCATAGAATTAAAC
TCCAACCAAGAATACCATGCTTTTGGAGAAAAATAAAAAATAAGTTCTGTAA

>Tn29166001 Ensembl prediction. Sequence extended by 456 nt in 5'-direction to reach initiation codon (marked green and blue, where green corresponds to the conserved domains, and blue corresponds to cytoplasmic loop).

ATGTACGCTCTGTACTTCACTATGATGGCTTCCAGATGCCACGTCTGGTCAAGTGTGAG
CAGTGGCCTTGCCCCAACAAAGGTGGACTGCTTCTATCTCCAGGCCGACGGAGAAAACGGTC

TTCACCATCTTCATGGCGGCTCGTCCGCCCTTTGTATGGTTCTTAATATTGCTGAACTC
GTCTACCTGATTGTCAAGGCTCTCGTGGGTTATCAGCCAGGTCCAAGCAGCGGAAACAG
AGATACGCTCGGGAGAACTTCCACCGGATCACATGCTTCTGGAGAACAAGAAGATGAG
AACATGTTTTCTCAGACCCAACAAGCAACAGAACCATGTGTTGA

>Tn29166001 Our suggested modification.

ATGCTTTGGGCCACACTTTACAGTCAGCTGGCGGGCTCAACAAACACTCCACCAGNCTG
GGGAAGGTCTGGCTGTCTGTCTCTTCTTCCGCGTCACTATTCTGGTTCTGGCCGCC
GAGAAAGTCTGGGGAGATGAACAGTCCGACTTTACGTGCAACACACAGCAGCCGGGCTGC
AAAAACGTCTGCTACGATCACTTCTTCCCGGTTTACACATCCGCTTGTGGTGCCTGCAG
CTGATCTTGGTGTGACCCCGGCCCTTCTGGTGGCCATGTACGTGGCTACAGAAAACGT
GGAGATAAGAGAAGTGTCTGGCTCCGCGGGGATGAGAAAGTGAAGGAGGCAGACCTG
CAGACCCTGAAGACGAAGCGTCTGCACATCACAGGCCCTCTGTGGTGGACCTACACCAGC
AGCTTGTCTTACAGACTGCTCTTTGAGGGGGGCTTCATGTACGCTCTGTACTTCTAT
GATGGCTTCCAGATGCCACGTCTGGTCAAGTGTGAGCAGTGGCCTTGCCTCAACAAGGTG
GACTGCTTACATCCAGGCCGACGGAGAAAACGCTTTCACCATCTTCATGGCGGCTCG
TCCGCCCTTTGTATGGTTCTTAATATTGCTGAACTCGTCTACCTGATTGTCAAGGCTCTC
GTGCGGTTATCAGCCAGGTCCAAGCAGCGGAAACAGAGATACGCTCGGAGAACTTCCAC
CGGGATCACATGCTTCTGGAGAACAAGAAGATGAGAACATGTTTTCTCAGACCCAACA
AGCAACAGAACCATGTGTTGA

>Tn19592001 Ensembl prediction. No modification.

ATGCTTTGGGCTGCTCTGTACAGCCAGCTGGCTGGAGTAAACCGCCACTCCACCAGTCTG
GGGAAAGTCTGGCTCTCTGTGCTCTTTATTTTCCGAGTCATGGTTTTGGTTGTGGCTGCT
GAGAGCGTTTGGGGGATGAACAGTCCGACTTACCTGTAACACCCTACAGCCTGGCTGT
GAGAAGCTCTGCTACGATCAGTTCTTCCCTGTCTCCACATCCGGCTCTGGTGTCTTCA
CTGTCTTTGTTCCACCCAGCGCTCTGGTGGCGATGTACGTGGCTACCGGAACAC
GGCGATAAGAAAGCTCTACAGAATTCTGGAAGAGTTGGGATCTTGAACACGGAAGGT
CCAGAGGAGCAGCTGGAGAGCCTCAGGAGCAGGAGGCTGCCATATCTGGCGCTCTGG
TGGACGTACGCTTACGCTTCTGTTCAGGCTCCTGTTGAAGGAGTTTCATGTACGCC
CTCTACGTGATTTACGACGGTTTCCGGATGCCACGCTTGGTGCCTGCGACAGTGGCCG
TGCCCCAACCTAGTGGACTGTTTCATCTCACGGCCGACAGAGAAAACAGTCTTACCGTT
TTCATGGCCACCTCATCATCCATCTGCATGCTCCTCAAGTGGCAGAGCTCGCATATCTT
GTTGGCAAGGCTGTACAAAGCCACAGCAGGAGTAAAGACGTGAAGAAGAGAAAGTGA
AGAAGACTGGTGCTGTTTGACAGGAGAGCGGTAAGCAGGTTGGAGGCTGCCTGGACTAAT
GCTGCTGTAGGTGGGAGGAGGAATTAGGAGGAGGATTCATCTCCTTCCACCTGCTGG
ATTATGAGCTTCTCATGGTGCCTTACGTCCATGTGGGAAAGACTGCAACAGGAAAA
CTGGCATATTTCAAGATGATTATAGTCAGTATTGTTGAAAATTGTTGTAA

>Tn31361001 Ensembl prediction. No modification.

ATGAAGTGGGCATTCTCCAGGGCTCCTCAGCGGGGTGAACAAGTACTCCACCGCCTTC
GGCCGAGTGTGGCTCTCCATCGTCTTCTTCAAGGTCATGGTTCTGGTGGCGGCG
GAGAAGGTCTGGGGCGAGCAGAGCAAGGACTTCAAATGCAACACGGCTCAGCCTGGTTGC
CACAAAGTCTGCTACGACTACTTCTTCCCGTGTCCACATCCGGCTGTGGGCCCTGCAG
CTCATCTTCTGTCACCTGCCCTCTCTCCTGGTGGTGTGACAGTGGCGTACCGGAGGAC
AGGGAACGGAACACAGGCTAAATTCGGCGAAACTGCCACCTATTTACCAGAACACT
GGGAAGAAGCGCGGAGGCTGTGGTGGACCACGTCCTCACTTTGGTCTTCAAAATCGCC
GTAGACGCCGTCTCTACCTTCTCTACCATCTATGAGGCTACGACTTCCCTTG
CTCATCAAGTGGCAGCAAGAGCCCTGCCCAACATAGTGGACTGCTTACGCTCGCCCC
ACCGAGAAAGCGCATCTTACCATCTTATGGTGGTACACAGCCTGGTCTGCATCTTCTC
TCCCTCTCGAGATCCTCTACCTGGTGGGCAACGCTGCCACGAGTGTTCAGAGCGGTT
CACGACTCCACCGCATCTGACCGGCCATCTCCAGCGGAACCAACATGATGGAGTCC
CGGGTCGCAAAAACAGCCCGAAAGTCTGGCGCCCTGTACAACCTCTGCCAAAGCGGAC
AGCCGGACAGCCAGGACTCTGCCAGACTCTGAAGTTCACCTGA

>Tn31360001 Ensembl prediction. Sequence extended 201 nt in 3'-direction to reach stop codon (blue).

ATGGACTGGAAGACCTTCCAAGCCCTCCTCAGCGGGGTGAACAAGTACTCCACCGCGTTT
GGGAGGGTCTGGCTGTGCGTGGTGTTCGTGTTCAAGGTCATGGTGTACGTGGTGGCGGCG
GAACGCGTGTGGGGCGATGAGCAGAAAGACTTCGATTGCAACACCAAGCAGCCGGGCTGC
GCCAACGTCTGCTACGACTACTTCTTCCCATCTCCACATCCGCCTGTGGGCCCTGCAG
CTCATCTTCTGTCACCTGCCGTCTTCTGGTGGTGTGACAGTGGCGTACCGGGATGAG
CGCGAGCGCAAGTACAGGGCCAAGCACGGCGATGAAAGCAAGCTGTACAACAACACGGCC
AAGAAACACGGCGGCTGTGGTGGACCTATCTGCTGAGCCTGTTTGAAGACGGGCATC
GAGGTGGCTTCTCTACATCTCCACCTCGTCTACGACAGCTTCTACCTGCCGAGGCTG
GTCAAGTGGCAGGTGGCGCCCTGCCCAACAGGTGGACTGCTACATCGGACACCCCACT
GAGAAGAAGGTCTTACATATCTCATGGTTGGCGCCTCGGCCCTCTGCATCGTCTCAAC
ATTTGCGAGATCATTTATCTCATCTCAAGCGCATTGCGCGGTGCGCAACAAGCACAGG
AGGCAC

>Tn31360001 Our suggested modification.

ATGGACTGGAAGACCTTCCAAGCCCTCCTCAGCGGGGTGAACAAGTACTCCACCGCGTTT
GGGAGGGTCTGGCTGTGCGTGGTGTTCGTGTTCAAGGTCATGGTGTACGTGGTGGCGGCG
GAACGCGTGTGGGGCGATGAGCAGAAAGACTTCGATTGCAACACCAAGCAGCCGGGCTGC
GCCAACGTCTGCTACGACTACTTCTTCCCATCTCCACATCCGCCTGTGGGCCCTGCAG
CTCATCTTCTGTCACCTGCCGTCTTCTGGTGGTGTGACAGTGGCGTACCGGGATGAG
CGCGAGCGCAAGTACAGGGCCAAGCACGGCGATGAAAGCAAGCTGTACAACAACACGGCC
AAGAAACACGGCGGCTGTGGTGGACCTATCTGCTGAGCCTGTTTGAAGACGGGCATC
GAGGTGGCTTCTCTACATCTCCACCTCGTCTACGACAGCTTCTACCTGCCGAGGCTG

GTCAAGTGCAGGTGGCGCCCTGCCCAACAGGTGGACTGCTACATCGGACACCCCACT
 GAGAAGAAGGTCTTACATACTTCATGGTTGGCGCCTCGGCCCTCTGCATCGTCTCAAC
 ATTTGGAGATCATTTTATCTCATCTCCAGCGCATTGCGCGGTGCGCCAAACAGCACAGG
 AGGCACCATCGCAATCCGGCCGTACACCCTCCAGACGAAGAACAACATGGACGACCCC
 TTCAGCAACCAAGGATGGAGCCCAAGCCAGGGCTGAAGGAAAGGCCCGCTCCTTT
 AACACCGCCTCCAAGTCCCATATAACATGGACAGCTTCCGGATGGCCGACAAGATCCGA
 GCCTCCGCCCAATCTGTCTTCATGA

>Tn16997001 Ensembl prediction. Sequence extended 177 nt in 3'-direction to reach stop codon (blue).

ACTTTCCAGCCGGGTGACCAACATATGCTACGACCACATCTTCCCCATCTCCACATC
 CGCCTGTGGGCGTGCAGCTGATTTTCGTACCTGCCCGTCGTGATCGTGATGGCTCAC
 GTCAAATTCGCGAAGGAAAGACGCCAAGTACGTGGAGCAGCACCACGGCTCCACCTC
 TACAGCAACCTGGCAAGAAGAGAGGGCGGCTGTGGTGGACCTATCTCCTCAGTCTGATC
 CTCAAAGCCGGATTTGATGCTTCTTTTACCTCCTGTACAGAATCTATCACGGTTAT
 GATCTGCCAGATTATCGAAATGTTTCGTGGAACCGTGTCCCAACACGGTGGACTGCTTC
 ATCAGCCGTCCACGGAGAAAAAGATCTTCATGCTCTTCATGGTCTGCTCCAGCGCTCTT
 TGCATTTTCATGTGATCTGCGAAATGTTCTATCTCGTGGAAACGCATCGCCAAA

>Tn16997001ic Our suggested modification.

ACTTTCCAGCCGGGTGACCAACATATGCTACGACCACATCTTCCCCATCTCCACATC
 CGCCTGTGGGCGTGCAGCTGATTTTCGTACCTGCCCGTCGTGATCGTGATGGCTCAC
 GTCAAATTCGCGAAGGAAAGACGCCAAGTACGTGGAGCAGCACCACGGCTCCACCTC
 TACAGCAACCTGGCAAGAAGAGAGGGCGGCTGTGGTGGACCATTCTCCTCAGTCTGATC
 CTCAAAGCCGGATTTGATGCTTCTTTTACCTCCTGTACAGAATCTATCACGGTTAT
 GATCTGCCAGATTATCGAAATGTTTCGTGGAACCGTGTCCCAACACGGTGGACTGCTTC
 ATCAGCCGTCCACGGAGAAAAAGATCTTCATGCTCTTCATGGTCTGCTCCAGCGCTCTT
 TGCATTTTCATGTGATCTGCGAAATGTTCTATCTCGTGGAAACGCATCGCCAAACGG
 GTGCAGAGCCACCGGAGAACAGCAGATTCTGTTTGTGACCAGCACGAACGACCAAC
 ATGGTTCCACCCAGATCCAGTATCGGGAGACTGACCCACCCGACGGGAGTCAGCTG
 AGCCTGGGCAGAGGGACAAGGTGAGGAGGAGGCCGTGACAACGACCTGTAA

>Tn17530001 Ensembl prediction in 2 exons (black/blue letters). The first 51 5' nt of the second exon (blue letters) is removed (marked gray). Sequence extended by 195 nt in 3'-direction to reach stop codon (marked blue).

ATGAAGTGGTCTGCCCTGGAGGCCCTGCTCAGTGGGGTCAACAAGTATTCACCTGTGTTT
 GGACGCGTCTGGCTGTCCATGGTCTTCTCCGCGTGATGGTGTGTTGTGGTGGCGGCT
 CAGAGGGTGTGGGGTGAGAGAGCAAGGACTTTGTCTGCAACACAGCGCAGCCGGGCTGC
 AACACGCTGTGTACGACAGCATCTTCCCCATCTCACACATCCGCCTGTGGGCCCTGCAG
 CTCATTTTCGTACGTGCCCGTGCCTGATGGTGGTGGGGCACGTCAAGTATCGGGAGAAG
 AAAGATTCCAGTACAGCACCTCACACCACGGGAAGCATCTGTACGCCAATCCTGGAAAG
 AAGCGCGAGGGCTGTGGTGGACCTACCTGATTCCAGTTTACCCCCATATTTAATGAAA
 ATATCACTAATTTCTTCCCGGTGAGTCTGATTTTCAAGGCCAGCTTTGACGCTGGTTTC
 CTGTACATCCTCTACCACTACGATGGTTACGACATGCCCGCCTGTCCAAGTGTTCC
 CTGGAGCCGTGCCCAACACGGTGGACTGCTTCATATCCCGGCCACCGAGAAGAAGATC
 TTCACCTCTTCATGGTGGTCTCCTCTGCCATCTGCATCCTGATGTGCATCTGTGAGATG
 ATCTACCTCATCTGCAAGCGCATCAGT

>Tn17530001 Our suggested modification (This suggestion gives an unorthodox splice-acceptor sequence (end of intron), but gives a high identity to the other fish orthologs)

ATGAAGTGGTCTGCCCTGGAGGCCCTGCTCAGTGGGGTCAACAAGTATTCACCTGTGTTT
 GGACGCGTCTGGCTGTCCATGGTCTTCTCCGCGTGATGGTGTGTTGTGGTGGCGGCT
 CAGAGGGTGTGGGGTGAGAGCAAGGACTTTGTCTGCAACACAGCGCAGCCGGGCTGC
 AACACGCTGTGTACGACAGCATCTTCCCCATCTCACACATCCGCCTGTGGGCCCTGCAG
 CTCATTTTCGTACGTGCCCGTGCCTGATGGTGGTGGGGCACGTCAAGTATCGGGAGAAG
 AAAGATTCCAGTACAGCACCTCACACCACGGGAAGCATCTGTACGCCAATCCTGGAAAG
 AAGCGCGAGGGCTGTGGTGGACCACCTGGTGGTGGTGGGGCACGTCAAGTATCGGGAGAAG
 GCTGGTTTCTGTACATCTCTACCACTACGATGGTTACGACATGCCCGCCTGTCC
 AAGTGTCTCCCTGGAGCCGTGCCCAACACGGTGGACTGCTTCATATCCCGGCCACCGAG
 AAGAAGATCTTACCCTCTTCATGGTGGTCTCCTCTGCCATCTGCATCCTGATGTGCATC
 TGTGAGATGATCTACCTCATCTGCAAGCGCATCAGTAACTCATTAAAGAGAAGGAACGAG
 GCAGAGAGGAGGCTGTTTGTCTAGCAGCACGAGATGACGCCGCTGGCACCACCGAGGTCA
 GAGCTGAGGTCCAATCGCCGATCAGGGTGGATCCAACAGCCTCAATCCAAGATCTTGGC
 GCCATCGCTGAGGACAAGCAACCGCTGAGAAACAACAGGTACACAGCGTAG

>Tn24102001 Ensembl prediction. Sequence extended by 249 nt in 3'-direction to reach stop codon; the 3 first nt considered as a part of second conserved domain (marked green), remaining corresponding to cytoplasmic C-terminal tail (marked blue).

ATGGGGGAGTGGGGCTTCTCGGTGGACTCTTCGACAGCCTCCAGGCTCACTCGCCCATG
 CTGGGTGCGTTCTGCTCATGCTCATCTTTTCGGATAGTGATCCTCGGAACGTG
 GCCAGCGACCTGTTTGAGGACGAGCAGGAGGAGTTGCTGCAACACTCTCCAGCCGGC
 TGCAAGCAGGTGTGTACGACATGGCCTTCCCCATCTCGCAGTACAGATTCTGGGTCTTT
 CACATCGTGTCTATCGCCACGCCCTTCGCTACTTCTCCTGTTTACACCATGCATCATCAC
 AACAAAGACAAGTGGCAATTCATCCAGGTACAGGGAAGACGTGCGTTTGAGGAGGCTC
 TATATTCTCAACGTGGTGTTCGGATTCTGGCCGAGGCCGCCCTTCTGGTGGGCCAGTGG
 CTGCTCTATGGCTTCAAGGTGGAGGCCAGTTCCTCTGACGCCCTTCCCTGCCCTAC
 ACCGTGGACTGCTTACCTCGCTCTGCGAGAGAAGACCATCTTCTCTGCTTCTACTTC
 GTCATCGGGGCAATAGCCGCCCTCTTACGCTGTGCGGAGCTCTTCACATCTCT

>Tn24102001 Our suggested modification.

ATGGGGGAGTGGGGCTTCTCGGTGGACTCTTCGACAGCCTCCAGGCTCACTCGCCCATG
 CTGGGTCGGTTCTGGCTCCTGCTCATGCTCATCTTTCGGATAGTGATCCTCGGAACCTGTG
 GCCAGCGACCTGTTTGAGGACGAGCAGGAGGAGTTTGCCTGCAACACTCTCCAGCCGGGC
 TGCAAGCAGGTGTGTTACGACATGGCCTTCCCCATCTCGCAGTACAGATTCTGGGTCTTT
 CACATCGTGCTCATCGCCACGCCCTTCGCTACTCTTCTCGTTCACACCATGCATCATCAC
 AACAAAGCAACTGCAAAATTCATCCCAGGTACAGGGAAGACGTGCGTTTGAGGAGGCTC
 TATATTCTCAACGTGGTGTTCGGATTCTGGCCGAGGCCGCCCTTCTGGTGGGCCAGTGG
 CTGCTCTATGGCTTCAAGGTGGAGGCCAGTTCCTCGAGCCGCTTCCCTGCCCTTAC
 ACCGTGGACTGCTTCACCTCGCGTCTCGCAGAGAAGACCATCTTCTCTGCTTCTACTTC
 GTCATCGGGGCCATAGCCGCCCTCTTCAGCTGTGCGGAGCTCTTCCACATCTCTGTAAG
 TGGTTCTGCGCCGCCGCCGAGCCCTCGAAGACAGAGGACTCGGGCATCAGCGATAACCTT
 CTCAACGTGAAGCAGGAGGAAGGGATCAAAAAGGAGAAGCAGCAGGAGAAGAAGCAGCAA
 GCTCCGACAGCAGAAGGCTGAAGAGAGGATCGGTGAGGAGCAGCTCCAGCAGGAAGAGC
 TCCGGCGGCTCCACAGGCACATCAGTGGCAAATACGTGAGCAGCAGGACTTTGATGGTG
 TGA

>Tn17091001 Ensembl prediction in 6 exons (alternating black/blue letters). The 2046 5' nt removed (5 first exons) to reach an initiation codon corresponding to other connexins (marked gray).

ATGGCCGAGATGACTGAAAAAGGAAGGAGCTGTCTTGGAACTACAGCCTCCATTCT
 CTCCCTCTGGATCTGGAGCTAAGGCTGCGAAGGCTGGACCTGTCCAATAACCTCATCAGA
 CAGCTGCACACACTGTGCTTTCGCCGCACTGAAGCAGCTCAACCTGAGCTTCAACCAGCTG
 CATCTCATTTCTGAGAGCGCTTTTGAACCTGACGCAACTTGAGGAGTTGAATTTGTCC
 AGAAATGCGCTCAGCAACAACCTTGGCAACAACACAAGAGCCCTCCGATCCCTGGGCAGA
 CTGAAGACTTTGGACTTATCAGTGAACAGTCTGAACGACAAAGCAGCAGAGCTGTACCTT
 CAGAACAATCTGTTCTTGATCATCTCAAGATGACCGGTAATCTTTAACAAGACTTTCA
 AACAAGATGTTCTCAGCAGAGTAAACCTGAGGTCCATTAGTATCGATAACAACCTAATA
 TCAGTGATCCAAAAGGGGACATTTGAACCGCTGAGGCACCTGGAGAGTCTAACTTAGCT
 CAAAATAACCTGGTATACATCTGCGATTTTAACTCCAACAGGTTAAATATCTAAATCTC
 AGTAGAAATCTGTTGAGTCTTTGTTACACATGAAAAACAACAGTTATACATGCTGGAA
 ATTCTTGATCTGAGTCACAATAAGCTCTGTACTTTCCAATTATTCGAAGACAGAACCAT
 TTGAGGTATCTTTATTTGAGAATAACATGATTGGCACCTTTGAAACGGAGGCAGCGATG
 GTATCGGAGGTCAACGCCCTTTATCGTGACATCACGGGGGATAGAGCAGTGTCAAAGAAT
 AACTTCCATCCAGCTGGAGACAGATGCCAGTGGTTTTCTAGACCTGAGCTATAACCCAC
 TTCATCTCTTTCCCGTAGAGACTTTAAGCTTTCTCTCATCTTAGAGATGCTAAATTC
 AGCCACAACCTGTCTACGTAAGCTCTCGTGGAATGTGAGACAAGATAACCGAGGCCGTCGT
 CGGCAGCTTTTCTTCAATGCCTTGAAACCCCTTGATATGACAGCAATGGGCTGGTGCCT
 ATTTACCCCTCTTCTCAACGCGCTCAAGCAGATAGAAAGCTTAAATCTGCAGGACAAC
 GCACTGCAACCATGTGCGCTGGTAGTCGTTTGCAAACTCTTCTGCTCGCCTGCAGCCT
 CTCAACACTTCTGCGTGGTGTGTTGCAAAATTAAGACCCCTTAAAGACCTGAATCTAAGA
 GAAACACCATACAAATCTCTATCCAAACATATTTAAAGAACTCTTTGCTCGCTTTA
 AACCTTGACAGGAATTCGCATTTGACCATACAGGAAGGTGCTCTGGAGGGAGTGACAGAA
 ACCCTCCAGCTAGTGGTTATCAGCGAGCTAAACCTGAGCAGCAACATCTCCTTACCCTGC
 CTGCCAGCGCTAACTCATCTGAACATATCTCACAACGGCTTGGACTCCATACCCAGCTCC
 TTCAGCTGCTCCCTCTGAGAGAACTCGACATCAGGAACAACCACTTAGTGGCTCTGGAC
 CGTTCCCTGACCCGCGTTTGGGCCTCGCATCTCCGCATCATGCACATCAGCGGAACTAC
 TTCAACTGCTCGCAGCTAGTGGCTGACCGTCTCCAGAGGGGAAGCCAAAGCTAATG
 ACTGGGGTGACGCGCAGGTGAGGGCGTGGTCCGAAGGGGGCGCAGACGTGCTCGGCTTCAC
 CAGAGCCGGTGTCCGCCCTACGTTGGGGAACCTGGACCGCAACCTAAACCTGTGCCAGATG
 GTGATGCTGTGAATGGGATCAGTTGGTCCGAGGAGGCGGTTGAGTGGTTCCACACGATG
 GTGCGCAACAGGACTCTATGCTAGACTGCATCCAGGGGACGTGAGGTTACTGTGGAG
 CTGTTCTTGAAAAGGGAAAAATTTGATCTATGAGGAGGGGTGCATCGTGTCTGAGA
 CTCTCCAGAATGGACATGCAAGGTGTACGAATGGTAGCCTCGCACGAAGAGACCTGGAG
 TCTAAATGAACCTGGGGAACTTTTACGCCCTCATCAGCGGCTGAACAGGCACTCAACG
 GGCATCGGAAGGGTTGGCTCTCCGTCTCTTCCGATCCTGGTGTGGTGGTGGTGGTGGT
 GCTGCCGAGAGCGCTGAGGCGATGAGAAGTGGGCTTCACTGCAACACCCAGCAGCC
 GGCTGCAACAGTGTCTGTACGACCAAGTTCTTCCCATCTCGCACATCCGCTGTGGGCT
 CTGCAGCTGATCTTGGTCTCACCCCGGCCCTGCTGGTGGCCATGCACGTAGCCACAGA
 CGCCACATCGACAAGAAGATCCTGAAGAGGGCCGGCCGTGGCACACCCAAAGACCTGGAG
 CAGATAAAGAACCAGAGGTTCCAGATCACTGGAGCTCTGTGGTGGACGTACATGATCAGC
 ATCATCTTCAGGATCGTCTTTGAGGTGGCTTTCTCTACATCTTCTACCTGATTTATCCA
 GGTTTCAAAATGGTGCCTGCTGGTCAAGTGTGACTCTTACCCTTGCCCCAACACCGTGGAT
 TGTTTTGTGCTCAGACCACTGAGAAAACCATATTTACAGTGTTCATGCTGGGGTCTCG
 GGGGTGTGCTGCTTCTGAACCTGGCGGAGGTGGTCTACCTCATCGGCCAGGCCGTGCAGA
 CAGTGCATCCGAGGCTCTGAAGAACTCCAAAGTCCCTGGATCAGTCAAAATTTGTC
 TCTTACAGGCAAAATGAGATCAACGAACGATACTGGACCATCCCTCAGGTCAAAGTTC
 AGTGTGACCAAAAAGAAGCCAGCTGA

>Tn17091001 Our suggested modification.

ATGAACCTGGGGAACCTTTTACGCCCTCATCAGCGGCGTGAACAGGCACTCAACGGGCATC
 GGAAGGGTTTGGCTCTCCGTCTCTTCCGAATCCTGGTGTGGTGGTGGTGGTGGTGGC
 GAGAGCGTCTGGGGCGATGAGAAGTGGGCTTCACTGCAACACCCAGCAGCCGGCTGC
 AACAGTGTCTGCTACGACCAAGTTCTTCCCATCTCGCACATCCGCTGTGGGCTGTGAG
 CTGATCTTGGTCTCCACCCCGGCCCTGCTGGTGGCCATGCACGTAGCCACAGACGCCAC
 ATCGACAGAAGATCCTGAAGAGGGCCGGCCGTGGCACACCCAAAGACCTGGAGCAGATA
 AAGAACCAGAGGTTCCAGATCACTGGAGCTCTGTGGTGGACGTACATGATCAGCATCATC
 TTCAAGATCGTCTTGAAGTGGCTTTTCTCTACATCTTCTACCTGATTTATCCAGGTTTC
 AAAATGGTGCCTGTGGTCAAGTGTGACTCTTACCCTTGCCCCAACACCGTGGATTGTTTT

GTGTCCAGACCCACTGAGAAAACCATATTTACAGTGTTTCATGCTGGGGGTCTCGGGGGTG
 TGGGTGCTTCTGAACCTGGCGGAGGTGGTCTACCTCATCGGCAGGGCTGCAGACAGTGC
 ATCCGAGGCTCTGAAGAACTCCAAAGTCCCTTGGATCAGTCAAAAATTTGCTCTTAC
 AGGCAAAATGAGATCAACGAAGTACTGACCATCCCTCAGGTCAAAGTTCAGTGTG
 ACCAAAAAGAAGCCAGCTGA

>Tn22232001 Ensembl prediction. Sequence extended by 72 nt in 3'-direction to reach stop codon (marked blue).

ATGAAC TGGGCATCGTTCTACGCCGTCGTCAGCGGCGTGAACAGACACTCCACGGGCATC
 GGCCGCATCTGGCTCTCCGTGCTGTTTCATTTCCGCATCCTGGTCCTGGTGGTTGCTGCA
 GAGAGCGTGTGGGGAGACGAGAAGTCGGGCTTACCTGCAACACCCAGCAGCCGGGCTGC
 AACAGCGTCTGCTACGATCACTTCTTCCCATCTCCACATCCGCCTCTGGGCTCTCCAG
 CTCATCCTGGTCTCCACTCCGGCCCTGCTGGTGCCATGCACGTGGCTCACCGCCGCCAC
 ATCGACAAGAGGCTCTACAAGCTGTGAGGGCGCACCAATCCCAAAGACCTGGAGCAGATT
 AAGACCCAGAAAATGAAAATCACAGGCGCACTCTGGTGACATACGTATCAGCTGCTT
 TTCCGCGTTATCTTTGAGGTGACCTTTATGTACCTGTTCTACATGATCTATCCTGGTTAC
 AAGATGATCCGGCTGGTGAAGTGTGACTCGTACCCCTGTCCCAACACGGTGGACTGCTTT
 GTCTCCAGGCCACAGAGAAGACTGTTTTACCGTCTTCATGCTGGCTGTGTGTCAGGGGTT
 TGTATTTTGCTCAACATTGCGGAGGTGCTCTTCTTGGTGGGGAAGGCCCTGC

>Tn22232001 Our suggested modification.

ATGAAC TGGGCATCGTTCTACGCCGTCGTCAGCGGCGTGAACAGACACTCCACGGGCATC
 GGCCGCATCTGGCTCTCCGTGCTGTTTCATTTCCGCATCCTGGTCCTGGTGGTTGCTGCA
 GAGAGCGTGTGGGGAGACGAGAAGTCGGGCTTACCTGCAACACCCAGCAGCCGGGCTGC
 AACAGCGTCTGCTACGATCACTTCTTCCCATCTCCACATCCGCCTCTGGGCTCTCCAG
 CTCATCCTGGTCTCCACTCCGGCCCTGCTGGTGCCATGCACGTGGCTCACCGCCGCCAC
 ATCGACAAGAGGCTCTACAAGCTGTGAGGGCGCACCAATCCCAAAGACCTGGAGCAGATT
 AAGACCCAGAAAATGAAAATCACAGGCGCACTCTGGTGACATACGTATCAGCTGCTT
 TTCCGCGTTATCTTTGAGGTGACCTTTATGTACCTGTTCTACATGATCTATCCTGGTTAC
 AAGATGATCCGGCTGGTGAAGTGTGACTCGTACCCCTGTCCCAACACGGTGGACTGCTTT
 GTCTCCAGGCCACAGAGAAGACTGTTTTACCGTCTTCATGCTGGCTGTGTGTCAGGGGTT
 TGTATTTTGCTCAACATTGCGGAGGTGCTCTTCTTGGTGGGGAAGGCCCTGCGGTAGGCAT
 TTACAGCACGCTGGAGACTCCGCTGTGGGAACCTGGATCCAGCAAAAACCTCTGCTTCCTT
 TAG

>Tn32001001 Ensembl prediction. Sequence extended by 3 nt in 3'-direction to reach stop codon (marked blue).

ATGGGGGAATGGACTATACTAGAGAGGCTCCTGGAGGCTGCTGTCCAGCAGCACTCTACT
 ATGATAGGAAGGATCCTACTAACCCTGGTGGTCATCTTCCGGATTCTAATCGTGGCGATA
 GTTGGAGAGACTGTCTATGATGATGAGCAGACCATGTTTGTGTTGTAACACCTTACAGCCG
 GGCTGCAACCAGGCGTGTACGACAAGGCGTTCCTCATCTCGCACATTAGATACTGGGTG
 TTTAGATCATCATGGTGTGCACGCCGAGCCTGTGCTTCATCACTACTCGGTGCACCAG
 TCGGCCAAGCAGAAGGAGCGGCGCTACTCCACCGTCTATCTGACCCTCGATAAGGATCAA
 GATTCAGTGAACGCGCAGAGAGCAAAAAGATAAAGAACACCATTTGTGAACGGAGTACTT
 CAGAACACGGAGAAGTCCACCAAAGAAGCCGAACCGGACTGTTTGGAGGTGAAAGAGATC
 CCCAATTCCGGCCATGAGAACTGCAAAAGTCCAAATGAGGCGCCAGGAAGGCATCTCCAGG
 TTCTACATCATCCAGGTGGTCTTCAGAAACGCGCTGGAGATCGGCTTCCTGGTGGGCCAG
 TACTTTCTGTACGGATTCAACGTGCGGTGTACGAGTGCGACCGCTACCCCTGCATC
 AAAGACGTCGAGTGCTACGTCTCCAGACCCACGGAGAAGACGGTGTTCCTGGTGTTCATG
 TTCGCCGTGAGCGGCTTCTGCGTGGTGTGAACCTGGCCGAGCTCAATCACCTGGGCTGG
 AGGAAAATCAAGACGGCCGTGCGGGGCGTGCAGGCCCGGCGGAAGTCCATTTACGAGATC
 CGAAACAAGGACCTGCCGAGGATGAGCGTGCCCAACTTCGGACGCACTCAGTCCAGTGAC
 TCCGCGTACGTG

>Tn32001001 Our suggested modification.

ATGGGGGAATGGACTATACTAGAGAGGCTCCTGGAGGCTGCTGTCCAGCAGCACTCTACT
 ATGATAGGAAGGATCCTACTAACCCTGGTGGTCATCTTCCGGATTCTAATCGTGGCGATA
 GTTGGAGAGACTGTCTATGATGATGAGCAGACCATGTTTGTGTTGTAACACCTTACAGCCG
 GGCTGCAACCAGGCGTGTACGACAAGGCGTTCCTCATCTCGCACATTAGATACTGGGTG
 TTTAGATCATCATGGTGTGCACGCCGAGCCTGTGCTTCATCACTACTCGGTGCACCAG
 TCGGCCAAGCAGAAGGAGCGGCGCTACTCCACCGTCTATCTGACCCTCGATAAGGATCAA
 GATTCAGTGAACGCGCAGAGAGCAAAAAGATAAAGAACACCATTTGTGAACGGAGTACTT
 CAGAACACGGAGAAGTCCACCAAAGAAGCCGAACCGGACTGTTTGGAGGTGAAAGAGATC
 CCCAATTCCGGCCATGAGAACTGCAAAAGTCCAAATGAGGCGCCAGGAAGGCATCTCCAGG
 TTCTACATCATCCAGGTGGTCTTCAGAAACGCGCTGGAGATCGGCTTCCTGGTGGGCCAG
 TACTTTCTGTACGGATTCAACGTGCGGTGTACGAGTGCGACCGCTACCCCTGCATC
 AAAGACGTCGAGTGCTACGTCTCCAGACCCACGGAGAAGACGGTGTTCCTGGTGTTCATG
 TTCGCCGTGAGCGGCTTCTGCGTGGTGTGAACCTGGCCGAGCTCAATCACCTGGGCTGG
 AGGAAAATCAAGACGGCCGTGCGGGGCGTGCAGGCCCGGCGGAAGTCCATTTACGAGATC
 CGAAACAAGGACCTGCCGAGGATGAGCGTGCCCAACTTCGGACGCACTCAGTCCAGTGAC
 TCCGCGTACGTG

>Tn26779001 Ensembl prediction in 2 exons (black/blue letters). Exon 1 of Ensembl prediction does not fit with the strong degree of conservation in the other Cx36 orthologs, and is therefore removed (marked gray). Our suggested exon 1 is an upstream sequence more similar to the expected exon 1 (marked green). Sequence extended by 3 nt in 3'-direction to reach stop codon (blue).

ATGATGAACGGCGTTTGTGGGGACAGTGAAGCAGACAGGCTAGTTTGGATCCTGCTGACA
 GTGGTGGTGATCTCCGTATCCTGATCGTGGCGATCGTTGGGGAGACGGTGTACGAGGAT

GAGCAGACCATGTTTCATCTGCAACACTCTGCAACCAGGCTGCAACCAGGCCTGCTACGAC
AAGGCCCTCCCCATCTCCACATCCGCTACTGGGTCTTCCAGATCATCTGGTGTGCACT
CCCAGTCTCTGCTTTATCACTTACTCCGTCCACCAAGTCAAGTCAAGCAAAAGGACGTCGC
TACTCTTCTCTATCCATTATGAAAGGGACTACGGGGGAAGGGACGGGACACGAAAG
CTCCGCAACATCAATGGAATTCTAGTCCAACATGGCGGCGATGGCGGAGGAGGAAAGGAA
GAACCAGAGTGCCTGGAGGTGAAGGAGATCCCCAACGCCCCGCGGGGCTCACTCATGGC
AAGAGCTCCAAGGTTCCGCGCCAGGAAGGCATCTCCCGCTTCTACGTCAATCAAGTGGTC
TTCCGAAACGCCCTGGAGATCGGATTCTTGGCAGGCCAGTACTTCTCTACGGCTTCAGC
GTGCCTGGGATTTTCGAGTGCGACCGCTACCCGTGTCTGAAGGAGGTGGAGTGTACGTG
TCCCGGCCACCGAGAAGACGGTGTCTGGTGTTCATGTTTGGGTGAGCGGCATCTGC
GTGGTGTCTAAACCTGGCTGAGCTCAACCATCTGGGGTGGCGCAAGATCAAGGCGGCCATC
AGGGGTGTCCAGGCCCGCAGGAAGTCCATCTGCGAGATCCGGAAGAAGGACATGGCGCAC
CTCTCCAGCGCGCCAACTGGGCCGTACGCAGTCCAGCGAGTCGGCTACGTG

>Tn26779001 Our suggested modification.

ATGGGAGAATGGACCATCTAGAGCGCTCCTGGAGGCTGCGGTGACAGCATTCTACT
ATGATTGGGAGGATCCTGCTGACAGTGGTGGTGTCTTCCGTATCCTGATCGTGGCGATC
GTTGGGGAGACGGTGTAGCAGGATGAGCAGACCATGTTTCATCTGCAACACTCTGCAACCA
GGCTGCAACAGGCGCTGCTACGACAAGGCCCTCCCCATCTCCACATCCGCTACTGGGTG
TTCCAGATCATCCTGGTGTGCACTCCAGTCTCTGCTTTATCACTTACTCCGTCCACAG
TCAGCTAAGCAAAAGGACCGTCGCTACTCCTTCTCTATCCATTATGAAAGGGACTAC
GGGGGAAGGGACGGGACACGAAAGCTCCGCAACATCAATGGAATTCTAGTCCAACATGGC
GGCGATGGCGGAGGAGGAAAGGAAGAACAGACTGCCTGGAGGTGAAGGAGATCCCCAAC
GCCCGCGGGGCTCACTATGGCAAGAGCTCCAAGGTTCCGCCAGGAAGGCATCTCC
CGCTTCTACGTCAATCAAGTGGTCTTCCGAAACGCCCTGGAGATCGGATTCTTGGCAGGC
CAGTACTTCTCTACGGCTTACGCGTGCCTGGGATTTTCGAGTGCGACCGCTACCCGTGT
CTGAAGGAGGTGGAGTGTACGTGTCCCGGCCACCGAGAAGACGGTGTCTTGGTGTTC
ATGTTTGGCGGTGAGCGGCATCTGCGTGTGTCTAAACCTGGCTGAGCTCAACCATCTGGGG
TGGCGCAAGATCAAGGCGGCCATCAGGGGTGTCCAGGCCCGCAGGAAGTCCATCTGCGAG
ATCCGGAAGAAGGACATGGCGCACCTCTCCAGCCGCCAACCTGGGCCGTACGCAGTCC
AGCGAGTCGGCTACGTGTGA

>Tn05142001 Ensembl prediction. Sequence extended by 264 nt in 3'-direction to reach stop codon (marked blue).

