

Title	Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences
Authors	Ivanov, Ivaylo P.;Firth, Andrew E.;Michel, Audrey M.;Atkins, John F.;Baranov, Pavel V.
Publication date	2011
Original Citation	Ivanov, I. P., Firth, A. E., Michel, A. M., Atkins, J. F. and Baranov, P. V. (2011) 'Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences', Nucleic Acids Research, 39(10), pp. 4220-4234. doi: 10.1093/nar/gkr007
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr007 - 10.1093/nar/gkr007
Rights	© 2011, the Authors. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited. - http://creativecommons.org/licenses/by-nc/2.5
Download date	2024-04-25 15:12:48
Item downloaded from	https://hdl.handle.net/10468/5024



Supplementary Dataset 1

Nucleotide and amino acid sequences of human genes of candidate or known non-AUG initiated N-terminal extensions. A) The new 42 candidates identified in this study listed in the same order as in table 1A of the main text. B) The 17 known cases that passed our qualitative test for conservation of the extension listed in the same order as in table 1B of the main text. C) The 12 known cases that did not pass our qualitative test for conservation of the extension listed in no particular order. Candidate initiation codons are highlighted in green. Out-of-frame AUG codons between the upstream proposed non-AUG initiators and the next in-frame AUG are highlighted in light brown. Upstream in-frame stop codons are highlighted in magenta. The stop codons of the main open reading are highlighted in red. In each case the nucleotide sequence is shown on top and the predicted translated amino acid sequence on bottom. Below that the predicted molecular masses of the protein initiated by the first in-frame AUG and the upstream non-AUG codon are shown. The non-AUG initiated extension is highlighted in light blue. The main open reading frame initiated by AUG is highlighted in yellow. Nucleotides in positions -3 and +4 relative to the first nucleotide of proposed initiation sites and matching the preferred (Kozak) consensus context are highlighted in gray. Comments on the extent of conservation of the non-AUG initiation and/or the conservation of the coding potential of the extension are highlighted in magenta. The GenBank accession number, the name of the gene and common gene synonyms are indicated.

A)

#1

NM_001042589

Initiated by AUG codon(s) in non-mammals. No EST in mammals has upstream AUG.
Likely initiated by a cluster of CUGs/UUG

TMEM8B transmembrane protein 8B

Synonyms: NGX6; NAG-5; C9orf127; MGC120460; RP11-112J3.10; TMEM8B
CGACGTCAGTCGAGGCCGCCGCCGGGGCTGGTTATGCCGGTCAGCGCAGCCGGAGTCGCCAGG
CCTGAACCTCTACCCAGCCTAGACTCAAGTCTGGGTTTAGCTGCCAGCCCTATTGCTGCTGTTGCTG
TTCTCTGTCTTGCCCCAGGGCTGGAGGCCTTCTCTGACTGATTACTCCACCTGCTCACCCGCAAGCT
GAGTCCTTCCGCTCCTTGCCAGCACCGAGCTCTCCACTTCCATGTTCCCTGAGGACACATTCTGGCTG
TTTGGAACCTCATCATCTTCAAGGAGCAAGGGGAACTTTGGGACCCTGCCAGACCAAAGTGTGACT
GTGTATTCCGGTCCGGGGCACCCCTGTATCAATCCCTGCATACACACTTCCAGGGGACACAGCTGT
GCCTGGGTTTCACGTACCCCTCAGCTGGACACTGCCAACCGCACCTCAGGCATCTTAACGTCAGCA
GCCCTTACCTGGGACTGGTCTGGCTGCCACCTTCCCAGGCCACGCCACATCTGTCAAGGGT
CTCCAGGATGAGTGTCACTCCTTCAGCCGAGCTGATTGTCGGCGTTGCTGGACGTCGCTGTGCT
GGTCTGGGGGGCCCTCAGAGCAAACCTCTCCCCACACATCGCTCAGCCCTGTACAAGGTCTTGTGCT
CCAGCTTACCTACAGGGTTTCAGCACAGCTGGTGTGTGGGGGGCGTGGGTATCTGCCTGCCCCCTG
TCACGTGCGTCTGCGTCCAAAGCCCCACCCCTGACAACACTCAAGCTCTGCGCTGTGGAGGTGCTCAAGG
ATGCCAGCTGGAGCTGGCACTGCCCTGGGGACTGGTCTACGTGCGTGTGGAAACATCATCCCGGG

GCCCTGGTAGGACCATCCGCTTCCAGCTGTGTGCAGTTGCAAGAGTGCCCACAGCCCCGGCTGCTCCGA
 GCCCTGGTCCCTGGAGCTGCATGAACATGCCAGTCCTGGCAACCAGCCACTGCCCGAGAACCGCC
 ATCCCTTGGAACCCCTGCGGAGGGCCTGGGACACGTCCTCTGTCCACTCTACATCTCTTGGCCAAGTGTGCCCTTCCC
 CTCTGCGCAACGAGCTGGACACCTCTGTCCACTCTACATCTCTTGGCCAAGTGTGCCCTTCCC
 CCTGAGCGCCAGCGTGTTCGCCATGAGGCTGTGCCAGTGTGGACAGTGGAGGCCTCAGCCTGGA
 GCTCCAGCTCAATGCGAGCTCCGTGCCAGGAAAACGTGACGGTGTGGATGCTTGACTCACAGGGTGC
 CCTTGAGCCTGGGGATGCAGCAGTGCACCTGTCAAAGAGTCCCTGGCGCTTCCCTCTGTCACT
 GCCACCACCAGGGTGGCCAGGCTGCGAATCCCATTCCGAGACGGGACCTGGTCCCTGGCCCTCCGCTC
 CCTGTGCCGGGTTGGGGCTCGGTTGCGGTGCGCAACCGACGGCGAGGTGCGGATGCCACCTTCC
 TGTCCCCATGCGTGGACGACTGCGGCCCTACGCCAGTGCAGCTGCGCACACACAATTATCTGTAC
 GCAGCCTGCGAGTGAAGGCCGGTGGAGAGGCTGGGCTGACCGACAGTGCAGATGCGCTCACCTATGG
 ATTCCAGCTGCTGTCACACTCCTGCTGCCTGAGCAACCTCATGTTCTGCCACCTGTGGTCCCTGGCCA
 TTCGGAGTCGATATGTGCTGGAAGCTGCACTGCACTACACCTCACCATGTTCTCCACGTTCTATCATGCC
 TGTGACCAGCCAGGCATCGTGGTTCTGCATCATGGACTACGATGTGCTGCAGTTCTGTGATTTCCCTGGG
 CTCCTTAATGTCGGTGTGGGTCACTGTCATTGCCATGGCTCGTTACAGCCGTTGCAAGCAGGTGCTGT
 ATTGCTGGGAGCTATGCTGCTGTCATGGCTCTGCACTGCACTGGACTCTGGAAACCTGCTTGG
 CCCAGTCTCTCCGCCCCTGGGATCTTGCCACAGCCTGGACAGTACGCAGCGTCCGCCGGCACTGCTA
 CCCACACCAGTGGGCCGCTGGCTTTACTTGTGCCCTGGCAGCCTATTGCAGGCAGTGGCTCCTGC
 TTTATGCTTTGGAGACCCGGGACAACACTACTTCTACATTACAGCATTGGCATATGCTCATTGCGGGC
 AGTGTGGCTTCTGCTGCCCTCGTCCAAGACTGACCAACGGGCTCCATCTGGAGGCCGGGGGGGG
 CTGTGGTTACCAAGCTATGCATCAACGAGCAGGAGGAGCTGGCCTCGGGCCAGGAGGGCACTGTC
 GCAGCATCTGCGCAGCTGAAGGGGCTTGGCCTGGGCTGAGGGATATGAATGCTCCTAGAGTTCT
 TTCTGGGGGTGTGGAGCCCTTCTAGAAGGAGACAGGCTGTATTCTTGAGGACATGGAGTCTTCTCAAGG
 ACACAAAACCTTCCAGGGACCTGGAGCCCTCCAGGACATGGAGAACCTCCCTGAGGGCTGGAGTCCCC
 CTGCATCATGGAGTCCCTTAAGGACTGGAGCCTATGCAGGCACAGAGTCCCTCAGGACCAAGGAGTCCC
 TCCTGCAGGTGTGGAGCCTTCTGGGATGCAGAGCCTCCAAAGACATGGATTCCCTCCAGGGAGACAA
 AGCCCTGTCAAGGAGCACAGCATTTCCAGAGGAGGTGGAGTCTATCTGGGAAACCAAATTCCAGATT
 TTCCCAGAGGCTCAGCAACTCTGGCCTCAGGCTCCTCCAGGAGGCAGCGTCTGGCTGTGCTGT
 GGAGGAGGGATTGCAGGATGGATGGAGCTGGACTGGGCTGTCTGGGCTGGTATCCTCGTTGATAC
 AGGTGGAGTCTGTTGCTCCAGTGTGATTGGTCTGGAGATGGTTCTGTGATGCCCTTTTCCCCTGG
 GTCAGGGGTGTGGGAAGTGGGAAGAGGGCCCTCAGGGAAATCAGCAGGGCTGATGGAGACTGTG
 GAGGCTTGTACACCGTACACAGTCTGATCCGCTCATCTGGCCCTGCATTCAATTCTAAACAGTT
 TCTAATGCCTATTCCCCAATTCCATTGAGCCGATTGCACTATCTGAGGGGTGTGTGTTGTTG
 TGTTATGTATGTATACGTATGCTGAGATGATTAATCAGTGAGACAGACTTCCTTATGATGCCACAC
 CCAACACAGAAGGAAGCCGAGGTCCCAGGAAATTGGAATAGCAGGTACACGTCTCAGGTGTGCAAGAAATA
 TCACAAGAATGTAACCTCTGTTGAGCCCCAAAAAAA

LLLLFSVLGPAGGLFLTDYSTCSPRKLSPFRSFASTEFLHFHVPEDTFLAVWNLIIFKEQGGTFGDHCPD
 QSVTVYFRSGAPPVINPLHTFPGDTAVPGVFSLTLNRTSGIFNVSSPLPGDWFLAAHLPQAHGHI
 SVKGLQDECQYLLQPQLIVRLLDVAVLVPGPSEQTLSPHNRSLAYKVFVPSFTYRVSAQLVCVGGRGV
 ACPLSLRLRPKAPPLHNSSSSVACGGASGCQLELALPPWGHWVYVRVETSSRGPGRTIRFQLCVRLQEC
 PQPGLLRALVPGAAAMPQSLGNQPLPPEPPSLGTPAEGPGTTSPPEHCWPVRPTLRNELDTFSVHF
 YIFFGPSVALPPERPAVFAMRLLPVLDGGVLSLELQLNASSVRQENVTFGCLTHEVPLSLGDAAV
 TCSKESLAGFLLSVSATTRVARLIRIPFPQTGTWFLALRSLCVGPRFVRCRNATAEVRMRT
 FLSPCVDDCGPYQCKLLRTHNYLYAACECKAGWRGWGCTDSADALTYGFQLLSTLLCLS
 NLMLPVPVLAIRSRYVLEAAVYTFTMFSTFYHACDQPGIVVFCIMDYDVLQFCDFL
 GLSLSMSVWTVIAMARLQPVVQVLYLLGAMLLS
 MALQLDRHGLWNLLGPSLFALGILATAWTVRSVRRRHCPYPTW
 RRWFYLCPGSLIAGSAVLLYAFVETRDNYFYIHSIWHMLIAG
 SGVFLPRAKTDHGVP
 SGARARGCGYQLCINEQEELGLV
 GPGGATVSSICAS

AUG initiation = 51.9 kDa
 CUG initiation = 84.0 kDa

+++++

#2

NM_001037335

The GUG appears conserved in most primates but not in galago where it is AUG as it is in all other mammals