ATGACGGAATGGACCCTGCTGAAACGCCCTCCTGGATGCCGTCCATCAGCACTCCACCATG
ATTGGCCGCTGTGGCTGACTGTGATGGTCATCTTTAGGCTGCTGGTAGTTGCTGTGGCA
ACGGAGGACGTGTACACGGACGAAACAGGAGATGTTTGTGTGCAACACGCTGCAGCCGGGG
TGCTCGACCGTCTGCTACGACGCTTTGCGCCCATCTCGCAGCCGCGCTTCTGGGTGTTT
CACATCATCAGCGTGTCCACGCCGTGCTGCTTTCATCATCTACACGTGGCACAACCTG
TCCAAGGTCCACAGCGCTGCTCAAAGGCACCCACGCGGCGCTGGCCAGCCGCTCGGCCAA
GACGTGCGCCAAAGCGGCAAGGAGGTCTTGGGAAAGGAGGCGGGGAAGGCAGGTGGCGAC
CGGGAGGTGTACCTCCAAGCTGCAGCTCGGACAGCTGCTCCGTCTCTCCACAAGCAC
CTCGGCCACAGCCTGGTGGACATTTTAGACGGCGTCCGCGCCCGTAGTTTGCGAAACGGA
GACCCGGCGACCTCCAGGCCATCCAGCTTACACCTTTAAAGACGGAAGCTCGGAGGGT
CTGGCGGTCTCTGGAGGAGTTCTGTCCAAATGCTACATCTTCCACGTGTGTCTACGCGCA
GCTCTGGAGGTGGGATTTGTGCGCGCCAGTGGAACCTGTTCCGATTGCAAGTGCCTGTC
CGCTTTTGTGTACCTCTCGCCCTGCAACAGCCCGTGGAGTGTACGTCTCCAGGCC
ACGGAGAAGACCATATTCTGATCTTCTATGTTTGTGTTGGCGTCTTCTGTATCTTCTC
AACCTGCTGGAATCAATCACCTGGGCTGGAAGAAGATCAGGCAGGCGGTGCGGCTG

>Tn05142001 Our suggested modification.

ATGACGGAATGGACCCTGCTGAAACGCCCTCCTGGATGCCGTCCATCAGCACTCCACCATG
ATTGGCCGCTGTGGCTGACTGTGATGGTCATCTTTAGGCTGCTGGTAGTTGCTGTGGCA
ACGGAGGACGTGTACACGGACGAAACAGGAGATGTTTGTGTGCAACACGCTGCAGCCGGGG
TGCTCGACCGTCTGCTACGACGCTTTGCGCCCATCTCGCAGCCGCGCTTCTGGGTGTTT
CACATCATCAGCGTGTCCACGCCGTGCTGCTTTCATCATCTACACGTGGCACAACCTG
TCCAAGGTCCACAGCGCTGCTCAAAGGCACCCACGCGGCGCTGGCCAGCCGCTCGGCCAA
GACGTGCGCCAAAGCGGCAAGGAGGTCTTGGGAAAGGAGGCGGGGAAGGCAGGTGGCGAC
CGGGAGGTGTACCTCCAAGCTGCAGCTCGGACAGCTGCTCCGTCTCTCCACAAGCAC
CTCGGCCACAGCCTGGTGGACATTTTAGACGGCGTCCGCGCCCGTAGTTTGCGAAACGGA
GACCCGGCGACCTCCAGGCCATCCAGCTTACACCTTTAAAGACGGAAGCTCGGAGGGT
CTGGCGGTCTCTGGAGGAGTTCTGTCCAAATGCTACATCTTCCACGTGTGTCTACGCGCA
GCTCTGGAGGTGGGATTTGTGCGCGCCAGTGGAACCTGTTCCGATTGCAAGTGCCTGTC
CGCTTTTGTGTACCTCTCGCCCTGCAACAGCCCGTGGAGTGTACGTCTCCAGGCC
ACGGAGAAGACCATATTCTGATCTTCTATGTTTGTGTTGGCGTCTTCTGTATCTTCTC
AACCTGCTGGAATCAATCACCTGGGCTGGAAGAAGATCAGGCAGGCGGTGCGGCTGAAG
GAGGACGAGGCGCCCTGGCAGGCTGCGCGGGGATAGGACGCGGATACGAGACCATCCCT
CCGGTCAGCCCTTTCGCCCAAGTCTTCCGGTATGAACGGCACCGCCCTCCGCCACTTTG
GACGTGGCGATGGGCCACAAACCGAGTGGGGTGCCTGGTCAACTGTGGCGGCACCCGA
GGGTGCGGAAAGGTCAAAGGAGAGAGGAAACACAAGGAGCTGAGAGGGTTCAAGCAGAGC
AGCGCAGAAGTCTGGATCTGA

>Tn35459001 Ensembl prediction in 3 exons (blue/black/blue letters). Predicted intron (between exon 1 and exon 2) of 29 nt included as a part of the reading frame (marked blue). Predicted intron (between exon 2 and exon 3) of 121 nt included as a part of the reading frame (marked blue).

ATGGGAGACTGGGGATTCTTGTGCTCCTTGTGGACAAAGTCCAGTCCCACTCCACCATC
ATCGGGAAGATCTGGATGAGCGTCTCTTCTGTTTCAGGATCATGGTCTGGGCGCCGGC
GCCGAGAGCGTCTGGGGCAGCAGCAGTCCGGTTCATCTGCAACACTCAGCAGCCGGT

TGCGAGAACGTCTGCTACGACTGGACCTTCCCCATTTCCACATCCGCTTCTGGGTCCTC
CAGATCATCTTCGTGTCCACGCCGACGCTGGTGTACCTGGGCCACGCCATGCATGTCATC
CACCAGGAGAATAAGCTGAGGGAGAAGCTGCAGAGCCCCGGCGGGAGCCGCTTCTCAAG
GTGCCCCAAGTACACCGACGAGAAGGGGAAGCTACCTGACCCAGCTGGTCTTCAAGATCCT
CATCGAGGCGGCCCTTATCGTGGGCCAGTACTACCTGTATGGCTTCATCATGGTGCCGAT
GTTCCCTTGCTCCAAGAAGCCCTGTCCCTTACCCTGGAGTGCTACATGTCCCGCCCCAC
CGAGAAGACCATCTTTCATCATCTTTCATGCTGGTGGTGGCTGCGTCTCCTGCTCCTCAA
CGTCATCGAGATGCTCTACCTCTGTGCACCAGGCTCAAATGCGCCTCCAGATCGCGGAC
GCAGAAGCTGACGGCGCCAAGGAGGAGATGAAGCTTCTGCATCACAACCTGA

>Tn35459001 Our suggested modification.

ATGGGAGACTGGGGATTCTTGTCTCCTTGCTGGACAAAGTCCAGTCCCACTCCACCATC
ATCGGGGAAGATCTGGATGAGCGTCTCTTCTGTTTCAAGATCATGGTCTGGGCGCCGGC
GCCGAGAGCGTCTGGGGCGACGAGCAGTCCGGGTTTCATCTGCAACACTCAGCAGCCCGGT
TGCGAGAACGTCTGCTACGACTGGACCTTCCCCATTTCCACATCCGCTTCTGGGTCCTC
CAGATCATCTTCTCGTGTCCACGCCGACGCTGGTGTACCTGGGCCACGCCATGCATGTCATC
CACCAGGAGAATAAGCTGAGGGAGAAGCTGCAGAGCCCCGGCGGGAGCCGCTTGTCTAAG
GTGCCCCAAGTACACCGACGAGAAGGGGAAGGTGAAGATCAAGGGGAACCTGCTGGGGAGC
TACCTGACCCAGCTGGTCTTCAAGATCTTCATCGAGGCGGCCCTTCATCGTGGGCCAGTAC
TACCTGTATGGCTTTCATGATGGTGGCGATGTTCCCTTGCTCCAAGAAGCCCTGTCCCTTC
ACCGTGGAGTGCTACATGTCCCGCCCCACGAGAAGACCATCTTCATCATCTTTCATGCTG
GTGGTGGGCTGCGTCTCCTCAACGTCATCGAGATGCTCTACCTCCTGTGCACC
AGGCTCAAATGCGCCTCCAGATCGCGGACGCGAGAAGCTGACGTGCGGCCAGAGCCCCGCC
GGCTGCTGGCCCCGAAATGGCCGACGCGGAGGACGCGCTCCATCAGAACCAGGATCAAC
CTGGAGCAGGAGCGCTGCCAGAGCGTCGGCGGGAGCCTGGACGCGCCAAGGAGGAGATG
AAGCTTCTGCATCACAACCTGA

>Tn35462001 Ensembl prediction. No modification.

ATGGGAGAGTGGGGTTTTCTGTCTCTCTGCTGGACAAAGTCCAGTCTCATTCTCCGTC
ATCGGGGAAGGTCTGGCTCAGCGTCTCTTCATCTTCAGGATCATGGTCTGGGAGCTGGA
GCTGATAAGGTTTGGGGCGACGAGCAGTCCAATATGATTTGTAACACCAAGCAGCCCGGC
TGTAAAAACGCTCTGCTACGACACGCGCTTCCCCATCTCGCACATTCGTTTCTGGGTCCTC
CAGATCATCTTCTGCTCCACGCCACGCTGGTGTACCTGGGACACGCTCTGCAGTGCATC
CACAAGGAAAAACAAGATGAGAGAGTACATGAAGACGCACACTCAGAGCAACCTGGCCAAA
TACCCCAAGTACACCGACGAGAAGGCCACGTTGGAGATCCGGGGCAACCTCCTGGGCACC
TACATGACTCTTCATCGTTTCCGCACTTCTGGAGATCGCCTTCATCGTGGGCCAGTAC
TACCTGTACGGCTTCATCATGGACCCCAAAGTGGTCTGCTCCCGGGCCCCCTGCCCTTC
ACCGTCAAGTGCTACATGTCTGTCGCCACCGAGAAGACCATCTTCATCTCTTCATGCTC
GCGGTCTCTTGCGCGTCTGCTGCTTAACTGGTGGAGATCTGCTACTTGGTGTGCTCC
CGCTCGAAGAAAAAGACCAAGACGCCCGCGCTCCGCGCTCGTCATTACCCACGGTTC
ACCAGCGAAAGCAAAAGCTGA

>Tn35458001 Ensembl prediction. No modification.

ATGGGGGAATGGGATTTGCTGGGCCGCTGCTGGATAAAGTGCAGACTCACTCCAGGTT
CTGGGCAAGATTTGGCTCAGCGTCTCTTCTGCTTCCGCACTCTGGTGTGCAGACGGCT
GCCGACAAGGTGTGGGGGACGAGCAATCGGACTTCATCTGCAACACTCAGCAGCCAGGC
TGCGAGAACGTCTGCTACGACCTGGCTTCCCCATCTCCACGTCCGCTTCTGGTTCCTT
CAGATCATTTGCCATAGCGACGCCCAAGCTGCTTACCTCGGCCACGCTCTTCAAGCCTGT
GTGTTCTCAGGAGTACAAAGTTCCCAAGTACGTCAAAGGCTCGGGCAAGATCAGATC
CGCGGGCGCCTCTCCGCACTTACACCTTCCACCTGATGGCCATGATCATCTCTGGAAGGC
GCCTTCATCGCCAGCCAGTACCTCTCTTGGCTTCTGCTGAGACGCGCTACGTGTGT
GAGCGCCACCCCTGCCCCCAAGGTTGACTGCTTCTGTCAGGCCACCGGAGAAGTCG
GTACATCATCTGGTTATGCTGGTGGCGGCCGTGGTCTCCCTGGCCCTCAGCTGGCCGAG
CTCTTCTACCTGGGCCTCAAAGCCACCAGGGAGTGCATGGCCAGGAGGAGGACTACAG
GTGACGCCCGTGACGCCGCCGTTTCGGGGAGAAAAGCCTTCAAATCTCCGATGAGATG
ATCCAGAACTGCATCAACCTGGAGCTGGAGCAGCTTAAAGAAAAGAGGTCCGGAGGGTC
GCCGGGGGGCCGAGGAGGTGCCAGTGTGCCCGGCCGCGCAACAGGAACAAGGGAGAG
GTCCACATCTGA

>Tn17534001 Ensembl prediction. G removed at pos 10 of predicted sequence to keep reading frame (marked gray; neighboring positions underlined in lower sequence). Sequence extended 142 nt in 5'-direction to reach initiation codon (marked green). Sequence extended 399 nt in 3'-direction to reach stop codon; 6 first nt considered a part of the second conserved domain (green), remaining corresponding to cytoplasmic C-terminal tail (marked blue).

ATGACGAGCGAAGCCAACCTTTGCTGCAACACGAATCAGCCTGGCTGCAACGAGCTGTGC
TACGACAAAGCCTTCCCCATCTCCCACTTCCGCTACTTTGTCTCCAGGTCATATTTGTG
TCCACGCCGACCATTTTCTACTTCGGATACGTGCTTTGAGAGTCAGAAGGATCAAAAAA
GACGCAGAGGAAGGTTTTGATAGAGGAAGTGTAAAAAGACAAACAGTCACTCAGAGGAA
GCGAGGAAAAAGCGGGAGGGCTGAAGAGGAAGCTCCCGAGGCACCTCGGCTGAAGGGCAGA
CTGCTGTGTGCATACGCCCTCAGCATCTTCTTAAAGGTCTCTGGAGGTTGGCTTCATG
TCAGGCCGTGATTTTCTACAAATGGCTTCTACATCGCAGCAAGTTCGAGTGTACAGG
AACCTTGTCTCACACGGTGGACTGCTTCTGCTCACGGCCACGGAGAAGACCATCTTC
GTGGTATACACTCAGGTCTGCTCCGGCGTCTCCCTGCTCTAACCTGCTGGAGCTGCTC
CACCTT

>Tn17534001 Our suggested modification.

ATGTCAAGAGGTGACTGGTCTTCTGGAGAACCTGCTGGAGGAGGGCCAGGAGTACTCG
ACAGGCATCGGCCGTGTCTGGCTACCGTGCTCTTCTCTTTTCGATGCTTTGTGCTGGGA
GCATCTGCAGAGTCGGCTGGGATGACGAGCAAGCCAACCTTTGTCTGCAACACGAATCAG

CCTGGCTGCACCAAGCGTGTGCTACGACAAAGCCTTCCCATCTCCACTTCCGCTACTTT
 GTCTCCAGGTCATATTTGTGTCCACGCCGACCATTTTCTACTTCGGATACGTCGCTTTG
 AGAGTCAGAAAGGATCAAAAAAGACGCAGAGGAAGGTTTTGATAGAGGAAGTGTAAAAAAG
 ACAAACAGTCACTCAGAGGAAGCGAGGAAAAGCGGGAGGGCTGAAGAGGAAGCTCCCGAG
 GCACCTCGGCTGAAGGGCAGACTGCTGTGTGCACTACGCCCTCAGCATCTTCTTAAAGGTC
 CTCTGGAGGTTGGCTTCATGTCAAGGCTGTATTTCTCTACAATGGCTTCTACATCGCA
 GCAAAGTTTCGAGTGTACAGGAACCCCTTGTCTCACACGGTGGACTGCTTCGTCTCACGG
 CCCACGGAGAAGACCATCTTCGTGGTATACACTCAGGTCGTCTCCGGCGTCTCCCTGCTC
 CTAACCTGCTGGAGCTGCTCCACCTTCTCCAGCTTGCCATCACTCACCAGGCTGGAGAAA
 CATTACCACGGTCGTCTGGAGACTACCTCCCCCGCAGAGCAGGGGACTGCGGAAGCT
 GCACGAATCCAAATGGAGGCTCGCAGTCCGGTAAGACAGGGAGCGACACTCACCTTCCA
 ACCCAGTGTGAGATGGAGGAGTCTGCCAATCCCTGCCAGAGTTTCGGAGAAGCAGGCATA
 GAACCAGGCATGAACCGCTCATCTGGAGAGACTGGGAACAGCCCTCCCCAGTTATGTG
 ACCTGCATCAAAGCCTCGAGGATGATGCATTACCCAGAGCCCATCAAAAAAACACACA
 GTCCACACCTCCAAAAACCAAGGCCGCTCAGAAGGGACATTCCAAACTCAGGCATTAC
 GTCTGA

>Tn13946001 Ensembl prediction in 2 exons (black/blue letters). Predicted 40 nt intron regarded as part of reading frame (marked green). 1 nt (A) removed from pos 252 of Ensembl predicted sequence (marked gray in upper sequence, neighboring positions underlined in lower sequence). 1 N added at pos 1121 of modified sequence to keep reading frame relative to Dr43 and Fr130494 (marked pink). Sequence extended by 32 nt in 3'-direction to reach stop codon (marked blue).

ATGGGTGACTGGAGTGTCTGGGTGCTGTGCTGGACAAGGTGCAGGCCTACTCCACCGCT
 GGAGGGAAGGTGTGGCTCTGTGCTCTTTCATCTTCCGGATCCTGGTGTGGGACTGCG
 GTGGAATCGGCGTGGGGGACGAGCAGTCTGCCCTCAAGTGCAACACGACGAGCCGGGC
 TGTGAGAACGTGTCTACGACAAATCATCTTCGTGTCCACACCGACCCCTCTGTACTTGG
 CTCATGTCTTCTACCTGAACAGGAAAGAACAGAAATTCAGCAAGATCGAGGAGGTGCTG
 AAGGCGGTCCAAACGATGGAGGCGACGTGGACGTCCCGCTGAAGAAAATTGAGATGAAG
 AAGCTGAAGTATGGCATTGAGGAGCACGGGAAGGTGAAGATGAAGGGAGCCCTGCTGAGA
 ACCTACATTGTGACGATCTTCTTCAAGTCGCTCTTGGAGTGGGCTTCTCGTGATCCAG
 TGGTACATGTACGGTTTCAGCCTGTCCGCCGTCTACACCTGTGAGCGGTCCCATGTCCA
 CACCGGTTGGACTGTTTCTGTCCCGTCCACCGAGAAGACAGTCTTCATCATTTTCATG
 CTGGTGGTGTGCTGGTGGTCTCCCTGCTGCTCAACATCATTGAGCTCTTCTACGTGCTCTT
 AAGAGGATCAAGGACCGGTGAAGGGCAAGCAGCAGCCGGCGCTCTACCCAGCGCCGGC
 ACCCTGAGCCCTGGGCCCCAAGGAGCTGTCCACCACCAAGTACGCCTACTACAAGGCTGC
 TCCTCACCCACCGCTCCGCTCTCACCCATGTCCCCCGGGCTACAAGACGGCCACGGGG
 GAGCGGGGACCGGCTCCGCTCCCGGAAGTACAACAAGCAGCCAGCAGCAGAACTGGGCC
 AACTACTCCACCGAGCAGAAGCGGCTGGGCCACACCGGCGCAGGAAGCACCATCTCCAAC
 TCCACGCCCAGGCTTCGACTTCCCGACGACACCCAGGAGCACAAGAAGATGCTCTCG
 CTGGCGGCCACGAGTGCAGCCGCTGGCGCTGCTGGATGCTCGGCCCTGCAGCCGCGCC
 ACAGCAGGCTGA

>Tn13946001 Our suggested modification.

ATGGGTGACTGGAGTGTCTGGGTGCTGTGCTGGACAAGGTGCAGGCCTACTCCACCGCT
 GGAGGGAAGGTGTGGCTCTGTGCTCTTTCATCTTCCGGATCCTGGTGTGGGACTGCG
 GTGGAATCGGCGTGGGGGACGAGCAGTCTGCCCTCAAGTGCAACACGACGAGCCGGGC
 TGTGAGAACGTGTCTACGACAAATCTCTTCCCATCTCCACGTCGCTTCTGGGTGCTG
 CAGATCATCTTCGTGTCCACACCGACCCCTCTGTACTTGGCTCATGTCTTCTACTGAAC
 AGGAAAGAACAGAAATTCAGCAAGATCGAGGAGGTGCTGAAGCGGTCCAAAACGATGGA
 GGCGACGTGGACGTCCGCTCAAGAAATTTGAGATGAAGAAGCTGAAGTATGGCATTGAG
 GAGCAGGGAAGGTGAAGATGAAGGGAGCCCTGCTGAGAACCACATTGTGACGATCTTC
 TTCAAGTCGCTCTTTGAGGTGGGCTTCTCGTGATCCAGTGGTACATGTACGGTTTCAGC
 CTGTCCGCGGTCTACACCTGTGAGCGGTCCCATGTCCACACCGGGTGGACTGTTTCTGT
 TCCCGTCCCACGAGAAGACAGTCTTCATCATTTTCATGCTGGTGGTGTGCTGGTGTCC
 CTGCTGCTCAACATTTGAGCTCTTCTACGTGCTCTTCAAGAGGATCAAGGACCGGGTG
 AAGGGCAAGCAGCAGCCGGCGCTCTACCCAGCGCCGGCACCCCTGAGCCCTGGGCCAAG
 GAGCTGTCCACCACCAAGTACGCCTACTACAACGGCTGCTCTCACCCACCGCTCCGCTC
 TCACCCATGTCCCCCGGGCTACAAGACGGCCACGGGGAGCGGGGACCGGCTCCTGC
 CGGAAGTACAACAAGCAGCGCAGCAGCAGAACTGGGCCAACTACTCCACCGAGCAGAAG
 CGGCTGGGCCACACCGGCGCAGGAAGCACCATCTCCAAGTCCACGCCCAGGCCCTTCGAC
 TTCCCGACGACACCCAGGAGCACAAGAAGATGCTCTGCTGGCGGCCACGAGTGCAG
 CCGCTGGCGCTGCTGGATGCTCGGCCCTGCAGCCGCGCCANAGCAGGCTGAGCAGCCGC
 GCCCGGCTGACGACCTGGAGCTGA

>Tn22263001 Ensembl prediction. Sequence extended by 321 nt in 3'-direction to reach stop codon (marked blue).

ATGGGAGACTGGTCCATTCTTGCCGCTTCTTAACCGAAGTTCAAATCATTCACGGTC
 ATTGGCAAGATATGGTGACCATGCTGCTCATCTTCCGCATCTTGTGGTAGCACTGGTG
 GGCGACGCGGTGTACAGTGACGAGCAGTCTAAGTTTACCTGCAACACCCCTCCAGCCTGGA
 TGTAACAACGTGCTGATGACACCTTTGCTCCCGTCTCGCACTTGCCTTCTGGGTCTTT
 CAGATTGTTCTTGTCTCCACACCTTCTATTTTCTACATCGTCTACGTTTCAAAAAGATC
 ACCAAGAATGAAAAGTTAGAGGTGAAAAGGTTGTAGTGATACCACGGTCTCCTACACCA
 TTCAAAGGGGGGAGGATCGAGGAGGAGATAAAGAGGCAATGCTGGAGACTGGTGGCCCT
 TATAACCAACCTATAACAATGAGGAGTGGAGTCTCAGGAGGATGAGTGTGAGGAGAGG
 AGCCAGCTGATGAGGAAATGAAAGAGGTGCGGAAAAGACCCAGCCAGCTCTCCAGATCAA
 GTGTTACTCATTTACATCATCATGTTCTGCTGCGCTCCATCATGGAGATCATCTTCTCT

ATTGGACAGTATTACCTCTTTGGATTTGAGGTGCCACATCTTTTCCGCTGCGACACCTAC
CCGTGTCCAAACAGAACCAGCTGCTTTGTCTCTCGAGCCACGGAGAAGACCATCTTCCTG
AACTTCATGTTTAGCGTCAGTCTTGGGTGCTTCATCTTGAACATCGTGGAGCTGCATTAT
CTCGGCTGG

>Tn22263001 Our suggested modification.

ATGGGAGACTGGTCCATTCTTGGCCGCTTCTTAACCGAAGTTCAAAATCATTCCACGGTC
ATTGGCAAGATATGGCTGACCATGCTGCTCATCTTCCGCATCTTGGTGTAGCACTGGTG
GGCGACGCGGTGTACAGTGACGAGCAGTCTAAGTTTACCTGCAACACCCTCCAGCCTGGA
TGTAACAACGTCTGCTATGACACCTTTGCTCCCGTCTCGCACTTGCCTTCTGGGTCTTT
CAGATTGTTCTTGTCTCCACACCTTCTATTTTCTACATCGTCTACGTCTTACAAAGATC
ACCAAGAATGAAAAGTTAGAGGTGAAAAAGGTTGTAGTGATACACCGTCTCCTACACCA
TTCAAAGGGGGGAGGATCGAGGAGGAGATAAAGAGGCAATGCTGGAGACTGGTGGCCCT
TATAACCCAACCTATAACAATGAGGAGTGGAGCTCTCAGGAGGATGAGTGTGAGGAGAGG
AGCCAGCTGAATGAGGAAATGAAAGAGGTGCGAAAAGACCCGACCCAGCTCTCCAGTCAA
GTGTTACTCATTTACATCATCCATGTTTCTGCTGCGCTCCATCATGGAGATCATCTTCCTC
ATTGGACAGTATTACCTCTTTGGATTTGAGGTGCCACATCTTTTCCGCTGCGACACCTAC
CCGTGTCCAAACAGAACCAGCTGCTTTGTCTCTCGAGCCACGGAGAAGACCATCTTCCTG
AACTTCATGTTTAGCGTCAGTCTTGGGTGCTTCATCTTGAACATCGTGGAGCTGCATTAT
CTCGGCTGGATTATATTTTACAGAGTGCTGCTCTCTGCATGCTGCACGTGCTGCAAGTCC
AACAGAGACCCGGTTCAGCAGGTGGAGTTGTATTTCGGACAACAACCCACTGCTGCTGGAG
CTCAAGCATTCACTGCGGGGAGGGTCTGCTCAGAGCCACCTCTGCCGTGACACGGGAC
AAAAGCAGCAGCGTCCCAATCAGGCCCCAGCTATCTCTTTTGAACAGACTCCACACTG
GAGTGCACCTTCGAAGCGGAACCCAGATGAAAAGGAACGCACAAAGGCAAGACTGCACAAA
ATCGGAAGAGGCAAAAAGTCATGGCTGTAA

>Tn26030001 Ensembl prediction. Sequence extended by 264 nt in 3'-direction to reach stop codon (marked blue).

ATGGGTGACTGGAGCTTCTGGGGAACCTTTCTAGAAGACGTCCAGGAACACTCCACCTCG
GTCGGGAAGGTCTGGCTCACCCTCTTTCATCTTCCGGATCCTGGTGTGGGCACGGCC
GCCGAGTCGTCTGGGGCGACGAGCAGAGCGACTTCTGTGCGACACCCAGCAGCCCGGT
TGCAACCAACGTCTGCTACGACAGCGCCTTCCCATCGCCACATCCGCTACTGGGTGCTG
CAGATCGTCTTCTGCTCCACGCCCTCCCTCATCTACATGGGTACGCCATGCACACCGTG
CGCCGGGAGGAGAAACAGCGGCCGAGGGAGCAGGAGGAGAGGGAGGCGAGGGGGAGCGC
GGAGACAGCTTGGAGGAGAAGGAGTCTCCAGCAGAAAGGAGAGCGAAAAGGCTCCGGCG
TCCGAAGGGAGCAGCCGCTGCGCCTGAGAGGAGCCCTGCTGCAGACCTACATACTCAGC
ATCCTGATCCGCACGGTGATGGAGGTGACCTTCATTGTGGTGCAGTACCTGATGTACGGG
GTCTTCTCAACGCCCTGTACCTGTGCAAGGCCCTGGCCCTGTCCCAACCTGTCAACTGC
TACATGTCCAGGCCACGGAGAAGAAGCTCTTCATCGTCTTATGCTGGTGTGGCCGGC
GTGTCCCTGCTGCTCTCGTGTGGAGCTCTACCACTCGGCTGGAGAGCCTCAAAAGG
TGTCTGCGCAAAAGCTGATGGAAAAGAGCAGCCGAGGACTGTGGCGGTGGCGGTGTG
GCGGCCCTGGAGCCCAACAGTCCGCCTCAGCCTTCTGTTTCTGCACGCCGCCCCAGAT
TTCAGCCAGTGCTGGCAGTCTCAGGTTCCATCAACGCCATCGCC

>Tn26030001 Our suggested modification.

ATGGGTGACTGGAGCTTCTGGGGAACCTTTCTAGAAGACGTCCAGGAACACTCCACCTCG
GTCGGGAAGGTCTGGCTCACCCTCTTTCATCTTCCGGATCCTGGTGTGGGCACGGCC
GCCGAGTCGTCTGGGGCGACGAGCAGAGCGACTTCTGTGCGACACCCAGCAGCCCGGT
TGCAACCAACGTCTGCTACGACAGCGCCTTCCCATCGCCACATCCGCTACTGGGTGCTG
CAGATCGTCTTCTGCTCCACGCCCTCCCTCATCTACATGGGTACGCCATGCACACCGTG
CGCCGGGAGGAGAAACAGCGGCCGAGGGAGCAGGAGGAGAGGGAGGCGAGGGGGAGCGC
GGAGACAGCTTGGAGGAGAAGGAGTCTCCAGCAGAAAGGAGAGCGAAAAGGCTCCGGCG
TCCGAAGGGAGCAGCCGCTGCGCCTGAGAGGAGCCCTGCTGCAGACCTACATACTCAGC
ATCCTGATCCGCACGGTGATGGAGGTGACCTTCATTGTGGTGCAGTACCTGATGTACGGG
GTCTTCTCAACGCCCTGTACCTGTGCAAGGCCCTGGCCCTGTCCCAACCTGTCAACTGC
TACATGTCCAGGCCACGGAGAAGAAGCTCTTCATCGTCTTATGCTGGTGTGGCCGGC
GTGTCCCTGCTGCTCTCGTGTGGAGCTCTACCACTCGGCTGGAGAGCCTCAAAAGG
TGTCTGCGCAAAAGCTGATGGAAAAGAGCAGCCGAGGACTGTGGCGGTGGCGGTGTG
GCGGCCCTGGAGCCCAACAGTCCGCCTCAGCCTTCTGTTTCTGCACGCCGCCCCAGAT
TTCAGCCAGTGCTGGCAGTCTCAGGTTCCATCAACGCCATCGCTCCATGGCTCCAC
CCCTTCAGCAACAGGATGGCGTGCAGCAGAACTCGGCCAACTTGGCCACCGAGCGCAT
CACAGCTCCGACAACCTGGAGGACGAGGCGGACTTCTGAGGATCCGATACGACCAAGCTG
CCCTCGGAGCTGCCCCGAGCTGCTGCGCGTGGCCCTCTGCACTGTGGCTTCATCAGG
GACAAACGGCGCCTGAGCAAGACCAGCGGGAGCAGCAGCAGACCTCGCCACGATGACCTT
GCAGTGTAA

>Tn18746001 Ensembl prediction in 7 exons. The predicted 81 nt intron between exons 6 and 7 is regarded as a part of the reading frame (marked blue).

ATGGCAGACTGGAGCTACTGGGAACTTCTGGAGGAGGTGCAGGAGCACTCCACCTCT
GTTGAAAGGTGTGGCTGACCATCTGTTTATCTTCCGTATCCCTGCTTGGGACCGCC
GCCGAGTCATCTGGGAGACGAGCAAGAAGATTTCAACTGTGACACCGAACAGCCAGGC
TGCGAGAACGTTTGTAGCAGCCAGCCTTCCCAATAGCGCATATACGATACTGGGTGCTG
CAGATTGTGTTGTGTCACGCCACGCTGATCTACATGGGCCACGCCATGCACAGGGTT
CGCAGGAGGAGAGAAGAGGAGGAACAGGAGGAGGAGGAAGGTGGGAGGGGAGAGGTGGAGAG
GAGGACCCAGGAGGAGGAGGAAGAGGAGGAGATGACGGCAGAGAAGAAGATAAGAAAGGA
GGGAAAGAAGTGCGGAGCAAGGAGAGAAGGAGAGCGGAGGTCTGTGCGCTTGAAGGGA
GCGCTGCTGCAGACCTATGTAAGTACTGATACGAAGCATCATGGAGGTGGTGTGTTT
CTCAGTCTTCCAGTATCTTCTGACGGGATCTTCTCACTCCCTGTATGTCTGCGAGGCG
TGGCCGTGTCCACATCCGGTGAAGTGTATGTCTCAGGCCAACAGAGAAAACGTGTTT

ATTGTGTTTCATGCTGGCTGTTTCTGCCGTCTCTCTGGTTCTCAGCGTGCTCGAGCTGCAA
CACCTGGCCTGGAGGCACTGCTGCAGGAAGGCGGTAGCTGCTAATGAGGCTCTCTGGGC
CGACAGCTCTCCTTGGCTGACCATCCTGTTTATCTTCCGTATCCTCGTGTGGGACCGCC
ATGATGGGCTCGACACACTTCTACCTCTGGCTTTCCCGAGCCACCACCTGATGGGCTGC
TACTCGCAGCGATGGCAAAAGAGCAATCCAGCCAGATCCAGGAGGACGCTACCTCAGG
AAGGACAATAACTGCTACGGGCCCGAGGAGGAAGATGAGCTGTCCGAGATTGAGAAT
GGGGGCTCCGACAGGCTGCTGCTTTGCCCGGCGGGGCTCTCAGTCAGAAAGGACAAGCGG
AGGTTACGAAAACAGCGGCACCAGCAGCCGAACAAGAGCGGACGACCTGTCGGTTTAA
>Tn18746001 Our suggested modification.

ATGGCAGACTGGAGCCTACTGGGAACTTCTGGAGGAGGTGCAGGAGCACTCCACCTCT
GTTGGAAAGGTGTGGCTGACCATCCTGTTTATCTTCCGTATCCTCGTGTGGGACCGCC
GCCGAGTCATCCTGGGGAGACGAGCAAGAAGATTTCAACTGTGACACCGAACAGCCAGGC
TGCAGAACGTTTGTACGACCGAGCCTTCCCAATAGCGCATATACGATACTGGGTGCTG
CAGATTGTGTTTGTCCACGCCAGCCTGATCTACATGGGCCACGCCATGCACAGGGTT
CGCAGGGAGGAGAGGAGGAGGAACAGGAGGAGGAAGGTGGGGAGGGAGAGGTGGAGAG
GAGGACCCAGGAGGAGGAGGAAGAGGAGGAGATGACGGCAGAGAAGAAGATAAGAAAGGA
GGGAAAGAAGTGGCGGAGCAAGGAGAGAGAGGAGCGGAGGTCTGTGCGCTTGAGGGGA
GCGCTGCTGCAGACCTATGTACTGAGTATACTGATACGAAGCATCATGGAGGTGGTGT
CTCAGTCTCCAGTATCTGTACGGGATCTTCTCCTCACCTCCCTGTATGTCTGCGAGGCC
TGGCCGTGTCCACATCCGGTGAAGTGTATGTCTCCAGGCCAACAGAGAAAAACGTGTTT
ATTGTGTTTCATGCTGGCTGTTTCTGCCGTCTCTCTGGTTCTCAGCGTGCTCGAGCTGCAA
CACCTGGCCTGGAGGCACTGCTGCAGGAAGGCGGTAGCTGCTAATGAGGCTCTCTGGGC
CGACAGCTCTCCTTGTCTCTCCACCATCAACCCACCTCCAGACTTCAGCCAGTGC
ATGATGGGCTCGACACACTTCTACCTCTGGCTTTCCCGAGCCACCACCTGGTGCACCAA
CAGAACTCCGAGAACATGGCCACCGAGAAAGCACAAAATCGCCGCCGCGCTCGAAGAGGCC
ACCCTCTCCAGATGGGCTGCTACTCGCAGGATGGCAAAAGAGCAATCCAGCCAGATC
CAGGAGGACGCTACCTCAGGAAGGACAATAACTGCTACGGGCCCGGAGGCAAGAGATG
AGCTGTCCGAGATTGAGAATGGGGGCTCCGACAGGCTGCTGCTTTGCCCGGCGGGGCT
CTCAGTCAGAAAGGACAAGCGGAGTTTCAGCAAAACAGCGGCACCAGCAGCCGAACAAGA
GCGGACGACCTGTGCGTTTAA

>Tn15227001 Ensembl prediction. The 12 first nt had low identity with the presumed orthologs (considering that this is the first conserved domain), and they were removed (marked gray). Sequence extended 372 nt in 3'-direction to reach stop codon; the first 12 nt as a part of the second conserved domain (green), the remaining corresponding to the cytoplasmic C-terminal tail (marked blue).

TCTTCCCTCTCAGGTAAGACCTGGTGGACTCTGCTGCTGGGTTTGCGCCTGAGCGTCCTG
CTGCTGCTGGGCTTACGCTCTTTCAGCGACGAGCAGGAGCGCTTCGCTGCAACACCATC
CAGCGGGCTGCTCCAAAGTGTGCTTTCAGCGCTTCGCTCCCGTTCGCTCTTCCGCTC
TGGCTCTCCACCTCGTCTCTGGCTTTCCTCATCTGCTCTTGGCACCTACGTGATG
CACCGGCTTCTGACCGCTCCGGGGAAGTGTCCCTTGCAGCGGAGCCAGCTCCAGGAGCC
GGAGGAGCGCGCTTCTACTGCGCGTACGTCCCGGTGGTGGTGGTCCGGATCCTTCTGGA
GTTGTTTTTCGGGGCCGGCCAGTTTCACTCTTCGTTTGTCTTTCCAAAGAGCTTCTG
TGCTACGAGGCCCTTCACCTCGGGGGTGGAGTGCTACATCTCCAGACGCACCGAGAAG
TCCCTCATGCTCAGTTTCATGTTGGGCTGGCCTCGCTCTCCATCTGCTGAGTTTGT
GATCTGCTGGGC

>Tn15227001ic Our suggested modification.

GGTAAGACCTGGTGGACTCTGCTGCTGGGTTTGCGCCTGAGCGTCTGCTGCTGCTGGGC
TTCAGCCTCTTACGCGACGAGCAGGAGCGCTTCGCTGCAACACCATCCAGCCGGGCTGC
TCCAACGTGTGCTTCGACGCTTCGCTCCCGTGTCCGCTTCCGCTCTGGCTCCTCCAC
CTCGTCTCTGGCTCTTCCCATCTGCTCTTTGCCACCTACGTGATGCACCGGCTTCTG
ACCGCTCCGGGGAAGTGTCCCTTGCAGCGGAGCCAGCTCCAGGAGCCGGAGGAGCGGC
TTCTACTGCGCGTACGTCCCGGTGGTGGTCCGGATCCTTCTGGAAGTTGTTTTCGGG
GCCGGCCAGTTTCACTCTTCCGTTTGTCTTTCCAAAGAGCTTCTGTGCTACGAGGCC
CCCTGCACCTCGGGGGTGGAGTGCTACATCTCCAGACGCACCGAGAAGTCCCTCATGCTC
AGTTTTCATGTTGGGCTGGCTCGCTCTCCATCTGCTGAGTTTGTGTTGATGCTGGGC
TCCGTGAAGCGGATGGTGAAGTGGAGGAGGAGGAGATGCTGGCGGAGGAGATGATC
AAAGGAGAACAAAGCAGCGTGATTACGGCGACGACCATGGCTGAAGACAGCGATAAAAGC
CCCGAGTCCAACAGTCCCGACAGCAGAGATGCTCAGGTGGACACACCTCCCACTCCACC
AGCACTCCAGCACTCTCTCGGTGGTCCCTCCAGAGCCGGGTCCGACCCCGCTGTCCCT
CGGCCTGACAGGGAACCACTGAGGGACCCAGCACCAGTGGGGGGAGGAAGCCGGCCAG
TACGGTCCAGCCGGGACAACCTCGGGCCAACAGTCTGACGGCGAGGCTCCAGACAGACGA
GCCTGGGTTTGA

>Tn16594001 Ensembl prediction. Sequence extended by 63 nt in 5'-direction to reach initiation codon (marked green). Sequence extended by 570 nt in 3'-direction to reach stop codon (marked blue).

GGGAAAGTGTGGCTCATAGTGATGATCTTTCTCCGTATCCTGACCTCCTTTTTGCCGGA
TACCCCTCTACAGGACGAGCAGGAGGACGATTGCTGTGCAACACCATCCAGCCTGGATGT
GCCAACGTCTGCTACGACCTGACTCCCCGCTCTCCCTCTTCCGCTTCTGGCTGGTCCAA
CTCATCACTTTGTGCTTCCCTACATCGTCTTTGTCTATCTACATCACTCAACAAGGTCTCA
AATGACCTCTGCGCACACCTGAACCTCTCGGGCCAGGTGAGAACCTCGCGGCTGTTCCAG
ATCCAGCAAGAGGACCTGGTGAGAAGATGGCGCTGAGAGGGGATCGGCTCGGTGCTTC
ACAGGAGCCTATATCTCCACCTGATGTTCCGAACCTTGTGGAGGAGGAGATTGGAGCT
GCTCACTTACTATGCTTTCGTTTCAACATCCCCCGGAGGTTCTGTGTCAACACCCACC
TGCAACCCAGGTGGACTGCTACGTGTCCAGACCCACCGAGAAGACTGTGATGCTCAGC

TTCATGCTCGGCGTGCCGCTCCTGTCCCTTTTTTAAACGTTTTGGATTTTATTAGCGCC
ATCAAGCGCTCTGTACCAAGAAGGGC
>Tn16594001 Our suggested modification.

ATGGCGGGATCAAGTACCTGTGAGGTCACTTTCATCTCTGTCAATCACAGCATCCCGCTG
ATGGGAAAGTGTGGCTCATAGTGATGATCTTTCTCCGTATCTGACCTCCTTTTTGCC
GGATACCCCTCTACCAGGACGAGCAGGACCGATTCTGTGTGAACACCATCCAGCCTGGA
TGTGCCAACGTCTGTACGACCTGTACTCCCCGCTCCTCTTCCGCTTCTGGCTGGTC
CAACTCATCACTTTGTGTCTTCCCTACATCGTCTTTGTCACTACATCATCCACAAGGTC
TCAAATGACCTCTGCGCACACCTGAACCTCTCGGGCCAGGTTCAGAACCTCGCGGCTGTTT
CAGATCCAGCAAGAGGACCTGGTGAGAAGATGGCGCTGAGAGGGGATCGGCTCGGTGC
TTCACAGGAGCCATATCTCTCCACCTGATGTTCCGAACCTTGTGGAGGCAAGATTTGGA
GCTGCTCATTACTATCTCTTCCGTTTCAACATCCCCCGGAGGTTCTGTGTCAACACCCA
CCGTGCACCAACCAGGTGGACTGCTACGTGTCCAGACCCACCGAGAAGACTGTGATGCTC
AGCTTCATGCTCGGCGTGCCGCTCCTGTCCCTTTTTTAAACGTTTTGGATTTTATTAGC
GCCATCAAGCGCTCTGTGACCAAGAAGGGCAAAAAGAAGTTGATGGTAGGGAAGATTTAT
GAGGAGGAGCAGTGTCTCTGTCAACGGGTGCGGCTCCGGACCAACAGACCCAAACCAC
TCGGTGGGTAAACAGAATCTAGAGGTGGAGGCTCAGGCGGGCGGTTTCCGGAAGAGGCAC
AACAGCAAGGGTTCTTGCAGAGGGTGTCTGTCCCTGTAGGGCAAGATCCACCCTCTCTC
GACCGTTCTTCTACGTCTTCCACGTTCACTTGGACCTCCAGGGTCCAACACAAATGGGAAC
AATGGCTACTCCCTTCCACAGGAGGATGTTTCAAGAAACAATGGCAGCGACGTGGCCCTC
TGCCCCCAGAGTCCATGGGACACCTAGATCCATTGAGTTAGCAAAACGAGGTGATTA
AAACCTCCTCTCCACCTAGGCGAGATCTGGGTTCTGTCTCAAGTGGGCGGCGGGGCC
CCTGGGAGACATTCAGCAATCTGTACCAAGAAGGTCGGCCAATTACACTGGTAGAGCTG
TCCAACGAGAGCTACGACAGTGAAGACGGGCAAGACAAAAGGTGAGAGTGGGTCTGA

>Tn19395001 Ensembl prediction. Sequence extended by 69 nt in 3'-direction to reach stop codon (marked blue).

ATGAGCTGGAGTTTCTGACTCGCTGTTGGAAGAAATTCACAACCATCCACGTTTGTG
GGCAAGATCTGGCTCACTGTCTGGTTGTCTTCCGCATCGTGTGACGGCCGTTGGCGGC
GAGTCCATCTACTACGATGAGCAGAGCAAGTTTGTGTGCAACTCGGGCCAGCCGGGCTGT
GAGAAGCTCTGCTACGACGCTTTGCTCCGCTGTGCGATGTCCGCTTTTGGGTGTTCCAG
ATCATTCTGGTGGCCACCCCTTCTGCTCATGTACCTGGGATATGCCGTCAACAAAATCGCT
CGCACAGAGGAGCAGGTGGGTGGGATGGGAGTGAGGGGATGTTTCGAGAGGAAGCTCAAG
AGAAAGCTGTATCTGGCAGACAGAAAGCAGCACAGAGGCATTGAAGAAGCTGAGGATGAC
CAAGAGGAAGACCTATGATCTATGAAATGGCAGAAAGTGGGAGCGACTGCAAGTGAAGAA
ACAAAAGGCAATGTTGTTGGAAGATAAGGTCAAGGTCCGCCACGATGGACGCCAGCGT
ATCAAAGAGGATGGTCTGATGCGTATTTATGTCTTCACTCTCGGCCGCTGCTTGCTG
GAGGTGGCTTTCTTGTGCGGGCAGTACGCCCTGTACGGATTGCTGTTCCCTTACCTAT
GTCTGTTCTCAGCTGCCCTGCCCCACAGCGTGGACTGCTTCTGTCTCCGCGCCACTGAG
AAGACCGTCTTCTCATATTATGTACATCGTCTCCCTGCTCTGTCTGATGCTCAACATC
TGGGAGATGCTTCACTGGGCATCGGCACCATCTGCGAGATCATTGTTCCCGCGGGGTC
CCCGAGGAGGAGCTGTACGGGCTGACACAAGCGAAAGAGCCCCACGCCAGAGAGGATTAC
AGCAGTACCCTTTCTCTGGAACGCGCCATCAGCTCCGCTGGGTACAACATCACAATC
AAGCCCCCAATGGTGCCGCGAGAACGCCACGATCAACCCCTACCGGTACCGACCTCACC
AGCGCGAAGATGGCATGCCGACAAAACACGCTAACATCGCGCAGGAGGAGCAGCAACAG
TACAACAATAACGACGAAACCTTTCGAGAGCCGGGATGGGAGATGACCGCACGCGTTCC
CATCACTCTCAGAACAGACTGGAGATGGACGCGTCAGCCACAGCCAGCCGAGGGCCAA
AACAAACAAGCCTACCGCGACCGCAACACCGCCAGGCTCCAAACATGCGTCCGGC
AAGGCTGACGCA

>Tn19395001 Our suggested modification.

ATGAGCTGGAGTTTCTGACTCGCTGTTGGAAGAAATTCACAACCATCCACGTTTGTG
GGCAAGATCTGGCTCACTGTCTGGTTGTCTTCCGCATCGTGTGACGGCCGTTGGCGGC
GAGTCCATCTACTACGATGAGCAGAGCAAGTTTGTGTGCAACTCGGGCCAGCCGGGCTGT
GAGAAGCTCTGCTACGACGCTTTGCTCCGCTGTGCGATGTCCGCTTTTGGGTGTTCCAG
ATCATTCTGGTGGCCACCCCTTCTGCTCATGTACCTGGGATATGCCGTCAACAAAATCGCT
CGCACAGAGGAGCAGGTGGGTGGGATGGGAGTGAGGGGATGTTTCGAGAGGAAGCTCAAG
AGAAAGCTGTATCTGGCAGACAGAAAGCAGCACAGAGGCATTGAAGAAGCTGAGGATGAC
CAAGAGGAAGACCTATGATCTATGAAATGGCAGAAAGTGGGAGCGACTGCAAGTGAAGAA
ACAAAAGGCAATGTTGTTGGAAGATAAGGTCAAGGTCCGCCACGATGGACGCCAGCGT
ATCAAAGAGGATGGTCTGATGCGTATTTATGTCCTTCACTCTTCCGCGCTGCTTGCTG
GAGGTGGCTTTCTTGTGCGGGCAGTACGCCCTGTACGGATTGCTGTTCCCTTACCTAT
GTCTGTTCTCAGCTGCCCTGCCCCACAGCGTGGACTGCTTCTGTCTCCGCGCCACTGAG
AAGACCGTCTTCTCATATTATGTACATCGTCTCCCTGCTCTGTCTGATGCTCAACATC
TGGGAGATGCTTCACTGAGGATCGGCACCATCTGCGAGATCATTCGTTCCCGCGGGGTC
CCCGAGGAGGAGCTGTACGGGCTGACACAAGCGAAAGAGCCCCACGCCAGAGAGGATTAC
AGCAGTACCCTTTCTCTGGAACGCGCCATCAGCTCCGCTGGGTACAACATCACAATC
AAGCCCCCAATGGTGCCGCGAGAACGCCACGATCAACCCCTACCGGTACCGACCTCACC
AGCGCGAAGATGGCATGCCGACAAAACACGCTAACATCGCGCAGGAGGAGCAGCAACAG
TACAACAATAACGACGAAACCTTTCGAGAGCCGGGATGGGAGATGACCGCACGCGTTCC
CATCACTCTCAGAACAGACTGGAGATGGACGCGTCAGCCACAGCCAGCCGAGGGCCAA
AACAAACAAGCCTACCGCGACCGCAACACCGCCAGGCTCCAAACATGCGTCCGGC
AAGGCTGACGCA
GACCGAGGCGGCAAGCAGCAGCAACACAGCAAAATACGGAGTCTGA
AAAGGCTCCGAGTGGATCTGA

>Tn09585001 Ensembl prediction. Sequence extended by 3 nt in 3'-direction to reach stop codon (marked blue).

ATGAGCTGGAGTTTCTCAGCGGCTGCTGGAGGAGATCCACAACCACTCCACCTTCTGTG

GGGAAGCTGTGGCTACCGTGCTCATCGTCTTCCGCATCGTCTCACCGCGTCGGGGGA
 GAGTCCATCTACTACGACGAGCAGCAAGTTCGTGTGCAACTCGGGACAGCCGGGCTGC
 GAGAACGTCTGCTACGACGCTTTCGCGCCGCTGTCCACGTCCGCTTTCGGGTTTTCCAG
 ATTATCCTGGTGGCCATGCCCTCCCTCATGTACATGGGCTACGCCATCAACAAGATCGCC
 AGGCTGGAGGAAGCCAAAGGAGGCGGGGCTTCGGCGGCCATCAGGACGGGAGGCGGAGGC
 TACACGCACAGAAAGCCAGGAAGATCTGTTTCGGAGCGCGGCAGCACCGGGGCATCGAG
 GAGACCGAGGAGGACGAGGAGGACGACCCCATGATCTACGAGGTCCCGAGGTGGAGGCC
 CCCAAGAGGCCCGGGACCCGCTGCAGCCACGCCAGACCCAAAGTCCGGCACGACGGG
 CGCAAGCGGATCAGAGACGAGGGCCTGATGCGGGTTTACGTGCTGCAGCTGGTACCCGT
 ACCGTGCTGGAGGCGGGCTTCTCGCCGGCCAGTATCTGCTCTACGGTTTCCGCGTGATG
 CCCGTGTTCTGTGTCTCGGGGAGACCGTGCCCCACAGCGTGGACTGCTTCGTGTGCGCG
 CCCACGGAGAAGACCATCTTCTGCGCATCATGTACGGCGTCACCGTCTTTGCCTCGTC
 CTCAACGTCTGGGAGATGCTCCATTTAGGGGTGGGCTCCATCTACGACATCCTCCGCCGC
 CGGCGGCCCGCCCCAGGACGATGAGTACCAGCTGGGCTTGTGGGCGCCAACGGAGGCC
 GTGGAGGGCTCCGTTCGGGGGACGCGGCCCGGAGCGGGTTCCGAAGGAGGGGTGGGCGGC
 GACGGGGCCGCGGACTACGTGGGCTACCTTTCTCGTGAACACGCCGTCGGCTCCGCC
 GGCTACACATCGTGGTCAAACCGAGCAGATGCCCTACACGGACCTGAGCAACGCCAAG
 ATGGCGTGCAAGCAGAACCAGGGCCAACATCGCCAGGAGGAGCAGCAGAGTTTGGTAGC
 AACGAAGACAACCTCCCGACGGGGGAGAAGCCCGCGTGGCTTTGAACAAAGACATGATC
 CAGCAGGCTCAGAGCAGCTGGAGGCGGCCATCCAGGCCTACAGCCAGCAGCACCAGGCC
 GAGGTGACGTGGGGGACAACCAAGGACGACAAACCCAAAGCAACATCATCCAGGCGCAG
 CCGCTGCTGCAGCCGAGCCTCAGAAGGAGCGCAAGCATAGATTCAAGCACGGGAAAGGA
 GGCAGCAGCGCAGGAGGACGAGCAGCAACAGCAGCAGCAGCAAGTCGGGGGAGGGGAAG
 CCTCCGTGTGGATT

>Tn09585001 Our suggested modification.

ATGAGCTGGAGCTTCCTCACGCGGCTGCTGGAGGAGATCCACAACCACTCCACCTTCGTG
 GGAAGCTGTGGCTCACCCTGCTCATCTCTTCCGCATCGTCTCACCGCCGTCGGGGGA
 GAGTCCATCTACTACGACGAGCAGAGCAAGTTCGTGTGCAACTCGGGACAGCCGGGCTGC
 GAGAACGTCTGCTACGACGCTTTCGCGCCGCTGTCCACGTCCGCTTTCGGGTTTTCCAG
 ATTATCCTGGTGGCCATGCCCTCCCTCATGTACATGGGCTACGCCATCAACAAGATCGCC
 AGGCTGGAGGAAGCCAAAGGAGGCGGGGCTTCGGCGGCCATCAGGACGGGAGGCGGAGGC
 TACACGCACAGAAAGCCAGGAAGATCTGTTTCGGAGCGCGGCAGCACCGGGGCATCGAG
 GAGACCGAGGAGGACGAGGAGGACGACCCCATGATCTACGAGGTCCCGGAGGTGGAGGCC
 CCCAAGAGGCCCGGGACCCGCTGCAGCCACGCCAGACCCAAAGTCCGGCACGACGGG
 CGCAAGCGGATCAGAGGAGGAGGCGCTGATGCGGGTTTACGTGCTGCAGCTGGTACCCGT
 ACCGTGCTGGAGGCGGGCTTCTCGCCGGCCAGTATCTGCTCTACGGTTTCCGCGTGATG
 CCCGTGTTCTGTGTGCTCGGGGAGACCGTGCCCCACAGCGTGGACTGCTTCGTGTGCGCG
 CCCACGGAGAAGACCATCTTCTGCGCATCATGTACGGCGTCACCGTCTTTGCCTCGTC
 CTCAACGTCTGGGAGATGCTCCATTTAGGGGTGGGCTCCATCTACGACATCCTCCGCCGC
 CGGCGGCCCGCCCCAGGACGATGAGTACCAGCTGGGCTTGTGGGCGCCAACGGAGGCC
 GTGGAGGGCTCCGTTCGGGGGACGCGGCCCGGAGCGGGTTCCGAAGGAGGGGTGGGCGGC
 GACGGGGCCGCGGACTACGTGGGCTACCTTTCTCGTGAACACGCCGTCGGCTCCGCC
 GGCTACAACTGCTGGTCAAACCGAGCAGATGCCCTACACGGACCTGAGCAACGCCAAG
 ATGGCGTGCAAGCAGAACCAGGGCCAACATCGCCAGGAGGAGCAGCAGAGTTTGGTAGC
 AACGAAGACAACCTCCCGACGGGGGAGAAGCCCGCGTGGCTTTGAACAAAGACATGATC
 CAGCAGGCTCAGAGCAGCTGGAGGCGGCCATCCAGGCCTACAGCCAGCAGCACCAGGCC
 GAGGTGACGTGGGGGACAACCAAGGACGACAAACCCAAAGCAACATCATCCAGGCGCAG
 CCGCTGCTGCAGCCGAGCCTCAGAAGGAGCGCAAGCATAGATTCAAGCACGGGAAAGGA
 GGCAGCAGCGCAGGAGGACGAGCAGCAACAGCAGCAGCAGCAAGTCGGGGGAGGGGAAG
 CCTCCGTGTGGATTAG

>Tn16136001 Ensembl prediction. Sequence extended by 156 nt in 3'-direction to reach stop codon (marked blue).

ATGAGCTGGAGCTTCCTCACAGTCTCCTGGAAGAGATCCACAATCATTCCACATTTGTG
 GGAAGTGTGGCTGACAGTGTCTATCATCTTCCGATTGTGCTCACGGCAGTCGGAGGC
 GAATCCATCTACTCGGACGAGCAGAGCAAGTTACCTGCAACACCAAGCAGCCGGCTGT
 GATAACGTATGCTACGATGCCTTCGCCCTCTCTCGCACGTCCGTTTTTGGGTTTTCCAG
 ATCATCATGATCTCCACCCCTTCGTGATGTACATGGGCTATGCTATTATAAGATAGCG
 CGGAGTTCCGATGAAGAGCGCAGAAAGCTCCACAGGCTTCGCAAAAAGCCCCACCACAT
 TCCAGATGGAGAGAGAACCATCACCTGCAGGGCGTCTTAGAGGAGGACGAAGACGACGAC
 GCTGAGCCCATGATCTATGAGGATACGCTAGAGGTTCAAGATGCCAAACAGGACGAGG
 AACAGCGGTAGCAAAAACCCACCAGAAATATGACGGCCGTCGAAAAATTATGAGGAAGGT
 CTAATGAGGATCTATGCTCTTCAGCTGATGTCAAGAGCTGTTTTGAAATTGCCTTCCTT
 GCTGGACAGTACCTCTGTATGGTTTTCTGTGTGAGTCCATCATATGTATGCAACAGGATC
 CCGTGCCACACAGGGTGGACTGTTTCATCTCAAGACCCACAGAAAAAATAATTTTCTC
 CTGATTATGTACGTGGTGAAGTGTCTCTGCTCGTAAACGCTGTGTGAGATGCTTCAC
 TTGGGAATCGGTACTTTCCGGGACACCTCCGCTGAAGAGGAACAGGGGCGACAGTCA
 TCCTACGGCTAGGCTTTTCTCGCAATATCCAGCGTCTCCTCAGGGTACAACCTTTGTG
 ATGAAAACAGACAAACCAAGCAGGATCCCAACAGCCTTATTGCCCATGAGCAGAACGTG
 GCCAATGTAGCTCAGGAGCACCAGTGATCAGCCAGACGAGAATCCCTCTGACCTT
 GCGAGCCTACACCGGCACCTAAGAGTTGCTCAAGAACAGCTCGATATGGCTTTTCAGACT
 TACCAAAACCAACAAACCAAG

>Tn16136001 Our suggested modification.

ATGAGCTGGAGCTTCCTCACAGTCTCCTGGAAGAGATCCACAATCATTCCACATTTGTG
 GGAAGTGTGGCTGACAGTGTCTATCATCTTCCGATTGTGCTCACGGCAGTCGGAGGC
 GAATCCATCTACTCGGACGAGCAGAGCAAGTTACCTGCAACACCAAGCAGCCGGCTGT
 GATAACGTATGCTACGATGCCTTCGCCCTCTCTCGCACGTCCGTTTTTGGGTTTTCCAG

ATCATCATGATCTCCACCCCTTCCGTCATGTACATGGGCTATGCTATTCAT AAGATAGCG
 CGGAGTTCCGATGAAGAGCGCAGAAAGCTCCACAGGCTTCGCAAAAAGCCCCACCACAT
 TCCAGATGGAGAGAGAAACCATCACCTGACAGGGCGTCTTAGAGGAGGACGAAGACGACGAC
 GCTGAGCCCATGATCTATGAGGATACGCTAGAGGTTCAAGATGCCAAACAGGACCAGGG
 AACAGCGGTAGCAAAAACCCACCAGAAATATGACGGCCGTCGAAAAATTATGCAGGAAGGT
 CTAATGAGGATC TATGTCTCTTCAAGTGTGCAAGAGCTGTTTTGAAATTGCCTTCCTT
 GCTGGACAGTAACCTCTGATGGTTTTCTGTGTCAAGTCCATCATATGTATGCAACAGGATC
 CCGTGCCACACAGGGTGGACTGTTTCATCTCAAGACCCACAGAAAAAATATTTTCTC
 CTGATTATGTACGTGGTGAGCTGTCTCTGCCTCGTGCTAAACGTCTGTGAGATGCTTCAC
 TTGGGAATC GGTACTTTCCGGGACACCTCCGCTGAAGAGGAACAGGGGCCGACAGTCA
 TCCTACGGCTACGCTTTTTCTCGCAATATCCAGCGCTCTCTCCAGGGTACAACCTTGTG
 ATGAAAACAGACAAACCAAGCAGGATCCCAACAGCCTTATTGCCCATGAGCAGAACGTG
 GCCAATGTAGCTCAGGAGCACCAGTGATCAGCCAGACGAGAACATCCCTCTGACCTT
 GCGAGCCTACACGGCACCTAAGAGTTGCTCAAGAACAGCTCGATATGGCTTTTCAGACT
 TACCAAAACCAACAAACAGCAGACGTCAGAACAGTATGCTGAGTGTGGAGGACAC
 ATGGCAGAACAAACAGAGTCAATGCAGTTCAAGAGAAGCAGGGCGCAAGGCCAAAATCA
 GCCACAGAGAAGGCCACGACCGTGGTAAAAAATGGAAAGAGCTCTGTCTGGATTAG

>Tn01153001 Ensembl prediction in 2 exons (black/blue letters). We suggest 67 nt in exon 1 as an intron (marked in gray in upper sequence, adjacent positions marked by underlining in lower sequence). This makes a good identity to the presumed orthologs (and close paralogs). 1 N added in pos 370 of modified sequence to keep the reading frame (marked pink). Sequence extended by 396 nt in 3'-direction to reach stop codon (marked blue).

ATGAGCTGGAGTTTCTGACGCGTCTGTGGACGAGATCTCCAACCACTCCACCTTCGTG
 GGGAGATCTGGCTGACCATTTTGTATCATCTTCCGCATCGTGCTGACGGCCGTGCGCGGT
 GAGACCATCTACTACGATGAGCAGAGCAAAATTTGTTTGAACACGACGAGCCCGGATGC
 GAGAACGTGTGCTACGACGCTTTCGCCCGCTCTCCACGTACGATTCTGGATCTTCCAG
 GTGAGCTGCGCCTCGTTGGATTCTCCACAGGACGGCGCGGGGCTGACGCTTGTGTCTCC
 TTGGCAGGTGATCTGATCACCACCCCAACCATCATGTACCTGGGCTTCGCCATGCACAA
 GATCGCACGATGAACGACAGCGAGTACCGCGTCTGCCGAAAGCCAAGAAGAAGATGCC
 CATAGTGAACCGCGGACCGCCCGGACTACGAGGAGGCGGAGGACAACGGCGAGGAAGAC
 CCCATGATCGCCGAGGAGATTGAACAAGAGAAGCCTGACAAAGCGGAGAAGGGGCCGGAG
 AAAAAAGCATGATGGCCGACGGCAATCCAGCGTGACGGCTGTGAAGGTCTACGTGTGC
 CAGCTGCTGTGGCGCTTCTCTTCGAGGTCGCCCTTCTCTTTGGCCAGTACGTCTCTAC
 GCTTTCGAAGTGACGCGCTTACGCTGTGCACCGCTCGCCGTCGCCCCACACGGTGGAC
 TGCTTCGTGTGCGCGCCACAGAGAAGACCATCTTCTGCTGGTCAATGTATGTGGTGTCC
 TTCTCTGCGCTGCTCCTCACCCTCTTTGAAATGCTCCACTTGGGGATCGCGCGCTCCGC
 GACACCTTCGCGAGGGCGTCCGCTCTCAACAGCGGGCCCTCGTCTGACGGCCCCACGT
 AGCATCGCCACGGCGCCGCCG

>Tn01153001 Our suggested modification.

ATGAGCTGGAGTTTCTGACGCGTCTGTGGACGAGATCTCCAACCACTCCACCTTCGTG
 GGGAGATCTGGCTGACCATTTTGTATCATCTTCCGCATCGTGCTGACGGCCGTGCGCGGT
 GAGACCATCTACTACGATGAGCAGAGCAAAATTTGTTTGAACACGACGAGCCCGGATGC
 GAGAACGTGTGCTACGACGCTTTCGCCCGCTCTCCACGTACGATTCTGGATCTTCCAG
 GTGATCTGATCACCACCCCAACCATCATGTACCTGGGCTTCGCCATGCACAAGATCGCA
 CGCATGAACGACAGCGAGTACCGCGTCTGCCGAAAGCCAAGAAGAAGATGCCATAGTG
 AACCGCGGANCCGCCCGGACTACGAGGAGGCGGAGGACAACGGCGAGGAAGACCCCATG
 ATCGCCGAGGAGATTGAACAAGAGAAGCCTGACAAAGCGGAGAAGGGCCCGGAGAAAAAG
 CATGATGGCCGACGGCGAATCCAGCGTGACGGCTGATGAAGGTC TACGTGTGCCAGCTG
 CTGTGGCGCTTCTCTTCGAGGTCGCCCTTCTCTTTGGCCAGTACGTCTCTACGGCTTC
 GAAGTGCACGCGTCTACGTGTGCACCGCTCGCCGTGCCCCACACGGTGGACTGCTTC
 GTGTGCGCGCCACAGAGAAGACCATCTTCTGCTGGTCAATGTATGTGGTGTCTTCTC
 TGCTGCTCCTCACCCTCTTTGAAATGCTCCACTTGGGGATCGGCGGCGTCCGCGACACC
 TTCCGACAGGGCGTCCGCTCTCAACAGCGGGCCCTCGTCTGACGGCCCCACGTAGCATC
 GCCACGGCGCGCGGGCTACACGCTACCATGAAGAAGGAGAAATTTGAAAGGACGGCTG
 AGGGACTCGCCCATGGGCGACTCCGGGAGGGAGAGCTTCGGTGACGAGGTCCTCATCC
 CGGGAACCTGGAGCGGCTGAGGAGGCACCTGAAGCTGGCCAGCAACACCTGGACAGGCC
 TACCAGGTTGAGGACAGGAGCCCTCGCGGAGCAGCAGCCCGAGGTGAACACGGCCGCG
 CAGACGGCCCGGAGCAGAACCAGCTCAACTTTGCCAGGAGAAGCAAGGAGAACCACAGC
 GAGAAAAGGTAAAGAAATGCTCAGGCGCCATCAGATGGAGGTGCGGCGACTATTACAGGTG
 GTGGTTTTGGCCAGCAGGCAGCGTTTTTCAGAGGCGTCTCTGTCGCTGTTAG

>Tn16902001 Ensembl prediction in 2 exons (black/blue letters). The predicted 19 nt intron regarded as part of reading frame (marked blue). Sequence extended by 4 nt in 3'-direction to reach stop codon (blue).

ATGAGTTGGAGCTTCTCACCAGCTGTTGGACGAGATCTCCAACCACTCCACCTTCGTG
 GGCAAAATCTGGCTCACCTCTCATCGTCTTCCGCATCGTCTGACGGCCGTGCGCGGC
 GAGTCCATATACTACGATGAACAGAGCAAAATTTGTGTGCAACACAAACAGCCCGGTTGC
 GAGAACGTGTGCTACGACGCTTTGCGCGCTGTGCGACATCCGCTTCTGGGTGTTCCAG
 GTGATCATGATCACCACGCCACCATCATGTACCTCGGCTTTGCCATGCACAAGATCGCC
 CGGATGGACGACAACGACTACCGGCCCGCGCCAGGAAGAGGATGCCAATCGTCAGCCGC
 GCGGCCAACCGGACTACGAGGAGGCGGAGGACAACGGCGAAGAGACCGGATGATTCTA
 GAAGAGATCGAGCCAGAAAAGGAGAAGGAGACCGCGGAGAAGCGGGCAAAAAGCACGAC
 GGCCGGCGTGGATCAAGCGCGACGGTCTGATGAAAGTTACGTGTTCCAGCTGCTGTG
 CGCGCCATCTTTGAAGCTCTTCTGTTGGGCGAGTACATCTCTACGGGCTGGAGGTG
 GCGCCCTCGTACGCTTTGACGCGCTCCCTGCCCCACACGGTGGACTGCTTTGTTTTCC
 CGTCCACCGAGAAAACCATCTTCTGCTCATCATGTACGCGCTCAGCGCGCTCTGCCTG

CTCTTCACCGTGCTGGAGATCCTCCACCTCGGCATCAGCGGCCTCCGGGACTGCTTCTGC
 GCCCCGCGGCCCCGCCGCCACCCCGCTCACTCGGCCCTGGCCAGCCAGAGTCTCTCC
 ATCTGCCGCCAGCCGCTCCGCTCCTCCGGGAAGGACCCCTTCGGGAAAGCTGGGCTTCAGG
 GACAACCTGGGGGACTCCGGCCGGGAATCCTTCGGGGACGAAGCTTCGTCGGGGAAGT
 GAGAGGCTGCGCAAGCAGCTGAACTGGCGCAGCAGCAGCTGGACATGGCCTACCAGAAC
 GAGGAAAGCAGCCCTCGCGCAGCAGCAGCCGGAGTCCAACGGCACCAGCGGCCGAGCAG
 AACCGACTGAATTTGCCCCAGGAGAAGCAGAGCGACAAAG

>Tn16902001 Our suggested modification.

ATGAGTTGGAGCTTCTCACCCGCTGTTGGACGAGATCTCAACCACTCCACCTTCGTG
 GGCAAAATCTGGCTCACCTCCTCATCGTCTTCCGCATCGTCTGACGGCCGTCGGCGGC
 GAGTCCATATACTACGATGAACAGAGCAAATTTGTGTGCAACACAAACCAGCCGTTGC
 GAGAACGTGTGCTACGACGCGTTTGGCCGCTGTGCGACATCCGCTTCTGGGTGTTCCAG
 GTGATCATGATCACACGCCACCATCATGTACCTCGGCTTTGCCATGCACAAGATCGCC
 CGGATGGACGACAACGACTACCGGCCCGCGCCAGGAAGGATGCCAATCGTCAGCCGC
 GCGCCCAACCGGCTACGAGGCGGAGGACAACGGCGAAGAACCCGATGATTCTA
 GAAGAGATCGAGCCAGAAAGGAGAAGGAGACCGCGGAGAAGCCGGGCAAAAAGCAGAC
 GGCCGGCGTCCGATCAAGCGCGACGGTCTGATGAAAGTTACGTGTTCCAGCTGCTGTCG
 CGCGCCATCTTTGAAGCCTCCTTCTGTTGGGCGAGTACATCTCTACGGGCTGGAGGTG
 GCGCCCTCGTACGTTTGCACGCGCTCCCTTGCCTCCACACGGTGGACTGCTTTGTTTCC
 CGTCCACCGAGAAAACCATCTTCTGCTCATCATGTACGCCGTCAGCGCGCTCTGCCTG
 CTCTTACCGTGTGAGATCCTCCACCTCGGCATCAGCGGCCTCCGGGACTGCTTCTGC
 GCCCGCGGCCCCGCCGCCACCCCGCTCACTCGGCCCTGGCCAGCCAGAGTCTCTCC
 ATCTGCCCGCAGCCGCTCCGCTCCTCCGGGCTACACACGGCCCTGAAGAAGGACCCCTC
 GGGAAAGCTGGGCTTCAGGGACAACCTGGGGGACTCCGGCCGGGAATCCTTCGGGGACGA
 AGCTTCGTCGGGGAAGTGGAGAGGCTGCGAAGCACCTGAACTGGCGCAGCAGCACCT
 GGACATGGCCTACCAGAACGAGGAAAGCAGCCCTCGCGCAGCAGCAGCCGGAGTCCAA
 CGGCACCGCGCGAGCAGAACCAGCTGAATTTGCCCCAGGAGAAGCAGAGCGACAAAG
 TGA

>Tn29168001 Ensembl prediction in 2 exons (black/blue letters). The 5' 124 nt (incl. exon 1) removed to achieve an initiation codon consistent with other connexins (marked gray). Two single nt removed to keep reading frame (marked gray in upper sequence, adjacent positions underlined in lower sequence). Sequence extended 669 nt in 3'-direction to reach stop codon (marked blue).

ATGGCCAGGGGCGGGGGCGGGGACGTGGCCACAGCGCGGTTGTGCGACAGATGCCTCC
 ATCAGTGAGGAGGCCCGGAGGCCCGGGCTGAACCTTCTCCGTAGGCTCGAAATCTTG
 AAGCATGGGCGACTGGATCTTTCTGGGGCGGCTGCTGGAGAAGCTCTAGGGAGATTCC
 ACCGTCATCGGCAAGTCTGGCTGACCGTCTCTTTCATCTTACAGGATCCTGGTGTGGGC
 GCGGCCCGCGAAGAGGTTGCGGGCGACGAGCAGTCCGACTTACCTGCAACACCCAGCAG
 CCGGGCTGCGAGAAGCTGTGCTACGACGAGGCCCTCCCATCTCGCACATCCGCTTCTGG
 GTGCTGACAGATCATCTTCTGTGTCACGCCACGCTCATCTACCTGGGCCACGTGCTGCAC
 ATCGTCCGCATGGAGGAGAAGCGCAAGGAGAAGGAGGAGAGATGCGCAAAGCCAACCGC
 TTCCAGGAGGAAGGAACTCCTTTACGAAACGGGGGGACGCAAGGAGCGGCGGCGAGG
 AAGGAGAAGCCGCCCATCAGGGACGAGCAGCGCAAAATCCGCATCAGAGGCGCGCTGCTG
 CGGACCTACGTCTTCAACATCATATTCAAACCCCTGTTTGAAGTGGGATTCACTTCTGGGC
 CAGTATTTCTGTACGGCTTCCAGCTGAGGCCCCGTGTACAAGTGTGCGGTTGGCCCTGC
 CCCAACACGCTGGACTGCTTCTCATCTCCCGGCCACCGAAAAGACCATTTTCACTTCTTT
 ATGCTTGTGGTGGCTGCGTGTCTCTTTTGTGAATTTGTTAGAGATCTATCACCTCGGG
 TGAAGAAGGTCAAACAGGGCGTGACC

>Tn29168001 Our suggested modification. 2 nt removed from underlined sequence.

ATGGGCGACTGGATCTTTCTGGGGCGGCTGCTGGAGAAGCCTCAGGAGCATTCCACCGTC
 ATCGGCAAAAGTCTGGCTGACCGTCTCTCATCTTCAGGATCCTGGTGTGGGCGCGGCC
 GCCGAAGAGGTGTGGGGCGACGAGCAGTCCGACTTACCTGCAACACCCAGCAGCCCGGC
 TGGCAGAACGTGCTGCTACGACGAGGCCCTCCCATCTCGCACATCCGCTTCTGGGTGCTG
 CAGATCATCTTCTGTCCACGCCACGCTCATCTACCTGGGCCACGTGCTGCACATCGTC
 CGCATGGAGGAGGAGGAGGAGGAGGAGGAGGAGATGCGCAAGGCCAACCGCTTCCAG
 GAGGAGAAGGAACTCCTTTACGAAACGGGGGGACGCAAGGAGCGGCGGCGAGGAAGGAG
 AAGCCGCCCATCAGGAGCAGACGCGCAAAATCCGCATCAGAGGCGCGCTGCTCGGAGCC
 TACGTCTTCAACATCATATTCAAACCCCTGTTTGAAGTGGGATTCACTTCTGGGCCAGTAT
 TTCTGTACGGCTTCCAGCTGAGGCCCTGTACAAGTGTGCGCGTTGGCCCTGCCCAAC
 ACGGTGGACTGCTTCTCATCTCCCGGCCACCGAAAAGACCATTTTCACTTCTTTATGCTT
 GTGGTGGCTTGGCTGTCTCTTTTGTGAATTTGTTAGAGATCTATCACCTCGGGTGAAG
 AAGGTCAAACAGGGCGTGACCACCAAGTTGTCCCCGACGGCGAGTCTGCTGCGCGGCTC
 AACATCGCGGAGCCCGAGTGTGTTGGCCCCGCCGCCAGAACTGCCCGCTCCAAGTTACCCC
 CCCGACTACACGGACGTGACGGCAGGACGCGGGCCTTCTGACGCCGTTGGCGCGGCCG
 GCGGTGCTTCCGGCGCGCGTCAAGATGGACGACCTCCAGCGGAGCCAGCCTCCCCAC
 CAGCCCCCTCCTCCTCCTCCTCTTCTCCCTCACTACTACATCAGCAACAACAAC
 AACCCAGGTTGGCCGCGCAGCAGAACTGGGCCAACCTGGCCACCGAGCAGCACTCGG
 GAGATGAAGGCCACCTCCCCCTCGCCCTCCTCCTCAACACCTCCCATGATGAGCAGCAG
 CAGCAGCCTGTTGATGCGGAGCTGCTCCCTCCCGCCACCAACCAACACCATCGCCGCC
 GCCGATGCTCCAGGAGCAGCAGCCGGGCTCCGCTCCAACGCAAGCAGCTGGGGTGGG
 GGAACCAACGAGCAGGAAGGAAGGCGCTCTCCACCAACAGGGTGGAGATGCACGACCT
 CCGCCCGCCCCGCGTGGACCTCGGCGACTCAGCCGAGCCAGTAAGAGCAGCAGCGTC
 AGAGCGAGGCCGAGCGACCTGGCCGTCTGA

>Tn19591001 Ensembl prediction. Sequence extended by 720 nt in 3'-direction to reach stop codon (marked blue).

>Tn19591001 Our suggested modification.

ATGGTGACTGGAGCTTTCTAGGGCGGCTGCTGGAGAATGCTCAAGAACACTCCACTGTG
ATTGGAAGAGTTTGGCTGACTGTCTCTTTATCTTCCGCATCTTGGTGCTGGGCGCAGCC
GCTGAGGAGGTTTGGGGTGATGAGCAGCTCCGATTTCACCTGTAAACACGACGAGCCGGCT
TGCAGAAACGCTCTGCTACGACGAGGCCCTTCCCCATCTCCACATATTCGTTCTCTGGGTGCTG
CAGATCATTTTTGTCTCCACGCCAACCCCTCATCTACCTGGGCCACGTGCTGCACATTGT
CGCATGGAGGAGAGGAGGAGGAGGAGGAGGAGCTCCGGAAGGACGAGGCCGGCAGG
GAGGACACGAGTCTCTTTTTCATATGAGATTAGCAACGGAGGAGGACGAGGTTGGCGG
AAAAAAGAGAGGCCCTATTTCGGGATGAACACGGGAAGATCCGATCTCCGCGGGGCGTTA
CTGAGGACC **TACATCTTCAACATCATCTTCAAGACTCTATTTGAGGTGGGCTTCATCTG**
GGGCAGTACTTCTCTATAGGCTTCCACCTGAGGCGGCTATATAATGTGGCCGCTGGCCC
TTCCGCGAACTGTGGACTGCTTCACTCCAGGCCACCGAAAGACAATTTTTATATCAT
TTCATGCTGGTGGTTGCGTGCTGATCTCTTGGCCCTCAACTGTTGGAATCTACCACTG
GGATGGAAGAAGGTCAAGCAGGGAGTACCAATGAGTTTGTCCCCGACGGTGAGTTGCTG
TTGAGGAGTGCCAAACAGCAGAGACGCGGAGAAAGATCCGTGAGCAGGCTTCTCCATCG
GTGCTTGAATGTTTGTGCACTTACTCTCAGCATGAATGTGGCAGGAAGCGGAGGGAATGAA
GGAGAGACTCTACAGTCCGCCCGGAGGCTCTCTGGCTGTGATGTATCACTCCGACCTCT
AAGATGGACGGCAGCGGCTTCCACCGGACGACCTCTTGTGGAAGGCCCTGCCCTGCTCT
TTTTGCGGCAGTAGTGACAAAGTGAGCCACGGGACGCTGACAGAAGTGGAGCAAACTGG
AGCAACAGTGGCACTGGAGCTCGAGAATCTCAATGGGAAAAAATCTCTCTACCTCTCTCC
CTTCCCTCCCTCCCAACT
ACAAACCTCTCGCTTCTCTCAAGGGGAGCAACACTCCATGTTCCCAACGCTGCCCTGCTCAT
GATCCCCTCTACGCTCTCACTCCAAAGGAGACCATGGAGGAGCCCTCTACTGCTCATGT
GACGCTCCACCCGACGATGTACCGGTGGTTACCAAGGCAGAGATGCACTGGCTCTGCT
TCTGCTGCCACGACATCGGAAGCCAAAGTCGGGCGAGCAGGAGCAGCGTCAGAGCAGC
CCGATGACCTGGCAGTGTAG

>Tn17092001 Ensembl prediction. Sequence extended by 378 nt in 3'-direction to reach stop codon (marked blue).