PRIC285 peroxisomal proliferator-activated receptor A interacting complex 285

Synonyms: PRIC285; PDIP-1; FLJ00244; KIAA1769; MGC132634; MGC138228

AGAATCGAAACTGAGAGCTCTGGCAGGCTCGCAGGGCAGGCAGCTCCAGGAGGGCTTCGAACCGTGGC
CAACAGTTCCAGTGGACTGCGTGGACCCGTGAGCTCAGGAGCCTCAGACGCCCTGGAGAGCCAAGCTG
GTGTTGAGGTGGCGCTCCAGGTCCACCCCTGCTGCCAACAGCCCCGGCCACCAGAGGCCGTCCCT
GGCCCGGCTGTGCGCCCTGGTGGACCTGTGTCTGGCTGCTCCCGCTGCCACCCAGGGCTCAATGAAAGCA
CCTACGTCCCTCGTAGGGTGGAGCATGACTGCTCCCAGGAGATCCTGCTGGCCCGTTAACGAGGCCACC
AAGAGCAAGGTCTGGCGCTGGTGGCTGCCGGCCACCTTCCAAGGGCCCTGTGCTACCAAGTCTGCCA
CTACTACAGCCCTGGCTCGCTGCCGGCCACCGAACCCGTGCACTTGTGCCAGTCAGGAG
CCCTGGCTGGACCTTCGAGCGTCAGCACAACTCCAGCCTATGGCTGAAGGCCGAGGAG
GGGGCCAGGGAGGGCAGGCCGGCGGCCAGCCTCTGCTACGGAGTTGGCGCCGCTTCGAGCTGCT
TTGCTCCCTTGCTCAGGCCGTGCCCCATGCACTGTGCGCCGAGGGCAGTGGCCCTGAGC
ACGGAGCCTGCCCTCCCTGGCCACGTGAGCGCCGAGGGCCGCCAGCAACAGTTGTGGTGGTGG
AGGCCGCGCCGGCCAGCCTCTGCTACGGAGTTGGCGCTGGGCAGCCGTGCTGGCG
TGGGAGTCCCGCTGCCAGTTGCACACAGCGCCGTGGAGATGGCTGTGAGGAGGCCGAGCAGCTGGGTG
GCCTCCAGCGGGGGACCTGTCACACCCCTGCCCTGATGGCGACGGCGCACGGCCCCCTGGCCAG
CCCCCTGGGGCCAGCTGTAUTGCCGCTGCTGGTACCTGCCACTCTCAGGAGGCCCTGAGAACCA
CTGCGCATCTCGAGCACGCCAGATGGTGGCCTTGACCAGGCCCTGGAGCACCGTCCCCAC
CCCCGGACTCTCCAAGTTCGAGCTGCTGCCAACGCCGTGACCTCTGTGAGTATGGGACGCCCTGACCAAG
GCACACTCAGCACAGGAGCTGCAAGGAGTGGGTCGGCGCACCGAGGCTGTGGAGCTGCCAGCG
CTGGCAGGACGGCTGGTGCCTACCAGGAGCGCTGCTGGCGAGTACCGCGAGCAGCAGTGGTCC
TTGTGCTGGCAGAGACCCCTGATGGAGTGCCTGTCACCTGCAACCAGGCCCTGATGTACCGAGGCCAGGAG
AGGAAGACCCAGTACAGCTGACGTTGCCGTCACCTGAGGAGGCCCTGCTACACGTGCCCTGCTGAA
GCAGGAGCCAGGAGCCGACTCTCTCTGGTGGCTCCGGCCAGGCCAGGCTTACGACGGGGTG
AGCGCTCCGTGCCCAGCTTCACTGCCACTGGCAGTCCAGGTGGAGTGCCTGTCAGGCTGCCCTGTTGGC
ACCTTTGAGCAATGGGTGGCTTCGACTTGGCCGCCGGCGCTGCTGCTACAAAGCTGGGCTGAGCT
GGGCCAGGGCGTCGCCAGGCCACGTGGTGCCTGGCGAGCAGACGCCCTGATGCCAAGTAC
ACACTGGCAACCAGGCCACGTGGTGCCTGGCGAGCAGACGCCCTGATGCCAAGTAC
AAGGGCCCTGCCCTGGCCCTGGAGTTCAACCGCAGCAGCTGCCCTGGGCCCATCTCACCAACCAACTA
TCGGCAGAGGATGCAACAGTTCTATGAGGAGGAGGCCCTGAGCAGCAGCTGGTGGCAAGCTGACCC
TGCAGGGCCAGGTGTTCTGAAGACGCCATTGCAAGACGCCAGCGCTGAACATGCTTCTGCCCTCCGG
GCACTGTACGCAGAGGTCCCCGTCCCCCTCCCTGATGCCAGACACAGACCAGGGCTTCTGCTGGGCC
GGCGGTCAAGCACGCCCTGGTGGCCCTGTACCTGCACCCGACAATACGGTGTGCTGAGGTGCCCTGGAGA
GGCAGGGCCAGCTCAGAGCAGCGCTGTTGGCTGCTGCTTCCGCCGCTGCTGCTTGGCCCTGGGCTGCA
CCTGAGGCCGCCCTGGTCCCTGGAGGTGCAAGTCCAGATTGACCCGATGACCTCCGCTCTGGCACAGGC
AGTGGACACACTGCCAGGGAGCTGGTGGCCGACTTGCACCTGCCCTGCCCAGACCTTGGT
CTGTCCTACCCCTGCCGGCTGGCAACCGCAAGCAGGAGCTGCCGTGGCGCTATGCCGGCTGGGCC
GGGGATGGGAGGCCGTGCCCCCGCTACTCATCTATGGCCCTTGGCACCGCAAGACCTACAGCTGGC
CATGCCCTCCCTGGAGGTCACTGGGAGGCCCTGAAACCAAGGTGCTCATGCAACACACCAACAGTGCC
CCGACATCTACATGCCGGAGTATTCCACAGCCACGTCAAGCGGCCACCCGAGGCCACTCTCTCC
GTGATGTACACGGACGCCGCGCTGAGCAGACGCCAGTCACGCTGAGTACTGTTGCCCTGACCGACGA
CCGCCAGGCTTCCGCCGCCACACGGGAGCTGGCGCCGACCGCTGGTGGTCACCACCACTCCC
AGGCCCGTGAAGCTCAGGGTGCCGGTGGCTTCTCTCCACATTCTCATGAGTGGCGGCCAGATGCTG
GAGTGCAGGCCCTCACCCCGCTGGCCTATGCCCTGCCACGGCACCCGCTGCTGGCGGCCAC
GCAGGTCACACCCCGCTGTTCACTGAGGAGGAGCTGAGGAGGCCACTAGGCCAGGTGCTGCA
TCCTGTGCTACCAGCAGGAGACTCACGAGGTGGCGGGCAGAGCCGCTGGTCTTCCACGAGAACTACCGC
TGCACGGACGCCATTGTCAGCTCATCTCGCGCAGTCTACGTTGCCAACGGCAACCCCATCCACGCCAG
GGGCAAGGTTCCGCCACCCCGGCACTACCCGCTCATGTTGCCACGTGGCGGCCAGGCCAGACCGGG
ACATGTCATGGCGCTGGCTGAATCTGGCTGAGATTGCGCAGGTGTCAGAGAAGGTGCAAGGAGC
AACACCTGGCCAGCTGCTGGCGGCCAGCAGAGGTGCTATGTCAGTGTGCTGGTGGGCCAC
CAGTGCAGTGGCAGGAGCTGAGGAGGCCACTAGGCCAGGTGCTGCGCAGTTGAGATCCTGC
CAGGGCCGAGTCCGGTGTGCTGCACTGTGCAACACCTGCCAGAGCCTGCTGCCCTGGGCA

CTGGCCCCCTGAGTTCTTCACCGACGCCCGGTGCTCAACACCGTCTGACCCGCAGCCCAGTCCCAGCTGGT
GGTAGTGGGGGACCCGTGGCCCTCTGCTCCTTCGGGGCTGGCAAGCTCTGGGAGAGCTTCATCCGTG
AGTGCCTGGAGCAGCACAGTGTCTGCCCCGAGGGCAGGGCTGTCCATGGAGCAGGTGAGCAGGGTGTGGCGCAG
AGACGGCGCTGGCTCCCCGAGGCACACAGGCTGGGCAGGGGAACCTGGGAGGTGCCCCAGAGCCAGT
AGGGGACCTGGCCGAGGAGCAGGGCTGTGGTGACGGCATGGTAAGGCAGAGCCGGAGATGAGGCTC
TGAGCCCAGCATCCCGTGACATCACGGAACCACAGCGCAGACGGAGGTGCGGCAGCACAGCAGGAGAC
GCAGTGAGGAAGACGTGGTCCCAGGGCTGTGCGGCAGGAGCGGCTGTGAGCAGGGCGTGGAGTCCAC
GGAGGCTGAGGATGCAGAGGCTGACTTCTGGCCGTGGGACGGGAGCTAAGCTGACGACGCCATCTAC
GGGAGCTTCTGGACGAGAGCAGAAGGTGATGGTGACCGTCGGGAGGACGGCTGTGGACACTGTGCC
AGGCCCAGTCCCTGCAGCAGGCCGGTACGCCACTGCTCTTGTGCCAGAGACCTCGAGCAGGGCTGCC
TGGACGATGCCTCCTGGCCCCATCCAGGTGAGGGCCCTGGACTGTGGGATGGCCTTCGCCGGGAT
GAGGTGCTGGTGAGCAGCTCCTTCGGGAGACAAGGGCCCGAGGGGCGGCTCGGGGCGCTGCTGGGCGT
GCTGAAGAGGAAGAGGCACGAGCTGGCTTGTCGCGCATGGACACGTGGGACCCGCGCATCATGGTCC
CCATCAATGGCTCCGTGACCAAGATCTCGTGGCCAGCTGAAGGACCATCGCAGGTCCCCATCTACAGC
CTCCGAAAGGCCGGCTGAGCGTGTGGGCTTGAGAGGCTACCGCCAGGCCGGCACAGCCGGCTCTT
CTGGTCAAATTGTCTGTGGCGCAAGGCTTCTACTACCGCTGGCATCGTCGGGAGGTGCTGCC
AGGCCAGCCTGGGAGCAGGGCTCCGATCCTCGGCTGGAGTACAGCTTGAGGGTGCCCCGTCGGAC
CAGGCCACCATACCAAGGTGCTGCAGAAATACCAACACGGAGCTTGCCGGGTTGCCGGCCGGAGAGGA
CTGCCGCGCTTCTTGACCTCACTGTGGACCCCCAGGGCGCTGCAACCTCGATGATGCCCTCAGTGTCC
GAGACCTGGTCCAGGTGCGAGGTGGCTGTGACATCACTGATGTGGCAGCTTCGTGCCAGGGACGGG
GTGCTGGACGTGGAGGCGGAAGGCAGGGCGCTGCGTTCTATGCCCGGGCAGGGAGCCAGTCCCAGTGT
GCCGGCAGCCTCTGCCAGGACGTCTCAGCCTCTGCCGGACCGCCTGCCATCTCCCTGTCC
TCACCATGGAGAAGGCCAGTGGCAGCTGAAGAGCCTGCGCTTGCACCCCTCCGTGGTCCAGTCTGACCC
CAGCTGCTACGAGGAGGGAGGAGGTGATCAGGCAGCACCCGGTGCCGGCGTGAGCTGCCGGCC
CCTGGACTCCGTGGACGCCCTGCGTGGCCGGCTGCTACTCTCTCGGCTGCTGCCGGACCCGCTGC
GGTCCGACTGCTCTATGAGCAGCCGGACGAGGACGGCACCCCTGGCTCCGCCGGCCACATCATGGT
AAGGAGTACATGATTCACTGTTAATAGGCTGTTGAGTTCTGGTGGCAGCGAGTGCACCGGGACGGT
CACGCCCTGCGGGTGCAGCCAGCACCCCGCAGCCAGCAGCTAACGGCCCTGTGAGAAGCATGGGACC
GGGTGCCCTGTCACTGCACCTCGGCCACCACCTGCACGGGGGGGAGTCCCCCGACACGCCGGCTG
CACCTCTGCCCTCCCTGGAAGCAGTCCAGTTGCTGCCGCACCCAGGACTACGAACAGATGGTGA
CTTGGTCACCACGGACATGCACCCATTCTGGCTCTGCAAGGCCGGACCTCCGCAAGGCCCTGGAGC
GCTCGGGCTTCGGGCGTGCCTGGGGGCCACCAGCAGCAGGGCGGCCACTACTCGCTGCAGGTGGACTGG
TACACGTGGGCCACCTCGCCATCCGCAAGGTACCTGGACGTGGTGTGAGCAGGGCAGATCCTGCTGGCGT
GGGCATGGGGCTCTGCCACTCTGCCAGGGACATCGATGGCTCTGCCAGGCCCTCAGCCCTCAGCACG
CACTTGGCCAGAGCTATCAGGGCGGGCGCGCAGCCTGCACCTGGCGTGCAGCTAACGGCCAGCCTCTG
GACAAGCTGGCTCGTGGTGGACGTGGAGGGCGGCTCCCGTGTCTCCCGCTGCTCTTCCCCAGCAACCG
GGAGACGCTGCTGACCCCTGCCCTACGGCTCCCTGCAAGTGGCCGAGCACCCACGCCCTGG
CAGGCCGGCGGGCTGCCCTGCGGCTCTGTGGCGGCCGTGTACTCAGGCCAGGGATCCAGCCGCCCTG
CCACTGCCCTGGCACTGTGCCGGACCCACACACCCCTGCCCTGGAGACGCCCTGTGGAAGCAGCTGCTGGA
GCTGGTGGAGCTGCAAGCCTGGCCGGAGGGCGGCTGCTCTCATCCAGGAGAAGGGCGAGGCGTCCAGCG
GGGAGCTGGTGCAGGTGCAGGGAGGCCACTGTGGCATTCTGGAGGTGGCCGGAGCTGGCAGTGG
GACACCCCTGCAGGTGCAGCTGGCACCCAGCCTGCAGCACGGCTCTGGTACCGAGCCCTCAGCTCTGGAC
GGTGGCACCGGGCTTCAGCCTCTGCCCTGGAGCACGTGGAGCGGCCGGAGACTGCTCTCAGGCCGTGT
ACCGGGCCCGAGGGACGGTACCGCGACGTGGATGAGTACGCCCTGCGTGTGGGAACCATCTGCCCTG
GAGTCGCCACCAGCGCGGTTGCCAGAATGACTCCGTACACTCAGCACCTGAGTGTCTCTGGAGGC
GTCACGGACGCCCGAGGGCAGCTGCAGGGCGCTTCCGCCCTGGAGGCCGCCCTCGAGGAGAACTGTG
CCGACATCAACTTCAGCTGCTGCTACCTCTGCATCCGGCTCGAGGGGCTGCCGGCTCCACGCCAGGCC
CGCCCTGGGCCAGCAGCCTGCCCTGGCTGAATGGTACCCCGCACGTATAACCTGGGTGGCCACGG
GCAGACGCAAGGACTGGGACCAAGGAGGCCGGAGACCCGGCAGGGCTCCAGACGGGTGCAACCTCTG
TCCACCATGGCATGGAGAAGGTTCCGAAGAGGTGCTGAGGCCGGCACCTGTTACCGTTGAGCTG
CTGCCCAAGCAGCTCTGATCTCGCAAGGAGGAAGCCGTGCGTGGACTAGAGGAGGCGTCCCGCTGGT
CACCAAGCAGCTGCCCTGGCCCTGCTCCCGCAGGCCCTCTGCAGACTGATCCCCAGCAGGTCTGGAGC
GGCAGACCTACAACATCCCCGGAGGCCAACAGCTGAACCCAGCCAGAACGTGGCGTCAAGGGAGGC
CTGGAGAAGCCTTCACGGTCAATTCAAGGGCCACCAGGTACAGGGAGACGATCGTGGCCTCCACATCGT
ATTCTGGTTTCAAAATCAAACCAAGGAGCAGGTGCAGCCGGAGGCCCCCGTGGGAGAAGCGGCTGG
GGGGTCCCTGCATCTGTACTGCCCTCCAACAAGTGGATGCTCTGGCAGGACTGCTCTGGAGA

AGGATGGAGCTGAAGCCCTCCGTGTACAGTGAGCAGGCTGAGGCCAGCGAGTCCCAGTGCCCGTGT
GGGCAGCAGGAAGCTGCTAGGAAGAGCCCCGGAGGGGAGGCCAACAGAGCTCAGGAGCATCACCC
TGCACCACCGGATCCGGCAGGCCAACCTACTCGTCGAAATCAAGGCTTGTACACCCGGCTGCAG
AGAGGGGAGCTCTCCAGGGAGGACTGGTCTGGTACAAGAAGGTCTTGTGGGAGGCTCGAAGTCGA
GCTGGACCGGCATGAGGTATCCTCTGCACCTGCTCTGTGAGCCTCTGCCAGCCTAAAATCCTGGACG
TGAGGCAGATCCTGTTGACGGAGGCAGGCATGGCACCGAACCTGAAACCTCATCCCCCTGGTGCAGTTC
CCACAGGCCAGAGAAGGTGGTCTTCTCGGAGACCACAAGCAGCTGCAGGCCGTGGTCAAGAATGAGCGGCT
GCAAAACCTGGGTCTGGACCGGTCTGTGAGCGGTACACAGAGGACGCACATATGCTGGACACTCAGT
ACCGCATGCATGAGGGCATCTGTGCCCTCCCCCTGTGGCCTTCTACAAGAGCAAGCTGAAGACGTGGCAG
GGCCTGAGGAGGCCAGTGTCTGGGCCACGCTGGCAAGGAGAGCTGCCCTGTACATTTGGCACCT
GCAGGGCCACGAGGGAGCCTGCTGGTGTCCACGGACGAAGGAATGAGAACTCCAAGGCCAACCTGGAGG
AGGTGGCTGAGGTGGTCCGTATACCAAGCAGCTGACCTGGGGAGGACCGTAGAGCCCAGGACATGCC
GTCCCTCACGCCCTACAACGCCAGGCCCTGTGAGATCAGCAAGGCCCTCGGCGAGAGGGCATTGCCGGGGT
GGCCGTGTCTCCATACCAAGAGCCAGGGAGCGAGTGGCCTATGTGCTGGTGAGCACCGTCCGCACCT
GTGCCAAGAGCAGCCTGGACAGCGGCCACCAAGAGCTGGCTCAAGAAGTTCTGGGCTTCGTTGTGGAC
CCCAACCAAGTGAATGTGGCTGTACGCCGGCCAGGAGGGCTCTGCTGAGGACCACTCCTCT
GCGCTGCTGCCCTCTGGCGTAGCCTCTGGACTCTGCGAGGCTCAGCAGACCTCGTGCCTGCCGGG
AGGTGCGCTCTGCAGGAGGCCAACTATGCCCTCTGAGATCAGGAGGCCCTCCACCTGCAAGGTGCCAGGACTG
GGAGGGAAAGTCCAGGGCCCCCCCACCCCTCCCCACCCCTCCGCATAGCCTGCCCTGGGCCAGCCTGCC
TCCTGGCCTCAGGATCTCCCTGGGCAGGAAGTGGTGGACTTGGAGGAGGAACGGCAAGGGCCGGTCCCC
ACACTCTTGTCTCCCTGGGCATCTGGGACAGGAACACCTCTCAGGGTGGCCAAGAGGGCTTCTGG
ACCAGTGAAGGAAGTTCTGCTGTCCCTGGGAAGTCTCAGAACAGGAAGGAACCCCTGTGGAGGCTGCC
TTTGACCTTGGACGCCACCTGTAATCACTCTGGATGGTCTGTGGAGTTCAAGGTTTATTTCTATTAA
ATGAACCTAAAGCAGCATTCTGCCGGCGTGGTGGCTCACCCATAATCCAGCACTTGGGAAGCCGAG
GCCGGCAGATCACCTGAGGTCAAGGAGTCAGGAGACCATCCTGCCAACATGGTAAACCCGCTCCACTAA
AAATACAAAATTAGCTGGCGTGGTGGGGCGCTGTAATCCCAGCTACTTGGAGGCTGAGGAGA
ATCGCTGACTCCAGGAGGCCAGGCTGAGCTGAGCTCTGATGGTACCACTGCACTCCAGCCTGGGTGACA
GAGCAGACCTGCTTAAAAAAACAAAACACAGCATCTGTTACTGGAACAGAGACCTCAGCC
CAAGCTCAGGACAAGGAGGCCCTGGAGAGGGGGCTTCTCCCACCATCACCCTGTGTCTTCAG
GGGCTGCTGGAGGGCAGCCAAAGCCAGGGGCCACCCGCCACCCACCCACCCACTCTGAAAT
GTGAAAAGCTGCACTCTTCTGTGGGCTGCGAGTGGCTGGGGTGGCTGGGGTGGCGATGCTT
CCACCTGGACTGAGTGTGGGGCAGGCTGCTGGCTGTATGTTCTGCACGCTCACCGCTTCAGTGC
GGCTCTGCTCCCTGCTGCCCTCCAGTCTGAGGGAGGGGCTTGGGGCACCTCTACATTCAATTAA
ATCTTGAATTATGTGATTAGATATTAATTAGATAAAACACTCTGAAAGCTCTCTCA

VAPPGSTLLPNSPAATRGPSSLARLCAVLDCLGCSRCTQRLNESTYVLRVEHDSREILLARFKQATKSK
VWRVVGCRPTFPRPLCYQVCHYYSPGLCRRHRNRCTFARSREEALWTFERQHNQLQRLWLKAEVQGSGAQ
GGAGRAADAILTEFGGRFELLCSLCFRCPICRVDPPQGCPEHGACPSLLAHVSAEGRRKQQFVVVRPR
PRAGQPAPAYCRFVGRGQPCWRGESRCQFAHSAVEAVWEAEQLGGLQRGDLLTPPAPDGDRTAPLGQPPG
AQLYCPACLVTCHSQEAFENHCASSEHAQMVAFDQALPWEHRSPPPGLSKFELCPKPDLCYGDACTKAHS
AQELQEWRRTQAVELRGQAAWDQGLVPYQERLLAEYQRSSSEVLVLAETLDGVRTCNCPLMYQAQERKT
QYSWTFAVHSEEPPLLHVALLKQEPEGDLSVAPGLPPGRLYARGERFRVPSSTADFQVGVRVQAASFGTFE
QWVVFDFGRRPVLLQKLGQLQGQGRGPCCRNLALGHPEEMERWHTGNRHVPGVERTAEQTALMAKYKGP
ALALEFNRSSVASGPISPTNYRQRMHQFLYEEEAAQQQLVAKLTLRGQFLKTALQTPALNMLFAPPGALY
AEVPVPSLMPDTDQGFLLGRAVSTALVAPVPAPDNTVFEVRLERRASSEQALWLLLPARCCLALGLQPEA
RLVLEVQFQIDPMTFRLWHQAVDTLPPEQLVVPDLPTCALPRPWSVPLRRGNRKQELAVALIAGWGPGDG
RRVPPLIYGPFGTGKTYTLAMASLEVIIRPETKVLICHTHNSAADIYIREYFHSHVSGGHPEATPLRVMY
TDRPLSQTDPTVTLQYCCLTDDRQAFTPPTRAELARHVVVTTSQARELRVPGFFSHILIDEAAQMLECE
ALTPLAYASHGTRLVLAGDHMQVTPLFSVARARAEEHTLLHRLFLCYQQETHEVARQSRLVFHENYRCTD
AIVSFISRHFYVAKGNPIHARGKVPPhRHYPLMFCHVAGSPDRDMSMASWLNLAIEAQVVEKVQEAYNTW
PSCWGGRERQRCICVVSHGAQVSALRQEIRRDLGQSVGSFEILPGRQFRVVLSTVHTCQSLLSPGALAP
EFFTDARVLNTVLTRAQSQLVVGDAVALCSFGACGKLWESFIRECVRHSVCPEGLSMEQVEQGVAQRR
WPPRGTQAGAAGNWEAAPEPVGDIAEEQAAVVTAMVKAEPGDEALSPASRDITATTQTEAAAAPAGDAVK
EDVVPGACAAGAAAAAGVESTAEDAEDFWPWDGELNADDAILRELLDESQKVMVTVGEDGLLDVARPE
SLQQARLYENLPPAALRKLLLHAEPERYRHCSFVPETFERASAIPLDDASSGPQVGRRLDCGMAFAGDEV
VQLLSGDKAPEGRLRGRVLGVLRKRRHELAFVCRMRTWDPRIMVPINGSVTKIFVAELKDPSQVPIYSLRK
GRLQRVGLERLTAEARHSRLFWVQIVLWRQGFYYPLGIVREVLPEASTWEQGLRILGLEYSRVPPSDQAT

ITKVLQKYHTELGRVAGRREDCRAFLTFTVDPQGACNLDDALSVRDLGPRCEVAHVITDVASFVPRDGVL
VEARRQAAFYAPGREPVPMPLPASLCQDVLSSLPGDRRLAISLFLTMEKASGQLKSLRFAPSQQSDRQLS
YEEAEVIRQHPGAGRELPARLDSVDACVVAACYFSRLLRRRLRSDCFYEQPDEDGTLGFRRAAHIMVKEY
MIQFNRLVAEFLVGSECTRTVTPLRWQPAPRSQQLKALCEKGDRVPLSLHLGHHLHGGGGSPDTRLHLL
ASLWKQVQFAARTQDYEQMVDLVTTDDMPFLAPAGRDLRKALERSAFGRCARHGQQQGGHYSLOQDWYT
ATSPIRRLDVVLQRQILLALGHGGSAYSARDIDGLCQAFSLQHALAQSYQRRARSLHLAVQLKAQPLDKL
GFVVDVEAGSRCFRLFPSNRETLPDPCPVYGSILQAEHPHALAGRPLRLWRRRVYSAQGSSPPLPLP
GTVPDPTHETALWKQLLVELVLRWPEAAALIQEKGAEASQRRELVQVQRSHCGHFLEVARELGSQDTL
QVQLGTSLQHGFLVPSPQLWTVAPGFSCLCLEHVERPGDCFSGRVYRAPDRYRDVDEYACVWEPCALESA
TGAVAENDSVTLQHLSVSWEASRTPQGQLQGAFRLEAAFLEENCADINFSCCYLCIRLEGLPAPTASPRPG
PSSLGPGLNVDPGTYTWVAHQQTQDWQERRADRQEAPRRVHLFVHHMGMEKVPEEVLRPGLFTVELLPK
QLPDLRKEEAVERGLEEASPLVTSIALGRPVPQPLCRVIPSRLERQTYNIPGGRHKLNPSONVAVREALEK
PFTVIQGPPGTGKTIIVGLHIVFWFHKSNEQEVQPGGPPRGEKRLGGPCILYCGPSNKSVDVLAGLLLRRME
LKPLRVYSEQAEASEFPVPRVGSRKLLRKSPREGRPNQSLRSITLHHRIRQAPNPYSSEIKAFDTRLQRGE
LFSREDLVWYKKVLWEARKFELDRHEVILCTCSAACASLKI LDVRQILVDEAGMATEPETLIPLVQFPQA
EKVVLGDHKQLRPVKNERLQNLGLDSLERYHEDAHLTDQYRMHEGICAFPSVAFYKSKLKTWQGLR
RPPSVLGHAGKESCPVIFGHVQGHERSLLVSTDEGNENSKANLEEVAVVIRITKQLTGRTEPQDIAVLT
PYNAQASEI SKALRREGIAGVAVSSITKSQGEWRYLVSTVRCAKSLDQRPTKSWLKKFLGFVVDPNQ
VNVAVTRAQEGLCLIGDHLLRCCPLWRSLLDFCEAQQTLPAGQVRVCRRPTMPS

AUG initiation = 294.6 kDa
GUG initiation = 322.3 kDa

+++++

#3

NM_001010858

Extension conserved from human to fish but non-AUG initiation conserved only in mammals. Many other human paralogs exist (incl. BRCA1) with the extension but initiated by AUG codon and part of a larger extension.