ATGGGGGACTGGAACCTGCTGGGGAACCTTCTGGAAAAAGCCAGGAGCACTCGACTGTG
GTGGGGCAAGTGTGGCTGACCGTCTCTCTTCATTTTCCGCATCTGATCCTGAGCGCCGCC
ACTGAGAAGGTGTGGGGCGACGAGCAGTCGGGCTTCCACTGTGACACCAACAGCACTGGT
TGCAGAAACGTCTGCTATGACATACATTTCCGCATCTCCACAGTCGGTTTCTGGGTGCTG
CAGATCATCTTTGTGTGCGACGCCGACGCTGATTTACCTGGGACACATTCTCCATCTGGTG
CGGATGGAGGAGAGCAGAGAAGGAGAAGGAAAAAGGACGCGAAGACTGTGACAGAGATGCG
GGTCTGCTGGTCTCCAAGCAAAAAAGCCCTGGTGAGGAGCAGAGAAGGGCAGAGTGCGC
CTGACGGGGGAGCTATTGCGCACATACGCTTTAACGTCGTCTTTAAACGCTGTTTGAG
GTGGGCTTCATCGTGGCTCAGTATTTCTTTATGGCTTTGAGCTGAAGCCGATGTATACA
TGCAACAGAGCCCTGCTCCAATGCTGGTGAACGCTATATTTCCGGGCCACGGAGAAG
ACCATCTTCATCATCTTCATGCTGGGCGTGGCCAGTGTATCTTCTGCTCCTGAATCTCAT
GAGATCTATACCTGGGCTTACCAAGTGCGCCAGGGTCTCACC

```
>Tn17092001 Our suggested modification.
```

ATGGGGGACTGGAACCTGCTGGGGAACCTTCTGGAAAAAGCCAGGAGCACTCGACTGTG
GTGGGCAAGAGTGTGGCTGACCGTCTCTTTCATTTCCGCGATCTGTACTCTGAGCGCCGCC
ACTGAGAAGGTGTGGGGCGACGAGCAGTCGGGCTTCACTGTGACACCAACAGCGCTGGT
TGCAGAAACGTCTGCTATGACATCACATTCCCGATCTCCACAGTCGCGTTTCTGGGTGTGT
CAGATCATCTTTGTGTGCGACGCCGACGCTGATTTACCTGGGACACATTCTCCATCTGGTG
CGGATGGAGGAGACAGCAGAGAAGGAGAAGGAAAAAGGAGCGCAAGACTGTCCAGAAAGCG
GGTCTGCTGGTCTCAAGCACAACAAAAGCCCTGGTGAGGAGCAGAGAAGGCGACAGTGC
CTGACGGGGGAGCTATTGCGCACATACGCTCTTTAACGTGCTCTTTAAACGCTGTTTGAG
GTGGGCTTCATCGTGGCTCAGTATTTCTTTATGGCTTTGAGCTGAAGCCGATGTATACA
TGCAACAGAGCCCCCTGCTCCAAATGTGTGTTAACTGCTATATTTCCGCGCCACGGAGAAG
ACCATCTTCATCATCTTCATGCTGGGCGTG6GCCAGTGTGCTCTGCTCTGAATCTCAT
GAGATCTATCACCTGGGCTTACCAAGTGCCGCCAGGGTCTCACCCTCAGGAGGCGGGAC
CTTCCCTCCGAGGGGATTCTCAAGGACCCAGCGTG6GCTCTGTG6CCTTTGCGCCCAAG
TAGACTGAGTACTTCCAGCAGACCAACCCGGTGACGCGGGCTACCCGCCCTG6CCGCT
TACAACCTCTCCC6GCTGTGACG6GCACCGAGTGCTCTTTTCATCTTACAACAGCAAG
GCAGCCTACAAGCAGAATAAGGACAACCTGCTGGTGAAAGGAGCAGCAGCAAGCGGAG
GAATGCGACCTGAAAGGAGAGAAGGATCGGGTTCTGCCCCGAGTCACTACGCAAGTCA
AGGCTAGCCGCGAGCGCCAAACACGGCAACAACAAGACTAGAATAGACGATCTGAAGATA
TGA

>Tn22231001 Ensembl prediction. No modification.

ATGGGGGACTGGAAC TTGTTGGGGAAGCTGCTGGAGAGTGCCCAGGAACACTCCACCGTT

GTGGGCAAAGTCTGGCTGACAGTGCTGTTTCATCTTCCGTATCCTGGTGCTGGGAAGTGCC
 GCTGAGAAAGGTGTGGGGAGATGAGCAGTCCGGCTTTACGTGCGACACCAAGCAGCCCGGT
 TGTGAGAACGTTTGTCTAGCAGCAAGACCTTTCCATTTCCACATCCGCTTCTGGGTGATG
 CAGATCATTTTTCGTCTCCACGCCACCCCTCATCTATTTGGGCCACATCCTTCATCTGGTT
 CGCATGGAGGAAAAAGAGAAACAGAAAGAGAAGGAGCTGGCAGCCAGAGTGAAAAACAG
 CAGCAGTTGCTTGGCAACAAGCCGAAAAAGGCCCAATTAAAGACAACCAGGGTCACGTG
 CGTTTGGCAAGGCGCCTGCTGCGAACTTACGTCTTCAACATCATCTTCAAGACCCTGTTT
 GAAGTGGCCTTTATTGTAGCTCAGTACTTCTCTATGTTTCGAGCTCAAGCCGATGTAC
 ACCTGCGACCGCTGGCCTTGCCCAACATGGTGAAGTGTACATCTCTCGACCCACTGAG
 AAGACGGTCTTTCATCTCTTCATGCTGGCGGTGGCTTGCATCTCTCTGCTGCTCAACCTG
 GTGGAAATGTACCATCTGGGATTCAACCAAGTGCACACAGGGCCCTTCGGTACAGCGATCA
 AAGACCAGAAAACAGTCTCCCAAGGCCCTCCACGAGCCCGTATGCCCTTTGTTCCAGT
 TACAACACTACACCGGTGACCCGTGAGTGCCGGAGCCGTTTCCGACCGACTCCAAGTAC
 AGCGTGACAGAGCCCGCTCCGCTTACAGCCCCACAGCAATAAGGTCGTCTACAAGCAG
 AACAGGGACAACCTGGCTGTGGAGAGGAAGGGAACCCGAGGACGAGGTCTGATGGAG
 AGGAAACCCACCTGCCCTGCCTTTGAGGGGTCTGCTGACAGTCAGCGCAGAAACAGTCAG
 TCAGCAAGCACAGCAAGAGCAGACTGGATGACCTAAAGATCTAA

>Tn26028001 Ensembl prediction. Sequence extended by 3 nt in 3'-direction to reach stop codon (marked blue).

ATGGGTGACTGGAGCTTTCTGGGTAATATTTTAGAGGAAGTGAACGAGCACTCTACGGTG
 ATCGGCCCGGGTGTGGCTCACGGTCCTCTTCATCTTCCGCATCCTCATCCTGGGCACGGCG
 GCAGAGTTTGTGTGGGGGAGCAACAGTCAGACTATGTCTGCAACACGACGACGCTGGC
 TGTGAGAATGTGTCTATGATGAGGCCTTCCCATCTCCACATCCGCTGTGGGTGCTG
 CAGATCATCTTTGTGTCCACGCCGTCTCTGGTGTACGTGGGTACGCTGTGCACCACGTC
 CACATGGAGGAGAAGCGCAAGGAGCGGGAGGAGGACAGCTCAGCCGGCAGCAGGAGCTG
 AGCGAGGAGACGCTCCCTTGGCCCCGACAGGGTAGCGTCCGACCAACCAAGGAGACC
 AGCACCAAGGGGAGCAAGAAGTCCGGCTGGAGGGCACCTGCTGAGGACCTACATCTGC
 CACATCATCTTCAAGACGCTGTTTGAAGTGGGCTTCGTGGTGGGCCAGTACTTCTGTAC
 GGCTTTTCGATTCTGCCGCTGTACAAATGCAGCCGCTGGCCCTGCCCAACACGGTGGAC
 TGCTTCGTGTCCGACCCACCGAGAAGACCGTCTTCATCATCTTCATGCTGGCTGTGGCC
 TGCGTCTCTCTCTTCTCAACTTTGTGGAGATTAGTCACTTGGGCCTGAAGAAGATTTCGC
 TTTGTCTTTTCGCAAGCCGGTGCCGCCCCGCGCCAGGGCGAGGGCTCGGCCCGCTCCCG
 GCCCGGGGCAAGAGTCTGCCCTCCCTCGCCGTGCCCTCCATGCAGAGAGTGAAGGGGTAC
 AGGCTGCTGGAGGAGGAGAAAGCTCCCAATAACTCACCTCTACCCACTGGCCGAGGTG
 GGCATGGAGGCCGGCAGAGGGAGCCCCCTTCCAGGGACTAGAGGAGAAGCCGGAGGAG
 GTGCTGCCCATGGAGGACATCTCAAGGTGTACGACGAGACTTGCCCTCCTACACCCAG
 ACCACCGAGACTGGGGGGGTGACACTACACGAGGAGGAAGAGGTAGAGGTAGAGGAGGAG
 CAGCCAGCCGAAGCAGAGAAGGAGGAGGTGGTTGTGAGGGAGGAGGACAGAGGTGGTG
 AATGTGGAGGGGCCAGAGCCGCGGAGGCCCGGATACGATAGAAGACACCCGACCGCTG
 AGCCGACTGAGCAAAGCCAGCAGCAGAGCCAGGTGAGATGATCTGACGGTA

>Tn26028001 Our suggested modification.

ATGGGTGACTGGAGCTTTCTGGGTAATATTTTAGAGGAAGTGAACGAGCACTCTACGGTG
 ATCGGCCCGGGTGTGGCTCACGGTCCTCTTCATCTTCCGCATCCTCATCCTGGGCACGGCG
 GCAGAGTTTGTGTGGGGGAGCAACAGTCAGACTATGTCTGCAACACGACGACGCTGGC
 TGTGAGAATGTGTCTATGATGAGGCCTTCCCATCTCCACATCCGCTGTGGGTGCTG
 CAGATCATCTTTGTGTCCACGCCGTCTCTGGTGTACGTGGGTACGCTGTGCACACGTC
 CACATGGAGGAGAAGCGCAAGGAGCGGGAGGAGGACAGCTCAGCCGGCAGCAGGAGCTG
 AGCGAGGAGCGCCTCCCTTGGCCCCGACAGGGTAGCGTCCGACCAACCAAGGAGACC
 AGCACCAAGGGGAGCAAGAAGTCCGGCTGGAGGGCACCTGCTGAGGACCTACATCTGC
 CACATCATCTTCAAGACGCTGTTTGAAGTGGGCTTCGTGGTGGGCCAGTACTTCTGTAC
 GGCTTTTCGATTCTGCCGCTGTACAAATGCAGCCGCTGGCCCTGCCCAACACGGTGGAC
 TGCTTCGTGTCCGACCCACCGAGAAGACCGTCTTCATCATCTTCATGCTGGCTGTGGCC
 TGCGTCTCTCTCTTCTCAACTTTGTGGAGATTAGTCACTTGGGCCTGAAGAAGATTTCGC
 TTTGTCTTTTCGCAAGCCGGTGCCGCCCCGCGCCAGGGCGAGGGCTCGGCCCGCTCCCG
 GCCCGGGGCAAGAGTCTGCCCTCCCTCGCCGTGCCCTCCATGCAGAGAGTGAAGGGGTAC
 AGGCTGCTGGAGGAGGAGAAAGCTCCCAATAACTCACCTCTACCCACTGGCCGAGGTG
 GGCATGGAGGCCGGCAGAGGGAGCCCCCTTCCAGGGACTAGAGGAGAAGCCGGAGGAG
 GTGCTGCCCCATGGAGGACATCTCAAGGTGTACGACGAGACTTGCCCTCCTACACCCAG
 ACCACCGAGACTGGGGGGGTGACACTACACGAGGAGGAAGAGGTAGAGGTAGAGGAGGAG
 CAGCCAGCCGAAGCAGAGAAGGAGGAGGTGGTTGTGAGGGAGGAGGACAGGAGGTGGTG
 AATGTGGAGGGGCCAGAGCCGCGGAGGCCCGGATACGATAGAAGACACCCGACCGCTG
 AGCCGACTGAGCAAAGCCAGCAGCAGAGCCAGGTGAGATGATCTGACGGTATGA

>Tn11616001 Ensembl prediction. 19 nt removed at 5'-end to achieve initiation codon consistent with other connexins. Sequence extended by 543 nt in 3'-direction to reach stop codon (marked blue).

CCAGAATGGCAGAGTGAAGATGGGAGACTGGAACCTCCTCGGAGGGATTTGGAGGAGGT
 GCACATTCACCTCCACCATGGTGGGAAGATCTGGCTCACCATTTCTGTTATTTTCCGCAT
 GCTAGTCTCTGGCGTGGCGGCGGAGGACGTGTGGAACGACGAGCAGGCTGACTTCATCTG
 CAACACCGAGCAGCCGGATGCAAGAACGTGTGCTACGACCGGGCTTTTCCATCTCCCT
 CATTTCGCTACTGGGTGCTGCAAGTTATTTTCGTGTCCTCTCCCTCGCTGGTTTACATGGG
 CCACGCTCTGTACAGACTGCGGGCCCTGGAGAAAGCGCGCAGAGGAAGAGAGCGCTGCT
 CCGGAAGGAGCTGGAGATGGTGGGCGTGATTTGGCCGAAGCTAGGAAGAGGATGGAGTG
 TGAGGTGAAGCAGCTGGACAGGCCAGGCTGAACAAAGCCCCGCTCAGGGGATCCCTGTT
 ACGCAGCTACGTGGCCACGTTCTCACTCGCTCCGTTTTCGAAGTGGCCTTCATGACGGG
 CCAGTACCTTCTTACGATTTCACCTCTACCCGCTTTTCAAGTGCAGCGGGATCCTTG

TCCTAATGCCGTGGACTGTTATGTCTCCAGGCCACAGAGAAAAGCCTTTTCATGGTCTT
CATGCAATGCATCGCCGCCATTTCCCTCTTCTGAACATTTTGGAGATCGTGCATCTGGG
TTACAAGAAGATTAAACGGAGCATCTTGGATCTTTGCCGTTACGGGATGAACCTGGAGGA
CGACTTTGCTGTTAAGGACAAAAGAGAATCTGTGCACAGTTGTGCACCGCTGCGGCCCG
GAAGATGACCATTACGTTTTACCGGGCGGATGACAACGTGCTGCA

>Tn11616001 Our suggested modification.

ATGGGAGACTGGAACCTTCTCGAGGGATTTTGGAGGAGGTGCACATTCACCTCCACCATG
GTGGGGAAGATCTGGCTACCATTTCTGTTCAATTTCCGCATGCTAGTCTCGGCGTGGCG
GCGGAGGACGTGTGGAACGACGAGCAGGCTGACTTCATCTGCAACACCGAGCAGCCGGGA
TGCAGGAACGTGTGCTACGACCGGGCTTTCCCATCTCCCTCATTGCTACTGGGTGCTG
CAGGTTATTTTCGTGCTCTCCCTCGCTGGTTTACATGGGCCACGCTCTGTACAGACTG
CGGGCCCTGGAGAAAGCGCGGCAGAGGAAGAGAGCGCTGCTCCGGAAGGAGCTGGAGATG
GTGGGCGTGGATTTGGCCGAAGCTAGGAAGAGGATGGAGTGTGAGGTGAAGCAGCTGGAC
CAGGCCAGGCTGAACAAGCCCCGCTCAGGGGATCCCTGTTACGCACGTACGTGGCCAC
GTTCTCACTCGCTCGTTGTGCAAGTGGCTTTCATGACGGGCCAGTACCTTCTTTACGGA
TTTCACCTCTACCCGCTTTTCAAGTGCAGCGGGATCCTTGTCTAATGCCGTGGACTGT
TATGTCTCCAGGCCACAGAGAAAAGCCTTTTCATGGTCTTACGCAATGCATCGCCGCC
ATTTCCCTCTTCTGAACATTTTGGAGATCGTGCATCTGGGTTACAAGAAGATTAAACGG
AGCATCTTGGATCTTTCGCCGTTACGGGATGAACCTGGAGGACGACTTTGCTGTTAAGGAC
AAAAGAGAATCTGTGCACAGTTGTGCACCGCTGCGGCCCGGAAGATGACCATTACGTTT
TCACCGCGGATGACAACGTGCTGCAAGGAACGGGCGTCCAAACAATATCGCGCCAACG
GTTCTTCTCTTCTGAGCGAGGCGTCCACTCAACTGGATCTGGAGGAATCCAGATGTGCG
TCCACGGCTCAAAGGCTGCAAGTGCCTGCTGACCGCGCGGCTGGTGAAGCGGCTCG
CCCTCCGTCGTTCTCCAGAGCGGGAAAAGCAGAGCAGCGGAAGTGGCAGCCCGGATTGT
CCAAGACGCCACCAAAGCCGGCACACGGATCCACCTTCCCGCGCTGCCGGCGAACGCC
CCCAGGAGGCTTGGAGGCTCGTTCTTTCAATGCGCCACAGTCTGGAGGGGAAAAGC
TCTGACACCGACTCACGCGAGAGCGCCAGGGAGAGCAGCGAACCGCAGGGAAAACCTGGC
GCCTGCCGCCACAGCCTCAGCTCTGCGGCAGAGTACCGGACGACTCCAGCGCGGGTCC
GTGCACAGCCCCAGGCTGCCTTCTTCTGCTGCAAAACATCAATAACAAGCAAACACAGC
AGCGGTGCGGCTCCAGATCTGCAAAATTTAA

>Tn10470001 Ensembl prediction of gene with unknown function on opposite strand. Connexin sequence (reverse complement) is underlined. 91 nt removed to reach the initiation codon (gray). The predicted 99 nt intron included as part of reading frame (marked blue), with last 3 nt as part of second conserved domain (green). Sequence extended by 20 nt in 3'-direction to reach stop codon (marked blue).

GCAGGTGCGGTCGGGGCCCGCGGCTTCCGGCGCTGCTCGCCGCGACACCCGCGGTTCA
GGGGCGACGCGCGCGGCTGTTGCTCTCGCGGCTGTGCCGAGACGCGGAGCTCATTTGCG
CGCAGAGTCCGGGCTCTGGGGCTTCCGGGCTTCTTACTGTACAGATGGAAGAGCATCC
TCGCCCCGGGGGGCCCCACCGCCCCCGCGCGCCCCCGCGCAGCTTATGCCGCCCGCGC
CGCCGAGGTAGGAGTCCCCCGCTGTCCGAGCCGTTGCCCTCCGGCAGCGTGGAGCAGTTC
CAGGGTGGACTCACCTCCGGGACTTCTTGGTGATCCCCGACAGCGTCGGGCAGGAGCT
GGCGTCGGGCTCAAGGAGTGGGATCTCGGGGTCACCGGGGCCCCCTCGGGCTCCTC
CTGCTTATCGGCGGGAAGGAGTGCCTTTGCTGCTGGTGTGTTGGAGTTACTGTTTTG
TTCCGTGGGGCTCGGCACGCGCTCCTTGCGCTCCTTGTGCCGATCCGACTTGGTGCCAGG
GTCCCCCTTGGATCGGCACGAAAGCCGAGGAGCGTTGAGGAGCGGGTAGGTGGGCCCAT
GCCGTGCTTCTCCAGCAGCGGTGTATCCGCACGGCGCGGTGGGAATGTTGTCTTTCC
CGCCACGGAGATCGGGGTAAAGTCCAAGATGCCCTTCTTACCCCTTCTTAAATCCCAGG
TGGATAATCTCAAGGAGACTGAGGAAGGGGAGATGCAAGCAATGGCTGTCATGAACATC
ATGAAAACCGTCTTCTCGTGGGTCTGGACACGAAGCAGTCCACCACGTTTCGGGCACGGC
TCTCTCTCGACTGTAAAGAGGTTTCAGGCGGTGTCCGTACAGGATGTACTGACCATC
ATGAAGCTGACCTCCACCACCGAGCGATTACGATGTGGGCCACTTCCGATCCACCGCT
TCCATCTCCCGCGCAACGCCACCTTCTGCAAGTGGCGCTCCTTCCAGAGCTCGTAGC
TGATAAATGGCGTGACCATGTAGACCAAAGAGGGGAGGACAAAAATCACCTGGAGC
ACCCAGTATCGGATGAGGAGATGGGAAGGCTGGTCTGAACACACGTTGCGACAGCCG
GGCTGCTCCGTGTTGAGATGAAATCGGACTGCTCGTATTCCACACGCTCTGCGGGC
ACCCCCAGCACCAGCATCGGAAGATGAACAGGATGGTCAGCCAGATCTTGCCACCATG
GTGGAGTGAATGTGGACCTCTCCAAGATTCTCCAAGGAAGTTCCAGTCTCCCATTTTT
TCAGATTAGCAGCGCTACGACATGTAACACACGTTGGATCACTTAAATACAATTTTA
TTGTAACGATTAAGCGTTCAATCATA

>Tn10470001 Our suggested modification.

ATGGGAGACTGGAACCTTCTTGGAGGAATCTTGGAGGAGGTCCACATTCACCTCCACCATG
GTGGGCAAGATCTGGCTGACCATCTGTTTCATCTTCCGGATGCTGGTGTGGGGTTCGCC
GCAGAGGACGTGTGGAATGACGAGCAGTCCGATTTTCATCTGCAACACGAGCAGCCCGGC
TGTGCAACGTGTGTTACGACCAAGGCTTCCCATCTCCCTCATCCGATACTGGGTGCTC
CAGGTGATTTTTGTGCTCCCTTCTTTGGTCTACATGGGTACAGCCATTTATCAGCTA
CGAGCTCTGGAGAAGGAGCGCCACTGCAAGAAGGTGGCGTTGCGCCGGGAGATGGAAGCG
GTGGATGCGGAACTGGTGGAGGCGAGGAAGAGAATCGAGAAAGAGATGAGGCGACTGGAG
CAGGGCAAACCTCAACAAAGCCCCCTGAGGGGCTCCCTGCTGTGACTTACGTGGCCAC
ATCGTAATCTGCTCGGTGGTGGAGGTACGCTTTCATGATGGGTACGTACATCTGTACGGA
CACCGCTGAAACCTCTTTACAAGTGCAGAGAGAGCCGTGCCGAACGTGGTGGACTGC
TTGCTGTCAGACCCACGAGAGAACGAGACGTTTTTCATGATGTTTCATGCAAGCCATTGCTTGC
ATCTCCCTCTTCTCAGTCTCCTTGAGATTATCCACCTGGGATTAAAGAAGGTGAAGAA
GGGCATCTTGACTTTTACCCGCATCTGAAAAGAGGACCCGGATGA

>Tn06204001 Ensembl prediction. Sequence extended by 675 nt in 3'-direction to reach stop codon (marked blue).

ATGGGGGACTGGAATTTATTAGGAAGCATTTTAGAAGAAGTCCACATTTCGACCATC
 GTGGGGAAATCTGGCTGACCATTTCTTCATTTTCCGAATGCTTGTCTTGGCGTTGCG
 GCTGAGGACGTCTGGGACGACGAGCAAAGCGAATTTGTTTGCAATACGGAGCAGCCCGGG
 TGCAAAAACGTGTGCTACGACAGGCCCTTCCCGTCTCCCTGATCCGTTACTGGGTCTTG
 CAAATCATCTTCGTGTCTCCCGTCACTAGTCTACATGGGACATGCGCTGTATCGCCTG
 AGGACTCTCGAGAAAGAGAGACACAAGAGGAAAGCCTGCCTGAAGGCTGAGCTGGAGGGC
 ACAGACCCTGTCCAGGAGGACCACAGGAGGATTGAGCGAGAACTCAGAAAGCTAGATGAA
 CAGAAGAGGGTGAGGAAAGCTCCTCTAAGGGGCTCCTTGCTGCGCACATACGTCCTCCAT
 ATCTTAACCAGATCTGTTGTGGAGGTGGCTTTTATTATAGGACAATGTGCTCTGTACGGG
 CTCGGGCTGTGCGCCCTGTACCGATGTACCAGACCGCCATGCCCAACACCGTCGACTGC
 TTTGTCTCTCGGCCTACAGAGAAGAATGTTTTCATGGTTTTTCATGCTGGTTATCGCTGGC
 GTTTCGTTGGCGCTCAACATTCTGGAGATCTTGCACTCTGGGTGTGAAAAGGATTAACAA
 AGTTTGTATGGATATAAATACAGAGAC
 >Tn06204001 Our suggested modification. Potential intron start underlined.
 ATGGGGGACTGGAATTTATTAGGAAGCATTTTAGAAGAAGTCCACATTTCGACCATC
 GTGGGGAAATCTGGCTGACCATTTCTTCATTTTCCGAATGCTTGTCTTGGCGTTGCG
 GCTGAGGACGTCTGGGACGACGAGCAAAGCGAATTTGTTTGCAATACGGAGCAGCCCGGG
 TGCAAAAACGTGTGCTACGACAGGCCCTTCCCGTCTCCCTGATCCGTTACTGGGTCTTG
 CAAATCATCTTCGTGTCTCCCGTCACTAGTCTACATGGGACATGCGCTGTATCGCCTG
 AGGACTCTCGAGAAAGAGAGACACAAGAGGAAAGCCTGCCTGAAGGCTGAGCTGGAGGGC
 ACAGACCCTGTCCAGGAGGACCACAGGAGGATTGAGCGAGAACTCAGAAAGCTAGATGAA
 CAGAAGAGGGTGAGGAAAGCTCCTCTAAGGGGCTCCTTGCTGCGCACATACGTCCTCCAT
 ATCTTAACCAGATCTGTTGTGGAGGTGGCTTTTATTATAGGACAATGTGCTCTGTACGGG
 CTCGGGCTGTGCGCCCTGTACCGATGTACCAGACCGCCATGCCCAACACCGTCGACTGC
 TTTGTCTCTCGGCCTACAGAGAAGAATGTTTTCATGGTTTTTCATGCTGGTTATCGCTGGC
 GTTTCGTTGGCGCTCAACATTCTGGAGATCTTGCACTCTGGGTGTGAAAAGGATTAACAA
 AGTTTGTATGGATATAAATACAGAGACGACGAGAGCGTGTGTCTGCTCCAAGAAAACTCC
 ACCGTGCAGCAAGTTTGCCTTCTTGCTAGTTCTTCCCTCAGAGGCTGGTGCAGCTGACC
 CAAGTCACTTGTCTCCGCTCTGCCCAACACTAATGCGACGAGTCTGTCCCATCAGAACCAG
 GAAGGGTCCGGAACGCCAACAGCATCCCTCACACGCGTGCGTGCCCATCCAGGGTGTG
 CAGCAGGTGGCACCGGCCGAGCAGCATCGCCTTTCGGGACTGAGGAAGCCGTCTGTGCAGC
 AGCGAGGAATCCAGCGAGCCTCACGTGAAGCCCCAGTACGCCGGCCCTCGAGCCACGCTC
 GTG6CCAGCCACATGGAGATCCC6GCGGCCCTAAAGAACCCAGCACGGAAGCAGAGCAGA
 GTCAGCATTATAAAGAGCTCAGTGACATGAGCGACTCTGCCGAGAGCGAGCCCCACCTC
 CCAGCCC6CAAATGTAGCTTTATGTCTCG6GGCCTGTGCGGACGGAAGCTGTCCAACCCA
 TCCGACAGCGCCGACAGCCGAGCGGAACAGACTCGGAAGCCAGCACCTCAACCAATCA
 GAGAGCTCAGTGGTGACCCACCGCCTCCAGCCAGTGGCAGAAGGATGTCCATGGTTAGT
 GGGCCGGGTTCAATTTTCCACCACAAAACCTAGACACACGTGA

Suppl. Figure 6 *Connexins* from the tunicates *Ciona intestinalis* (Ci) and *Halocynthia pyriformis* (Hp). The *Ciona* sequences were mainly collected from the *Ciona* database (<http://genome.jgi-psf.org/ciona4/ciona4.home.html>), and partly from Ensembl, and the *Halocynthia Cx47* (Hp47) was taken from GenBank. When *Ciona* database predictions do not have a corresponding prediction in Ensembl, or the *Ciona* database predictions are not identical to the Ensembl predictions, we mainly use the gene identification number from the *Ciona* database. The corresponding Ensembl gene identification is given in parenthesis (Ci14820 should be read as ENSCING00000014820). The conserved domains are marked in yellow. Alternating black/blue sequences are exon predictions copied from Ensembl. In contrast to sequences given above in Suppl. Figs. 1 to 5, we here fully rely on the predictions given by the databases. Three gene predictions (Ci02449, Ci15999, Ci15377) in the present Ensembl *Ciona* genome database (v39, JGI2, Ensembl genebuild June 2006) are not included, as the sequences are very short. The longest of these (Ci15377) encodes a second conserved domain-like sequence.

>Ci0100141929 (Similar to Ci14820)

ATGTCCTGGCATTCTTGGAAACGGTGGGTCGAGAAGGCCAACCAACATTCTACATTAGTA
GGCAAGTTCTGGATTACATTTCTTATCGTTTGTGCGGATGGTGATCGTTGCTTCTGTTGGT
GATCGGGGTGTACACGATGAACAAAGTGAATTTAAATGCAACACACTTCAACCTGGGTGT
ACGAATGTATGTTTCAACAGATTCTCACCAGATATCCCATTTACGATTTTGGAGTTTCAA
ATATTGTTTCGTGGCAACTCCTTCGTAATGTTTCGTTGTTTATTCTGCTCATAAATCCAAG
CTGAAAAAACGTCAGTATCGCTTTTGATTCAAACCTCGAAACCATGGATCAGTGCATCG
CGTGTTCATTGGCAATGGACACGATTCCACTTAAGAAATGAAAAGTCAAAGAAAGTAAGC
GATTTTCGTGACGAAAAGTTTAAAGCCAACCAATCCTTTGGCAATGAAGGCTCATATACCA
GCATTTGCGCAACACTTTCGTTGGTATTAGCCAATGTGTTTGTTCGCTTTTGCATCGAA
GTTGGTTTCTCGTTTTCGAAAGTTCTTCTCTTCGGATTTCACGTTCTGAATTGTACAAA
TGTGAACGCCGGCGGTGTCCCAACACCGTCGATTGCTTTATATCTCGACCAATGGAGAAG
ACGATCTTCTTATGGTTCATGTTTCTATACAGCTGCATATGTGTAGCACTCAATCTAGTA
GAGTTTGTGTTATCTGGTATACGCCTATATAAAGAAGCAAGCACAGAAACGGGTGTCTGT
TCTAAGAGTAAAAAGAGGTTTCACTTCATCCTAATAAAGATAGTGGACTCAGTCGGAAT
GGATACACCGTTTACGAGGAGGGTGGATGAAACACGACATCGATCCTGTTGAACCTGAT
CTTGCTGAATATCAATTTAGGCAAAATGAATGACGACACTTCTTCTATCGGAAGTAGGCCCT
CCAAATATCGGGTGGGGTGGAGGTACACTGGACGCCAAGTTTCTCAACGTAAAGTTATT
GTCAGCGGTTTCAATGCTTTCACTGGTTCTGCACGGGATGCAAGCAGAGGTGACACTTCT
ACAAGAGATATGATGGCGTTTGGAGCTCGATTCCGATGACGAACGAGACATGATTGAAATG
GAGATAGAAAGATAGAACAAAGACGAAATGGATTTCCAACAAGATAATGATGCAATCCT
GAAGATGGTTTGGATGGGAACGACCAAGTAGATGCAGGGGATAATTGCATTTAA

>Ci0100145841 (similar to Ci08747)

ATGTCCTGGCACTTCGTAGAAGCTTGGTTACAAAAGGTCAACCAACATTTCGACCCTCATT
GGAAAGTTTGGATCACTTTCCTTATCGTGTTCGAATCGTGGTTGTGTCCAGTGTGGGT
GATCGTGTGATTTCGGACGAGCAAAAGCGAGTTTAAATGTAATCAATGCAGGTTGGGTGT
CCAAATTTATGTTTCAACAAGTTTCTCCGATATCCACATTCGATTCTGGTCTGTTTCAA
ATCATCATGGTGACTACGCCCTCAATCGTGTTCATGGTGTATTTCGGCTCATCAGATAAAA
AAGAAACCGCCCAACAAGGAAAAATCTTCGGTCACTGGTCCACCCCGTCACCGACCAAA
AAAGCAGAGACTGCTGCTGAAGCTATACAGCTCGAAGAGGAAGATGAGAAAAAGAAAAGA
AAAAAATCAACTATGCGGCTACGTCACTGGCAACACGAATGAACAGCAAGACTCCAGG
AAATTACGTCAATCTGTAATCCAGGAGATGTTTGGGGAAGAAAAGCGGCTTTTTCTCAG
CATTATGTCCTCTATACTTCTTCCGTAAGTATGAGAACTTTAGTGGAGATTCTATTTATT
GTACTTTCAGTATGCTCTGTTTGGATTCCGAGTATCAGAGCTCTACAAATGTACTGGCGAC
CCTTGTCACCAACATAGTGGATTGTTTGTATCCAGACCGATGGAGAAGACTATATTTCTT
TGGTTTATGTTTGTGATAGCCGTTATATGCCTTGCTCTAAACGTCGGTGAAATGTATTAT
CTCTTATGGCAGGTATGTGTTTCGTATTTTAAAGTAATGGGTTAA

>Ci0100142776 (similar to Ci03663)

ATGTCGTGGCATTGATCGAGAAAATACCTTGACGAAAGTTTCGAAGCAATCTACGTTTGT
GGGAAAGTTTGGGTGACTTTGCTTTTGTGTTTTCGAATAATTGCTGTGACTAGAAATAGGG
GATACAGTGTATGCTGTGATGAACAAGCGGCATTCAAATGCAACACTCGACAGCCGGGTTGC
GAGAATGTTTGTCTCAATAAGTTTTCCTCATATGAGATTTTGGGGTTTTCAG
ATAATTATTGTTGCAACACCAGCAATGATGTTTCATGGTTTATGCCGCACACACAGTTTCA
AAGCTACCAGGACCATGCAAGTTATCGGGCAGGCGTATAGATGCAGGTGTCGTCAGAAAA
AGACGAAGACGAGCTTTGATTAACCTGCCACAAAGTTTGTGAGAGTTGCTTCGACGGG
GCAGGCAGAGCGGCCAGTCACAAGTTGCTGAAAACCTTGAATCTCCGTCACGTTTTCG
TCGCAAGCAACAGCAGGCTTTCAACTCACCAAGAAAACGCTCAATTTCAAGTTGTATC
GCTTCCGCAATGTTTCGAGCGCCAGCGGCTTGTACGGGTTCAAGACAAAGATCGTGC
TTATGGAGCAACTGAACGCAAGATTGCGATTGCAATTTTCGTGCAAGTGCCTGCTACGAATG
CTACTGGAAATAGCGTTTCTGGTTCTGCAGTATAAATATTTAGTTTTAAGGTCCAGAG
CTTTACAGTGCCTTCAATGGCCTTGTCCAAACATGGTTGATTGCTTCATATCAAGACCA
CAAGAGAAATCCACGATGCTCATATTCATGTTTGGGGTCAACGCGTTGTGATTGTACTT

ACATGTTTTGAACTTATCAAACCTTXXTTACAGCGTTAAGAATATGATGAAAACCTTCAAAA
 AGAGATATGGTTGGAGAGGATAGAAAACGATCATCAAACGAACCACTTTCATAAACTT
 TTCGACTCATAGTCATCCAGAACATTGTAATAATATCTTCAGGCAATCAAGTGATGACGTC
 GATAGAAGCTCGTCATCGGATTCTTCATCAAGCAATAGTAGCAGTTCTGACGCATCAGCG
 TTGGCAGAAGCTTCGTTTAATTCGGAAGAACTCAGTGATATGTTGGATGAAACTGTGGAG
 GTGGTTGAATATAATGATCAAGCTGAAGCAAATGAGTTGTTTGAATTCTCTGTGTTTG
 TTGAAATGTTCTGATTTTGAATCCGTTTTTTTGTGA

>Ci0100132051

ATGAGTTGGGCTCTCTTGGAGAAGTTGTTGGGTGAAGTAGCCACCCATTCAACGCTGATC
 GGGAAATATTGGTTCTCGTTCTTATTCTGTTTCAGGGTGATTGTAGTTCTAAGCGTCGGG
 GACAAGGTATACAGCGACGAGCAGTCAAAGTTTGTGTGCAACACTTTGCAACCTGGATGC
 GACAATATTTGTTTCAACCAATTTAGTCCAATATCCCATATAAGGTTTGGGCGATGCAG
 ATTCTCCTCGTTTCCCTACCAAGTTTATATTCATAGTATGGGCGATGCATTCGCTAAGT
 ATAACGAAGGAGAAACGAGAAAAGTTGCAAAAAGACCCGGATGTAAACCCCACTACGT
 CACAGATACTCAGATGCACCGCCCAATTATAATTCAACGGCGGCCAAGAATGAAAACGAA
 ATCGTTTCAAAATCGAAGAAAAGTCGTTTCGAGTAACCGAAGAAGCGAAGCGATATATAAC
 AATGAACCTCCGGGTTATAACAAAGGGCTTCAGTAGCTGAGGTCGTTCCCGTGCAAACCT
 GTTATCCCTAAAGTTGATTGGAAGCAAATAAAGACCCGCCATATTGGCCGCATACGTT
 CTACATATTTACTTCCGAATGGCGATCGAAGCTTCATTGCTTTGCAATATATTCTG
 TTTCAATTCTACGTCCTGAGATGTATAGATGCCAAGGCTACCTTGTCCACAACGTGTT
 GAATGTTTCATATCACGACCCATGGAGAAAAACGTGTTCTGATATTCATGTACGTCATA
 TCCGGAGTTTGCCTCGTATGGATTTCATGGAGGTCAATTAATTTGGATTAGGAAGGTA
 AGTGTGGGTTGTTTTGGTCAAATATCACAGAAGATCTTCAGTTCTTGGTTGTCGCATTTCG
 GAGATAAAATCTCATTAGATAAACAACCTGCTGTTTTATAACCGGACGAGGTACCTGGGT
 GATTTTTTCGAATAA

>Ci0100147424 (similar to Ci13396)