RNF187 ring finger protein 187

Synonyms: RNF187

TCTAGCGAGGTGACAGCGTAGAACCAAGGTGCGCGTCCCCGGCGTTGGCGTCTCGCCTGTTGCTGGTCTC
CGTCCGGTGC CGCCGCCGTCTAGGTCTCCGGCCCTCCCAGCCGCTCTGCGCCCTTGCCGGCCCCGCCGCC
CGCACGCCCTGGCGCTCCCTCGGGGCCCCGCCGAGGCCGCGCTGCGCCCTGTGCCAGCGCGGCCGGAA
CGGTGCGCGCGACTGCGGCCACCGCTCTGTCGGCGTGTGGCGCTTCTGGGCCAGGGAGGACGG
CCCTTCCCGTGC CCCGAGT GCGCCGACGACTGCTGGCAGCGCGCGTGGAGGCCGGCAGGCCCGCTCAG
CCGCCGCCCTCTGGCGCTCGAGGAGGCCGCCGCCGCCGCGACGGGCCGCCAGCGAGGCCGCC
TGCAGCTGCTGTGCCGCCAGGCCGCCGCTCTGCCGCCCTGCCGTATGGCTGCGGGCCCCGAGCG
CCCGAGTGGGACCGCGCTGGAGGAAGGGCTGCCAGGGAGAACAAGGGCTGTGGAAATCATGAG
AAAGGACTTGAATGACGCCGGGACCTGCATGGCAGGCAGAGTCAGCAGCTGCAGTGTGGAAAGGGACACG
TGATGGACCGTAGGAAGAAGGCACTGACCGACTACAAGAAGCTGCCCTTCTTGTGGAGGAGGAGG
CATTTCCTGCAGGGAGCTGAGAAGGAGGGCTCCCTGAGGACGAGCTGGCTGACCCACTGAGCGTT
CAGGTCACTGCTGCAGGCCGGTCTCGGAGCTGGAGAAGAACATCGAACCTGGCCCTCAGCATGCTG
AGTGAATGGCGCCAACCGTGGCAGTCCCAGAGCTGGAGGCAGGAGGATGGATCCTCATCTCCATGGGAAGT
GTCAGCGTGTGGCTGCCAGGGAAAGCGTGGCAGGCCCTGCCCTGGTCCATCTACATAGTTGCGTGTTC
AACAAATGTCCATTATCCTTCACCCCGAGGCCTGTTGGGGCTGCAAACACCTCCCGTAGAGGCTGG
CCTGAGGACCCCTCCACCTGTGCCCTGCCCTGTCAGTCCACGCCATCCTCGTCATCCAGGGACCCAGACCC
GGCAGCTCTGGGTCTGCCCTCCCTGGTCAACCCTGCTCACCTCGTCATCCAGGGACCCAGACCC
TGCACCTTCATGTGGGCCACAGATCCTGGCAGGTACCTGAGGTGCAATTGAGTGTGGATTGGGG
TTAGCATCCAGAAAGAAGAATGCGCATGACGCTCTGTGAAGGCTGGAACTCAGGTCTCAGGGAGAGAAA
GAAGACTGGATTGCACTTGATGCCTCTGAGGAGGCCGGCCCCCTTGAGGTGGCGTGGGCCGGCC
AGCCTTATCCAAGTCGCTCTGTCACCTCCCCCTCCCTGGCCCCCACCCACTCCTGTGCCTCCAGGAGC

CCTCCCTGTGCTCACCGCAGAAGGAAGCCTTTCTGTTCCCTGGGTGAGGGGGCTGGCA
 GGTGGCTAACCCATTAGCATCTCCAGGCCATCGTCTCATCTGCTTATCTCTAGCTCTT
 CCCTCCTCCCATTCTCTTAGTGTAAAGCTGTAGCTAGCTCAGTCAGGGACAGGATTGCAAAGTCGGGACATAGAT
 CTTGTGTAGATAAAATTAGTCGACAGAAAACACTCAGCACTGGGACAGGATTGCAAAGTCGGGACATAGAT
 GCAGACAGTTGAGATTGGGATAGCCGGCTGTGAGCGGTGCCATTCCAGATGAAGCCTTCAG
 CCCTCTGAGTCCCCGGCCCTGGTGCATGTCTGTGAGTTGACCTGCCAGCGTGTGGCTGGCTCAAT
 GCTGAATAAAAGTGGTTGTGTAGCTCGTTGCTCGTCCGTGTCCACCTGGCTCTCCCCCTGC
 CCTGGCACCCCTCAGTGTCAAAGGAAACTTCCTCGTACACGTCTAAAGCATGGTGAGGAGGACTTG
 TTGGGACCATTGAGATGGGTTGGGACCTTCTGGGCTGGGGGAGATGGGCTCCACCCGACGT
 AGCAGGGCAGGGGTTGGAGGAGCAGTATAGGGTCATGGGGAGATGACTGTGAGGAGACATC
 AGGGCTGAGGGGCTCTGGCTAAACCCACCTCACAGAGTCAGTGTGAGGAGCAGACAT
 TGGCTGCAAACGGTCAGAGAGGAACCCAGTCAGTACATTGAGGGTGGTCAGATATTATGGTAACCAA
 TTAGGGTCTGCTAAACTGGATTCTACAAGAAAGGCAAAGAGGGCCCTAGGAGAAGATTCCAGAGCCT
 GGCCAGAGTTGGCAAGTAGAGAATCTTGTAGCAGCAGCCAACAACATCCGACCCCTGAGACCTCCAGTT
 TGTCTTCTCACTGTCTCCGCTGCTGAGTCAGTGTCTGTCATCCCTGAGCATCCCTGCCCTGCCCTGCACA
 CCTGTGATGCTGCCGGACAGTCCTGATGGCAGAGTCCTCACATCAGTGTCTCCACATCACCAGG
 TCCGACAGTGGCTCACCATCCTCACCTAACCTAGCTGACCAGCAACATCCCACCTGTCAATCACAAACCT
 CTTCTATTAAGAAAATTATATATTATGGGCACAGTG
 CGTCCTGCTGGCTCCGGTCCGGCCGTCTAGGTCTCCGGCCCTCCCCAGCC

LALPAGPAEAACALCQRAPREPVRADCGRFCRACVVRFWAEDGPFPCEPADDWQRAVEPGRPLSRR
LLALEAAAAPARDGPASEAALQLLCRADAGPLCACRMAAGPEPWEPRWRKALRGKENGSVEIMRKD
LNDARDLHGQAESAASAVWKGHVMDRRKALTDYKKLRAFFVEEEEHFLQEAEKEEGLPEDELADPTERFRS
LLQAVSELEKKHRNLGLSMLQ

AUG initiation = 14.6 kDa
 CUG initiation = 26.2 kDa

+++++

#4

NM_001136108

CUG initiated extensions in mammal; in non-mammals extension is initiated by AUG
 In *Bos taurus* the longest AUG initiated in-frame CDS is only 9 codons long!

R3HCC1 R3H domain and coiled-coil containing 1

Synonyms: DKFZp564N123; R3HCC1

CTCGGGGCGCTGGCCCTGGGACGCCAGGGCGCTGCGACGCCAGAGGCCGGCTCTCCACCT
GTCACCTGGGCCCTCTGCTGGGATGGTGTCTCCTCTCCTCAGCCGAGAATGACTTCGTCCACCGGAT
 CCAGGAGGAACGTGACCGCTTCTGCTGAGAAGCAGCTGTCAAAGGTTCTCTTCCCCCACTCTCCA
 GTCGCCCTCGGTACCTGATCCATAGAACAGCAGAGAATTGATCTCTGAGCAGCTCTCCGTTGGGAG
 GGCTGGAAGAGGAGGACGGTCATCTGTCACCAGGACATCAGGGTACCCAGTTCGGATGGCCTCTGGGCC
 CTGCCGCGCTCCTGCCCTGCCAGCAGGTACCAAGGTCTCGGCCATCTAACCAACCAAGGAGCAGCTG
 CGGTTCCCCGAGGTGCCGGCTGCCGGTGGTACGTGGACGCAAGCCTGACAGCCTTGTATGTGCC
 CGGGTGTGCGCAGGCAGGAAGAATGGGGCTGACCTCTACCTCGGTGCTCAAGAGAGAGGCCAGCTGG
 CAGGGACCCAGAACAGAGCCTGGAGATGTGGTGTGGAGACCCAACCTGTACAGGGACTCCGTGCTGA
TGACTCAGGGAACAGAGGACCTAAAGGGCCCAGGACAAAGGTGTGAGAATGAGCCACTGCTGGACCCTGT
 GGCCCTGAGCCTCTGGGGCTGAGAGTCAGTCAGGGAGGGAGACATGGTGGAGATGCCACACGGTTGG
 GTCCACCTGCAGCTAGACCTGGAAAAGGGGAAGGGAGAGTCAGTGTGGAGAAGAGGCTGGCAGAGGAGG
 AAGAGGAGCAAGAGGAGGTGGAAGAGGATGGCCCAGCAGCTGCTCGGAGGACGATTACAGTGAGCTGCTG
 CAGGAGATCACAGACAACCTGACGAAGAAGGAGATTAGATAGAGAAGATCCATTGGACACATCCTCCTT
 CGTGGAGGAGCTGCCCTGGAGAGAAGGACCTTGGCCACGTGGTAGAGATCTATGACTTGAACCAGCGCTCA
 AGACGGAGGACCTGCTGGCAACGTTCTGAGTTCAAGAGAAGGGGTCAGGATTCAAGTGGGTGGATGAT
 ACTCACGCACCGCATCTTCCCTGCTGCCAGCTGCCAGGGAGTTCTCGGTGCT

CAAGATCCGGCCCCTCACACAGGAACCAAGCAGTCAAAGCTCAAAGCCTTGCAGAGGCCAAACTCCTGC
GTCTGGTGAAGGAGAGGCCACAGACAAATGCGACTGTGGCCGGCTGGTGGCCGGCCCTGGGACTC
CAACACAAAAAGAAAAGAGCAGCCTGCTGTCGGGGTCCGCTGCCGCCCTGAAGGCCTGGAGACCCAACCTGGC
CTGGATCTCGTCCCAGCTAGCTGGCCCCAACACCATAAGCCTCACAGACGCCAGAGCAGCCCCGC
ACCACCCCTCGAGCTTACCATGGGGTGTGGTGGGTTAGTTAGTCCAGAAATGGAGAAAAAATAAAAAA
CTCACGTTGTTCTAATGTGAAAAAAAAAAAAAA

LALLCLDGVLSSAENDFVHRIQEELDRFLQLQSKVLLFPLSSRLRYLIHRTAENFDLSSFSVGEGW
KRRTVICHQDIRVPSSDGLSGPCRAPASCPSRYHPRPISNQAAAVPRGARAGRWRGRKPQPLYVPRV
LRRQEEWGLTSTSVLKREAPAGRDPPEPGDVAGDPNSDQGLPVLMQGTEDLKGPGQRCENEPLLDPVGP
EPLGPESQSGKGDVMEMATRGSTLQLDLEKGKESLLEKRLVAEEEDEEEVEEDGPSSCSEDDYSELLQE
ITDNLTKEIQUIEKIHLDTSFVEELPGEKDLAHVVEIYDFEPALKTEDLLATFSEFQEKGFRIQWVDDTH
ALGIFPCLASAAEALTREFSVLKIRPLTQGTKQSKLQALQRPKLLRLVKERPQTNAVARLVARALGLQH
KKKERPAVRGPLPP

AUG initiation = 28.3 kDa
CUG initiation = 49.1 kDa

+++++

#5

NM_003760

AUC initiated extension in mammals - In other vertebrates extension is initiated by AUG
EIF4G3 eukaryotic translation initiation factor 4 gamma, 3

Synonyms: eIF4GII; EIF4G3

TCACGGTCCGCTGCTCGTTGAGTGAAGAAAATCACCAGCATCGCCTGAGCCCCCTACCGAGAAGGGCG
CCGCTTCCTCCGGGGAGGGGATAAAAGATCCCCCGCCGCCATGAGGATATTGCCGTGAAAGGCACA
GCGACTGCAGCAGGAACCGGACCCGGCACCGGAGCGGGCGGCCGGCGCAGCAGCGTACCGCCTCCTCA
CCCGGCGCGCAGCAGCGGGCGGCCGGCGGCCGGCGCAGCGGGTACCGCCTCCTCA
CGAACATCAGGGCCTCCAGACTCAGGCCTTAAACAATTCTAGAGGACCTGTGCAACAAACCTCTTGAG
GATCGAATCTTCACTCCGCTGTCTCAGCAGTCTACAGCACGGTAACACAAGTGGAAAGACAGCCGGAAAC
CCCTACCCATCCCTTATTCAAGCAGATGAAATAAACAAAGGGGATCCAAATCTTGGCAACGCCCGG
GACATGATCGTCCCTGGACTCTCTCAAACCCCTTATCCCTCTGGACAGAATGCAAGTCCAACACGCTG
GTATACCTCAAACCCCTCAGACAATGAAATTCAAAACCTCAAACCCGTTCTCCGTTTCCAGAGGCCTCA
AATACAGCCTCCTAGAGCTACCATCCGAACAGCAGTCCTCCATTGCTCTGGTGCACAGACACCCACTG
CAGTGTACCGAGCTAATCAGCACATCATGATGGTTAACCATCTGCCATGCCGTACCCAGTGGCCAGGG
CCTCAGTACTGTATACCACAGTACCGTCATAGTGGCCCTCTTATGTTGGCCCCCAACAAATATCCAGT
TCAACCACGGGGCAGGTCTTTTATCCTGGACCAGGACCTGGGACTTCCCAATGCTTATGGAACGC
CTTTTACCCAAAGTCAGCCGGTGTATCAGTCAGCACCTATCATAGTGCCTACGCAGCAACAGCCGCTCA
GCCAAGAGAGAGAAAAAACTATAAGAATTGGATCAAACCAAGGGAGGTAAAGACATAACAGAGGAGAT
TATGTCGGAGGTGGCAGCAGAAATCCTACTCCACCCATAGGAAGACCCACGTCCACACCTACTCCTCCT
AGCAGCTGCCAGCCAGTCCCCGAGCACAGCCCTGTGGTTATGGGACTGTGGAGAGCGCTATCTGCT
GCCAGCACCCCTGCACTGCAGTAGGCAGAGCAAGAGGAGAAGCCAAACCAAGATCCAGTGTAAA
GTCTCCTCCCCAGTCCTAGGCTAGTCAGTGGAGAGAAGAAAGAACAGAAGGCCAGACATCTGAAA
CTACTGCAATAGTATCCATAGCAGAGCTCCTCTGCCTCCATCACCTACCAACTGTTCTCTGCTCGA
AGTACAATTGCAAGCCCCACCTCTCTGCTCTAGTGCACCAACAAATTCAACACTGCTATAGATGACAG
ATGTGAACCTCATCCCCAAAGAGAACACAATTCTATACCCAGCCTCACATCTGCACAGAAACATCAG
ACCCCTTACCAACAAATGATGATGATATGCAAGAAACCTGTAGTGTAGCACCTAAATGATATT
CCACTGGTTCTAGTACTAACCTAATTAAATGAAATAATGGAGTTAGCGAAAAAATTATCAGCCACGGAGAG
CATTGGAAATAGTAAACAGGAAGTATTGCCATTGACTCTGAATTGGAGATTCTGAAATCCACTCCT
AAGAAATGAAACTGGAGTGATCCAGCTCCATACCCCTTCCACAGTTCTCTTCTCCAACTCCT
CCAACCTCCAGCTTCTCCCTCACACTCCAGTCATTGTTCTGCTGCCACTACTGTTAGTTCTCC
GAGTGCTGCCATCACAGTCCAGAGAGTCTAGAGGAGGACGAGGACATAAGAACTTGCCTAGTGAAGATG
AAAAGAGATTCAAGAACAAATAGAGGTTAGAAGCAGATGGCAAACAGAAGAGATTGGATTCTCAAAAC

TTAAATTCAAGAAGGAGCCCTGTCCCAGCTCAAATAGCTATAACTGTACCAAAGACATGGAAGAACCAAAGATCGGACCCGAACCACTGAAGAGATGTTAGAGGCAGAATTGGAGCTAAAGCTGAAGAGGGAGCTTCCA
TTGACAAAGTACTTGAATCTGAACAAGATAAAATGAGCCAGGGGTTCATCCTGAAAGAGAACCCCTCTGAC
CTAAAAAAAGTGAAGCTGTGGAAGAAAATGGAGAAGAAGCTGAGCCAGTACGTAATGGTCTGAGAGTGT
TTCTGAGGGTGAAGGAATAGATGCTAATTCTAGGCTCCACAGATAGTCTGGTCTGAGGGTACATTCCAT
TTAAACCAGAATCTGGAAAGCCTACTGATACTGAAGGTAAGAAGCAGTATGACAGGGAGTTCTGCTGGAC
TTCCAGTTCATGCCTGCCTGTATACAAAACCAGAGGGCCTGCCTCTATCAGTGATGTGGTCTTGACAA
GATCAACCAACCAAATTGCCAATGCGAACTCTGGATCCTCGAATTTCGCTCGAGGACAGACTTACAC
CAGCCTTGTGATTTGGAAGGAGACACCTGGAGAAGAGGCGTACCTTGTGAAATGTTGGTCACGA
AGATCTCAACCTGCCAAAGAAGAGAACCCAGAAAGATCATCACAGTTCTGTAAGAAGATGTACACCT
GAAAAGGCAGAAAATGCCTGGAAGCCAAGCCAAAACAGAGACAGCCAGCGATGATCCCAGAAACATTA
AAACCCAGGAGCTTTAGAAAAGTCTGAAGTATCTAAATAATTGACACCACAGATGTTCAATCAACTG
ATGAAGCAAGTGTCAAGGACTTACTGTTGACACAGAGGGAGCGCTGAAAGGAGTTATTGACCTGGTCTTG
GAAGGCATATTGATGAACCCAGTTCTCTGTGGCTACGCAAACATGTTGTCATGCTAGTAACGCTGAAAG
TACCCATGGCAGACAAGCCTGGTAACACAGTGAATTTCGGAAGCTGCTACTGAACCGTTGCCAGAAGGAG
TTTGAAAAGATAAAAGCAGATGATGATGTCTTGAGAAGAAGCAGAAAGAAACTTGAGGCTGCCAGTGTCC
AGAGGAGGAGACAAGGCTCATGATGAACGGAAAGCCAGGACAAGGCCGGAGATCCATTGGCA
ACATCAAGTTATTGGAGAACCTTTAAACTCAAAATGCTGACTGAAGCCATCATGCATGACTGTGTGGTG
AAGCTGCTAAAGAACCATGATGAGAAGAACCTGGAGTGCCTGTGCTGCCGTGTCACCACATTGGCAAAGA
CTTGGACTTGAAAAGCAAAGCCACGTATGGACAGTACTTTAATCAGATGGAGAAAATTGTGAAAGAAA
GAAAACCTCATCTAGGATTGGTCTGCTCAAGATGTTAGACCTAAGGCTGTGCAATTGGTATCT
CGAAGAGCAGATCAAGGGCTAAACTATCGAACAGATTCAAAGAGGCTAAATAGAAGAACAGAACAGA
GCAAAGGAAGGTCCAGCAACTCATGACCAAAGAGAACGAGGAGTGTCCAGAGAGTGGACGAAGGTG
GGTGGAACACTGTACAAGGGCCAAGAACAGTCGGTACTGGACCCCTCAAATTCTAAATCACTAAG
CCTACAATTGATGAAAAAATTCAAGCTGGTACCTAAAGCACAGCTAGGCTGGGAAAAGGCAGCAGTGG
TGGAGCAAAGGCAAGTGAGACTGATGCCATTACGGTCAAGTGCTCCAGTTAAACAGATTCTGCCCTGC
AACCTCAGCACCCCTCAGGTCCACGCCATCCACGCCGTAGAGTTGATTCCGAAGGACCTAACTAGT
CGTGGACTATGGCAGGGAGAACATGACAAGCCCTTCCATCTGCAACAGCTGGCCAAATACTTCT
GAGGGTGGCAGCAGTAAAGACCTGCTAGACAATCAGTCTCAAGAACAGCAGCGGAGAGAGTGTGAGA
CCGTGAAGCAGCTCACAGGAGGTGTGGATGTGGAGAGGAACAGCACTGAGGCTGAGCAGAACAAAGG
GAGTCAGCAAACACAGAAATTCAAGCAATGTCAGCTCATGACAAGGCTGATTACAGAACAGGAACTGG
GAGGAAGTCGAAATCTATCATTGATGAATTCTACACATTAATGATTAAAGGAAGCCATGCACTGTG
AAGAGCTGAATGCCAGGGCTACTACATGTTTGAGAGTGGAGTGGAGTCCACCTGGAAAGGAGC
CAGATCACCAGGGATCACATGGCCAATTACTCTATCAGCTGGTACAGTCAGAAAACACTCAGAACAGGA
CTTTTCAAGGTTTCAGAAACTTTGAATTGGCAGATGACATGGCATTGATATTCCCCATATTGGT
TGTACCTTGTGACTGGTACCCCATGTTAAAGAACGGTGGAAATCTGCATGAGAACATTACCATAGAA
TTTAGCAAACCTTTACTTCCTGTTGGAGAGCTGGGAGGCTTATGGAGGGAGGCTGACCTCAGCTGG
ACAAATGAGCATAAGAACGGTGGAGCCTTATGGAGGGAGGCTGACCTCAGCTGGAGGACTTTACCA
AAGGAGAACATGTACATAATTTCAGGAGCAGAACGGTGGACTTACAGAGTCTGACAGTCCCTGTTCC
TCTGAAGCACTTCAAAAGAACAGACTGCTGCCAGAGAGCTGTATAAGCGACTCGAGAACACTCATTATTGA
GGACAAAGCGAATGTAACAGATCTTGACTGGTAGAGGCTAATCTAGACGAAATCCAGATGAGTTCAC
CTACATTCTTAGAGCTTAATGACTGCTGTTGAAAGCAGCTATTAGCCACTTCTACCTCAGA
GTGGACACTGCTGTTATCAAGCAGAGAGTGCCATCTACTCAAGTACCTAGACTCAGATACAGAACAGGA
ACTGCAAGCACTTATGCACTACAAGCATCGATAGTAAACTGATCAACCTGCCATTGCTGCCAGTGT
TTTTGATTGTCTATATGACGGAGGAGGTGATCTCCGAGGATGCCATTACAAATGGGAGAGCAGCAAGGAC
CCTGCAGAGCAGAACGGAGGGCTGGCTGAAATCTGTCACGGCATTCTCACGTTGGCTGCCAGGAGC
AGAACAGGAGTGTGAGGATAACTAAACCTCAAATACACAAAATGAAACAAAAGAACAAATTAAAGTATT
TTTAAAAAGTTCACTGCTTCGCCAATCACAGTCAGCAAGGCCATTCTGCAGAACCCCCACGTGTG
CACGAGTGGAGAGGGAAAGAGAAAAAAAGGTGATCATGGAGGAAAAGGTACTGGATAAAAGTAAACCT
CAAACCTTAGGGGGAGCACTAAACAAACATGTATTATTATAGAAAATATTCTGTTTAATC
TTTCTTTAAACAAGGACTCATACTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
TTTATTAAAGGACTGCAAATGCCAGTGTAAATTGCAAGTCTGTAACAAACTTGATATAAG
AAAAGCAGAGAACAAATTCCCTCCCTCAAGATGCAACCTCATGTTGTTAAGGTATAGCATTAGT
CCAGATTGAGAACAGTTGGGTGAACAAAGGTAAGAACAGATTTTTTTGGCATCAAATCTTCTGC
CTGCCTCTCAGCTGCTCAGAAAATTAAAAACATAATGAAATCAAACACATAACATTGAAACA
GAAGGAAATGCTGTGGACCACAGAACCTCAAGAACATTGTTAAAAAAAAAGTGTACCCGTGAGAAAAGT
ACTCTTAATACTCTGAAATCTTAGAGCAACTTAAGGCTGTAAATACATAGAACAAATTTAAAAAA

ACAAAAAAGAAATTGACTCAGTACTATTCTTTCACTTGAAAATATAAAGAACAAATAAGACAAACAT
TGCAAGTTAAAAAAAAAAAAAAAAAAAAAA

IHRHRLSPATEKGAAASSGEGDKDPPPAHEDIAVKGTATAAGTGGTGAACAVPPHPAAAAAAAA
AAAAAAAAVPPHPNIRALQTQAPQQIPRGPVQQPLEDRIFTPAVSAYSTVTQVARQPGTPSPYSAHE
INKGHPNLAATPPGHASSPGLSQTPYPSQNAGPTTLVYPQTMNSQPQTRSPFFQRPQIQPPRATIPN
SSPSIRPGAQTPTAVYQANQHIMMVNLPMPPVQGPQYCIPQYRHSGPPYVGPPQQYPVQPPGPFPY
GPGPDFPNAYGTPFYSQPVYQSAPIIVPTQQQPPAKREKKTIRIDRNQGGKDITEEIMSGGSRNPT
PPIGRPTSTPTPPQQLPSQVPEHSPVVYGTVESAHAASTPTAASDQKQEEKPKPDPLVLSPLVRLVL
SGEKEQEGBTSETTAIVSIAELPLPPSPTTVSSVARSTIAAPTSSALSSQPIFTTAIDDRCELSSPREDT
IPIPLTSCTETSDPLPTNENDDICKEKPCSVAPNDIPLVSSTNLINEINGVSEKLSATESIVEIVKQEVL
PLTLELEILENPPEEMKLECIPAPITPSTVPSFPPTPPASPPHTPVIVPAAATTVSSPSAAITVQRVL
EEDESIRTCLSEDAKEIQNKIEVEADGQTEEILDQNLSRRSPVPAQIAITVPKTWKKPKDRTRTEEML
EAELELKAEELSIDKVLESEQDKMSQGFPERDPSDLKKVKAVEENGEAEAPVRNGAESVSEGEgidans
GSTDSSGDGVTFPFKPESWKPTDEGKQYDREFLLDFQFMPACIQKPEGLPPISDVLDKINQPKLPMRT
LDPRILPRGPDPFTAFAFGRQTPGGRGVPLLNVGSSRSQPGQRREPRIITVSVKEDVHLKKAENAWKPS
QKRDSQADDPENIKTQELFRKVRSLNKLTPQMFNQLMKQVSLTVDEERLKGVIDLVFEKAIDEPSFSV
AYANMCRLVTLKVPMAKPGNTVNFRLLLNRQCKEFEKDADDVFEKKQKELEAASAPEERTRLHDEL
EEAKDKARRRSIGNIKFIGELFKLKLMLTEAIMHDCVVKLLKNHDEESLECLCRLLTIGKLDFEKAKPRM
DQYFNQMEKIVKERKTSRIRFMLQDVIDLRLCNWSRRADQGPKTIEQIHKEAKIEEQEEQRKVQQLMTK
EKRRPGVQRVDEGGWNTVQGAKNSRVLDPSKFLKITKPTIDEKIQLVPKAQLGSWGKGSSGGAKASETDAL
RSSASSLNRFSSALQPPAPSGSTPSTPVEFDSRRTLTSRGSMGREKNDKPLPSATARPNTFMRGSSSKDLLD
NQSQEEQRREMLETVKQLTGGVDVERNSTEARNKTRSAKPEISAMSAHDKAALSEEELERKSksiiDEF
LHINDFKEAMQCVEELNAQGLLHVFRVGVESTLERSQITRDHMQQLYQLVQSEKLSKQDFFKGFSETLE
LADDMAIDIPHILYLAELVTPMLKEGGISMRELTIIFSKPLLPVGRAGVLLSEILHLLCKQMSHKKGAL
WREADLSWKDFLPEGEDVHNFLLEQKLDIFIESDSPCSSEALSKKELSAEELYKRLEKLIIEDKANDEQIFD
WVEANLDEIQMSSPTFLRALMTAVCKAAIIADSSTFRVDTAVIKQRVPILLKYLDSDTEKELQALYALQAS
IVKLDQ PANLLRMFFDCLYDEEVISEDAYKWESSIONGKVALKSVTAFFTWLREAEEEDN