ATGGCGTGGCACATTGTAGAAAAGCTGTTGGAACAAAGTGGCTGAACACTCGACGCTTGTT
 GGAAAGTTTTGGGTGACTTTCTTCTTCGTGCTTCGATTCTTGATGGTGGTTTCAATCGCT
 GATACCGTGTTTGGTGACGAGCAGGGGTCCTTCGTGTGCAATACACTGACGCTGTTGT
 GCGAATATTTGTTTCAATGATTTTGTGCAATTTCCCTTATCAGGCTATGGGCTTTACAG
 ATCCTCAGCGTGTCTCTCCACAATAATCTTCATGGTTTATACGGCCATAAGATGACA
 AAGATCAGCCAGGCTAAGAACTTCGGCAAGATGAAATGGACAAGAAAAGAAAATGAGA
 GAAGAAGCGAAGAGTTAAAGGAAGAAGCGAGGAAGAAGGCAATGGCGGACAGAAGAAG
 AGATTGGAAGAATCTGGTCTCCAGTCGTCGCAACTAACTTGACTCCGATACTCCGTCA
 CGACTCATGCTTGCTTACGTGGTTTCAGGTCGTGTGCAAGCTTGTGTTGAGATCGGCTTC
 ATGGTTCTGCAAGTTCAATATATATGATTTAAGTTCTGGGTTCCAGAGTTATTTAAGTGT
 GCAAGATGGCCATGCCCTAACGTTTGTGACTGCTTTGTGTCAAGGCCAAAAGAGAAGACG
 ATTATGTTGTGGGTGCTGTTTGTACCGGGCTTATAATGATCGTGCTCAACCTGATTGAG
 TTATACCATCTCGGTGCAAGGAAGATTCTAGAAGCATGGAAGCAAGGGGGCAAGATATT
 ACAAAGGAATACAGAGTATCCAGCAACATGCCGCACTTTGGCACCCATGGCTGCGAGTCC
 GGACAGAAAAGAGTTGACATGCGTGTGATGATGAATAA

>Ci0100135673 (similar to Ci13649)

ATGGCGTGGCATTGCTGGAAAAGATAGTAGAACATTCTGCTGCTCATTCCACATTATTA
 GGAAAATGGTGGATCGCATTCATGTTTGTATTTGATATAGTTGTGCGGAGCATTGGG
 GATGCTGTATATTCTGATGAACAGACGAGTTTGTGCAATACAAGACAACAGGGTGC
 AAGCAGATTTGCTACAACCTATTGCTCCGATATCTCACATTCGGTTTTGGGCTTTCCAA
 ATTCTTGGTACCGGTCTTCCAAGCGTGATCTTTATCATTTTTCGCACTCACAACTTGCG
 GTACTCCCAACCAAGCCATGGGAAGGACAGCCATAATGAGGTGTCGACACAACTTCA
 CCCAAACATACAACAGCAACGCGACTTAGAAAACGACAGTTAAGTTCCAAACATTGCGGA
 AGTTCAAATTGTTCCAAGCCAAGCTCATACAAGATGAAGCTGTTTTCTGATGCGACA
 GCACCCCTCCACCGGAATATGACGGGTTGAGACGAACAAAGCAGTTGGTTCTCAGAGCT
 TATATCTTGAACGTAATAACTCGTTGCATTATTGAGGGTGGGTTCTTGTATTTACAGTAT
 TATCTTTTATAATTGAACGGTCCCGGAGTTCTTCGAATGCATCCGATATCCCTGCCCAAA
 ATCGTGGAGTGCTACATATCGAGACCAATGGAAAAGACCGTCTTCTTTACTTTATGTTT
 GCGGTGTCAGCGCTGTGATTGTCATAAACATCATTGAGATTCAATTACCTGGGGTGGAA
 AGATTAAGCGAGCATTATTCGCAACAAACACAGCAGCTGCCGCTGCTCAATCATT
 GGAGCATTTCCTGTTGAAATTGACTTTCCACCGAAGCGTGACCGCGTTTCAGAGAAAAC
 AGAGAGCCCCCAACCAGACCTTCAATGGGTGCGATCATCACCGACGACCCGAAGCGGCA
 CCACAGGGTATTATATACCTGACGTCGACCCACCTTCAGAGAAATACGCTCTCCCGCA
 AACGGATTCCCAACATAAGAATGCGGGCTGCGAGCTTTACTACGTCAGAGAGTGACGTC
 ACAACCGATGCAACAATGTCGTCAATGGGAAGTGTTCGAAATGTTATGGAGGAAAACGTT
 CGGGTATATTTGTGA

>Ci0100132172 (similar to Ci03773 in previous version of Ensembl [v36 and earlier])

ATGGCATATTATCCACAGTTTGTGTTGAACAAAGTTCGACTTCATTCTACTTTCCAGGGAA
 GCTATGGATTATAATGATGTTTCATATTCGTATCGTGGTTGTGGCAAGGATAGGAGATAT
 GGTCTATCACGATGAACAGTCCCACTTCGTTTGAACACGTTATCACCGGTTGTTTCAAA
 CGTTTGTTCACGCAATTTAGCCCGATATCTCAACTACGTTTCTGGTCATTGATGGTTTT
 GGTCTGATCTTCCACAGGGGTTTTATTCTGCATCTACGCAAGCCACAAGATATACACGC
 TTATGTTAAGTTCCCTCGCCAGGAAGAAGGAGCGCCCAAGAGTCGCACTTGGGACAC
 GGCTTCGGAAATTGACCGCGGGTCTGAATACGGTCGTCCAATCGCGAAGCGAAAAGGAG
 GAAAGGCGATTCTCCCTGCGTTACCGGCCACCCACCGCCCACTTTACGTTGT
 TAAGCGCTAAACTCGAAAAATGAAAAAATGACAGAAGACGAAAACCGAAATTTTC
 TCGCCGCAACCCGACGCTTATAGAGCGTATTGGTGGCACGTGGTCATTGCAACTGTCCT

AGAGATTGGGTTCTTGGTTGGTCAGTATTATTTGTACGGTTGGTTCGTACCTGAATTGTT
TGAATGCAAGAGATGGCCTTGCCCCAAAACCGTGGATTGCTTCGTGTCAAGGCCGATGGA
GAAAACCTTGCTTTTGGTGGCTCATGTTTGGTTTGGGTGCCATTGCTTGTGTGTTAAGCAT
TGGGGAGTTCTGGGCACTTGGTTGGGAAAGTTCAAGTATGCT

>Cigrail.250.16.1 (similar to Ci08464 in previous version of Ensembl[v36 and earlier])

ATGTATTGTTTGTAGCCTTAAGAACATACAGTGTTTATATATCAACGGAAATCGCATTTC
AGTGGTTTTATCATTTGCTGCTACAATTGGTTAATCATGGTAGCAAGCATTGGGGACACT
GTGTACGGCGATGAACAATCGTCATTTAAATGCAACACGTTACAACCGGGTTGTGAAAAC
GTTTGTCTATAACCATTTTGTCTCCCATCTCCAGATCAGGTTCTGGGCTGTGCAGATTCTC
TTCGTTTGGCACTCCAAGTATTATGTTTATAATCTATGCAATGCATAGGATGACAACAATT
CCACCAAGCCATGCGAAAAGTAGTTTCAGATGGAGCTGACGACGAATCAGAAAAACCAAAG
AAGATCGAAACAGCGTCGTCAAGTACTTCTTCTCTAAAGCTGGACTTCCCGGTTATGAC
GAAATAATGCACAACGGGAAGCAGAGTGGGCAAGGTAAAAACGGCAAACTACTCAAGAAA
GTAGAGGAGGCAACGATCAACCAGAAAGTGGAAATAAGTTTGACGGATCTCGTAGTGAAAGT
ATTTATGTAAACCGACCACAAAAGAAACGCAACGCAAAAGCCCAGCCAAGAAAAGT
CACCAAGCTAAACGAAAAGCAGTTTACCAGCAAGATGGCAAAACAGAAATCGTTAAAAACA
CCAGGAATCGTTCCGCGCTTATGTGGTTCAAGCGTTTCTTAGAACGGCTATAGAAGCTGTT
TTCTTATATTTACAGTACGACATGTACAAGTTCGATGTTCTCGAATTGTATGTTTGCAAT
GGCTTTCTTGCCCCAGAACCGTGGAAATGTTTTGTATCAAGACCCATGGAGAAAACCTCTG
TTTCTGCGATTCTGTAGCCCATGACAATCGTATCTTTGCTCTTGAACCTGATTGAAATC
TACCAGTTGATATCAAGTGGGTACGACGGTCGTGCTGTCGCTCTGGCAACTCTAATGCA
GTGAGCTTGGATGAGGTACCAGTGGAGGAGGTTGTTATACAGAGTGGGGAAGATGGG
GCGATCTTTAGAAGAGCAGCAATGCAGAACACGTTCCAGATTGGGACCAATTCCCCGC
TCCTATGACAATTCTTGGGGGAGTCGTCCGGCGACGAGAGCATGGACAGTACTATCGT
TTTTACGGTCGCTAA

>Ci0100142372 (similar to Ci17404 and Ci15377)

ATGGGTTGGCATATTGTGCAAAAGTTGGTGGAAGGAGTAGGATCTCATTCAACTCAGATT
GGAAAATATTGGATTACTTCTTTCTTCTGTTGAGAAAATTTTACATCTTGTCTTCCC
ACAGAGGCGACATGGAGTGGGGATCAAGGCGGGATGGTTTGCAATACCTTGTCTCTGGT
TGTAAGAACGTTTGTTCACGAATTTAGTCCAATTTCTCCTATCCGTTACTGGTTTTTT
CAAATCTTGCCGTGTGCTTGGCGTCGATAATATTCAGTATTTACACAACCCACAAGTTA
GCGCTGGTCACAAGAGCTATTCGGCAAAAGAAAAGATGGATCGCCAAGGAAAAGCTAAAA
AAAAAAGATGACAAGACGACGAAGAAGTTGTTGCCACGAACTACAACAGATCATCCGCT
AAACTCTTCTTAGCATATTTCTTTATGGTTTTTATGAGATTGGTAATCGAAGTTGGATTTC
ATGGTTGGCCAATACTTTCTCTATACCTACAAGTTCGTTGTACCTGAGTTGTGGCAATGT
AGTCACTGGCCTTGTCCCAATGTGGTCGATTGCTTCGTTTCAAGGCCGGAAGGAGAAGACA
ATTATGATCTGCACGTTTACGTCACCTGCTGTTGTCATCATGGTCGTTCTCAATGTGATCGAG
TTATATCACATCGGTGGCAACGTTGGTAAAGCTTGGAGCAATAGAAGTAATGATATTACG
AAGAAGTGTGTTGGGAAAAGTGC CGGCCCTAATTTGAGGCCCGCGGTGGCATGAGACAT
TGGGTGCCGGGGGAGAGGGTATGTACCAACAATATCCGAGTGGTTACCCCGCTGCTGTA
GGGATACCTACACTCTATACAAGCTACGGTAGATCTCGACGTAAAGGTAGAAAATCAGGC
TTACCAAGAACCAGATTGGCGGGGACGAATACTATTACTGA

>Ci09598 (Ensembl gene id; identical to Ciona database genewise.21.185.1)

ATGTCGTGGAAATATTTGGACGATGTTCTTCTCTCGTCTCAGTAAGAAGTCTGACTCTGCTC
GGGAAATATTGGATTGCGACTGTTTTGCTTCTGAGACTTGTGGTGTGACGGTGTGTTGGT
GAAAGTGTGTTGGAGCGACGAACAACGAGAGTTTACATGCAACACTAAACAACCTGGTTGC
GAAAATGTGTGTTTAAATACATTTGCTCCTGTTTCAAAATGTTCAATGTGGGGGTTGCGAG
ATTCTTGCTGCTGCTGTTCTTCTGGAGCTTTATACTTTGCTACGTCATGCACGTGGTTTCC
CGAAAAAAGCTTCTGGTGATGACATTTTCTGAAAAACGACAAAGTCAAAACATCCGAT
CTTTCCACAGCCGCAACGAACGAGTCCGCACAGCGAACAAGGTGGAAAGTACTTACA
AACCATGTTATACCATCTGAGCTCCATAAAACGAGAGGACAATGCATTTGGCGGCTTTAC
CTGCTGCAGATATTAGTCAATGCTTATCGAATGCGTTTTTATAATGCTGCAGTACACC
GTTTACTTCTACAGATGACAGTACCTCAGGTGTTCCGGTGCCAAAGTTTGGCCATGCCCA
CACACCGTGGATTGCTTCATTAGTCGACCACAGGAGAAAACCGTTTTCTTCCGCTACATG
TAGCTCATTTCCGCTTCTCGATTCTTCTCAGCATCGGGGAGTTGTTATACATCGGCTGG
ATTGGCGCAAGACGAGCTTGCAGTCACAGGAAGGCACCAAGCTCACGGGAAAGCCCTT
CACCTAAATACATTGACGAATGGAGAACTCGTGATATCAAAGTGATGCCTCGAGGTGCT
TTATTGGATCGTGTCTTCTCATCAGAATGCAGCGACTCGCAACCCCCAACCCCATATCT
AAACATGAGGATTACGTCGATATGCAATCCAACAACCTCATCAGTTATTTGTAG

>Ci0100135879 (identical to Ci07798)

ATGTGGCATCTTTTAGAAAAGATTTCTGAACAGGTGACCAAATACAGTACTTTCTCTGGG
CAGCTGTGGTACATTTCTGTTTTCTGATTACAGACTGATTGTAGTCGTAAACGATCGGCGGC
GCAGTTTACGGTGATGAACAGTCTGCATTCAAATGTAGTACAGATGTCGTGGGTGCCAG
AACATATGCTACAAACCAATTCGCAAGGTGTCATATGAGGTTTTGGGCGTTTCAAGTTG
CTGGCCGTCGCTACCCCGTCAGTATTCTTCTTCTACTCAATGTATATGTCTGGTGAA
ATTCAGAAAGCTAAAAAATCTGAAGATCATCTTCGACAAGTGAAGAAGCGGACCAAGGAT
CATTTGCTACCGAAAAACGCTCAAGACGAGAAAGACTTCAAGAAAATGTCAAAACGTCGC
AAGAGCATCGGGCACGTAAAAAGTCCGAAAAATTTATCTGGAACGAATGAAAGAGATT
CCATACACCAACAAAATCCATTACGATATTACTTGAGCATCATCATCCGCTTATAATG
GAAGGAGTTTTCTATACATTGGCTACACCCTGTTTAAACTCATGGACCCCAATGCCAA
CACACTCCCACAGACCCCATGTGTCCAGGCAGCAACAACCAATTTACTTTTTATGGATA
AGAGTGCCAGCAATGTACAAATGCGAGGGTAATCCAGGAAGTACGATATACACAGCTTGT
ACCATGCACGTTAGAGATAATGGATTCTGCTTCTGCTTGTGCTCGACCTTATGAAAAG

ACCGTCTTTCTCTGTACATGAGCATTCTAAGCGGAATTTGCTTCGTAATTTCAAGCTCTC
GAACTTGTTATTGTCTCAATTCGACGATTCTGGAAAAACGCAAGCCCACCCACAAGCAT
CTCTTCGGTAACAGCAGTAGCCGACCACGAAACGAGTCTGTATTTTATTCTACTGCCCG
ACAGCTCCAAGCAAACTTTGAGGAAGGAAATTTCTACCCATCTTTAGCTGTACCTGTC
CGCATAGCGAACGGCAAAATACCAAGGGTACCAAAACAGCGACACAGCGACAAAGAGAGTC
CAGATTCATGAGCAAGACCAACATAACTTGGAAAAGGAAAACATGTCCGAAACTGCATCA
GTTAAAAGCTATGTTTAA

>Ci0100132269

ATGTTTAAAGTGGAACTTTTCAATAATGAGCTTCCGAGATATGGTCTCAACTATGCAACCTTT
GCTGGGAAGGTTTGGTACGTACTGGTCTATTTGTTTCAGGCTAGTAGTGGTCTGTCTCGCTG
GGTACCGCTGTGTATTACGATGAGCAGGAGCAATTTTCGATGCAGCACTGTGATTGTGGGT
TGTGAAATGGTATGCTATGACGAGTTCCGCAAGATCTCCACCTCCGGTTCTGGGCATTT
CAAATTCGGCTCTTAATTTCCGGTTGTGTTGTTCCACTTCTACGCTGTTTTATGAAAGC
GGAAAAATAAAGAACAGGAAGCGAAAGAATAGAAACAAGGAAAAATAGGTCGTGTGAAAA
TTAAACTATCCATCGATGACTCGAGGAAATATAAAGAAGAGAAAAGGACACGGAAGCGG
TTTGATACACAACGAATACGGACGGACAGAAGCCATAGTAATATCATGGAAGATAGAAATT
GCATACCTTTGTTGCTACTTTACTTAGAGGATTGTTAGAAGCTATGTTCACTTTCTTCGCT
TTTATTTTGTATGCTCAAACTTATGCAAACTGGTTCGTATACTTTGTTGGACAGATT
TGGGTTGAAGTACCACGATGTATCAGTGCAGGGGCGAGTGCATATATGTCGCATGTGAA
CAACATATGCTACTGGAACCGAGGTGGCTACGTACCTTGTGGGTGTCAAGGCCATGG
GAGAAGACAATATTCGTCGCTACATGAACATTATGAGCGCGGTTTGTGGTGTCTGACA
TTCATGGAGTCTGTCTATCTACCCACGAAATTTGTGCAAAATATTTTACCAAGAAACAA
AAAAGAACAGTTCTTATGTGAGTTCACAGAATACACATAAGCTCAAAGCATTGAAAGCT
GCCCTGAAGGACTACCCGCCGGGAAATTTCTATCACATTTCAAAGCAAACAAGTTTCAGA
TTCGCAACTCATTTCTCTCTGAAGAGTGTAGTCATGAGAAAGGAGTATCTTAAGTATC
ACCAAGATATAA

>Ci0100142421 (similar to Ci07796)

ATGTGGCCATTACTACAAAAATCTCGGAACAGGTAAACGAGGTACGAAACCGTTGTAGGG
CGGTTGGGCTACTTGTGTTTCTCTACCGCTTATAGTGGTAATTACTGTGCGGGGC
GCAGTGTACGGAGATGAGCAGTCTGCTTTTAAATGTAGTCAAAAAATCCAGGGTGTGAG
AACGCTTGCTATAACAGCTTTGCCCAAATATCTCATTTACGGTTTGGGCATTTTCAGATA
CTTGGCATTACAACACCGACTGTGTTTTCAACTTATACTCGAGCTTTGTTCTGGCGAG
ATCCGTAATTAATCACTGTGAGGAACAACACGAAATGTGGAAGATAAAAGAGACGGG
CAACTTACACTAACTAACACATTGTGAGAGAGGAAATCGGAAAAATGACGCGAAGAAAA
AGAGGGATTGGTAACCTGAAAGTGAACTCGTTTACTCTGACAACGAAATAAAGAGGTT
CCAGTTACGAACAAAATCAAAGCAGCGTACTACGTGAGCGTATTATTGCGATTGGTTTTA
GAAGGGGTTTCACTTATCTTGGATATGTTTTGTACAATATCCCGATATATGCGTCTTGT
GTCGGCCCAAATGCAACGCAAGATCCACTATTGCAAGTGTATTATGGATGAAAGTGCCT
TCAGAATACAGATGCATTGGCGGCCAGGATCCGAAAGTTATAACATATGTAACATGCAT
TATGGGCTAATGCTGAACGAGGCTACGTTCCCTGTATAACTTCAAGAGCAACTGAAAAG
ACAGTCTTTCTCGTTATCTGACCATCTGAGTGTCTGTTTTCATACTATCTGCTTTG
GAACTTATTTATCTGACTATCAAATCCATTTGGAAAAAAGCTCAAATAACAGCAGATCC
AACCTGCAAAAAATCAATACACCTGAGCAGTTCTATTTCCCTTCGCCACCCCAATAGCC
ACAGCACCTAGCAAACTGTACGACGAATTCAATTCTTCTGGTCTGGTTGTACAGGAAAA
AATGCGAGCAAGGAGCTGAGAAGAAGTGAGTCGAAAGATGGCGACAGAGAGTCGTGTAA
TCGTTGCGTAGCTACGTGTGA

>Ci0100132251

ATGTGGAACATATTGAAAAAGTGCTGAACAGGTTTCGTTATATGCAACTTTGACCGGG
CAAGTTTGGTACTTGTCTGTATATCTCTTTCGTTTGATCGTCTGTTGTTACTATCGGCAGC
TCCGTATACAGCGACGAACAGTCGGCGTTCCGTTGCAGCACAAAAGTAATAGGATGCGAC
AACGTTTGCTATGACAAGTTTGCTAAGATCTCATATTCGGTTCTGGGCATTTCAATTG
CTTGCCCTCAGTGCTCCAGTGATCGTCTTCCATTTCTACACGATCCACGTTAAAGGTAGG
ATCGAAAAACTGAAAGTGAAAGAGCAAGAGGCGATAGAAAGGGAGGAGCAATACCAAAAA
GAGCTGACAAAGACGGAAGACACGATGTATGCGAACACTCAACTTGAAGAGATAACTCTT
ACACGCGGAATCAAAGTCGCTTACTTTCTCTGTGCTCACACGATTCTGTATGGAGATT
ATATTTCTCTTCTTGCCTATGACTTGTTCGATTCTGCTGAATACCCGCATACACCTAAC
GATGGCCGAATAGTGTGGGCGTGTGGCCAGCATATGCTGCCAGGTAATTCCGATGCTTAT
ATTCCGTGTTGGGTTTCAAGGCCGTGGGAGAAAATATTTTCTCAGGTACATGAACACA
CTCAGTGTCAATTTGCTTTTCTTGCATTGCGGAAATATCCAAATAGGTATACGCGGG
GGACTTGAGTTGGCCAAGAAAAAGTCAAAGCCTCGCCAAGAATTACCTGAAATGAAATG
CTTCTTTACAGCGGATACAAACCTAAACCGATACTAAAAACCCCAAGCATTGGTACCCTC
GCCACGGAACAGAAAGCCAGCATCACACCCCGACCCACAACACCACCAACAGAAACA
TCAACAGCTAACACCGAGACAAGTCCGAAGGAGCTTTTAGATGCACCGCGGAAGGAAAGT
GGCAGCAAGCACACTGCAATCAAAGGGCAACGACCCAGCAATCGAAGTGATCTTGATCAT
AATGAAAGGCGATCCATAACAATGGACGATCTAAAAACAGCAATTTTATTAGGTATGTAT
ACCAGCATCGTTACTTTTCATACGCATCACATCGGCTGTGTGAACAATAATTCTACTTATC
TTTCTTCCACAGCGAGAAAGGAAAGCCAAAGCCAAACACAGCTACACGGAGACCTAGCA
AAAACAACGAACACAGCAGCAAAAACTCACTTTCAAATTCATACGGTACCCAATCAC
TTAACCGACCAAAGAAAGTGA

>Hp47 AY380580

ATGGCGTGGCATATACTACACGGTCTGTTGGAACAGGTACAGAGTTCAATCAACATTTCCG
GGCAAGTTATGGATCATTGTAAATGTTTCATATTTGCAATCGTAGTTGTCGCGAGGATAGGA
GACAACGTGTATCACGATGAACAGGCAAAATTTGTTTGAACGTGCTGACACCTGGTTGT

GAAAAATGTTTGTTCATAGATTCAGTCCAATATCACAACTCAGATATTGGTCACTCATG
ATACTGGTTGTGTCTACTCCAGCTATATTATTTTTTCTCTACGCGACTCACATAATCTAT
CACGCGGGAATCAAATATCCATCGGAAGAAGATGAAGAAAAAAGGAGATTATGACGACA
CGATCAGATCAGTCGTATATTCATCCGCCAAGTCTAAAACGGCGGAGAGATGAGGGATTT
TTACCGAACGGGACAACGGGTCCATTGTTCTGAAGTTAAAAAGATCGGGAACTCGGACGGG
AAATACAAAACAGCTAAATTCGGTAACAACAGGCGAAGTGACGATATCAGACGCAGACGT
CATGTAATACGTCATGACGGTGACGTCAAGAGAAAATTATCAACGGACAATAATGAATCG
AATGAGGAAGTGATATAGATTACAGAAACAAAGTTCCAATTTTCGTGGATCCTCACGGA
CAACGAGTTGAATTCATGAAAGAAGTTCACAGGAATCGCGCGGAGGATGACGCGATGGCA
GAAAGAAAATTCGGAAGAGGATGTTGGAACATCCCGAACTATACAGAGCTTATTGGTGG
CACGTGTTTTCTACGAACATTATGCGAGGTTGGATTCTTGTTGGCCAATATTACTTATAC
GGTTGGTCAGTTCATGAATTATATGAATGTAATTACTGGCCTTGCCCCAAAACAGTCGAT
TGCTTCGTGTCAAGACCACAAGAGAAAACATGTTTGTTATGGCTCATGTATGTTCTAGGT
ATAATCGCATGTCTACTAAGTGTGCGCCGAGTTCTGGACTCTTGGTATCACTCGACTCAAG
TACGCGTTCCTATGTTGTAGCCACGACGAGTCGAGAGAAGAACGAAGACGAAGGCTTGCA
TTACTCGAGACACATTCAACTCATCACGATAAAGAGAATGACCCACTTCATCCAAGACTT
ATCACAGGAGCAATGAGCTTTGACAGCTTGAGTAGCGAGACTACTGTTGAGAGCGTATGA

Suppl. Figure 7 Coincidence of SNPs in human *Cx43* with sequence differences of *Cx43* and *Cx43p1*. The functional human *Cx43* (Hs43, NM_000165) and human *Cx43* pseudogene 1 (Hs43p1, NG_003029) were aligned by ClustalW, and SNPs were marked into the Hs43 sequence in the line called Hs43-SNP. The identity line (*) indicates identities and differences between Hs43 and Hs43p1. The GeneBank SNPs are marked in **yellow**, and the standard degeneration code is used (Y, C+T; R, A+G; M, A+C; S, C+G; W, A+T). The following dbSNP numbers were used (in sequence from 5' to 3'): 2227886, 2227873, 2227875, 2227878, 2228966, 2227889, 2228960, 2228961, 2228962, 2228963, 2228964, 2228965, 2227888, 2227887, 17653265, 2227879, 2227880, 2228974, 2227881, 2227884, 2228975, 2227882, 17852395, 2227883, 2227885.

```

Hs43p1      ATGCGTGACTGGAGCACTTTAGGCAAACCTCTTGACAAGGTTCAAGCCTACTTAACTGCT 60
Hs43        ATGGGTGACTGGAGCGCCTTAGGCAAACCTCTTGACAAGGTTCAAGCCTACTCAACTGCT 60
Hs43-SNP    ATGGGTGACTGGAGCGCCTTAGGCAAACCTCTTGACAAGGTTCAAGCCTACTCAACTGCT 60
            ***      ***** *      *****

Hs43p1      GGAGGGAAGGTGTGGCTGTCACTTTTCATTTCCAAATCTGCTGCTGAAGACAGTG 120
Hs43        GGAGGGAAGGTGTGGCTGTCACTTTTCATTTCCGAATCTGCTGCTGGGGACAGCG 120
Hs43-SNP    GGAGGGAAGGTGTGGCTGTCACTTTTCATTTCCGAATCTGCTGCTGGGGACAGCG 120
            *****      *****

Hs43p1      GTTGAGTCAGCCTGGGGAGATGAGCAGTCTGCCTTTCGTTGTAACACTCAGCAACCTGGT 180
Hs43        GTTGAGTCAGCCTGGGGAGATGAGCAGTCTGCCTTTCGTTGTAACACTCAGCAACCTGGT 180
Hs43-SNP    GTTGAGTCAGCCTGGGGAGATGAGCAGTCTGCCTTTCGTTGTAACACTCAGCAACCTGGT 180
            *****

Hs43p1      TGTGAAATGTCTGCTATGACAAGTCTTTCCCAATCTCTCATGTGCGCTTCTGGGTCCTG 240
Hs43        TGTGAAATGTCTGCTATGACAAGTCTTTCCCAATCTCTCATGTGCGCTTCTGGGTCCTG 240
Hs43-SNP    TGTGAAATGTCTGCTATGACAAGTCTTTCCCAATCTCTCATGTGCGCTTCTGGGTCCTG 240
            *****

Hs43p1      CAGATCATATTTGTGTCTGTACCCATACTCTTGACCTGGCTCATGTGTTCTATGTGATG 300
Hs43        CAGATCATATTTGTGTCTGTACCCACACTCTTGACCTGGCTCATGTGTTCTATGTGATG 300
Hs43-SNP    CAGATCATATTTGTGTCTGTACCCAYACTCTTGACCTGGCTCATGTGTTCTATGTGATG 300
            *****

Hs43p1      CAAAAGGAAGAGAACTGAACAAGAAAGAGGAGGAACCAAGGTTGCCCAAACCTGATGGT 360
Hs43        CGAAAAGGAAGAGAACTGAACAAGAAAGAGGAAGAACTCAAGGTTGCCCAAACCTGATGGT 360
Hs43-SNP    CAAAAGGAAGAGAACTGAACAAGAAAGAGGARGAACCAAGGTTGCCCAAACCTGATGGT 360
            *      *****

Hs43p1      GTCAATGTGGAGATGCACTTGAAGCAGATTGAGATAAAGAAGTTCAAGTATGGTATTGAA 420
Hs43        GTCAATGTGGACATGCACTTGAAGCAGATTGAGATAAAGAAGTTCAAGTACGGTATTGAA 420
Hs43-SNP    GTCAATGTGGAATGCACTTGAAGCAGATTGAGATAAAGAAGTTCAAGTAYGGTATTGAA 420
            *****

Hs43p1      GAGCATGGTAAGGTGAAAATGCAAGGGGGGTTGCTGCGAACCTACATCATCAGTATCCTC 480
Hs43        GAGCATGGTAAGGTGAAAATGCGAGGGGGTTGCTGCGAACCTACATCATCAGTATCCTC 480
Hs43-SNP    GAGCATGGTAAGGTGAAAATGCRAGGGGGTTGCTGCGAACCTACATCATCAGTATCCTC 480
            *****

Hs43p1      TTCAAGTCTATCTTTGAGGTGACCTT - - - GCTGATCCAGTGGTACATCTATGGATTCAGC 537
Hs43        TTCAAGTCTATCTTTGAGGTGGCCTTCTTCTGATCCAGTGGTACATCTATGGATTCAGC 540
Hs43-SNP    TTCAAGTCTATCTTTGAGGTGRCCTTCTTCTGATCCAGTGGTACATCTATGGATTCAGC 540
            *****

Hs43p1      TTGAGTGCTGTTACACTTGCAAAAGAGATCCCTGCCACATCAGGTGGACTGCTTCCTC 597
Hs43        TTGAGTGCTGTTACACTTGCAAAAGAGATCCCTGCCACATCAGGTGGACTGTTTCCTC 600
Hs43-SNP    TTGAGTGCTGTTACACTTGCAAAAGAGATCCCTGCCACATCAGGTGGACTGYTTCCTC 600
            *****

Hs43p1      TCTTGCCCATGGAGAAAACCATCTTCATCATCTTCATGCTGGTGGTGTCTTGGTGTCC 657
Hs43        TCTCGCCCAACGGAGAAAACCATCTTCATCATCTTCATGCTGGTGGTGTCTTGGTGTCC 660
Hs43-SNP    TCTYGCCCAAYGGAGAAAACCATCTTCATCATCTTCATGCTGGTGGTGTCTTGGTGTCC 660
            ***      *****

Hs43p1      CTGGCCTTGAATATCATTGAACCTCTTCTACGTTTTCTCAAGGGCGTTAAGGATTGGGTT 717
Hs43        CTGGCCTTGAATATCATTGAACCTCTTCTATGTTTTCTCAAGGGCGTTAAGGATCGGGTT 720
Hs43-SNP    CTGGCCTTGAATATCATTGAACCTCTTCTAYGTTTTCTCAAGGGCGTTAAGGATYGGGTT 720
            *****

```

Hs43p1	AAGGGAAAGAGCGACCCCTTACCATGCGACCACTGGCCCGCTGAGCCCCGCCAAAGACTGT	777
Hs43	AAGGGAAAGAGCGACCCCTTACCATGCGACCACTGGTGCGCTGAGCCCTGCCAAAGACTGT	780
Hs43-SNP	AAGGGAAAGAGCGACCCCTTACCATGCGACCACTGGTGCGCTGAGCCCYGCCAAAGACTGT	780

Hs43p1	GGGTCTCAAAAATATGCTTATTTCAATGGCTGCTCCTACCAACTGCTCCCCTCTCGCCT	837
Hs43	GGGTCTCAAAAATATGCTTATTTCAATGGCTGCTCCTACCAACCGCTCCCCTCTCGCCT	840
Hs43-SNP	GGGTCTCAAAAATATGCTTATTTCAATGGCTGCTCCTACCAACYGCTCCCCTCTCGCCT	840

Hs43p1	ATGTCTCTTCCTGGGTACAAGCTGGTTAATGGCGACAGAAACAATTCTTCTTGCCGCAAT	897
Hs43	ATGTCTCTTCCTGGGTACAAGCTGGTTACTGGCGACAGAAACAATTCTTCTTGCCGCAAT	900
Hs43-SNP	ATGTCTCYTCCTGGGTACAAGCTGGTTAMTGGCGACAGAAACAATTCTTCTTGCCGCAAT	900

Hs43p1	TACGACAAGCAAGCAAGTGAGCAAACTGGGCTAATTACAGTGCAGAACAAAATCGAATG	957
Hs43	TACAACAAGCAAGCAAGTGAGCAAACTGGGCTAATTACAGTGCAGAACAAAATCGAATG	960
Hs43-SNP	TACRACAAGCAAGYAAGTGAGCAAACTGGGCTAATTACAGTGCAGAACAAAATCGAATG	960

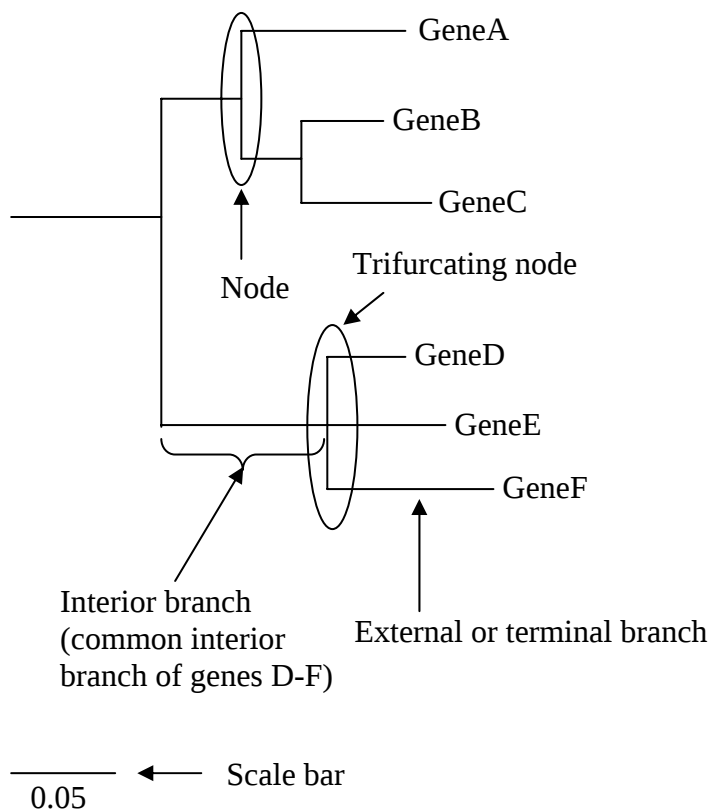
Hs43p1	GGGCAGGCGGGAAGCACCATCTCTAACTCCCACGCACAGCCTTTTGATTCCCCGATGAT	1017
Hs43	GGGCAGGCGGGAAGCACCATCTCTAACTCCCATGCACAGCCTTTTGATTCCCCGATGAT	1020
Hs43-SNP	GGGCAGGCGGGAAGCACCATCTCTAACTCCCAYGACAGCCTTTTGATTCCCCGATGAT	1020

Hs43p1	AACCAGAATTCTAAAAAAGTAGCTGCTGGACATGAACTACAGCCACTAGCCATTGTGGAC	1077
Hs43	AACCAGAATTCTAAAAAAGTAGCTGCTGGACATGAACTACAGCCACTAGCCATTGTGGAC	1080
Hs43-SNP	AACCAGAATTCWAAAAAAGTAGCTGCTGGACATGAAYTACAGCCACTAGCCATTGTGGAC	1080

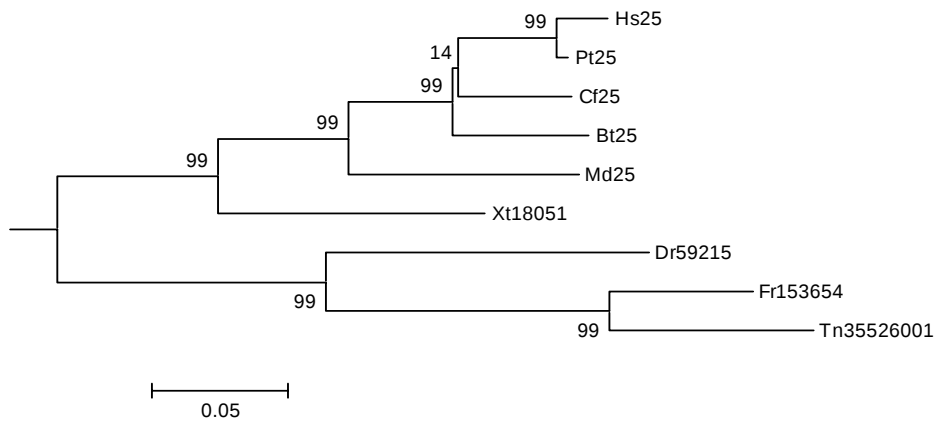
Hs43p1	CAGCAACCTTCAAGCAGAGCCAGCAGTCGTGCCAGCAGCAGGCCTCAGCCTGATGACCTG	1137
Hs43	CAGCGACCTTCAAGCAGAGCCAGCAGTCGTGCCAGCAGCAGACCTCGGCCTGATGACCTG	1140
Hs43-SNP	CAGCRACCTTCAAGCAGAGCCAGCAGTCGTGCCAGCAGCAGACCTCGGCCTGATGACCTG	1140

Hs43p1	GAGATCTAG	1146
Hs43	GAGATCTAG	1149
Hs43-SNP	GAGATCTAG	1149

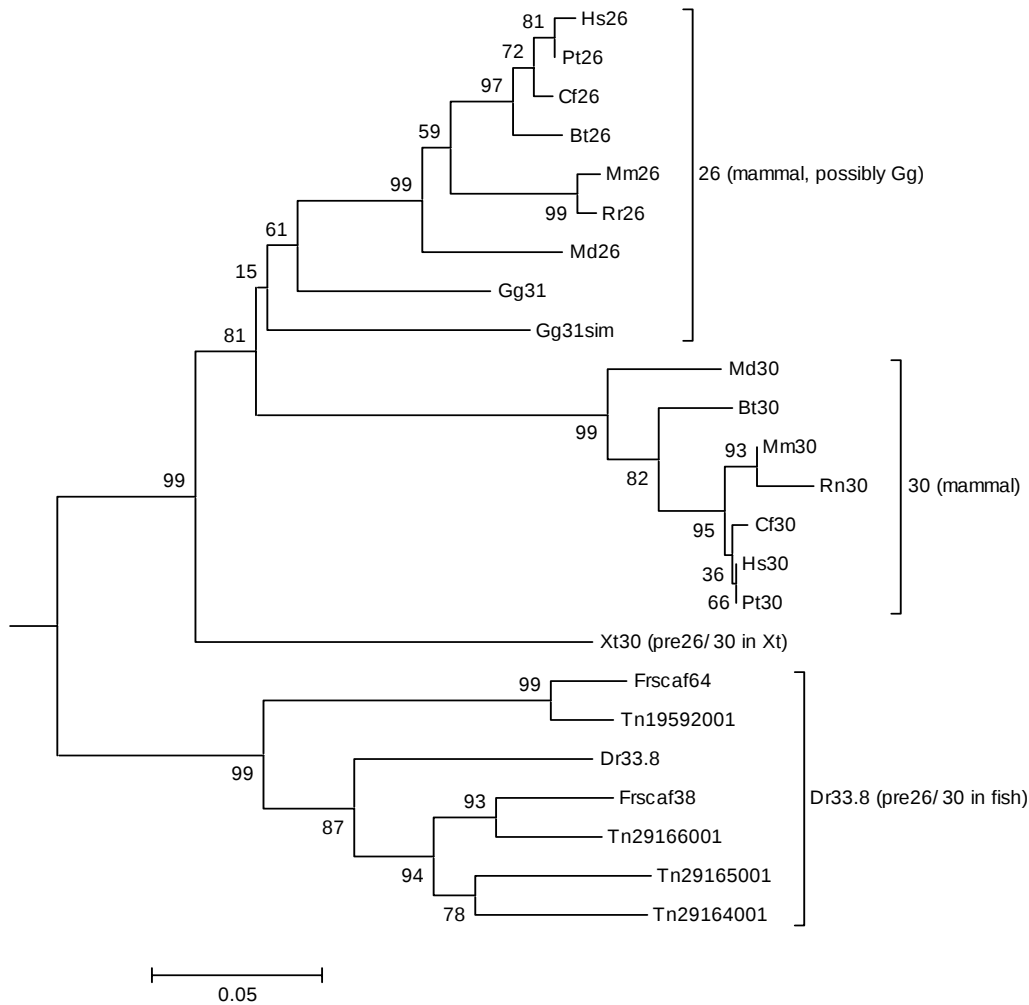
Suppl. Figure 8 Schematic explanation of terms used in the text when describing a branch from a phylogenetic tree. This branch indicates that the six present-day genes A to F originate from a common ancestral gene. If this is a tree constructed from amino acid sequences, the scale bar indicates the branch length that corresponds to a change in 5 % of the amino acids in the sequences used. As different genes may evolve at different rates (i.e., the selection pressure differs among the genes), the branch lengths may differ for genes originating (seemingly) at the same time, e.g, genes D to F. In the traditional rectangular style used in the present work, the deeper parts of the tree to the left represent the calculated evolutionary pathway to reach the present genes indicated to the right.