AUG initiation = 176.7 kDa
AUC initiation = 195.0 kDa

+++++

#6

NM_006375

Extension conserved from human to fish but non-AUG initiation conserved only in mammals

ecto-NOX disulfide-thiol exchanger 2 (ENOX2), transcript variant 1

Synonyms: APK1; tNOX; COVA1; ENOX2

GTGTGTGAGAGTCAGCCTGGGCCAGAGTGGCGGGCTAGGCCTGGAGCGCAGGGCGGGCATTCGGAA
GAGAGTGGCGCGCGTGGGGCGCCCGAGGGCACTACGAGCCCAGGAACCTGCTGGAGGGAACTGATT
GCGGCTGTATAAGGACTAGCGCTTCGGTAGCCGGGAGCTGGAGGGAACTGATT
AGAACACAAAGAAGTGTGATGGA
TTGAACCAGCTGCCCTTGACCCCTGCTAACAAACACGAGCCCCCTGCAG
TTTGGTAGTGCCAGTGGCCCTCTGGTCACAGAAGGCCTCATTGAGAATGGAGGGAACTCAAGCAAGAAAG
AAAGAGAACAAATACTCCTCAGCCGATAACAGTAGAACGTTGACACTGCAATGACACTAC
CTATGCTGATCCAACACTGCATGGGCCACAGCAATGAATAATCTTGGAAATGGCACCGCTGGATTGCCGGA
CAACCAATTTCACCTGACTTTGATCTGCTCTTGAATGACTGGAATTCCACCAATAACTCCAATGAT
GCCTGGTTGGGAATAGTACCTCCACCAATTCTCCAGATATGCCAGTAGTAAAGAGATCATACACTGTA
AAAGCTGCACGCTTCCCTCAAATCCAATCTCCACCTCCTGCAACCCGAGAAAGACCACAGGATGC
AAAACAGTATTTGGGGGGCTGCCGAAAGGGACAGAGCAAATCATTGGAAGTTTCGAGCAGTG
TGGAGAGATCATTGCCATTGCAAGAGCAAGAAGAACTCTGCCACATTGCTTGTGAGGAGTACATGG
TGGACAAAGCCCTGTATCTGCTGGTTACCGCATTGCGCTGGCTCTAGTACTGACAAGAAGGACACAGGC

AGACTCCACGTTGATTCGCACAGGCTCGAGATGACCTGTATGAGTGGAGTGTAAACAGCGTATGCTAGC
CAGAGAGGAGCGCCATCGTAGAAGAATGGAAGAAGAAAGATTGCGTCCACCATCTCACCCCCAGTGGTCC
ACTATTCAAGATCATGAATGCAGCATTGTTGCTGAAAAATTAAAAGATGATTCCAATTCTCAGAAGCTGTA
CAGACCTTGCTTACCTGGATAGAGCAGGGAGAGGTCAACCGCTGAGGCCAATAACTCTACTCCATGAT
CCAGTCGGCCAACAGCCATGTCCGCCGCTGGTGAACGAGAAAGCTGCCATGAGAAAGATATGGAAGAAG
CAAAGGAGAAAGTCAAGCAGGCCCTTCTGGAATTCTCATTCAATTGAGCAGATAGTGGCTGTGTACCAT
TCCGCCTCCAAGCAGAAGGCATGGGACCCTCACAAGGCCAGCGGAAGAACATCAGCGTGTGGTGC
ACAAGCTGAGGAAATTGCAACATTCTATAATGATGAATTAAATGGAATCAGGCGAGAAGAAGAAATGGA
TGTCTGATGATGAAATAGAAGAAATGACAGAAAGAAACTGAGGAATCAGCCTTAGTATCACAGGCA
GAAGCTCTGAAGGAAGAAAATGACAGCCTCCGTTGGCAGCTCGATGCCAACCGGAATGAAGTAGAACTGCT
CAAGCAAGAACAGGAAAGTCCACAGAGAAAGATGACCTAACAAAGAACAGCAGCTGAAACTCCTGCAAC
AAGCCCTGCAAGGAATGCAACAGCATTCTACTCAAAGTCCAAGGAGAAATACAAAAGAAAGACTGAACTT
GAAAAACTCAAAGATGACAAGTTACAGGTGAAAAAAATGTTGAAAATCTTAAAGAAAAGGAAAGCTGTGC
TTCTAGGCTGTGCGCTCAAACCAAGGATAGCGAATACCCCTTTGAGAAGACCATGAAACAGCAGTCCTATCA
AATCTGAACGTGAAGCACTGCTAGTGGGATTATCTCCACATTCTTCATGTTCACCCATTGGAGCAAGC
ATTGAATACATCTGTTCTACTGCAACGCTTGTATAATAAGATCTGCAACAGCGATGTGGAGTGTCTCAT
GGGTAGACTCCAGCATACCTCAAGCAGGAAATGACTGGAGTTGGAGCCAGCCTGAAAAGAGATGGA
TCTGTTGCTTCAGGGCTTGAAGCTGACCTAAATCTCTTGCTAACAAACTTGGGATCCTGAAGATAAATA
TGTGTTGACAAGCATAGAAAGTGATTATATTAAATGTTTAACTGTTTCAAGTGGAAAGTTCTTGAATTGTC
AGTTCACTCCTGAAAATCTTGTAGTTAAATAAGGATCCTAGGACAGCACCTCGAACTACAGGCC
AGAGAAATTGCCTCAAACCAAGTGTAACTTCCCTCCCTTGTCAATTGTTGTCTTAAATATT
GCAAAAGTCTGATGCTAAACAGTATTGGAGTTTCAAGTGTCTGACTACTGTTGAGACCTTGGT
TTTTTAAACACTGTTACTGAAATGTTGATGATTGTGATTGTTCTAAACTCTCTTAC
ATTAATGTTGTTACTGGTAAAGGCATGAGAGCAGCACTAAGTCTCTGTAACTGCCATTGCTTCA
ATCCCCAGTAGACAGTAAATAACACATCAGTGTCTTAGAAGGTGCTGACCAGGTTCACCTT
AAACGACAAAGCATGGTTGTGGCTTTGCAAATTACTATGAACCAAAGTTGACAATGTC
TATTTCTCTAACATATCACATTAAAGATCTGTTCAAGATTGAAAAGTACATCTAGATGTTACAG
AAAGCAAGTATCCAGTATGACTGGCATGTGTCATGCTATTCAAATCACTTGTAAATAGTCTGTT
AGGAGGGCATGTTCAAGTTCTGTGAATTAAATGCTCATGTGTTGGCACACACGCACAAACACACA
CGCACGCACACAGTGGCAGAAGGGATTATATTAAATATTCTTCCCTCTGGCTTCTACAGTGT
TCCCTTGCTTCTGTGTCAGTGTGTTGAATTGCAAACCGAGTACTGCTGTAAATACTATGTTACTTCAT
GCTGAATGTTGCAAAGACTTGATATAAGTATTAAATAGTAAATGAATCAAATAAATGAGCTAGGGT
TTGTGAGGCTTCTACAAATAGGTCAAGCTCCACCTGGAGTGGCAATTGCCAGAGACACCTGGTAGTGC
ATCGGCAAATCGCAATGGCAGCATGTGAGTGGACCATTCAAAGACTCTGCTGGTGGAAAGTAAACAGAG
AGGATGGAGGTTGGGCGAATGTCCTGAGGCAAGAGATGGTCTTATTGTGTTGGTGGTTGTGGT
TTATAATAATGCAAGCATACCTCCCTGAGTCTCAATTGAAGATAAAGAATGTA
CAATGGAGAGTATTTCACAAAAAATACTTGTAAATGAGATGCCAGTAGTGTCAAAGTTGATT
GATAAAATTCTTTTATACCTCAGTTGTGTCCTGTTTAATGACTTACGCTCTAACGAAACT
GTAGTTATCTCAGTTCTCCTTGGTTACTAGAATGTTGAAAAGATGCCAGTCTGTCTGACA
GGAAACAGGGTTCACAGCAGCCCATTGCTGAGGAAACTGGCTCCCCCTGAAGCACCCTGCTG
CAGCAGGAAGCTCAGGTTAATTTCACACTAGCTGCTCACTGATGCATCTCTCATCA
TACGGAAAGGCATTGTTGGGCTTGTGAAACCTAACGAAATTAAATATCTATGAA
CATTTACAGTTATTCTGTTAAATAAACTAAATTAAATTAGGGAGAGCAGTAGGAAAAGAGCTAA
TGCATGCGGGGTTAATACCTAGGTGATGGGTTGAGGTGCAAGCAGGAAACCATGG
CACACGTTCACCTTACATGTACCCCGAACCTTACTAA
GTAACAAACCTGCACATCTGCACATGTACCCCGAACCTTACTAA

LNQLPFDPANNNEPLQFGSASGPLVTEGLIENGESSKKRKRTNTPSADNSRTLNVDSTAMTLPM
SDPTAW
ATAMNNLGMAPLGIAGQPILPDFDPALGMMTGIPPIPPMMPGLIVPPPPIP
DPMVVKEIIHCKSCTLFPP
NPNLPPPATRERPPGCKTVFVGGLPENGEQIIIVEVFEQC
GEIIIAIRKSKKNFCHIRFAEYMDVKALYLS
GYRIRLGSSSTDKDTGRLHVDFAQARDLYEWECKQRMLAREERHRR
MEEERLRPPSPPPV
VHYS
DHECS
IVAEKLKDDSKFSEAVQTLTWIERGEVNRRSANNFYS
MIQSANS
HVRRLVNEKA
AHEKDMEEAKEKF
KQA
LSGILIQFEQIVAVYHSASKQKA
WDHFTKA
QRKNIS
WCKQAE
EI
NIH
DEL
MGIR
REE
EM
MS
D
E
IE
MTET
K
TEE
E
S
AL
V
S
Q
A
E
A
L
K
E
E
N
D
L
R
W
Q
L
D
A
Y
R
N
E
V
E
L
L
K
Q
E
Q
G
K
V
H
R
E
D
D
P
N
K
E
Q
Q
L
K
L
L
Q
Q
A
L
Q
G
M
Q
Q
H
L
L
K
V
Q
Q
E
Y
K
K
K
E
A
E
L
K
L
K
D
D
K
L
Q
V
E
K
M
L
E
N
L
K
E
K
E
S
C
A
R
S
L
C
A
S
N
Q
D
S
E
Y
P
L
E
K
T
M
N
S
S
P
I
K
S
E
R
E
A
L
L
V
G
I
I
S
T
F
L
H
V
H
P
F
G
A
S
I
E
Y
I
C
S
Y
L
H
R
L
D
N
K
I
C
T
S
D
V
E
C
L
M
G
R
L
Q
H
T
F
K
Q
E
M
T
G
V
G
A
S
L
E
K
R
W
K
F
C
G
F
E
G
L
K
L
T

AUG initiation = 66.6 kDa
ACG initiation = 73.0 kDa

++++++

#7

NM_176677

UUG/CUG initiation in mammals, birds and reptiles (and maybe all vertebrates)