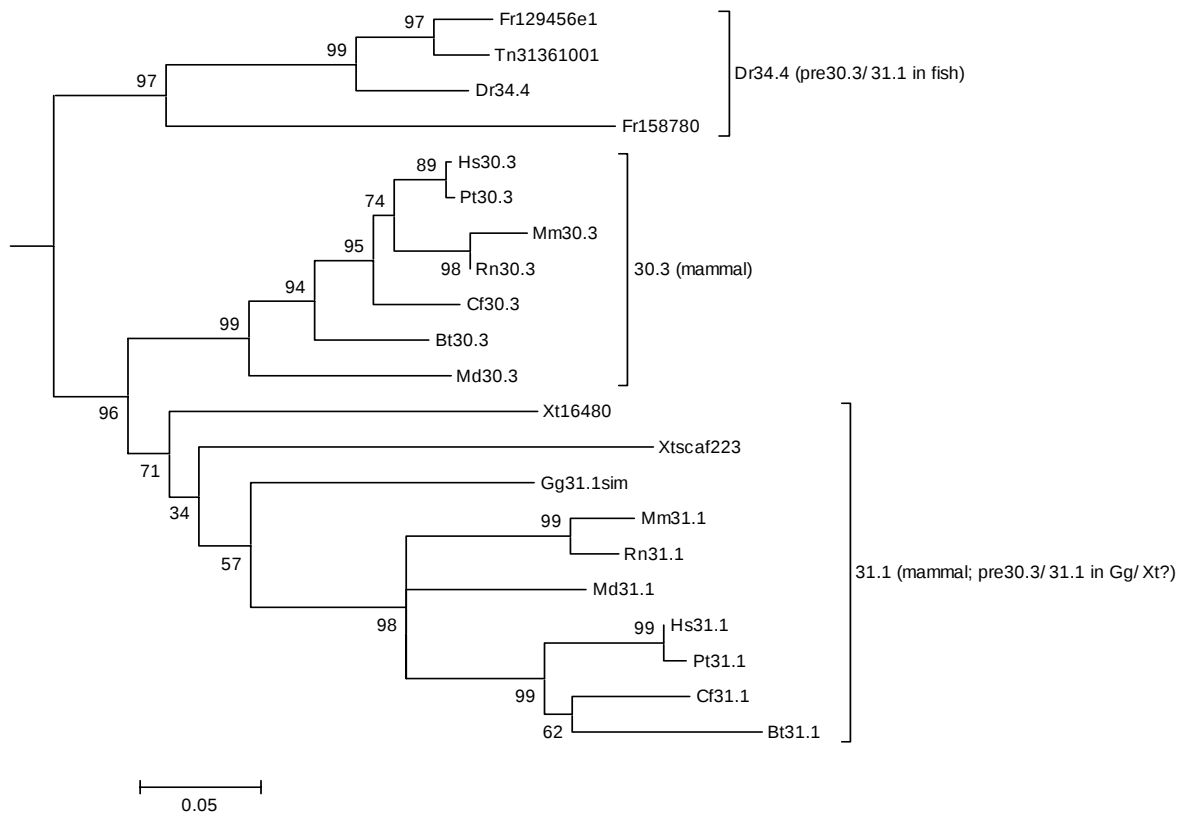
Suppl. Figure 9 Expanded view of all branches (triangles) of the phylogenetic tree shown in Fig. 2 of the manuscript. Common to all Figures A-W: The number of amino acid substitutions per site is indicated at the lower left. The statistics is shown as interior branch test. For simplicity and easier reading, the abbreviation "Cx" is not used. E.g., *Homo sapiens connexin25* (in Suppl. Figure 9A immediately below) is thus abbreviated Hs25 instead of HsCx25. All nonmammalian sequences used to build the tree are found in Suppl. Figures 1-5. Note that only the conserved sequences were used in these analyses.



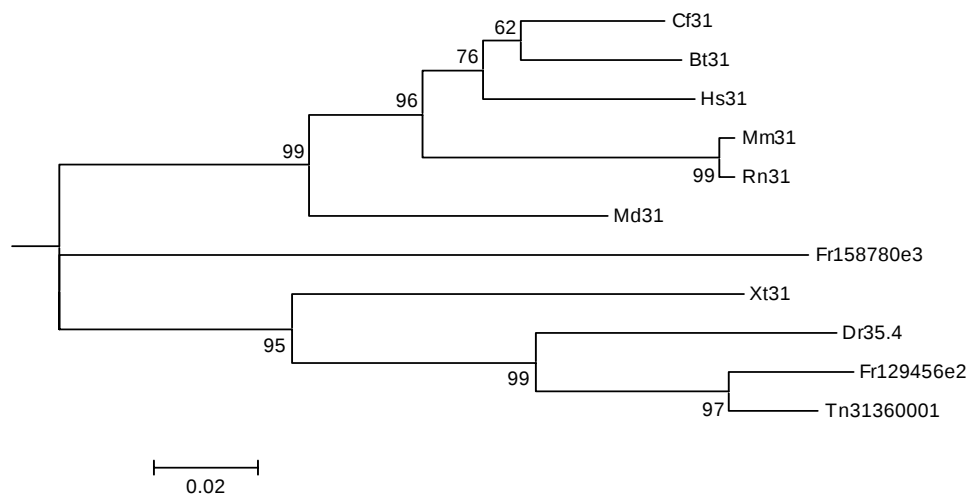
Suppl. Figure 9A Expanded view of the Cx25 group. This is an example of a branch with unique sequences in each of the species, including fishes.



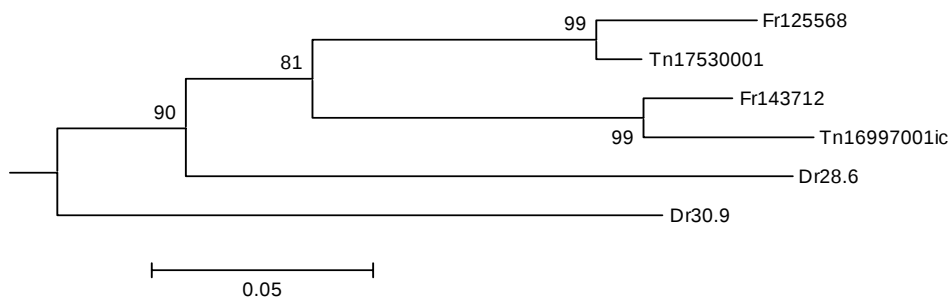
Suppl. Figure 9B Expanded view of the Cx26/30 group. Note that all fish *connexins* and the *Xenopus* sequence are grouping outside the sequences of both Cx26 and Cx30 orthologs from other vertebrates, indicating that a gene duplication, generating separate Cx26 and Cx30 sequences, may have occurred in the vertebrate lineage after the separation of amphibia. It is not clear whether the two chicken sequences are due to a separate gene duplication event, or that the bird lineage diverged immediately after the gene duplication, and before one of the sequences had gained the characteristics of Cx30. Note also that this is an example of a relatively complex relationship. When man:*Tetraodon* are compared, there is a 2:4 orthology relationship, but the two sequences in man (Cx26 and 30) are paralogs, and the four sequences in *Tetraodon* (*Tn19592001*, *Tn29166001*, *Tn29165001*, *Tn29164001*) are also paralogs.



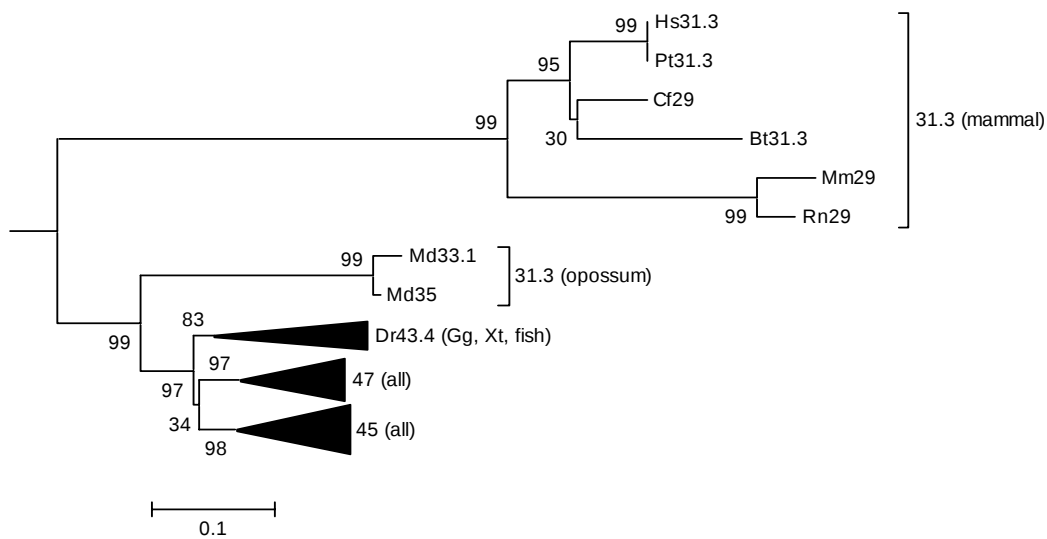
Suppl. Figure 9C Expanded view of *Cx30.3/31.1* group. Note that all fish connexins are grouping outside the sequences of both *Cx30.3* and *Cx31.1* orthologs from other vertebrates, indicating that a gene duplication, generating separate *Cx30.3* and *Cx31.1* sequences, may have occurred in the vertebrate lineage after the separation of fishes, or later.



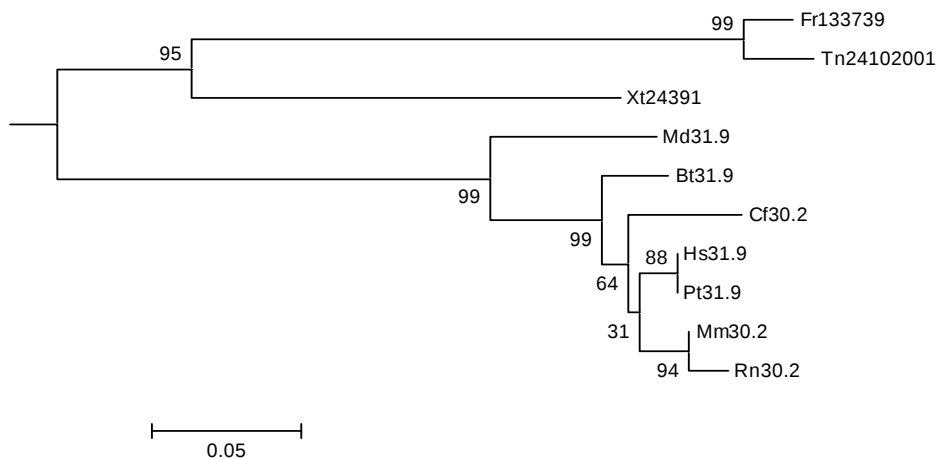
Suppl. Figure 9D Expanded view of the *Cx31* group.



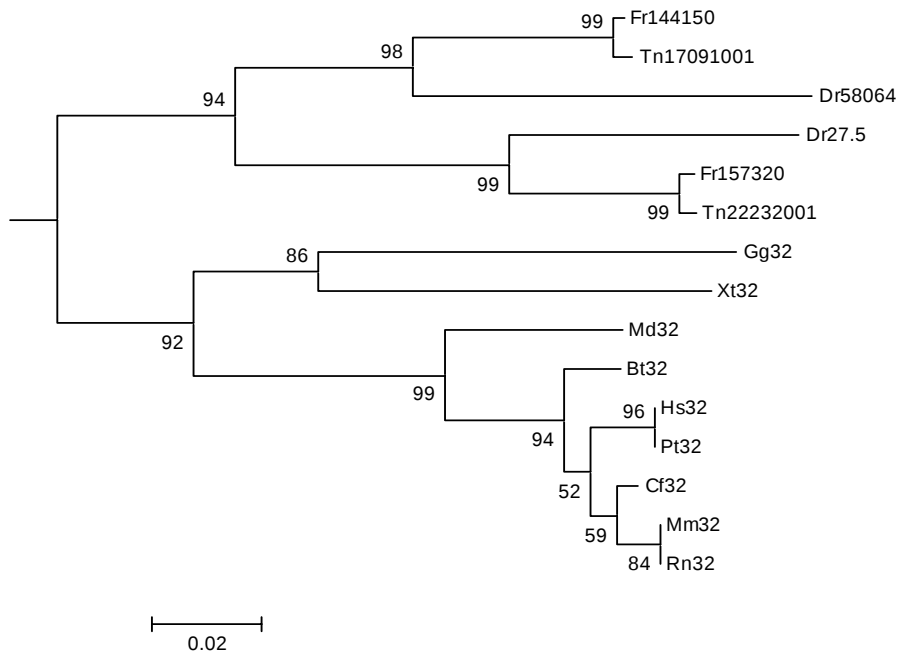
Suppl. Figure 9E Expanded view of the *Dr28.6* group. This branch only contains fish sequences, and may represent a connexin that originated in fish, or that was lost in the remaining vertebrate lineage after the separation of the fishes.



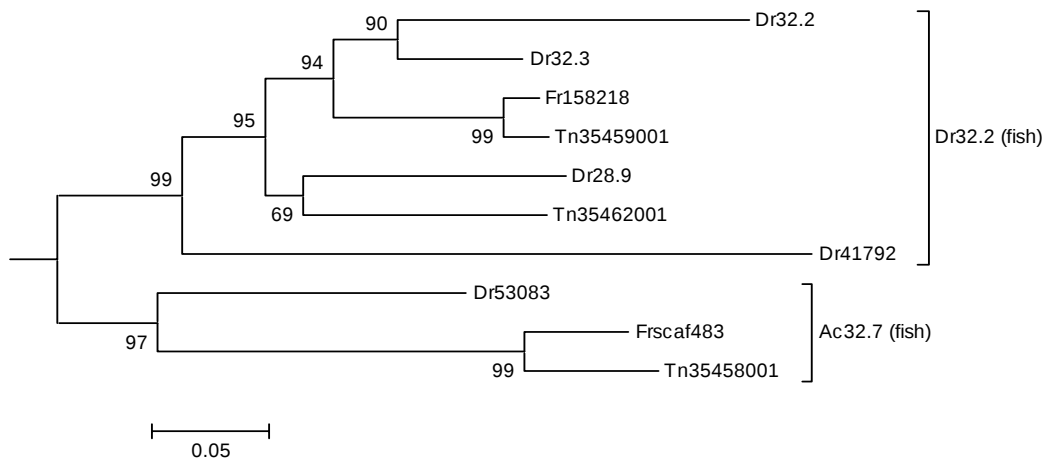
Suppl. Figure 9F Expanded view of the *Cx31.3* group. This branch only contains mammalian sequences, but in most models the two opossum sequences (*Md33.1*, *Md35*) split off from the interior branch leading to the *Cx45/47/Dr43.4* groups, as indicated here.



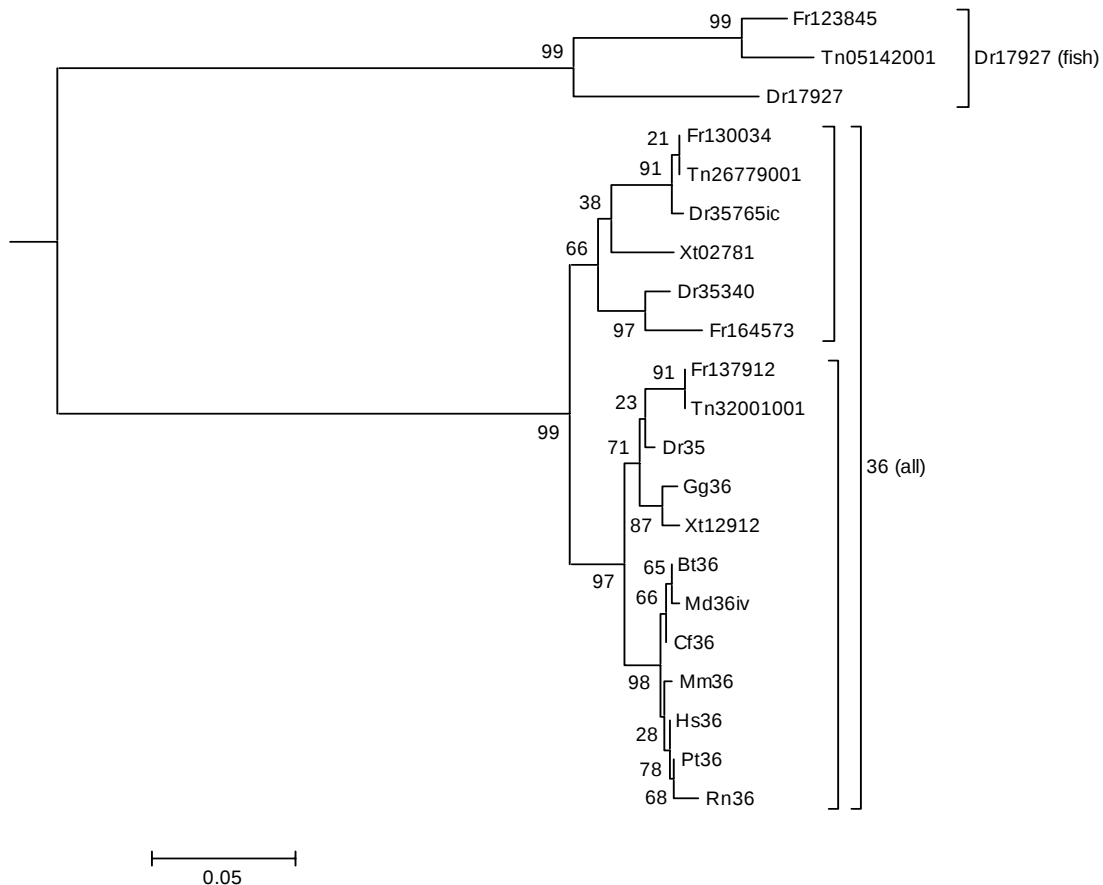
Suppl. Figure 9G Expanded view of the *Cx31.9* group.



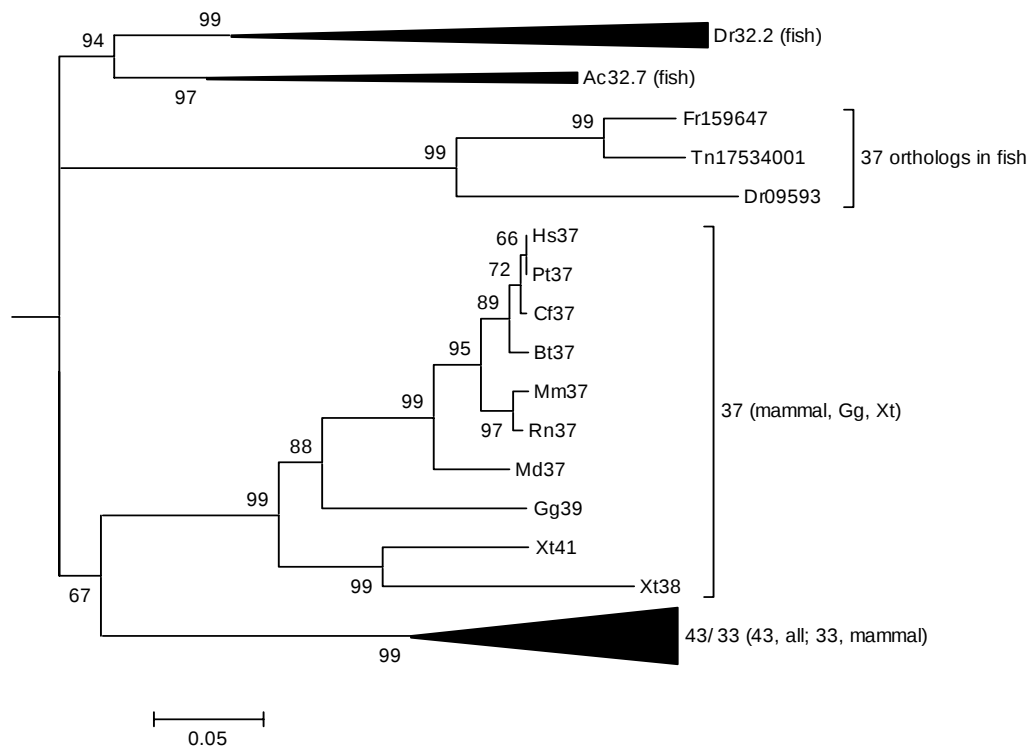
Suppl. Figure 9H Expanded view of the Cx32 group.



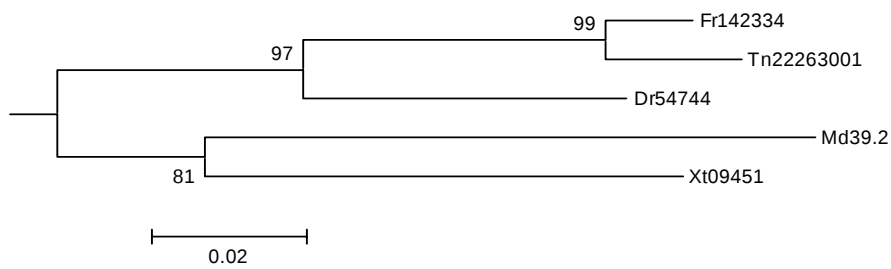
Suppl. Figure 9I Expanded view of the Dr32.2 and Ac32.7 groups. The naming Ac32.7 indicates that this is the orthologous group of Atlantic croaker Cx32.7. This branch only contains fish sequences, and may represent one or two *connexin* sequence that originated in fish, or that were lost in the remaining vertebrate lineage after the separation of the fishes.



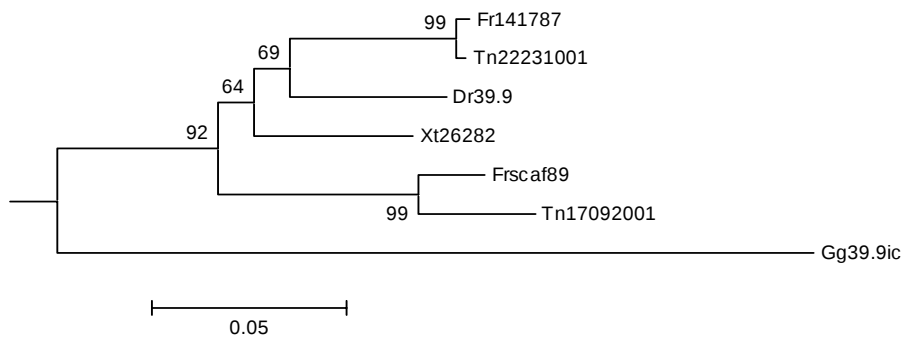
Suppl. Figure 9J Expanded view of the *Cx36* and *Dr17927* groups. The *Dr17927* group always splits off from the interior branch leading to the *Cx36* orthologs, and is only found in fishes. It may represent a connexin sequence that originated in fish, or that was lost in the remaining vertebrate lineage after the separation of the fishes. Note the two branches of the *Cx36* group itself, where one branch does not contain any sequences from birds or mammals. This may represent separate *Cx36*-like sequence that was lost after the divergence of the amphibia. Thus, the last unique common ancestor of fishes and amphibian (and thereby also mammals) may have possessed three *Cx36*-like genes, namely, the precursors of (i) the *Dr17927* group; (ii) the upper *Cx36* group (e.g., *Xt02781*), and (iii) the lower *Cx36* group (e.g., *Xt12912* and the mammalian *Cx36* sequences). Skate *Cx35* (O'Brien et al., 1996) locates itself in the lower *Cx36* group (results not shown).



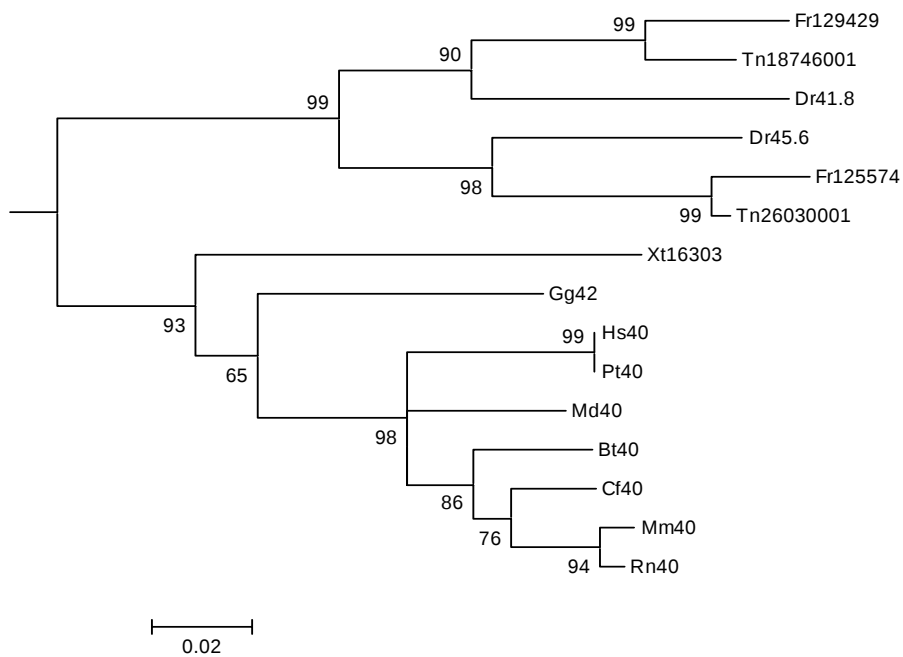
Suppl. Figure 9K Expanded view of the *Cx37* group. Note that a duplication has occurred in amphibia, generating two *Cx37*-like sequences, *Cx41* and the archetypical *Xenopus* *Cx38*. See text for more details on *Xenopus* and fish sequences.



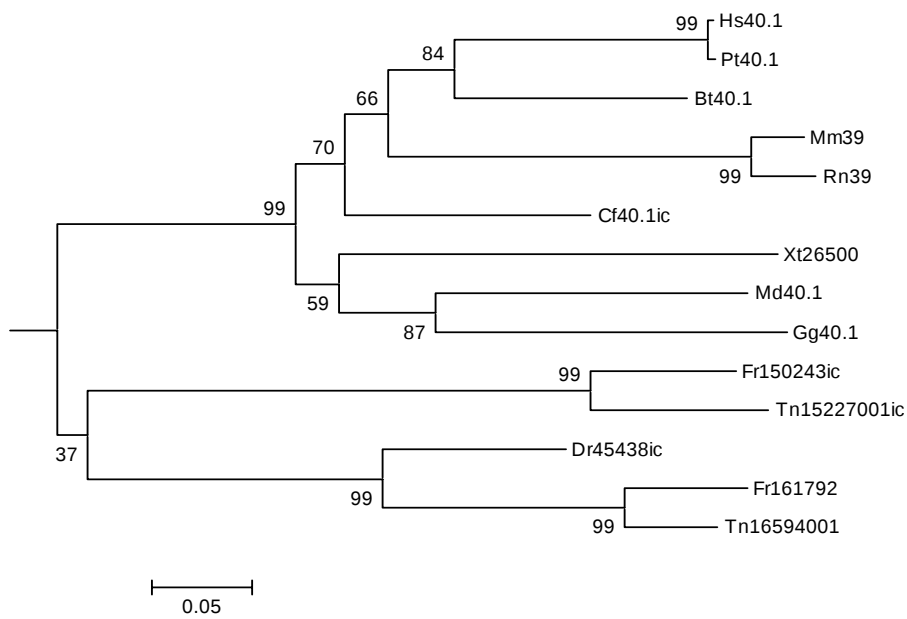
Suppl. Figure 9L Expanded view of the Cx39.2 group. Among the mammals, this connexin has as yet only been found in opossum, but single orthologs are found in *Xenopus* and in fishes. Thus, this sequence may have been lost from the remaining mammals after the divergence of the marsupials.



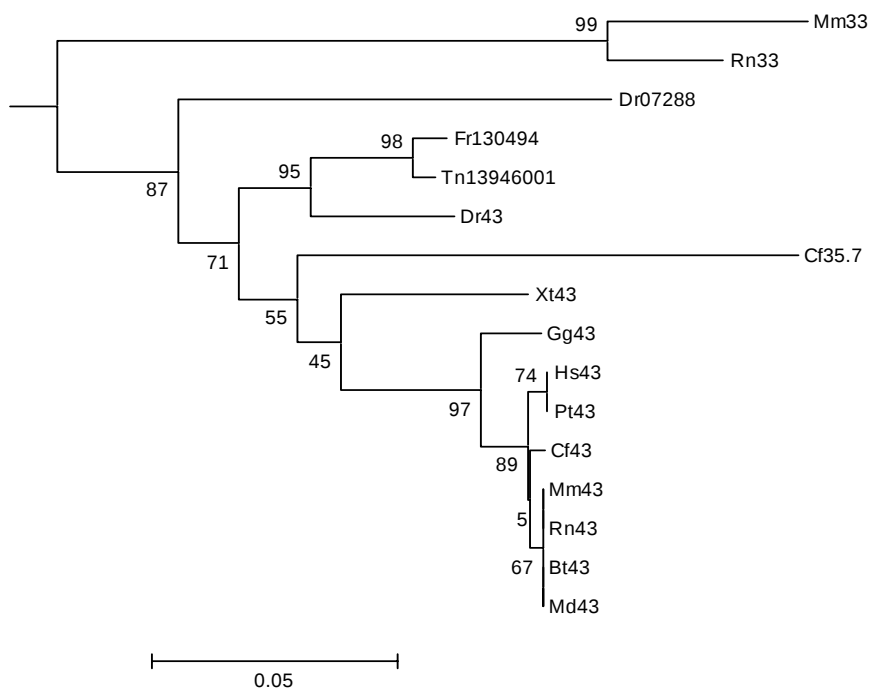
Suppl. Figure 9M Expanded view of the Cx39.9 group. Orthologs are found in fishes, amphibia and birds, but not in mammals. It may therefore represent a connexin that was lost in the lineage leading to mammals.



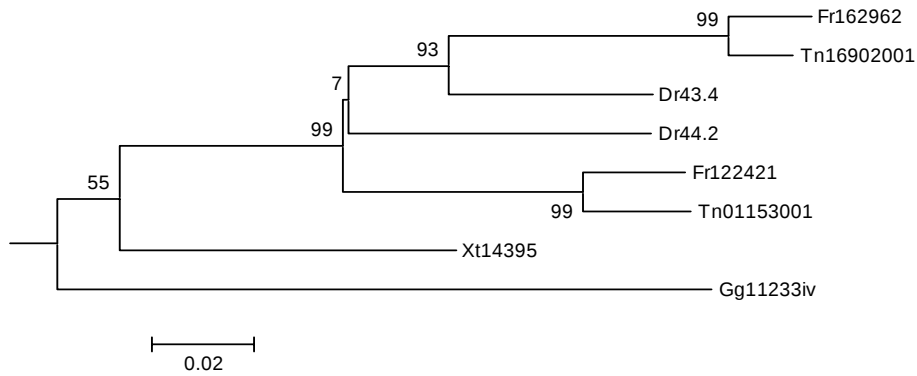
Suppl. Figure 9N Expanded view of the Cx40 group.



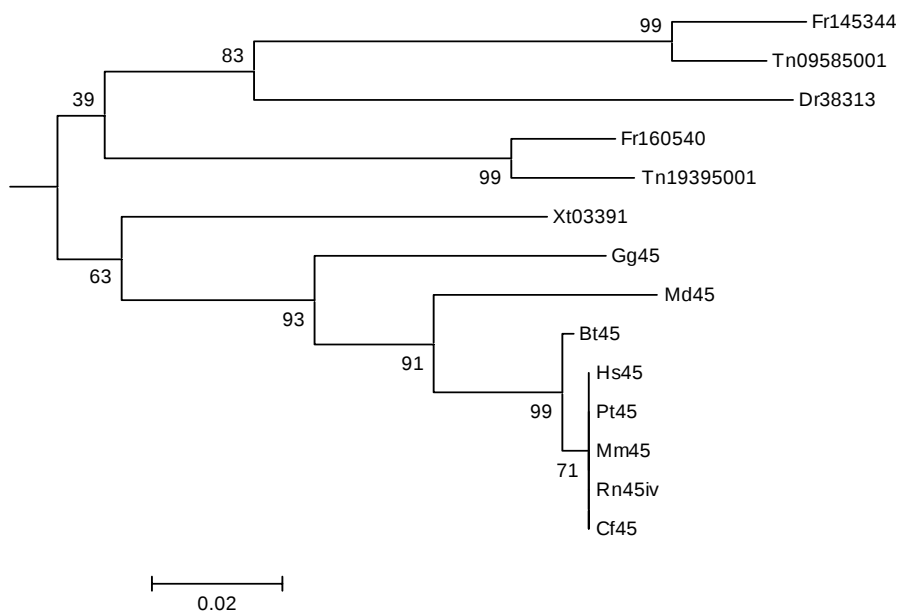
Suppl. Figure 9O Expanded view of the *Cx40.1* group.



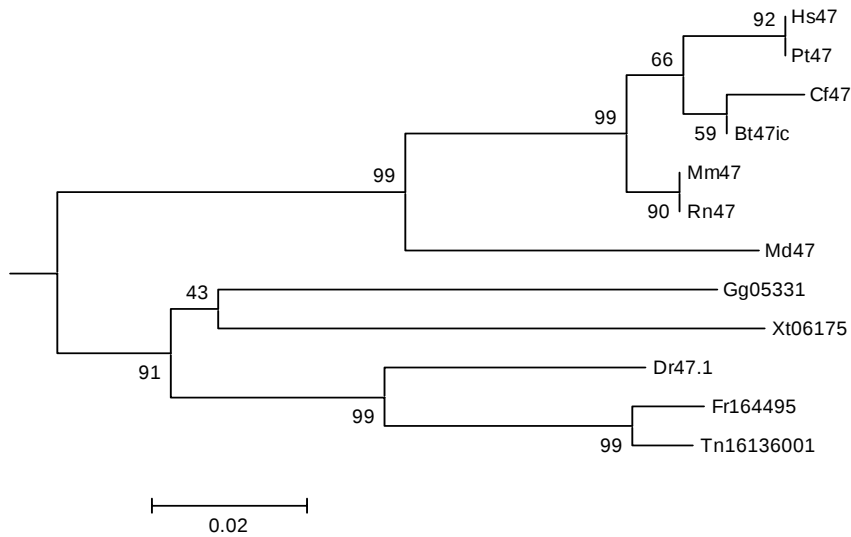
Suppl. Figure 9P Expanded view of the *Cx43* and *Cx33* group. Pseudogenes are not included.



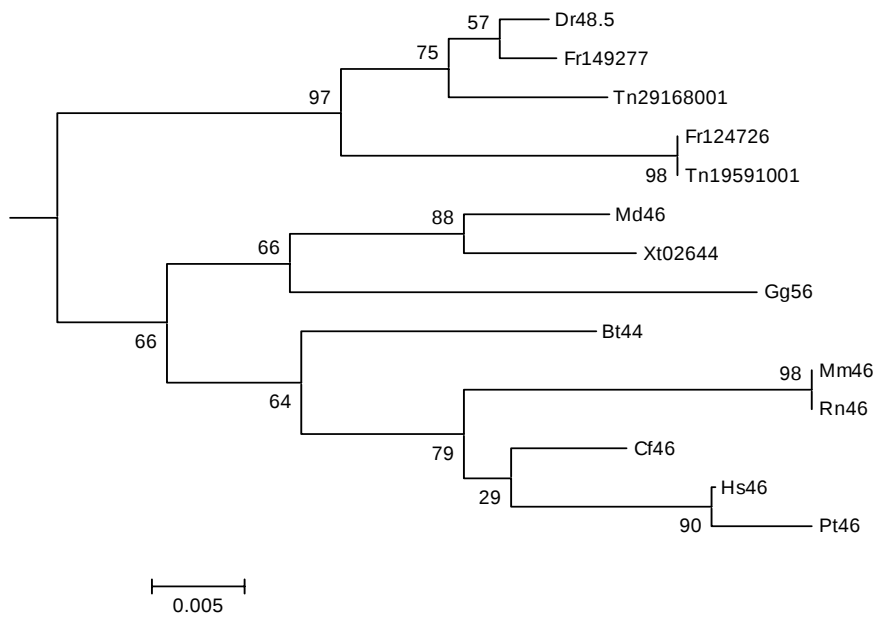
Suppl. Figure 9Q Expanded view of the *Dr43.4* group. This group is found in fishes, amphibia and birds, but not in mammals. It may therefore represent a connexin that was lost in the lineage leading to mammals.



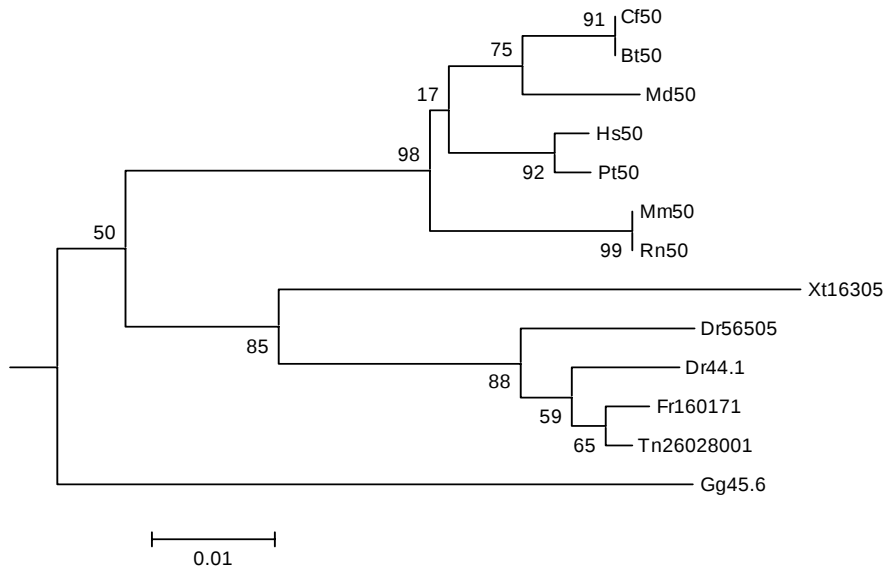
Suppl. Figure 9R Expanded view of the *Cx45* group.



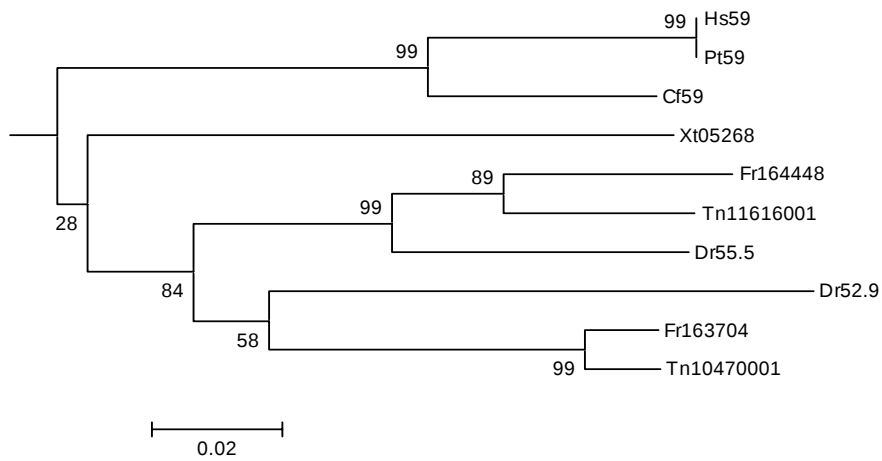
Suppl. Figure 9S Expanded view of the *Cx47* group.



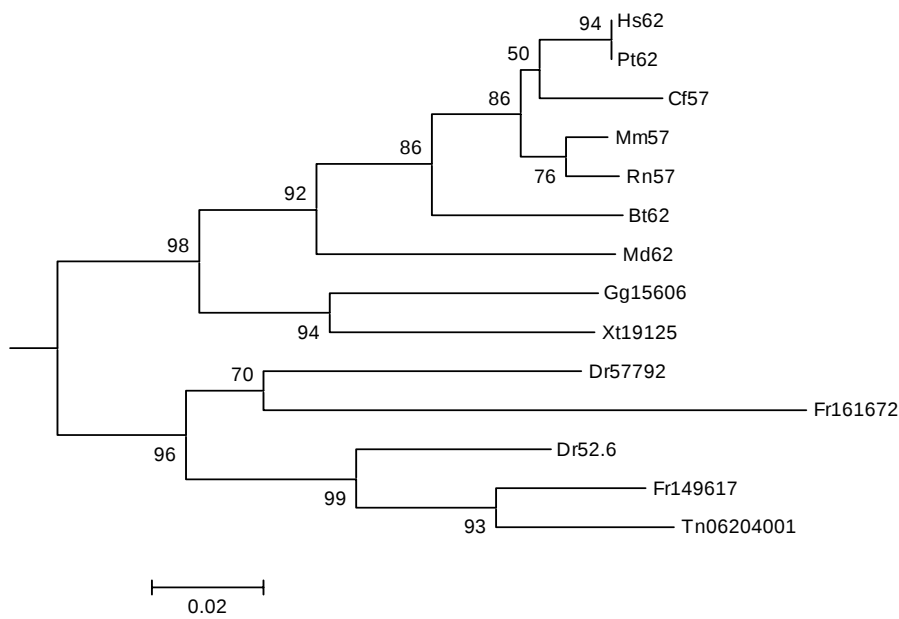
Suppl. Figure 9T Expanded view of the *Cx46* group.



Suppl. Figure 9U Expanded view of the Cx50 group.

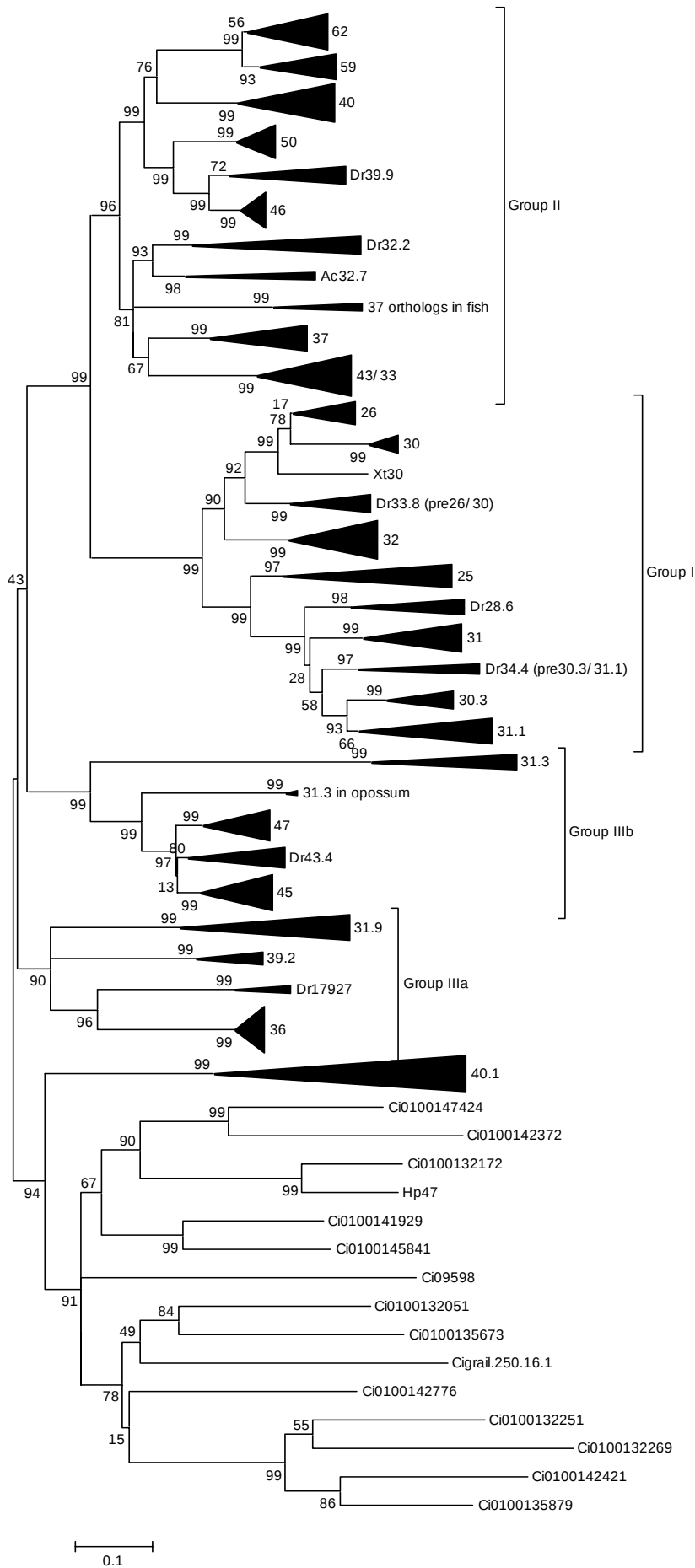


Suppl. Figure 9V Expanded view of the Cx59 group.

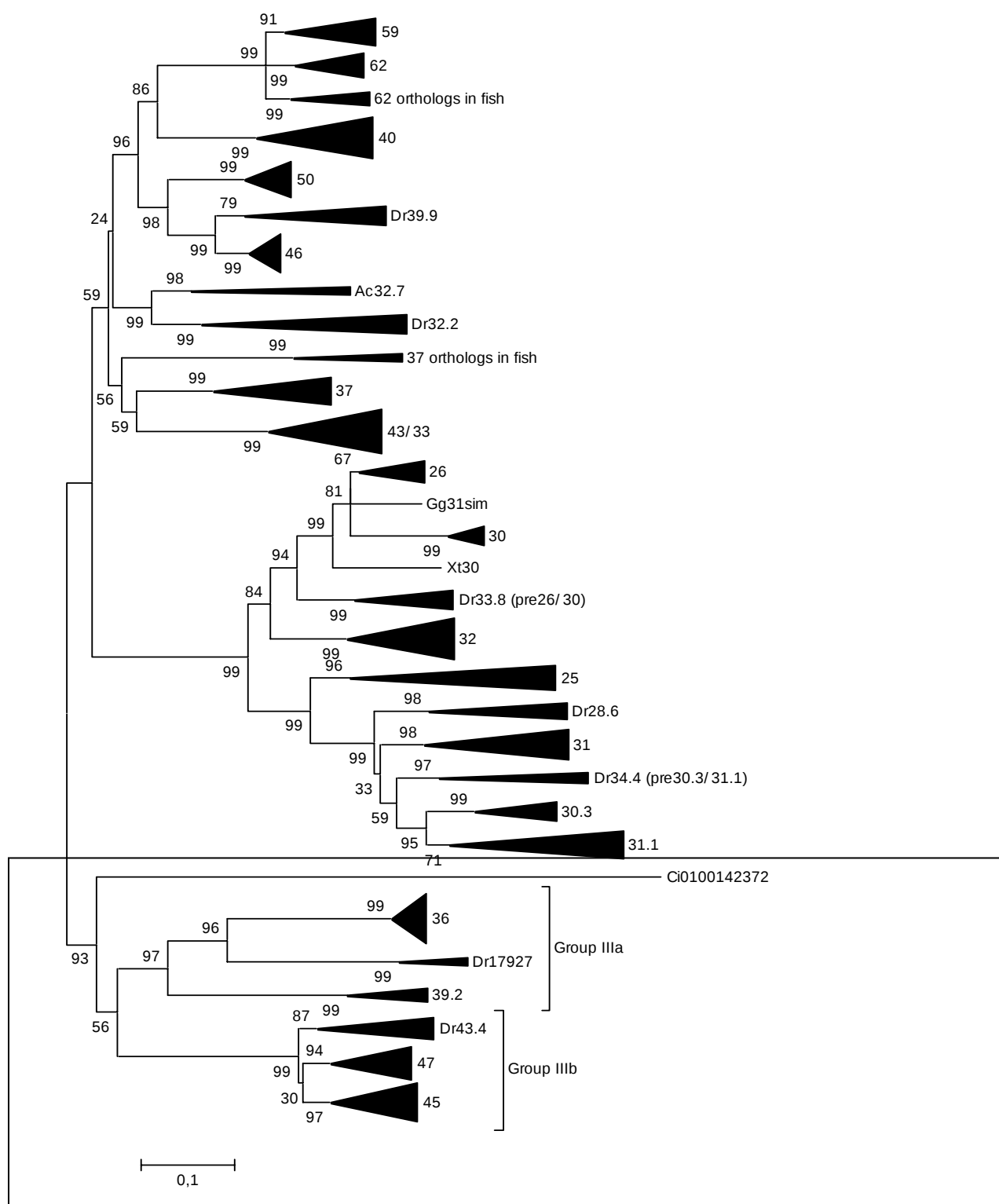


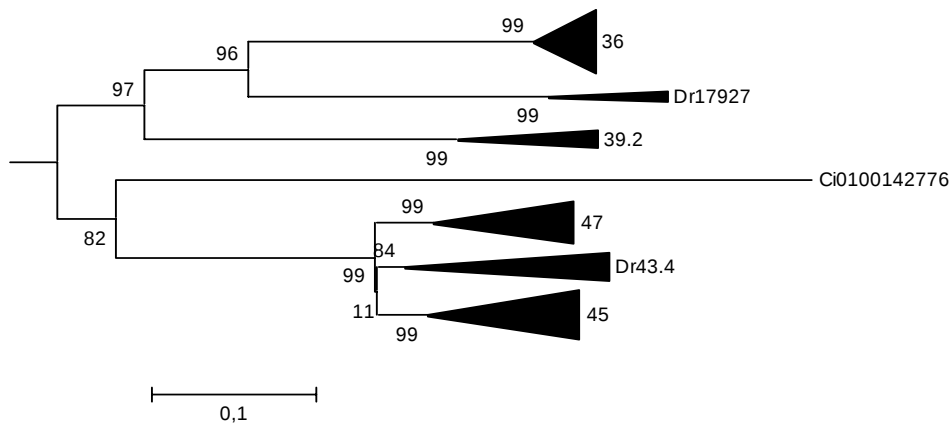
Suppl. Figure 9W Expanded view of the Cx62 group. See text for more details.

Suppl. Figure 10 (next page) Phylogenetic tree illustrating long-branch attraction of tunicate connexins with vertebrate *Cx40.1*, which is one of the vertebrate connexins with the less resemblance to the tunicate sequences. The tree was obtained with the minimum evolution equal input model, and interior branch statistics (500 replications) are indicated at the nodes. A total of 305 sequences were included in this analysis.

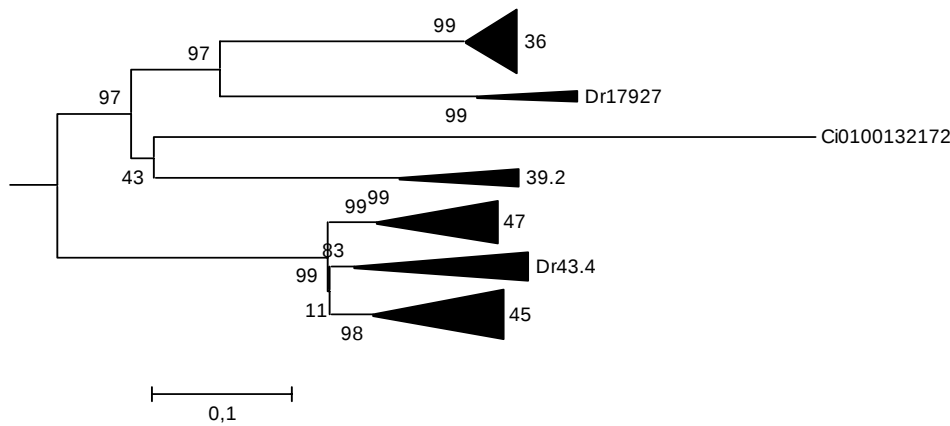


Suppl. Figure 12A-D The grouping of some *Ciona connexins* together with vertebrate sequences. All sequences in the Cx31.3, 31.9 and 40.1 groups were removed before the analysis. **(A)** The complete and compressed of which the framed part is shown in Fig. 3 of the manuscript. The tree was obtained with the neighbor-joining equal input model, and the interior branch statistics (500 replications) are indicated at the nodes. The frame also corresponds to the branches shown in Figs. 12B-D (shown on next page), using other *Ciona* sequences.

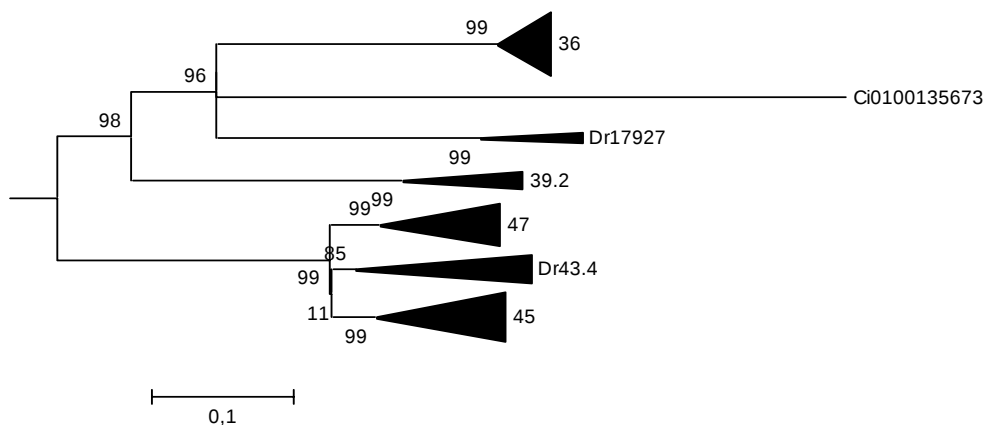




Suppl. Figure 12B The grouping of *Ciona* Ci0100142776 together with Cx43.4/45/47-like sequences. The tree was obtained with the minimum evolution equal input model with interior branch statistics.



Suppl. Figure 12C The grouping of *Ciona* Ci0100132172 together with Cx39.2-like sequences. The tree was obtained with the minimum evolution equal input model and interior branch statistics.



Suppl. Figure 12D The grouping of *Ciona* Ci0100135673 together with Cx36-like sequences. The tree was obtained with the minimum evolution equal input model and interior branch statistics.

Suppl. Table 1 Genomic positions of chicken (*Gallus gallus*, Gg) *connexins* according to WASHUC 1 (Ensembl v39) as of July 1 2006. Please note that the locations and Ensembl gene identifications may change upon future updates of the genomic database. If the coding sequence is divided into two exons, they are indicated by E1 and E2.

Present naming	Chr	From	To
GgCx31	1	170911388	170912179
GgCx31sim	1	170931895	170932572
GgCx31.1sim	23	3968170	3969060
GgCx32 ^a	4	2148920	2149604
GgCx36	5	E1: 29166156 E2: 29166720	E1: 29166226 E2: 29167563
GgCx39	23	3959185	3960216
GgCx40.1	2	E1: 12201560 E2: 12197873	E1: 12201623 E2: 12199059
GgCx42	1	86332113	86333222
GgCx43	3	60233089	60234234
GgCx45	27	1636036	1637220
Gg05331	2	1852747	1853994
Gg11233iv ^b	7	22887665	22888591
GgCx56	1	170948963	170950498
GgCx39.9ic ^c	4	2147867 2147311	2148039 2147778
GgCx45.6	1	86378805	86380007
Gg15606	3	73757453	73759048

a: Ensembl sequence is partial, and lacks the 155 5'-end nucleotides. The full-length sequence was assembled according to NM_204371.

b: Ensembl sequence is partial, and lacks unknown number of nucleotides in 3'-end.

c: Ensembl sequence is partial, and lacks the parts corresponding to EC1, TM2 and intracellular loop. Thus, the sequence is found divided in two contigs, but the final full-length sequence will probably be into one exon, assuming that the Ensembl assembly is correct.

Suppl. Table 2 Databank entries of chicken (*Gallus gallus*, Gg) *connexins* as of July 1 2006. Please note that the Ensembl gene identifications (Ensembl v39, WASCHUC1) may change upon future updates of the chicken genome.

Present naming	Ensembl	GenBank	Mammalian ortholog
GgCx31	ENSGALG00000017136	NM_204931 Cx30 AF072699 Cx31	Cx26/30
GgCx31sim		XM_425641 similar to Cx31	Cx26/30
GgCx31.1sim		XM_425784 similar to Cx31.1	Cx30.3/31.1
GgCx32	ENSGALG00000005541	NM_204371 Cx32	Cx32
GgCx36	ENSGALG00000009845	NM_204582 Cx36 AY700221 Cx35.1	Cx36
GgCx39	ENSGALG00000002532	NM_206984 Cx37 AY206667 Cx39	Cx37
GgCx40.1 ^a	ENSGALG00000007065	XM_418567 similar to Cx40.1	Cx40.1
GgCx42	ENSGALG00000015487	NM_205504 Cx40 M35043 Cx42	Cx40
GgCx43	ENSGALG00000014873	NM_204586 Cx43	Cx43
GgCx45	ENSGALG00000000997	NM_205503 Cx45	Cx45
Gg05331 ^b	ENSGALG00000005331	XM_418503 similar to Cx45	Cx47
Gg11233iv	ENSGALG00000011233	XM_422042 No Cx identification	-
GgCx56	ENSGALG00000017137	L02838 Cx56 XM_425642 Cx46	Cx46
GgCx39.9ic		XM_426260 similar to Cx39.9	-
GgCx45.6	ENSGALG00000015488	NM_204997 Cx50 L24799 Cx45.6	Cx50
Gg15606 ^c	ENSGALG00000015606	XM_419834 No Cx identification	Cx62

a: XM_418567 appears to predict the presence of an intron approx. at the border of E2/TM3.

b: XM_418503 is based on the presence of a potential intron with splice site approx. 30 nt upstream of the stop codon in *Gg05331*.

c: 5'-half of *Gg15606* sequence predicted as a part of a very large hypothetical protein.

Suppl. Table 3 Database positions of *Xenopus tropicalis* (Xt) *connexins* according to JGI4.1 (Ensembl v39). The genomic locations were not available as of July 1 2006. Please note that the locations, scaffold number and Ensembl gene identifications may change upon future updates of the genomic database. If the coding sequence is divided into two exons, they are indicated by E1 and E2.

Present naming	Scaffold_	From	To
Xt18051	76	1010439	1011182
XtCx30	329	932516	933262
Xtscaf223	223	779856	780659
Xt16480	223	769967	770773
XtCx31	223	721530	722333
Xt24391	582	84039	84959
XtCx32	117	799086	799880
Xt12912	37	E1:2494520 E2: 2499228	E1: 2495365 E2: 2499299
Xt02781	461	E1: 4848 E2: 22628	E1: 4919 E2: 23461
XtCx38	223	711752	712741
XtCx41	223	699090	700148
XtCx43	172	206147	207286
Xt09451	352	306412	307473
Xt16303	622	484737	485831
Xt26500	133	E1: 1559752 E2: 1563443	E1: 1560798 E2: 1563516
Xt03391	155	1334158	1335336
Xt06175	471	556763	557980
Xt14395	1022	216202	217372
Xt02644	329	892263	893507
Xt26282	117	804549	805922
Xt16305	622	E1: 579784 E2: 581601	E1: 581028 E2: 581680
Xt05268	403	330204	331766
Xt19125	599	447419	449110

Suppl. Table 4 *Xenopus tropicalis* and *Xenopus laevis* genome databank entries. XtDBv3.0, *Xenopus tropicalis* database v3.0 (<http://genome.jgi-psf.org/Xentr3/Xentr3.home.html>). Please note that the Ensembl gene identifications (Ensembl v39, JGI4.1) may change upon future updates of the *Xenopus* genome.

Present naming	Ensembl	XtDBv3.0	<i>X. tropicalis</i> GenBank	<i>X. laevis</i> GenBank	Mammalian ortholog
Xt18051	ENSXETG00000018051	73000004	-	BC055269 No Cx identification	Cx25
XtCx30	ENSXETG00000026112	327000024	CR848317:119 No Cx identification	BC043797 Cx30	Cx26/30
Xtscaf223 ^a		134000033 ^a	-	-	Cx30.3/31.1
Xt16480	ENSXETG00000016480	134000034	-	-	Cx30.3/31.1
XtCx31	ENSXETG00000016481	134000003	-	AY057997 Cx31	Cx31
Xt24391	ENSXETG00000024391	670000005	-	-	Cx31.9
XtCx32	ENSXETG00000026283	482000004	BC06715 No Cx identification	Y00791 gap junction subunit	Cx32
Xt12912	ENSXETG00000012912	199000005	-	-	Cx36
Xt02781	ENSXETG00000002781	572000001	-	-	Cx36
XtCx38	ENSXETG00000016484	134000007	-	BC068650 Cx38	Cx37
XtCx41	ENSXETG00000026789	134000008		U26256 Cx41	Cx37
XtCx43	ENSXETG00000025411	42000005	NM_203525 Cx43	BC073407 Cx43	Cx43
Xt09451	ENSXETG00000009451	370000011	-	-	Cx39.2
Xt16303	ENSXETG00000016303	351000006	-	-	Cx40
Xt26500 ^b	ENSXETG00000026500	137000002 ^b	BC062485:184 No Cx identification	-	Cx40.1
Xt03391	ENSXETG00000003391	316000023	-	-	Cx45
Xt06175	ENSXETG00000006175	451000007	-	BC042343 Similar to Cx45	Cx47
Xt14395	ENSXETG00000014395	1262000002	BC066131 No Cx identification	AY057998 Cx43.4 BC044270 similar to gj prot BC042343 Similar to Cx45 BC073513 Similar to Cx45	-

Xt02644 ^c	ENSXETG00000002644	327000021	-	-	Cx46
Xt26282 ^d	ENSXETG00000026282	482000011 ^d	BC074597 No Cx identification	BC043801 Similar to gja3 AY792358 Cx40.4	-
Xt16305	ENSXETG00000016305	351000007	-	BC046682 No Cx identification	Cx50
Xt05268	ENSXETG00000005268	442000006	-	-	Cx59
Xt19125	ENSXETG00000019125	604000005	-	-	Cx62

a: Scaffold number according to the present Ensembl genebuild (v37). According to

XtDBv3.0, the scaffold number was 134.

b: Sequence not predicted as a gene in the previous Ensembl genebuild (v36 and earlier).

According to XtDBv3.0, the scaffold number was 137.

c: Sequence according to the present Ensembl genebuild (v37). Previously, a part of this sequence was exon 2 of the gene prediction with this number, the remaining part of the present sequence was predicted as introns. Exon 1 of the previous *Xt02644* gene prediction corresponded to *XtCx30*.

d: Sequence not predicted as a gene in the previous Ensembl genebuild (v36 and earlier).

According to XtDBv3.0, the scaffold number was 482.

Suppl. Table 5 Genomic positions of zebrafish (*Danio rerio*, Dr) *connexins* according to Ensembl assembly WTSI Zv6, Ensembl v39, as of July 1 2006). Please note that the locations, scaffold number and Ensembl gene identifications may change upon future updates of the genomic database. If the coding sequence is divided into two exons, they are indicated by E1 and E2.

Present naming	Chromosome	From	To
Dr59215 ^a	20	126182	126946
DrCx33.8	9	15964611	15965414
DrCx34.4 ^a	17	1626843	1627742
DrCx35.4 ^a	17	1618334	1619248
DrCx28.6	17	15926098	15926812
DrCx30.9 ^a	19	37233052	37233855
DrCx27.5	5	16625164	16625886
Dr58064	14	12216203	12217030
DrCx35 ^a	20	E1: 11406257 E2: 11408277	E1: 11406327 E2: 11409120
Dr35340	5	E1: 40284811 E2: 40262777	E1: 40284741 E2: 40263605
Dr35765ic	15	6929027	6929872
Dr17927	7	11499226	11500197
DrCx28.9 ^a	20	48153500	48154261
Dr41792	20	48147145	48147888
DrCx32.2	20	48131463	48132296
DrCx32.3 ^a	20	48161088	48161945
Dr53083 ^a	20	48164669	48165565
Dr09593 ^a	19	37318638	37319663
DrCx43	20	48123936	48125081
Dr07288	17	15294422	15295486
Dr54744	15	42480120	42481184
DrCx45.6	9	26170674	26171876
DrCx41.8	1	55406487	55407599
Dr45438ic	24	5074406	5075614
Dr38313	3	12201754	12203187
DrCx47.1	2	2701898	2703127
DrCx44.2 ^b	6	55000611	55001777
DrCx43.4	9	E1: 45975548 E2: 45972964	E1: 45976676 E2: 45972977
DrCx48.5	9	15969645	15970949
DrCx39.9	5	16617062	16618123
DrCx44.1	1	55376449	55377624
Dr56505 ^a	9	E1: 26215140 E2: 26217153	E1: 26215210 E2: 26217269
DrCx55.5 ^a	16	54063976	54065472
DrCx52.9	2	43708904	43710330
DrCx52.6	20	26545290	26546690
Dr57792	17	E1: 15163965 E2: 15162447	E1: 15165345 E2: 15165345

a: Repeated once as exact or nearly exact (>98% id.) copy within ~1 Mnt of given positions.

b: Exact or nearly exact copies (>94% id.) found another 6 times on the same chromosome.

Suppl. Table 6 Zebrafish (*Danio rerio*) genome databank entries.

Present naming	Present Ensembl (Zv6)	Previous Ensembl (Zv5)	Previous Ensembl (Zv4)	Genbank	Mammalian ortholog
Dr59215	ENSDARG00000059215 (Cx28.8 ^b)	43971	10877	NM_001025544 Similar to gjb4	Cx25
Not incl.	ENSDARG00000059187 =59215				
DrCx33.8	ENSDARG00000042707	42707	21889	NM_212825 Cx33.8	Cx26/30
DrCx34.4	ENSDARG00000059220	42865	09863	NM_207169 Cx34.4	Cx30.3/31.1
Not incl.	ENSDARG00000059127 =59220/Cx34.4				
DrCx35.4	ENSDARG00000042866	42866	02087	NM_001017685 Cx35.4	Cx31
Not incl	ENSDARG00000040782 =42886/Cx35.4	40782 =Dr42866/Cx35.4			
DrCx28.6	ENSDARG00000003925	03925	93925	NM_001007212 Cx28.6	-
DrCx30.9	ENSDARG00000028833	28833	28833	NM_001007288 Cx30.9	-
Not incl.	ENSDARG00000055471 =28833/Cx30.9				
DrCx27.5	ENSDARG00000035533	35553	02284	NM_131811 Cx27.5	Cx32
Dr58064	ENSDARG00000058064 (Cx31.7 ^b)	37518		BX470260:64220 (rc ^a) No Cx identification	Cx32
	Absent in present version of Ensembl	37500 =Dr37518			
DrCx35	ENSDARG00000058009	37049	31311	NM_194420 Cx35 AF462040 Similar to Cx35.1	Cx36
Not incl	ENSDARG00000057914 =58009	43693 = Cx35			
	Absent in present version of Ensembl	43642 = Cx35			
Dr35340	ENSDARG00000035340 (Cx34.1 ^b)	35340	14503	XM_696083 Similar to Cx34.7 XM_679173 Similar to gja9	Cx36
	Absent in present version of Ensembl	32464 =Dr35340			
Dr35765ic	ENSDARG00000035765	35765		XM_684226 Similar to gja9	Cx36
Dr17927	ENSDARG00000017927 (Cx35.8 ^b)	17927	17927		-
DrCx28.9	ENSDARG00000041797	41797		NM_01007324 Cx28.9	-
Not incl	ENSDARG00000041741 =Dr41797/Cx28.9	41741 =Dr41797/Cx28.9			
Dr41792	ENSDARG00000041792 (Cx28.1 ^b)	41792		CR352322:18982 (rc) No Cx identification	-
	Absent in present version of Ensembl	41743 =Dr41792			
DrCx32.2	Not predicted	21696		AY340236:2480	-

				Cx32.2 NM_01030210 Novel Cx family protein	
DrCx32.3	ENSDARG00000041787	41787		NM_199612 Cx32.3	-
Not incl	ENSDARG00000041740 =Dr41787/Cx32.3	41740 =Dr41787/Cx32.3			
Dr53083	ENSDARG00000053083 (Cx34.5)	41782		NM_001030200 Novel Cx family protein	-
Not incl	ENSDARG00000053059 =53083	41739 =Dr41782			
Dr09593	ENSDARG00000009593	09593	09593	XM_680312 No Cx identification	Cx37
Not incl.	ENSDARG00000055468 =09593				
DrCx43	ENSDARG00000041799	12589	12589	NM_131038 Cx43	Cx43
	Absent in present version of Ensembl	41799 =Dr12589/Cx43			
Dr07288	ENSDARG00000007288	07288	07288	XM_688906 Similar to Cx43	Cx43
Dr54744	ENSDARG00000054744	09334	09334	XM_679941 Similar to Cx50	Cx39.2
DrCx45.6	ENSDARG00000040065	40065	01989	NM_001007213 Cx45.6	Cx40
DrCx41.8	ENSDARG00000036688	36688	20587	XM_693925 Cx41.8	Cx40
Dr45438ic	ENSDARG00000045438	45445		BX004842:24668 (rc) No Cx identification	Cx40.1
	Absent in present version of Ensembl	45438 =Dr45445			
Dr38313	ENSDARG00000038313	38313	02681	XM_679922 Similar to Cx45	Cx45
DrCx47.1	ENSDARG00000006961	06961	06961	NM_001004574 Cx47.1	Cx47
DrCx44.2	ENSDARG00000034207	34207	34207	NM_131810 Cx44.2	-
Not incl.	ENSDARG00000034129 =Dr34207/Cx44.2	34129 =Dr34207/Cx44.2			
Not incl.	ENSDARG00000052453 =Dr34207/Cx44.2				
DrCx43.4	ENSDARG00000007099	07099	07099	NM_131069 Cx43.4	-
DrCx48.5	ENSDARG00000021889	21889	21889	NM_207642 Cx48.5	Cx46
DrCx39.9	ENSDARG00000004082	04082	04082	NM_212826 Cx39.9	-
DrCx44.1	ENSDARG00000015076	15076	15706	NM_131809 Cx44.1 AF304050 Cx44.2	Cx50
Dr56505	ENSDARG00000056505	07859	07859	XM_692792 Similar to Cx44.1	Cx50
Not incl.	ENSDARG00000056529 =56505				
DrCx55.5	ENSDARG00000004403	04403	31820	NM_131812	Cx58

				Cx55.5 XM_688923 Similar to Cx55.5	
Not incl.	ENSDARG00000053302 =04403				
DrCx52.9	ENSDARG00000039845	39845	02232	NM_207093 Cx52.9	Cx58
DrCx52.6	ENSDARG00000034930	34930	34930	NM_212819 Cx52.6	Cx62
Dr57792	ENSDARG00000057792 (Cx52.7 ^b)		04403	XM_695977 Similar to Cx52.6	Cx62

a: rc, reverse complement.

b: No connexin name is available in GenBank for this sequence. The name given in parenthesis is used by Eastman et al., 2006, and some of these are also mentioned in Ensembl zebrafish Zv6.

Suppl. Table 7 Genome databank entries and genomic database positions of pufferfish (*Fugu rubripes*, Fr) *connexins* according to FUGU4.0 (Ensembl v39). Chromosomal locations were not available as of July 1 2006. We include the gene identifications in a previous Fugu genebuild, as some of those gene predictions have now been removed, while we still consider the sequences as connexins. Please note that the locations, scaffold number and Ensembl gene identifications may change upon future updates of the genomic database. If the coding sequence is divided into several exons, they are indicated by E1, E2, etc.

Naming Ensembl v36/37	Previous name Ensembl v35 and earlier	Scaffold Ensembl v36/37	From	To	GenBank accession no.
Fr153654	Fr153654	229	200916	201707	
Frscf38 ^a	Fr137409	38	1498637	1499401	
Frscf64	Fr124726 ^e	64	E1: 166710 E2: 166140 E3: 165367 E4: 165087	E1: 166880 E2: 166303 E3: 165483 E4: 165297	
Fr129456e1 ^b	Fr129456	23	99869	100660	AY739096 193088:193879 (no id.) AF411956 49418:50209 "connexin"
Fr158780 ^c	Fr158780	61	E1: 983046 E2: 983668	E1: 983581 E2: 983968	
Fr129456e2 ^b	Fr157125	23	104039	105046	AY739096 197257:198264 (no id.) AF411956 53589:54596 (no id.)
Fr158780e3 ^b	Fr137254	61	985141	985881	
Fr143712	Fr143712	34	E1: 1688045 E2: 1687406 E3: 1686977	E1: 1687875 E2: 1687707 E3: 1687319	
Fr125568	Fr125568	12	E1: 1275573 E2: 1274985	E1: 1275964 E2: 1275418	
Fr133739	Fr133739	3	1584130	1584871	
Fr144150	Fr144150	89	478047	478847	
Fr157320	Fr157320	6	1296110	1296832	
Fr137912	Fr137912	10	E1: 1871843 E2: 1872557	E1: 1871913 E2: 1873400	
Fr164573	Fr150812ic	6	E1: 1474887 E2: 1484693	E1: 1474957 E2: 1485557	
Fr130034 ^f	Fr130034ic	123	E1: 157832 E2: 169961	E1: 157902 E2: 150810	
Fr123845	Fr123845	487	55345	56472	
Fr158218	Fr158218	483	25964	26822	
Frscf483	Fr159440	483	E1: 23210 E2: 23444	E1: 23338 E2: 23636	

			E3: 23944	E3: 24457	
Fr159647	Fr125574ic	12	1263036	1264157	
Fr130494	Fr130494	268	229702	230868	
Fr142334	Fr142334	6	944370	945476	
Fr125574	Fr137519	103	959114	960262	
Fr129429	Fr129429	39	E1: 155177 E2: 154849 E3: 154414 E4: 154263 E5: 154042 E6: 153461	E1: 155245 E2: 155013 E3: 154706 E4: 154328 E5: 154186 E6: 153950	
Fr161792	Fr139615ic	127	E1: 776431 E2: 775174	E1: 776495 E2: 776351	
Fr150243ic	Fr150243ic	232	146359	147357	
Fr145344	Fr145344ic	15	1890603	1892060	
Fr160540	Fr148221	29	1136574	1137920	
Fr164495	Fr136128	68	1100512	1101768	
Fr122421 ^d	Fr122421	2150	E1: 5695 E2: 5382 E3: 4581	E1: 5934 E2: 5607 E3: 5335	
Fr162962	Fr121008	197	143008	144216	
Fr149277	Fr149277ic	38	1489367	1490749	
Fr124726 ^e	-	64	E1: 161778 E2: 160999 E3: 159968	E1: 161846 E2: 161561 E3: 160781	
Fr141787	Fr141787	6	1299270	1300334	
Frscf89	Fr144151ic	89	475053	476129	
Fr160171	Frscf387ic	103	977920	979170	
Fr164448	Frscf797	35	683284	684639	
Fr163704	Fr139587	45	135658	137139	
Fr149617	Fr149617	143	397020	398471	
Fr161672	-	400	838	2286	

a: Repeated once as nearly exact copy (99% identical) on same scaffold.

b: e1, e2 or e3 indicates that our suggested modification contains exon1, 2 or 3 of the Ensembl gene prediction; the remaining part of our suggested modification contains sequences predicted as introns by Ensembl.

c: Our suggested modification of this sequence contains exon 2 and 3 of the Ensembl prediction.

d: Identical (partial) sequence with gene identification *Fr163054* found on scaffold 3296.

e: Different sequences is predicted for the same gene identification number by Ensembl v36-37 and by Ensembl v35 and earlier.

f: Nearly identical sequence (*Fr150812*) found on scaffold 3625.

Suppl. Table 8 Genome databank entries and genomic positions of green spotted pufferfish (*Tetraodon nigroviridis*, Tn) *connexins* according to TETRAODON7 (Ensembl v39) as of July 1 2006. Please note that the locations, scaffold number and Ensembl gene identifications may change upon future updates of the genomic database. If the coding sequence is divided into several exons, they are indicated by E1, E2, etc.

Present naming	Chromosome or scaffold	From	To	GenBank accession no. ^f
Tn35526001	14	525273	526034	
Tn29165001	2	17702027	17702782	CR732917 (no id.)
Tn29164001	2	17698375	17699265	
Tn29166001	2	17704915	17705714	CR634537 (no id.)
Tn19592001	3	E1: 3723041 E2: 3722464 E3: 3721701 E4: 3721427 E5: 3720218	E1: 3723211 E2: 3722616 E3: 3721849 E4: 3721633 E5: 3720488	
Tn31361001	10	3080243	3081067	
Tn31360001	10	3076083	3076949	
Tn16997001ic	10	E1: 812112 E2: 812519	E1: 812422 E2: 812861	
Tn17530001 ^a	21	E1: 3956344 E2: 3956831	E1: 3956733 E2: 3957271	
Tn24102001	2	4470829	4471671	
Tn17091001	1	2009333	2010133	
Tn22232001	7	5690749	5691471	
Tn32001001	10	E1: 7900070 E2: 7900777	E1: 7900140 E2: 7901620	
Tn26779001	16	E1: 7782425 E2: 7768423	E1: 7782495 E2: 7769272	
Tn05142001	7762	71509	72669	
Tn35459001	14	1391299	1392159	CR721381 (no id.)
Tn35462001 ^b	14	E1: 1379792 E2: 1378832	E1: 1379920 E2: 1379443	
Tn35458001	14	E1: 1394877 E2: 1394528 E3: 1393595	E1: 1395005 E2: 1394692 E3: 1394152	
Tn17534001	21	3967733	3968819	CR693532, partial (no id.)
Tn13946001 ^c	14527	277225	278391	
Tn22263001	7	6046895	6048004	
Tn26030001	2	9879288	9880436	
Tn18746001 ^d	17	E1: 9215749 E2: 9215446 E3: 9215244 E4: 9214689 E5: 9214461 E6: 9213880	E1: 9215817 E2: 9215610 E3: 9214948 E4: 9214754 E5: 9214609 E6: 9214354	
Tn15227001ic	15	E1: 1114542 E2: 1113839	E1: 1114791 E2: 1114500	
Tn16594001	Un_random	E1: 133586493 E2: 133586651	E1: 133586556 E2: 133587846	
Tn19395001	14613	143750	145090	
Tn09585001	3	13116172	13117629	

Tn16136001	15random	1146258	1147514	
Tn01153001 ^e	3763	E1: 1303 E2: 1013 E3: 158	E1: 1544 E2: 1235 E3: 942	
Tn16902001	2	16087097	16088239	
Tn29168001	2	17710713	17712124	
Tn19591001	3	E1: 3718094 E2: 3717363 E3: 3716160	E1: 3718162 E2: 3717925 E3: 3716988	
Tn17092001	1	2012147	2013229	
Tn22231001	7	5687622	5688686	
Tn26028001	2	9862228	9863481	
Tn11616001	13678	2514	3923	
Tn10470001	21	541667	542431	
Tn06204001	8682	9059	10480	

a: Intron is modelled to give a sequence as similar as possible to its orthologs.

b: Identical sequence also found approx. 10 kb upstream.

c: Identical sequence (the first 300 nucleotides) found on scaffold 16340. Ensembl predicts this as an independent *connexin* (*Tn36946001*).

d: Identical sequence corresponding to exon 1, 2, and first part of exon 3 found on scaffold 4618. Ensembl predicts this as an independent *connexin* (*Tn01689001*).

e: Nearly identical sequence (last 800 nt) found on scaffold 6765. Ensembl predicts this as an independent *connexin* (*Tn03189001*).

f: None of these accession numbers identify the sequence as a *connexin*. All of them are named "*Tetraodon nigroviridis* full-length cDNA".