FLJ36208 hypothetical protein FLJ36208

Synonyms: FLJ36208; NHLRC4

CAGCCACCCAGCCCCGAAGCAGGGCCCAGCGAGACAGGGCCCAGCCCCAACGCATGACCCAGCTGCAGAGAA
GCTGAGACGGTATTGGACTGACTGGACTCTGGCTCCAGGTGTCTGCACCTCACATGGACCCACCTGG
AGACTCCACGGCTCCCGCAGCTCCGGCTGCTGGGTGGATCCGGGTGCCCGAGGGCTGTGGGAGGGC
CCCGGGGCGCTGCACTGCTCCCCGATGGCCTGCTTCCCTCACGGCCGGGCTGCACCCCTGTGTCCACGTG
CTAGATCTGGAGGGACGCCCATCTGCCTCCTGCCCTGCCGACTCCGGGAGCAGGGCTTCGTTCCAGA
GGACGTGGCTGTGACAGCGTCAGGGCTGTGGTGGTCAGCGATCCCATCCATGGGCTGTCCATGCACTCC
AGCACACAGCCGGACCCGGGGCACTGGGTGACAGTGGCACCTCCTGTCTCCCCGAGGGCTGGCT
GTGGATGCCCTAACGCCCTCTGGTGACGGACTACTTGCTGGGCTGTGCACAGCTTCTCGTTGGGTCC
TGCTTGGGAGCCCCCTGGCCCCAGCCTCCATGCTGGGTCTGGAGGGCCCTGCTGGTGGGCCAGGGCCTG
ATGGGGGCCTTGCTGTGAGTGAGGAGTTGGGATGTGAGGTGTTGGCAGTGCCTGCCAACCCCTGGGC
TCCCTGGGGGCTGGACGGGCACACTTGCCTGCCAGCGGGCATCTGCTCCAACCTCAGAGGGCAATGT
TATTGGGAGACAGCAGAGGGGCCAGGTGACCCCTGTTCCCCGGCTGGGCCACCCATCTGCCTGGT
CAGAGGGGCTGGCAGCCCTGGGAGTGGCCTGTGCACCCAGGGCAGCTCTGGTGGCTGATGCCAAG
GACAACCTCCATCAAGGTGTAACCAGGGCTCAAGGAGCTGGCTGACCTGAGGCTGGGTGGAGCAGCCCTC
CTGTGCCTGAGGCCAGCTCCAGGCCCTGGATCACCGGGAGAACCTCAGGATGGGTGGAGCCTCA
GGCTATGGCATTGCCTGCTGATGCCAGCACACCTGGCTGGCCCTGGCTGGCTCGAGTTCTCCTG
CTGGTGAGGCTCCGGATCTCAGGAGCAGCCCTGAGTCTGCTTCCAGGCTGCCAGGCCCTGCAGCC
TCCCCAGCCAGGGCTGCTCTGCTGTCCTGGCCATTCAAGTGCCTGGCCCTGCAATTGCCCCCACACCC
CCTCAGGCCCTGTGCCTGGACTTGGGCTGGCAGCTGAAGCCTTGAGATCCTGGCCAGTGCCTGG
CAGCTAGGCAGACTCTCCACCAGGTGCCCTGCCAGGCCTCTAACCGGGGAGACAGGCAGGGAGGG
TGTGGCTGGCTGGCTGGGGGGGGCGCCTGGGCAGGGGTGTGGCCCTAAATGTCCCACCTCAGAG
GGACCTAGAGTCTGAGCCTCAGTAGCTCTGGGCTGGCAGAGGTAAAGGGGAGGCAACCTGGAGT
GTCTGGAGGCCATGGCTGGCTGAACCTGGATGCCCTTCTCCGCTCCATGAATGAAAGCTGTCTG
GGCCTTCATTCTGCAGACAGGGACAAACAGCTCCATGCTGTTGCTCTCCAGTGCAGCCGTGCTGGGAGG
GTCTGGGGAGCTCCTACAAGGAGAGACTCCTGCTGCTTGGAAAATGAGAAAAAAATAGGGTCTAAC
CTCTCTCCCATTAACTGAGGGAAATGAGGCGTGAAGAGGAGAGCGTCTGGTTACTCCGTGGGTCTG
GGGTCCAGGGAAAGGGCTGTATGGGGAGGGAGCTGGAGGGGACGGTGTCTGGCTCTACCCCTGTGGG
GGGGAGGTGGGCTCCCCTGTATCACAGGACATCCCCCTGAGAGGTCCCTCATATGTCTGGTCTGTGG
GTGGGGGACTAACTGCGCAATGTAGTTAGGTGCTCAATAACGGAGTTGCCGTAAAAAAAAAAAAAA
AAAAAA

LDPPGDSTGSRQLLLGWIRVPAGAVGGPRGLHCSPDGLLFLTAGAAPCVHVLDLEGRPICLLPCRTPGSG
AFVPEDVAVTASGLVVSDPIHGAHALQHTARDPGGHWTVGTFLSRGLAVDALNLLVTDYLPGAVHS
FSLGPWEPLAPASMLGLEGPCWVGPGPDGGLAVSEFGDVRLFGSARQPLGSLGGWTGHTFGCPAGICSN
SEGNVIVADEQRROVTLFPRAGPPICLVSEGLGQPLGVACAPQGQLLVADAKDNSIKVYQGLKELA

AUG initiation = 12.6 kDa
CUG initiation = 28.6 kDa

++++++

#8

NM_153756

AUA initiation in humans - in all others, including gorilla and chimp, there is AUG in place of the AUA

FNDC5 fibronectin type III domain containing 5

Synonyms: FRCP2; FNDC5

GC GGCCGCCGGCGCCGGAGCGCGTCCCCCTGCGCCGCCCGGGCTGCCGGCGAGGAGCCACCATAC
ACCCCGGTGCGCGAGCGCCTGGCGCCCGCGCCCGCCTGTGGCTGGCTGCGTCTGC
TTCGCGCTGGTGCAGCGGACAGTCCCTCAGCCCCAGTGAACGTCACCGTCAGGCACCTCAAGGCCAAC
TGCAGTGGTGAGCTGGGATGTTCTGGAGGTGAGGTTGTCATCGGATTGCCATCTCCCAGCAGAAGAAG
ATGTGCGGATGCTGCGCTTCATCCAGGAGGTGAACACCACCCGCTATGTGCCCTCTGGACCTGGAG
GAGGATACGGAGTACATAGTCCACGTGAGGCCATCTCCATTAGGGCAGAGCCCAGCCAGCGAGCCTGT
GCTCTCAAGACCCCGGTGAGGCTGAGAAGATGCCCTCAAGAACAAAGATGAGGTAACCATGAAAGAGA
TGGGGAGGAACCAACAGCTGGGACAGGGTAGTGCATCGTCGTTCTCATGTGGGAGGT
GTCATTGCCCTCTCTGCCGCCAGTATGACATCATCAAGGACAATGAACCCAATAACAACAAGGAAAAAAC
CAAGAGTGCATCAGAAACCAAGCACACCAAGAGCACAGGGCGGGGGCTCTCCGAGCAAGATA
CTTTTCAGTGCCTGCCCTCAGCAGCTAAGAAGACAGACTAGAGAATGTGAGAGGATCTCATGGTTCTGA
TGATGATTATCCAACAAACATCTGGCCCTCTACATCTCTCCCTCATCTCTGTACCCCTGGCTTA
CTGTCCTCTCTGGCGCACTTCTGAAGCCTTATTAACTCCATCTCAGAACGACCTCAACAATGT
CAGTGGCTGAGGCTGCACTCAGAGGGATGACTGCTGGGGTAGACTGGGTGCCAGGGCCATGGGCCAGG
ACCCAGTCTGGCATTCACTGAGTGAAGGCTGGGTTGAAAAGGCAAAAGACAAGACATCCA
GGCAGGGCTCTCTTTCTTCCACAAGGGACAAGAGCTGGCTTATTAGGCTACAGCCCTGCTGCTGCT
CCCTTCCTCTCTGTCCCTGTTCCAGCCTTGCAAGAAGGCTATTACAATTAGGCTGGCTGCTCTC
ATTTTTCTCCAGTTCCAACACAAGCCCTCAGTGAACATCATTGAAGCGTGAUTGCCTGTGAGGG
AGAAGGATTCCATTTCCTCAGCTGGTCCCCAGGCCACGGGACAGGGAGAGGGACAAGTGCAGCAG
TGGGGAGGAGGCACAGCTAGCTGCACAGTTCTCTCTGTCTAGTCAGATGAAGGAGGCTGCACT
ACAAACCAAATTCTGCAAAAAAAATAAGCCACAAACTAAAAGGCCCTGGCCCATTCTGAAAAG
GCAAAGCTGCATGAGACACAGCCTCTGCCCTCGCTCTGGACTGGCTTCTTGAGAAAATGC
ACAAAGCCCTGGGAGATGACAAGCACAAGGACTGACTCAAGCTGTCTTCAGACCAAGGAACATCAGAG
AAGCTGTGGGCTGCCAGGCAGGATCATGGCTGCCATCAAGCCTTCTGGATCCAGCCATCAAGGA
CATGTTGTGGTGTGACACTTGCAGCGTGAAGATGTTACCTGGTTGCTCTTTGGAAAACA
AAAATCAGAAGGCTGCATTCTAGAGGGCAGAGAAATTCCCCGAAGACTGAGCTGGTTGCCTGCATCCT
ATCTTCTTGACCTTATGACTGAAAGATCATCAGTTGGAAGGTACTGGTCCAATTATTAGGAAGTAT
CTCTTGAGTTCAGAAATGCTAGCTGGACAATGAAAAGTCACATCACAGCTGGCATTCTGGGGCTAC
CAAAACACCCCTCTGGAGTAGAAGCTGGAAAGGCAGGCCATTGCCACAGGAAGAGCA
GCTCTGGCTATCACCACAGGCCCTGGGTCTTCATATCTGCCATCTCATCCAGGGTCCATGAAAGTTA
CCCAGGGCTCTCATGTCCCTCTAGAGCCTGAGTGGTGTGAGGTGACAGGTCTCTCTCCACTGCCCT
TTCTGGTTAAAAAAATGGTGTGATGAGGGAGGTAGACTCTCCCTAGGACTGACGAGTTACGGCTGC
CAGATGCCATGGGAAGAGGTGGACATCTGCATCTCATTGGTGTCAAGGATGGGTGGGAGAAC
ACACCTAGTGCAGCCTGGTACTCAGTAAATATTGTTGAAATGAATGATAAGAGCATTGGTCCCAGGCC
AGAGAGCCAGAAGCCATCACCAATGACCGCCCCCTCTCCGGTCTACAAGAGCTCTCAAGGCTGGGTCT
GCCACCACTCTGCTTGCCTAACAGTGTGACAGCACTGGGGAGAGACAGGATAAAGGGCAGATGTCAGCA
ATACTAAGGGCTCCTCATGGGAGGGCATGAGGCTCCACTCATTGTCTGTGACTTCCATCCCTGCTGAAT
GGGGCTGCAAGGCCAGGCTCTTAGGGGAGAGGTCTTACCTCTGATCCACTTAGAGCAATAACCAC
TTAAATGTAAGAACAAATGAAAAGGCAAAAAAAAAAAAAAA

IHPGSPSAWPPRARAALRLWLGCVCFALVQADSPSAPVNVTVRHLKANSAVVSDVLEDEVVIGFAISQQK
KDVRMLRFIQEVNTTRSCALWDLEEDTEYIVHVQAISIQQSPASEPVLFKTPREAEKMASKNKDEVTMK
EMGRNQLRTGEVLIIVVLFMWAGVIALFCRQYDIIKDNEPNNNKEKTKSASETSTPEHQGGGLRSKI

AUG initiation = 15.5 kDa
AUA initiation = 23.6 kDa

+++++++++++++

#9

NM_006688

Conserved in vertebrates

C1QL1 complement component 1, q subcomponent-like 1

Synonyms: CRF; C1QRF; C1QL1

GCGAGGCAGGAGCCGGCTGGGCTCCGCAGCGCAGCCAGCGCAGCGCGGCCCGGGCCGGCCCCAT
GCCCGCAGCCCCGGGACCGTCTTGAGCGCGGGCGCTAGCCCAGGCTCCCGCGGGCCGAAAGACGCTGCTAGCGGG
CCCCGACGGCGCGGCCGGGCGCTCCCGCTCCAGGCTCCCGCGGGCCGAAAGACGCTGCTAGCGGG
GCCGCGGGTGTGGTGAATGCTGCTGGTGTGGTGTGCTATCCCGTGTGGTGAAGCTCGGGCGGCCGG
AGGCCACTATGAGATGCTGGCACCTGCCCATGGTGTGCGACCCCTACCCCGCGCGGGGCCGGCGCCG
GCGCGGGACCGACGGCGGACGCCCTGAGCGAGCAGAGCGCGGCCCGCCTCACGCTGGTGCAG
GGCCCCCAGGGGAAGCCGGGGCGACCGGCAAGCCCGCCCTCCGGGGCTCCGGGGACCCAGGTCTCC
CGGCCCCGTGGGGCGCCGGGGAGAAGGGTGAGCCAGGAAGCGGGCCCTCCGGGGCTGCCGGCGCG
GGGGCAGCGGCGCATCAGCACTGCCACCTACACCACGGTGCCGCGGTGGCCTTACGCCGGCTCAAG
AACCCCCACGAGGGTTACGAGGTACTCAAGTTGACGACGTGGTACCAACCTAGGCAACAACACTACGACCC
GGCCAGCGGCAAGTTACGTCAACATTCCCGCACCTACTTTTACCTACCATGTCCTCATGCGCGGG
GCGACGGCACCAAGTATGTGGCAGACCTCTGCAAGAATGGCAGGTGCGGGCCAGTGCTATTGCCAGGAC
GCGGACCAAACACTACGACTACGCCAGCAACAGCGTGATCTGCACCTGGACGCCGGGACGAGGTCTTCAT
CAAGCTGGATGGAGGCAAAGCACACGGGGCAACAGCAACAAATACAGCACGTTCTGGCTTCATCATCT
ACTCCGACTGAAGCTCCCCACGTCTCCCTCCACCCACGTCCCTCACCGCCGGGGTCCCTCCGGGGCGGG
AGACGATGACTCGCCCTCGCCACCCGCTCGTGCCTGCCGGCCCTCCCCGGTATGACGCCCGGCCGTG
CTCAACACCGCTGGCCACAGCTAGGCCCTCCACCGGCTCGCTGCAGAGCCGGCCAGCGCGCCCTGT
CCCCGTGCCAGGGAACCGGGGTGACGCCCGCCAGCCGCGCTATATATTGTACAATAGGACTGTT
TACTGCCACCTCCGCTGCCAGCCCACCCAGCTGGGAGAGGTGCGGGCGGGTTGCTTCCTGCG
CTCTGAGATGAGCTGCCCTCGGCTCCCTCCGGTGGCGCGCCGGGGAGGGGGAGTTGGGGCTGGAT
AGCTTCCAGCACCTCAGAGCCCCGCCGGTGTGCCCGTGTGACCAAAGTTATAATAAAACATTT
CACCCCGAAAAAAA