Suppl. Table 9 d_N (n) and d_S (s) values for full-length mammalian *connexins* as calculated by the Kumar method. Deviations (SE) were calculated by bootstrap analysis (1000 replications).

Cx	Hs-Pt	Mm-Rn	Hs-Mm	Hs-Cf	Hs-Bt	Hs-Md
25	n 0.0079±0.0040 s 0.0096±0.0092			n 0.0580±0.0115 s 0.2470±0.0415	n 0.0551±0.0116 s 0.3068±0.0480	n 0.1155±0.0171 s 0.6525±0.0893
26	n 0.0026±0.0025 s 0.0260±0.0103	n 0.0042±0.0029 s 0.1651±0.0317	n 0.0383±0.0090 s 0.5211±0.0724	n 0.0077±0.0038 s 0.5088±0.0707	n 0.0301±0.0091 s 0.6124±0.0821	n 0.0409±0.0100 s 0.6947±0.0981
29/31.3/35	n 0.0018±0.0018 s 0.0066±0.0047	n 0.0659±0.0117 s 0.1110±0.0244	n 0.2398±0.0264 s 0.4942±0.0653	n 0.1052±0.0156 ^a s 0.2915±0.0425	n 0.1503±0.0201 ^a s 0.4202±0.0561	n 0.5760±0.0475 s 1.0838±0.1489
31.3/33.1						n 0.5251±0.0444 s 0.9945±0.1220
30	n 0 s 0.0223±0.0106	n 0.0038±0.0026 s 0.2200±0.0368	n 0.0229±0.0066 s 0.6286±0.0791	n 0.0101±0.0041 s 0.4037±0.0554	n 0.0405±0.0096 s 0.9991±0.1668	n 0.0564±0.0106 s 0.8797±0.1239
30.2/31.9	n 0.0015±0.0014 s 0.0081±0.0054	n 0.0068±0.0034 s 0.1118±0.0217	n 0.0689±0.0117 s 0.3233±0.0441	n 0.0730±0.0134 s 0.2748±0.0437	n 0.0702±0.0115 s 0.2532±0.0380	n 0.1916±0.0221 s 0.7698±0.1125
30.3	n 0.0111±0.0052 s 0.0108±0.0060	n 0.0202±0.0069 s 0.2026±0.0332	n 0.0897±0.0159 s 0.4949±0.0589	n 0.0674±0.0121 s 0.4393±0.0577	n 0.0940±0.0130 s 0.3444±0.0468	n 0.2151±0.0262 s 0.9721±0.1289
31		n 0.0168±0.0058 s 0.1586±0.0283	n 0.1000±0.0145 s 0.4327±0.0573	n 0.0970±0.0153 s 0.3561±0.0465	n 0.0788±0.0136 s 0.3067±0.0439	n 0.1452±0.0199 s 0.5580±0.0801
31.1	n 0.0048±0.0026 s 0.0274±0.0117	n 0.0342±0.0079 s 0.2596±0.0410	n 0.1469±0.0201 s 0.5480±0.0731	n 0.1156±0.0169 s 0.4903±0.0629	n 0.1203±0.0168 s 0.4082±0.0556	n 0.1708±0.0243 s 0.8459±0.1126
32	n 0 s 0.0137±0.0067	n 0 s 0.1050±0.0220	n 0.0061±0.0030 s 0.2980±0.0425	n 0.0064±0.0033 s 0.2722±0.0372	n 0.0081±0.0037 s 0.2344±0.0361	n 0.0429±0.0094 s 0.6394±0.0732
33		n 0.0942±0.0134 s 0.1197±0.0251				
36	n 0 s 0.0217±0.0096	n 0.0059±0.0029 s 0.0970±0.0208	n 0.0112±0.0038 s 0.3424±0.0415	n 0.0194±0.0058 s 0.1862±0.0285	n 0.0071±0.0035 s 0.2454±0.0346	
37	n 0.0030±0.0021 s 0.0133±0.0069	n 0.0118±0.0040 s 0.1790±0.0261	n 0.0555±0.0089 s 0.3848±0.0450	n 0.0238±0.0060 s 0.3445±0.0423	n 0.0292±0.0066 s 0.2889±0.0367	n 0.1167±0.0156 s 0.7625±0.0827
39/40.1	n 0.0167±0.0045 s 0.0097±0.0049	n 0.0641±0.089 s 0.1853±0.0258	n 0.3809±0.0326 s 1.0607±0.1253		n 0.3209±0.0294 s 0.7106±0.0740	n 0.4080±0.0318 ^b s 1.3853±0.1816

40	n 0.0028±0.0019 s 0.0081±0.0046	n 0.0354±0.0070 s 0.1945±0.0296	n 0.1004±0.0130 s 0.5696±0.0721	n 0.0470±0.0081 s 0.3162±0.0385	n 0.0676±0.0096 s 0.3864±0.0437	n 0.1639±0.0188 s 0.8184±0.0921
43	n 0.0013±0.0013 s 0.0125±0.0054	n 0.0013±0.0012 s 0.1836±0.0270	n 0.0123±0.0036 s 0.5499±0.0576	n 0.0149±0.0038 s 0.2968±0.0367	n 0.0101±0.0034 s 0.3131±0.0393	n 0.0151±0.0047 s 0.8122±0.0804
45	n 0.0010±0.0009 s 0.0073±0.0042	n 0 ^a s 0.1356±0.0238	n 0.0154±0.0046 s 0.2711±0.0328	n 0.0130±0.0039 s 0.2705±0.0337	n 0.0180±0.0046 s 0.3058±0.0359	n 0.0344±0.0066 s 0.6286±0.0644
46	n 0.0033±0.0024 s 0.0428±0.0106	n 0.0245±0.0055 s 0.1376±0.0207	n 0.1579±0.0162 s 0.5896±0.0586	n 0.1623±0.0167 s 0.4598±0.0479	n 0.2120±0.0203 s 0.5028±0.0564	n 0.2186±0.0213 s 1.1341±0.1405
47	n 0.0019±0.0013 s 0.0108±0.0053	n 0.0111±0.0033 s 0.1809±0.0214	n 0.0800±0.0090 s 0.6276±0.0576	n 0.1115±0.0134 s 0.4365±0.0479	n 0.1119±0.0131 ^c s 0.5544±0.0536	n 0.2592±0.0289 s 1.1641±0.1839
50	n 0.0053±0.0024 s 0.0084±0.0051	n 0.0095±0.0030 s 0.1369±0.0200	n 0.0584±0.0087 s 0.5006±0.0521	n 0.0549±0.0081 s 0.3503±0.0400	n 0.0695±0.0097 s 0.3590±0.0407	n 0.1076±0.0131 s 0.7236±0.0699
57/62 ^d	n 0.0030±0.0016 s 0.0117±0.0054	n 0.0292±0.0050 s 0.1796±0.0238	n 0.1047±0.0116 s 0.3842±0.0383	n 0.0914±0.0094 s 0.2282±0.0251	n 0.0947±0.0101 s 0.2465±0.0272	n 0.2905±0.0225 s 0.6950±0.0626
59	n 0.0027±0.0015 s 0.0018±0.0017			n 0.1128±0.0109 s 0.2352±0.0268		
Mean	n 0.0037 s 0.0143	n 0.0231 s 0.1612	n 0.0939 s 0.5012	n 0.0627 s 0.3373	n 0.0836 s 0.4104	n 0.1944 s 0.8534
Median	n 0.0027 s 0.0108	n 0.0118 s 0.1651	n 0.0745 s 0.4978	n 0.0580 s 0.3162	n 0.0695 s 0.3444	n 0.1639 s 0.8122

a: Not complete in 3'-end.

b: Not computable by Kumar method. Modified Nei-Gojobori method used with R=1.28 (calculated by Tamura 3-parameter or Tamura-Nei method).

c: Not complete in 5'-end.

d: Modelled with reading frame in two exons.

Suppl. Table 10 Selection pressure ($d_N/d_S \pm SE$) on full-length mammalian *connexins*.

Cx	Hs-Pt	Mm-Rn	Hs-Mm	Hs-Cf	Hs-Bt	Hs-Md
25	0.8829±0.8919 ^e			0.2348±0.0610	0.1796±0.0471	0.1770±0.0357
26	0.1000±0.1040	0.0254±0.0182	0.0735±0.0201	0.0151±0.0078	0.0492±0.0163	0.0589±0.0166
29/31.3/35	0.2727±0.3348 ^e	0.5937±0.1678 ^e	0.4852±0.0835	0.3609±0.0751 ^a	0.3577±0.0676 ^a	0.5315±0.0852
31.3/33.1						0.5280±0.0787
30	0 ^e	0.0173±0.0122	0.0364±0.0115	0.0250±0.0107	0.0405±0.0118	0.0641±0.0151
30.2/31.9	0.1852±0.2124 ^e	0.0608±0.0326	0.2131±0.0464	0.2656±0.0645	0.2773±0.0616	0.2489±0.0463
30.3	1.0278±0.7469 ^e	0.0997±0.0378	0.1812±0.0387	0.1534±0.0341	0.2729±0.0529	0.2213±0.0398
31		0.1059±0.0412	0.2311±0.0454	0.2724±0.0558	0.2569±0.0576	0.2602±0.0516
31.1	0.1752±0.1208 ^e	0.1317±0.0369	0.2681±0.0512	0.2358±0.0459	0.2947±0.0575	0.2019±0.0393
32	0 ^e	0	0.0205±0.0105	0.0235±0.0125	0.0346±0.0167	0.0671±0.0166
33		0.7870±0.1994 ^e				
36	0 ^e	0.0608±0.0326	0.0327±0.0118	0.1042±0.0350	0.0289±0.0148	
37	0.2256±0.1965 ^e	0.0659±0.0243	0.1442±0.0286	0.0691±0.0194	0.1011±0.0262	0.1530±0.0263
39/40.1	1.7216±0.9857 ^e	0.3459±0.4827	0.3591±0.0524		0.4516±0.0626	0.2945±0.0449 ^d
40	0.3457±0.3059 ^e	0.1820±0.0454	0.1763±0.0319	0.1486±0.0314	0.1749±0.0318	0.2003±0.0322
43	0.1040±0.1133 ^e	0.0071±0.0066	0.0224±0.0070	0.0502±0.0142	0.0323±0.0116	0.0186±0.0061
45	0.1370±0.1463 ^e	0 ^a	0.0568±0.0183	0.0481±0.0156	0.0589±0.0166	0.0547±0.0119
46	0.0771±0.0592	0.1781±0.0481	0.2678±0.0383	0.3530±0.0517	0.4216±0.0622	0.1928±0.0304
47	0.1759±0.1481 ^e	0.0614±0.0196	0.1275±0.0185	0.2554±0.0416	0.2018±0.0306 ^b	0.2227±0.0431
50	0.6310±0.4779 ^e	0.0694±0.0241	0.1167±0.0212	0.1567±0.0292	0.1936±0.0348	0.1487±0.0231
57/62 ^c	0.2564±0.1808	0.1626±0.0352	0.2725±0.0406	0.4005±0.0603	0.3842±0.0590	0.4180±0.0497
59	1.5000±1.6436 ^e			0.4796±0.0717		
Mean	0.4115	0.1556	0.1714	0.1922	0.2006	0.2138
Median	0.1852	0.0694	0.1603	0.1567	0.1936	0.2003

a: Not complete in 3'-end.

b: Not complete in 5'-end.

c: Modelled with the reading frame in two exons.

d: Not computable by Kumar method. Modified Nei-Gojobori method used with R=1.28 (calculated by Tamura-Nei method).

e: While most of the comparisons showed statistically significant ($p < 0.01$) purifying selection pressure, the marked comparisons did not.

Suppl. Table 11 d_N (n) and d_S (s) values for the conserved domains of mammalian *connexins* as calculated by the Kumar method. Deviations (SE) were calculated by bootstrap analysis (1000 replications).

Cx	Hs-Pt	Mm-Rn	Hs-Mm	Hs-Cf	Hs-Bt	Hs-Md
25	n 0.0106±0.0052 s 0.0117±0.0113			n 0.0448±0.0115 s 0.2460±0.0463	n 0.0476±0.0117 s 0.3072±0.0551	n 0.0857±0.0157 s 0.6680±0.1068
26	n 0.0034±0.0033 s 0.0334±0.0135	n 0.0053±0.0036 s 0.2051±0.0423	n 0.0412±0.0105 s 0.5149±0.0846	n 0.0077±0.0044 s 0.5342±0.0816	n 0.0182±0.0078 s 0.5492±0.0862	n 0.0400±0.0110 s 0.7133±0.1171
29/31.3/35	n 0 s 0.0103±0.0074	n 0.0362±0.0096 s 0.1358±0.0339	n 0.1830±0.0256 s 0.5583±0.0932	n 0.0650±0.0129 s 0.3320±0.0556	n 0.0949±0.0165 s 0.4452±0.0686	n 0.3614±0.0400 s 0.8725±0.1305
31.3/33.1						n 0.3892±0.0426 s 0.8999±0.1393
30	n 0 s 0.0270±0.0138	n 0.0055±0.0038 s 0.1944±0.0414	n 0.0050±0.0034 s 0.6061±0.0976	n 0.0022±0.0022 s 0.4356±0.0729	n 0.0207±0.0080 s 0.9862±0.2172	n 0.0310±0.0103 s 0.9877±0.2775
30.2/31.9	n 0 s 0	n 0.0048±0.0034 s 0.0949±0.0266	n 0.0161±0.0064 s 0.2424±0.0471	n 0.0247±0.0079 s 0.1932±0.0447	n 0.0217±0.0074 s 0.2109±0.0458	n 0.0643±0.0157 s 0.7456±0.1487
30.3	n 0.0028±0.0027 s 0.0111±0.0074	n 0.0131±0.0069 s 0.2119±0.0423	n 0.0431±0.0120 s 0.4489±0.0708	n 0.0360±0.0106 s 0.4165±0.0741	n 0.0527±0.0120 s 0.3493±0.0620	n 0.0842±0.0164 s 0.7205±0.1170
31		n 0.0034±0.0033 s 0.1878±0.0412	n 0.0455±0.0105 s 0.5069±0.0818	n 0.0429±0.0119 s 0.3209±0.0572	n 0.0356±0.0109 s 0.2422±0.0488	n 0.0681±0.0152 s 0.4242±0.0746
31.1	n 0.0023±0.0022 s 0.0273±0.0150	n 0.0198±0.0069 s 0.2750±0.0528	n 0.0770±0.0154 s 0.5052±0.0795	n 0.0645±0.0156 s 0.4324±0.0673	n 0.0770±0.0176 s 0.3958±0.0655	n 0.1082±0.0223 s 0.8309±0.1392
32	n 0 s 0.0169±0.0102	n 0 s 0.1216±0.0295	n 0.0097±0.0048 s 0.3256±0.0605	n 0.0102±0.0050 s 0.2859±0.0520	n 0.0129±0.0056 s 0.2641±0.0500	n 0.0412±0.0118 s 0.5876±0.0919
33		n 0.0288 0.0083 s 0.1123 0.0311				
36	n 0 s 0.0254±0.0145	n 0.0055±0.0036 s 0.0496±0.0186	n 0.0027±0.0026 s 0.3293±0.0596	n 0.0027±0.0027 s 0.1272±0.0334	n 0.0027±0.0027 s 0.1940±0.0402	n 0.0027±0.0027 s 0.6615±0.1113
37	n 0 s 0.0109±0.0078	n 0.0056±0.0038 s 0.1945±0.0382	n 0.0154±0.0060 s 0.2860±0.0517	n 0.0028±0.0027 s 0.2956±0.0539	n 0.0079±0.0044 s 0.2583±0.0474	n 0.0401±0.0122 s 0.6157±0.0930
39/40.1	n 0.0023±0.0023 s 0.0103± 0.0074	n 0.0260±0.0078 s 0.1755±0.0369	n 0.2123±0.0319 s 0.9475±0.1666		n 0.1342±0.0224 s 0.4969±0.0741	n 0.6017±0.0706 ^a s 1.0348±0.1466
40	n 0	n 0.0050±0.0035	n 0.0301±0.0095	n 0.0335±0.0108	n 0.0438±0.0113	n 0.0561±0.0136

	s 0.0056±0.0056	s 0.1557±0.0337	s 0.4736±0.0827	s 0.2848±0.0554	s 0.3500±0.0604	s 0.7200±0.1125
43	n 0 s 0.0163±0.0096	n 0 s 0.1559±0.0367	n 0.0028±0.0027 s 0.4463±0.0731	n 0.0050±0.0034 s 0.2434±0.0466	n 0.0028±0.0027 s 0.2057±0.0444	n 0.0058±0.0041 s 0.7226±0.1140
45	n 0 s 0.0109±0.0073	n 0 s 0.0993±0.0298	n 0 s 0.3150±0.0579	n 0 s 0.2472±0.0464	n 0.0028±0.0027 s 0.3499±0.0613	n 0.0291±0.0098 s 0.6997±0.1580
46	n 0.0050±0.0049 s 0.0426±0.0199	n 0 s 0.1621±0.0361	n 0.0146±0.0060 s 0.4080±0.0758	n 0.0106±0.0053 s 0.3320±0.0637	n 0.0243±0.0078 s 0.3318±0.0645	n 0.0216±0.0081 s 0.9945±0.2414
47	n 0 s 0.0106±0.0072	n 0 s 0.1556±0.0347	n 0.0149±0.0060 s 0.4694±0.0808	n 0.0077±0.0044 s 0.3481±0.0659		n 0.0547±0.0169 s 0.9443±0.2725
50	n 0.0027±0.0026 s 0.0054±0.0053	n 0 s 0.1702±0.396	n 0.0132±0.0059 s 0.4647±0.0793	n 0.0106±0.0052 s 0.3280±0.0617	n 0.0134±0.0058 s 0.2919±0.0544	n 0.0170±0.0075 s 0.6851±0.1141
57/62	n 0 s 0.0055±0.0051	n 0.0071±0.0054 s 0.1748±0.0389	n 0.0115±0.0054 s 0.3331±0.0608	n 0.0143±0.0058 s 0.2095±0.0420	n 0.0312±0.0094 s 0.2124±0.0439	n 0.0605±0.0136 s 0.5786±0.0902
59	n 0 s 0			n 0.0360±0.0094 s 0.1803±0.0394		
Mean	n 0.0015 s 0.0148	n 0.0087 s 0.1632	n 0.0410 s 0.4545	n 0.0222 s 0.3049	n 0.0358 s 0.3578	n 0.1081 s 0.7554
Median	n 0 s 0.0109	n 0.0053 s 0.1662	n 0.0152 s 0.4568	n 0.0106 s 0.2956	n 0.0230 s 0.3195	n 0.0554 s 0.7203

a: Not computable by Kumar method. Modified Nei-Gojobori method used with R=2.17 (calculated by Tamura-Nei method).

Suppl. Table 12 Selection pressure ($d_N/d_S \pm SE$) on the conserved domains of mammalian *connexins*.

Cx	Hs-Pt	Mm-Rn	Hs-Mm	Hs-Cf	Hs-Bt	Hs-Md
25	0.9060±0.9814			0.1821±0.0580	0.1549±0.0471	0.1283±0.0312
26	0.1018±0.1070	0.0258±0.0183	0.0800±0.0243	0.0144±0.0085	0.0331±0.0151	0.0561±0.0180
29/31.3/35	0	0.2666±0.0971	0.3278±0.0714	0.1958±0.0508	0.2132±0.0495	0.4142±0.0771
31.3/33.1						0.4325±0.0820
30	0	0.0283±0.0205	0.0082±0.0058	0.0051±0.0051	0.0210±0.0093	0.0314±0.0137
30.2/31.9	0	0.0506±0.0385	0.0664±0.0294	0.1278±0.0505	0.1029±0.0416	0.0862±0.0272
30.3	0.2523±0.2957	0.0618±0.0348	0.0960±0.0307	0.0864±0.0297	0.1509±0.0436	0.1169±0.0296
31		0.0181±0.0180	0.0898±0.0253	0.1337±0.0441	0.1470±0.0539	0.1605±0.0456
31.1	0.0842±0.0929	0.0720±0.0286	0.1524±0.0388	0.1492±0.0429	0.1945±0.0549	0.1302±0.0346
32	0	0	0.0298±0.0157	0.0357±0.0187	0.0488±0.0231	0.0701±0.0229
33		0.2565±0.1025				
36	0	0.1109±0.0836	0.0082±0.0080	0.0212±0.0219	0.0139±0.0142	0.0041±0.0041
37	0	0.0288±0.0203	0.0538±0.0231	0.0095±0.0093	0.0306±0.0179	0.0651±0.0221
39/40.1	0.2233±0.2750	0.1481±0.0543	0.2241±0.0518		0.2701±0.0605	0.5815±0.1070 ^a
40	0	0.0321±0.0235	0.0636±0.0229	0.1176±0.0443	0.1251±0.0388	0.0779±0.0225
43	0	0	0.0063±0.0061	0.0205±0.0145	0.0136±0.0135	0.0080±0.0058
45	0	0	0	0	0.0080±0.0078	0.0416±0.0169
46	0.1174±0.1274	0	0.0358±0.0161	0.0319±0.0171	0.0732±0.0275	0.0217±0.0097
47	0	0	0.0317±0.0139	0.0221±0.0133		0.0579±0.0245
50	0.5000±0.6875	0	0.0284±0.0136	0.0323±0.0170	0.0459±0.0216	0.0248±0.0117
57/62	0	0.0406±0.0322	0.0345±0.0174	0.0683±0.0309	0.1469±0.0537	0.1046±0.0286
59	0			0.1997±0.0680		
Mean	0.1150	0.0600	0.0743	0.0765	0.0996	0.1307
Median	0	0.0288	0.0448	0.0357	0.0881	0.0740

a: Not computable by Kumar method. Modified Nei-Gojobori method used with R=2.17 (calculated by Tamura-Nei method).

Suppl. Table 13 d_N (n) and d_S (s) values for the variable domains of mammalian *connexins* as calculated by the Kumar method. Deviations (SE) were calculated by bootstrap analysis (1000 replications).

Cx	Hs-Pt	Mm-Rn	Hs-Mm	Hs-Cf	Hs-Bt	Hs-Md
25	n 0.0083±0.0082 s 0			n 0.1118±0.0392 s 0.2544±0.1169	n 0.0839±0.0364 s 0.3254±0.1277	n 0.2441±0.1054 s 0.6036±0.2872
26	n 0 s 0	n 0 s 0.0404±0.0304	n 0.0273±0.0176 s 0.5617±0.2585	n 0.0075±0.0075 s 0.4142±0.1824	n 0.0776±0.0343 ^a s 1.0763±0.4502	n 0.0426±0.0230 s 0.6669±0.2787
29/31.3/35	n 0.0048±0.0049 s 0	n 0.1319±0.0296 s 0.0543±0.0319	n 0.3705±0.0718 s 0.3765±0.1090	n 0.1955±0.0417 s 0.2177±0.0653	n 0.4604±0.0813 s 0.4468±0.0993	n 1.1079±0.5586 s 1.6416±0.8334
31.3/33.1						n 1.2603±0.6610 s 1.6416±0.8334
30	n 0 s 0.0113±0.0119	n 0 s 0.2885±0.0941	n 0.0613±0.0213 s 0.6801±0.1994	n 0.0268±0.0124 s 0.3289±0.0914	n 0.0824±0.0342 s 1.0397±0.4280	n 0.1108±0.0284 s 0.7358±0.2017
30.2/31.9	n 0.0038±0.0038 s 0.0696±0.0264	n 0.0102±0.0073 s 0.1426±0.0433	n 0.2050±0.0369 s 0.4714±0.1025	n 0.2056±0.0442 s 0.5179±0.1227	n 0.2220±0.0404 s 0.4380±0.0943	n 0.5479±0.1354 s 0.9975±0.3551
30.3	n 0.0279±0.0152 s 0.0104±0.0102	n 0.0352±0.0166 s 0.2166±0.0648	n 0.1922±0.0439 s 0.5975±0.1391	n 0.1343±0.0322 s 0.4837±0.1129	n 0.1841±0.0342 s 0.3361±0.0839	n 0.5812±0.0921 ^a s 3.1681±1.1194
31		n 0.0439±0.0178 s 0.1050±0.0373	n 0.2200±0.0440 s 0.3066±0.0765	n 0.2162±0.0401 s 0.4360±0.1099	n 0.1722±0.0390 s 0.4507±0.1157	n 0.3242±0.0703 s 0.8828±0.2674
31.1	n 0.0094±0.0067 s 0.0278±0.0204	n 0.0623±0.0199 s 0.2365±0.0711	n 0.3211±0.0885 s 0.6479±0.2241	n 0.2212±0.0498 s 0.6305±0.1689	n 0.2081±0.0400 s 0.4337±0.1202	n 0.2967±0.0765 s 0.9358±0.2846
32	n 0 s 0.0088±0.0087	n 0 s 0.0805±0.0328	n 0 s 0.2627±0.0646	n 0 s 0.2533±0.0639	n 0 s 0.1877±0.0499	n 0.0451±0.0148 s 0.7463±0.1645
33		n 0.2194±0.0354 s 0.1420±0.0442				
36	n 0 s 0.0176±0.0133	n 0.0064±0.0046 s 0.1528±0.0411	n 0.0223±0.0083 s 0.3603±0.0658	n 0.0416±0.0137 s 0.2639±0.0511	n 0.0129±0.0075 s 0.3110±0.0638	
37	n 0.0065±0.0046 s 0.0153±0.0112	n 0.0189±0.0077 s 0.1629±0.0370	n 0.1047±0.0188 s 0.5159±0.0923	n 0.0490±0.0125 s 0.4029±0.0759	n 0.0547±0.0124 s 0.3209±0.0602	n 0.2235±0.0593 s 0.9122±0.2731
39/40.1	n 0.0299±0.0083 s 0.0093±0.0065	n 0.1044±0.0163 s 0.2103±0.0392	n 0.7161±0.0856 s 1.1682±0.2166		n 0.5810±0.0723 s 0.9334±0.1619	n 0.6622±0.3391 s 1.7833±0.7420
40	n 0.0055±0.0038	n 0.0671±0.0144	n 0.1779±0.0463	n 0.0602±0.0124	n 0.0914±0.0154	n 0.2833±0.0892

	s 0.0106±0.0078	s 0.2346±0.0460	s 0.7098±0.2288	s 0.3414±0.0588	s 0.4235±0.0663	s 0.9993±0.3130
43	n 0.0024±0.0024 s 0.0093±0.0063	n 0.0024±0.0023 s 0.2086±0.0382	n 0.0206±0.0065 s 0.6584±0.0961	n 0.0101±0.0045 s 0.3717±0.0616	n 0.0165±0.0058 s 0.4265±0.0738	n 0.0231±0.0083 s 0.9081±0.1424
45	n 0.0017±0.0017 s 0.0044±0.0044	n 0 ^b s 0.1700±0.0373	n 0.0277±0.0082 s 0.2341±0.0390	n 0.0182±0.0058 s 0.2911±0.0481	n 0.0300±0.0073 s 0.2670±0.0428	n 0.0385±0.0094 s 0.5879±0.0841
46	n 0.0020±0.0020 s 0.0438±0.0137	n 0.0440±0.0099 s 0.1234±0.0265	n 0.2883±0.0341 s 0.7583±0.1126	n 0.3039±0.0307 s 0.5351±0.0738	n 0.3822±0.0419 s 0.7320±0.1061	n 0.3566±0.1629 s 1.5031±0.6782
47	n 0.0033±0.0023 s 0.0101±0.0070	n 0.0189±0.0059 s 0.1989±0.0309	n 0.1312±0.0167 s 0.7324±0.0839	n 0.1937±0.0235 s 0.5121±0.0728	n 0.1756±0.0239 s 0.6270±0.0775	n 0.4469±0.0458 s 1.1628±0.2034
50	n 0.0080±0.0040 s 0.0142±0.0085	n 0.0161±0.0051 s 0.1163±0.0246	n 0.0928±0.0144 s 0.5383±0.0716	n 0.0927±0.0144 s 0.3799±0.0532	n 0.1123±0.0170 s 0.4125±0.0581	n 0.1839±0.0235 s 0.7723±0.1005
57/62	n 0.0047±0.0027 s 0.0154±0.0082	n 0.0422±0.0081 s 0.1839±0.0295	n 0.1642±0.0187 s 0.4162±0.0469	n 0.1368±0.0152 s 0.2358±0.0439	n 0.1299±0.0153 s 0.2529±0.0336	n 0.4656±0.0408 s 0.8041±0.0947
59	n 0.0041±0.0023 s 0.0027±0.0026			n 0.1588±0.0156 s 0.2762±0.0363		
Mean	n 0.0064 s 0.0148	n 0.0433 s 0.1615	n 0.1746 s 0.5554	n 0.1149 s 0.3761	n 0.1620 s 0.4969	n 0.3702 s 1.0158
Median	n 0.0041 s 0.0104	n 0.0189 s 0.1629	n 0.1477 s 0.5500	n 0.1118 s 0.3717	n 0.1123 s 0.4265	n 0.2900 s 0.9102

a: Not computable by Kumar method. Modified Nei-Gojobori method used with R=2.00 (Hs-Bt Cx26) or 0.50 (Hs-Md Cx30.3) (calculated by Tamura-Nei method).

b: Not full length.

Suppl. Table 14 Selection pressure ($d_N/d_S \pm SE$) on the variable domains of mammalian *connexins*.

Cx	Hs-Pt	Mm-Rn	Hs-Mm	Hs-Cf	Hs-Bt	Hs-Md
25	∞			0.4395 $\pm$ 0.2540	0.2578 $\pm$ 0.1508	0.4044 $\pm$ 0.2598
26	0	0	0.0486 $\pm$ 0.0385	0.0181 $\pm$ 0.0198	0.0721 $\pm$ 0.0439	0.0639 $\pm$ 0.0436
29/31.3/35	∞	2.4291 $\pm$ 1.5276 ^a	0.9841 $\pm$ 0.3428	0.8980 $\pm$ 0.3305 ^b	1.0304 $\pm$ 0.2925 ^b	0.6749 $\pm$ 0.4829
31.3/33.1						0.7677 $\pm$ 0.5604
30	0	0	0.0901 $\pm$ 0.0410	0.0815 $\pm$ 0.0440	0.0793 $\pm$ 0.0463	0.1506 $\pm$ 0.0565
30.2/31.9	0.0546 $\pm$ 0.0584	0.0715 $\pm$ 0.0556	0.4349 $\pm$ 0.1228	0.3970 $\pm$ 0.1270	0.5068 $\pm$ 0.1429	0.5493 $\pm$ 0.2380
30.3	2.6827 $\pm$ 3.0098	0.1625 $\pm$ 0.0908	0.3217 $\pm$ 0.1048	0.2777 $\pm$ 0.0929	0.5478 $\pm$ 0.1704	^d
31		0.4181 $\pm$ 0.2254	0.7175 $\pm$ 0.2295	0.4959 $\pm$ 0.1552	0.3821 $\pm$ 0.1308	0.3672 $\pm$ 0.1368
31.1	0.3381 $\pm$ 0.3459	0.2634 $\pm$ 0.1156	0.4956 $\pm$ 0.2192	0.3508 $\pm$ 0.1228	0.4798 $\pm$ 0.1618	0.3171 $\pm$ 0.1264
32	0	0	0	0	0	0.0604 $\pm$ 0.0239
33		1.5451 $\pm$ 0.5417				
36	0	0.0419 $\pm$ 0.0321	0.0619 $\pm$ 0.0257	0.1576 $\pm$ 0.0602	0.0415 $\pm$ 0.0256	
37	0.4248 $\pm$ 0.4326	0.1160 $\pm$ 0.0541	0.2029 $\pm$ 0.0514	0.1216 $\pm$ 0.0386	0.1705 $\pm$ 0.0502	0.2450 $\pm$ 0.0980
39/40.1	3.2151 $\pm$ 2.4178 ^c	0.4964 $\pm$ 0.1207	0.6130 $\pm$ 0.1352		0.6225 $\pm$ 0.1329	0.3713 $\pm$ 0.2450
40	0.5189 $\pm$ 0.5237	0.2860 $\pm$ 0.0831	0.2506 $\pm$ 0.1038	0.1763 $\pm$ 0.0473	0.2158 $\pm$ 0.0496	0.2835 $\pm$ 0.1259
43	0.2581 $\pm$ 0.3117	0.0115 $\pm$ 0.0112	0.0313 $\pm$ 0.0109	0.0272 $\pm$ 0.0129	0.0387 $\pm$ 0.0152	0.0254 $\pm$ 0.0100
45	0.3864 $\pm$ 0.5464	0	0.1183 $\pm$ 0.0402	0.0625 $\pm$ 0.0224	0.1124 $\pm$ 0.0327	0.0655 $\pm$ 0.0185
46	0.0457 $\pm$ 0.0478	0.3566 $\pm$ 0.1109	0.3802 $\pm$ 0.0722	0.5679 $\pm$ 0.0971	0.5221 $\pm$ 0.0949	0.2372 $\pm$ 0.1523
47	0.3267 $\pm$ 0.3211	0.0950 $\pm$ 0.0331	0.1791 $\pm$ 0.0307	0.3782 $\pm$ 0.0707	0.2801 $\pm$ 0.0515	0.3843 $\pm$ 0.0779
50	0.5634 $\pm$ 0.4394	0.1384 $\pm$ 0.0527	0.1724 $\pm$ 0.0352	0.2440 $\pm$ 0.0510	0.2722 $\pm$ 0.0563	0.2381 $\pm$ 0.0434
57/62	0.3052 $\pm$ 0.2391	0.2295 $\pm$ 0.0574	0.3945 $\pm$ 0.0632	0.5802 $\pm$ 0.1258	0.5136 $\pm$ 0.0912	0.5790 $\pm$ 0.0850
59	1.5185 $\pm$ 1.6923			0.5749 $\pm$ 0.0943		
Mean	0.6258	0.3506	0.3053	0.3078	0.3234	0.3214
Median	0.3267	0.1384	0.2268	0.2777	0.2722	0.3003

a: Statistically significant positive selection pressure ($p=0.037$).

b: Not full length.

c: Statistically significant positive selection pressure ($p=0.028$).

d: Not computed due to very high d_S (3.17).

Suppl. Table 15 Nucleotide distance ($\pm$ SE) for full-length coding sequences, introns and 3'-UTRs for some mammalian *connexins* as determined by the Tamura-Kumar modification of the LogDet method. Deviations (SE) were calculated by bootstrap analysis (1000 replications).

		Hs-Pt	Mm-Rn	Hs-Mm	Hs-Cf	Hs-Bt	Hs-Md
30.3	Cds	0.0111 $\pm$ 0.0037	0.0726 $\pm$ 0.0099	0.1937 $\pm$ 0.0165	0.1434 $\pm$ 0.0144	0.1589 $\pm$ 0.0152	0.3695 $\pm$ 0.0261
	3'-UTR	0.0456 $\pm$ 0.0156	0.1059 $\pm$ 0.0618	0.4858 $\pm$ 0.0603	0.5326 $\pm$ 0.1101	0.3858 $\pm$ 0.0594	1.3913 $\pm$ 0.4315
36	Cds	0.0052 $\pm$ 0.0022	0.0301 $\pm$ 0.0055	0.0999 $\pm$ 0.0102	0.0648 $\pm$ 0.0080	0.0683 $\pm$ 0.0085	^a
	Intron	0.0117 $\pm$ 0.0035	0.0643 $\pm$ 0.0077	0.3610 $\pm$ 0.0246	0.2984 $\pm$ 0.0196	0.2849 $\pm$ 0.0205	0.7442 $\pm$ 0.0464
	3'-UTR	0.0190 $\pm$ 0.0039	0.0562 $\pm$ 0.0069	0.2889 $\pm$ 0.0175	0.2080 $\pm$ 0.0135	0.2310 $\pm$ 0.0159	0.4525 $\pm$ 0.0266
37	Cds	0.0063 $\pm$ 0.0026	0.0607 $\pm$ 0.0078	0.1446 $\pm$ 0.0132	0.1087 $\pm$ 0.0108	0.1065 $\pm$ 0.0108	0.2552 $\pm$ 0.0180
	3'-UTR	0.0062 $\pm$ 0.0035	0.1320 $\pm$ 0.0174	0.4756 $\pm$ 0.0410	0.3735 $\pm$ 0.0328	^b	1.4018 $\pm$ 0.5174
39/40.1	Cds	0.0168 $\pm$ 0.0039	0.1072 $\pm$ 0.0109	0.5443 $\pm$ 0.0354	^a	0.4324 $\pm$ 0.0282	0.5513 $\pm$ 0.0311
	Intron	0.0101 $\pm$ 0.0022	0.1427 $\pm$ 0.0111	1.4253 $\pm$ 0.1538	0.7612 $\pm$ 0.0433	0.8514 $\pm$ 0.0477	1.4864 $\pm$ 0.2227
	3'-UTR	0.0183 $\pm$ 0.0051	0.1855 $\pm$ 0.0147	1.1341 $\pm$ 0.1135	^b	1.0955 $\pm$ 0.1065	1.4713 $\pm$ 0.4363
40	Cds	0.0047 $\pm$ 0.0020	0.0777 $\pm$ 0.0091	0.2076 $\pm$ 0.0156	0.1143 $\pm$ 0.0110	0.1470 $\pm$ 0.0124	0.3067 $\pm$ 0.0174
	3'-UTR	0.0099 $\pm$ 0.0029	0.1532 $\pm$ 0.1111	0.4631 $\pm$ 0.0308	0.2560 $\pm$ 0.0220	0.2865 $\pm$ 0.0217	1.0794 $\pm$ 0.0821
43	Cds	0.0052 $\pm$ 0.0022	0.0501 $\pm$ 0.0068	0.1300 $\pm$ 0.0120	0.0834 $\pm$ 0.0090	0.0803 $\pm$ 0.0080	0.1654 $\pm$ 0.0139
	3'-UTR	0.0058 $\pm$ 0.0019	0.1082 $\pm$ 0.0088	0.2242 $\pm$ 0.0145	0.1190 $\pm$ 0.0089	0.1268 $\pm$ 0.0098	0.3877 $\pm$ 0.0201
45	Cds	0.0034 $\pm$ 0.0018	0.0340 $\pm$ 0.0058	0.0831 $\pm$ 0.0085	0.0827 $\pm$ 0.0084	0.0912 $\pm$ 0.0092	0.1639 $\pm$ 0.0126
	3'-UTR	0.0027 $\pm$ 0.0027	^b	0.1468 $\pm$ 0.0181	0.0987 $\pm$ 0.0157	0.1039 $\pm$ 0.0157	0.1997 $\pm$ 0.0237

a: Not full length.

b: Lacking.

Suppl. Table 16 Nucleotide distance ($\pm$ SE) for full-length Cx43- and Cx31.3-like (pseudo-)genes as determined by the Tamura-Kumar modification of the LogDet method. Deviations (SE) were calculated by bootstrap analysis (1000 replications).

	Hs-Pt	Hs-Bt
43p1	0.0115 $\pm$ 0.0030	-
43pX	0.0035 $\pm$ 0.0018	0.2941 $\pm$ 0.0196
31.3	0.0036 $\pm$ 0.0020	-
31.3p1	0.0117 $\pm$ 0.0036	-