IAPTRPPGPALPRLRAGRKTLLAAAAGVVMLLVLVVLIPVLVSSGGPEGHYEMLTGCRMVCDPYPARPGP
AGARTDGGDALSEQSGAPPSTLVQGPQGPRTGKPGPPGDPGPPGVGPPGEKGEPKGPGLPG
AGGSIASTATYTTVPRVAFYAGLKNPHEGYEVLFDDVVTNLGNNYDAASGKFTCNIPGTYFFTYHVLMR
GGDGTSMWADLKNGQVRASAIAQDADQNYDYASNSVILHLDAGDEVFIKLDGGKAHGGSNKYSTFSGFI
IYSD

AUG initiation = 26.4 kDa
AUU initiation = 29.4 kDa

+++++

#10

NM_145008

ACG initiated extension conserved in vertebrates

YPEL4 yippee-like 4 (Drosophila)

Synonyms: FLJ30213; MGC102723; MGC138324; YPEL4

AGTGTATGATGCAGTTCCACAACACACAGCCACATTCACCCACAGACCGAGGTACAGAACGAGAGACAAC
CTCTGCCCTCCAGCAGCTGCCAGCTTGAGCCCCAGTCTTGAGCCCCAACTACCCTCCCCCACC
CCCCCATCCCCCTCCAATTGAAGGAGGGAAAGAGAAGAGAGATGAGCAGAGAGATTGAGAGATTG
AGAGAGAGAGAGAGAGATAGACGGAGATCTCTGGAGCAGACCTCAAGGTGACTTCTATTCTATCTGGTC
TCGTCTGGGGGGCCCTGGCGGGCAGCCCCAACACTTCTCTGCCCTGCTTGAAACACGGCTCTAGCCAACC
TGCTCCGCTGCTCACCTGCGACCGTCTGCGGGGGCTGACGGGCCAGCCCCCTCCAGCCACCAGGGC
ATTGTCTCCAGCCGTCATGCCAGCTGTGACCCGGTCCGGCCCTGCCTGCCTCCCCACCAAGACTTT

CCGCAGCTATCTGCCCGCTGTCAACGCACTTACAGCTGTCCACTGCCGTGACACCTGGCAAACACG
ATGAGCTTATTCCAAGTCCTCCAAGGGAGCCATGCCGAGCCTACCTGTTAACCTCGTGGTCAACGTG
GGTTGGGGCCAGCTGAACAGCGCCTCTGCTCACGGGCTTACACTCGTAGCTGACATTCTGTGAGAG
CTGCAAAACCACACTGGCTGAAATATGAGCAAGCTTTGAGACGAGCCAGAAGTACAAGGAAGGGAAAT
ACATCATTGAAATGTACACATGGTAAGGACAACGGCTGGGACTGA~~G~~GGGCTCAGGCAGGGTGTGCCCT
CCTCCGCATGCCCTCCCTCCCCACGGCCCTGCCAAGCAGTCTAACAGCATGAGTACTGCCCAACCC
GGGGAAACCTGGCTCCAACCAACCCCTCCCCTGCCCTACCATATCCACTACCAGGCACCCTTAGAAC
GGGGTCTGGGGTACCCAGGGTGTAAAGGCTCAGGAGTGGGAGCAGTCAGGGAGAGACAGAACTGGG
GAAAGGGATGGTTGTGGGTCTTCTGTTCCAAAGATCCTGAACATGGAAGCGATGGCAGGGCATAGACTCA
GGCAGAGAGGATTGTGGGAGGAATCCGTTTGCTCCACCTCTTTGAGTGAACAGAGGACAAACCTGG
GTCACAGGGCAAGTAGATCATGGACACAGAACAGCAGATGAGAAAAGACTTGGGTTGGAGTGAATTCTG
GTCTCAGACACCCAGGAGACCAGACTCTGAGGATGAAGTCTACCCCTATTTGAGGAAAGGACTT
GAGTCAGGGAAAATCAAATCCCAGGGCTGGGAAATAGTAAAATAATCAAAGGTTTCCATTTCAC
CACTTGTAGTTATCTGGCACTGAAGAGGCACTTCAGTATCTAACCTTGCCATTGGGTTGGGTTGG
GACAGCTGCTCGCGAACAGCCCTAGTCGGCTGCCAGAGTAAGCAGTCTTATGGGCTTCTGAG
GCCAGTCACTGCTCCTGGGACCCAGTCCCTGGAGGGAGGTGGAAAATCAGTGTACGGGCCAGTCT
TCCCCTGGCTGCCACCAGCGAATGAAACTTTGTATGATACATAAAAGTGTGAGTCTATTAAATAAAA
AGGGAAAAAGCAACTGAAAAAAAAAA

TALANLRCFTCDRLCGGCTAPAPPQHQIVLQPVMPSCDPGPACLPTKFRSYLPRCHRTYSCVHCRA
HLAKHDELIKSFQGSHGRAYLFNSVVNVCGPACQRLLLTGLHSVADIFCESCKTLGWKYEQAFETSQK
YKEGYIIEMSHMVKDNGWD

PROBABILITY of export to mitochondria: 0.3552 with extension

PROBABILITY of export to mitochondria: 0.2821 w/o extension

AUG initiation = 14.3 kDa
ACG initiation = 17.9 kDa

+++++

#11

NM_182528

Non-AUG initiation conserved in vertebrates

C1QL2 complement component 1, q subcomponent-like 2

Synonyms: CTRP10; C1QTNF10; C1QL2

GCTGATGACATGAGGGCTCCGTCTCCGAGTGTGGCAGCGCGCTGCTCGCCCTCCGCCGCTCAGC
CCCGACTCTTACGTCAAGGTAGCGGGTCCCCCTCCGCCGGAGCCAGCGAACAGCGAGAGAGCACA
GCAGAGCGCGCCGCGAGCGGGGCGCCTCACTCGCCTAGGAGCCCCACTAAACCCAGCGGAGCGGAGCC
TGGCGGGAGGCAGCGCCGCGAGCCAGCGCCGAGCGCAAGCAGACTCCCCGGCAGCGCAAGCACTCC
CTGGCCGGCGCCGGACCCCTCGGGGCCGGATTCGTCTGTGCCCCACGTATCGCGCGGGCGTCCGC
GGCTCTCCGACCAGCCCCAGCGCGTGGGCCAGGCTCAGAGTCCCCGCGCTCGCTGAAGTAGTTG
GGTAGCCGGGCTGGGGTCGCCACGTCGGGGCGCGCAGGACCCGCGAGCCGGTCCCCGAGCGCGGG
GAGCGGGGCCGCCCGCCGCCACCATTACCTCCCCGGCGCAAGGAGAGCTGGTGGGGTGCCTC
CCGGCTGTGGCAGCGCGGGCGTGCCTGCCCTGGCGCGTACTCTTGCCATGCGCTCGGGC
TGCTCATCGCCGTGCCGTGCTGCGAGCGGCCCGAGGCCCGCGCACTATGAGATGATGGGCACC
TGCCGCATGATCTGCAGCCCTACACTGCCGCCTGGAGGCGAGCCCCGGGTGAAAGGCGCAGCCACC
CGGACCCAGCACCGCCGCCCTGGAGGCTCATGCAGGACCTCAGCGCCAACCCCTCCTCCCTTCCAGG
GACCCAAGGGCGACCCGGCGACCGGGCAAGCCAGGGCGGGGGCCCCCTGGAGAGGCCGGCCGCT
GGACCCAGGGGCCCTCCGGGAGAGAAGGGCGACTCGGGCGGCCGGCTGCCAGGGCTGCAACTGACGGC
GGGCACGGCCAGCGCGTGGGTGGCGGGGGCGGGTAGGTGGCGATTCCGAGGGTGAAGTGA
CCAGTGCCTGAGCGCCACCTCAGCGCCCCAAGATCGCCTCTATGTGGGTCTAAGAGCCCCACGAA
GGCTATGAGGTGCTGAAGTTCGATGACGTGGTACCAACCTCGGCAATCACTATGACCCACCGGGCAA

GTTCAGCTGCCAGGTACGCGCATCTACTTCTTCACCTACCACATCCTCATGCGGGCGGCACGGCACCA
GCATGTGGCGGACCTCTGCAAGAACGGGCAGGTCGGGCCAGCGCATTGCACAGGACGCCACCAGAAC
TACGACTACGCCAGTAACAGCGTGGTGTGCACTGGATTCAAGGGACGAAGTGTATGTAAAGCTGGATGG
CGGGAAAGGCTCACGGAGGAATAACAAGTACAGCACGTTCTCGGGCTTCTCTGTACCCGGATTAGG
GGCGCGGGGGTGGAGGGGGCTGCGAGGCCGGCTGCGAGGCCGGCTCCGCCGGCGCGCTCCGGCAAAG
GCCACTCTGATTATAACACTCCTGACATCTCCTTGAAAAGACAAATCCCTGCGTCCTCCCTGCC
GCTCCTGGCCTCAGTGCCTGCGACCCACCACGCTCAGGGCTGTGCTCTGGTCTCCATCCCCATCC
CAAGGGAGGAAGGGACGCCGGAGCCCTTGAGGCGCGCACAGACTTGCAAACCTGATTAGACTGGACAG
GCAGGGCCGGGAGGCTGCCCTCCTCAGACAGCCTCCCAGTGCCTAGAAGCGGAGGGCTCCGGGCC
GCCAGGGAGGTAGGCCAGAGGGAGCGGGCTTCTCTGGGCGTCTTGTGACCCGAAATACTGTGC
AGATTTCCCTGTCCATCAGCAAAACCCACCCAGCAGAATTCCAGCAAACAGAAAATTCACCTCTCCA
CACCGCATTCCCTCTGACTCAGACTCACCGCGATGCATTAAATTATGTTTAGACTATG

ITSPGGKEELVAVASRLWQRRRACLAAGVLLAMALGLLIAPLLLQAAPRGAAHYEMMGTCRMICDPYT
AAPGGEPPGAKAQPPGPSTAALEVMQDLSANPPPFIQGPKGDPGRPGPKPGRGPPGEPGPPGPRGPPGEK
GDSGRPGLPLQLTAGTASGVGVVGGAGVGGDSEGEVTSALSATFSGPKIAFYVGLKSPHEGYEVLFDD
VVTNLGNHYDPTTGFSCQVRGIYFFTYHILMRGGDGTSMWADLCKNGQVRASAIAQDADQNYDYASNSVV
LHLDSGDEVYVKLDGGKAHGNNNKYSTFSGFLYPD

AUG initiation = 29.5 kDa
ACG initiation = 33.0 kDa

+++++

#12

NM_000314

CUG initiation conserved in mammals

PTEN phosphatase and tensin homolog

Synonyms: BZS; DEC; MHAM; TEP1; MMAC1; PTEN1; 10q23del; MGC11227;
PTEN

CCTCCCTCGCCCCGGCGGGTCCCCTCGCTCGCTCGCTCCGCCCTCCCGCTCCCTGGTCTCCGAGGCC
CCGGGCTCCGGCGCGGGCGGGAGGGGGCGGGCAGGCCGGGGCGGTGATGTGGCGGGACTCTTATGC
GCTGCAGGATACGCGCTGGCGCTGGGACGCGACTGCCTCAGTTCTCTCTCGGAAGCTGCAGCC
ATGATGGAAGTTTGAGAGTTGAGCCGCTGTGAGGGAGGGGGCTCAGGCAGGGAGATGAGAGACGGCG
GCGGCCGCGCCGGAGCCCTCTCAGCGCCTCTGAGCAGCCGGGGCAGGCCCTCGGGAGGCC
GGCCTGCGCGGGCGCAGCGCGCGTTCTCGCCTCTCTCGTCTTTCTAACCGTGAGCCTCTCCT
CGGCTCTCTGAAAGGGAAGGTGGAAGCCGTGGCTGGGCGGGAGCCGGCTGAGGCGCGGGCGGG
CGGCACCTCCGCTCTGGAGCGGGGGAGAAGCGCGGGCGGCCGCGGGCTGAGCTCCAGG
GAGGGGTCTGACTGCCCTGTCACCATTCCAGGGCTGGAACGCCGGAGAGTTGGTCTCTCCCTCTAC
TGCCTCAACACGGCGGGCGGCCGACATCCAGGGACCCGGCGTTAAACCTCCGTCC
GCCGCCGACCCCCCGTGGCCCGGCTCCGGAGGCCGCCGGAGGCCAGCCGTTGGAGGATTATCGC
TTCTCCCATTCCGCTGCCGCCGCTGCCAGGCCCTGGCTGAGGAGAACGCCAGTCGCTGCAAC
CATCCAGCAGCCGCCAGGCCATTACCGGGCTCGGGCTCAGAGCCAAGGCCAGGCCAGAGCGAGGGCA
TCAGCTACCGCCAAGTCCAGGCCATTCCATCTGCAAGAAGGCCAGGCCACCAGCAGCTCTGCCATC
TCTCTCTCTTTCTCAGCCACAGGCTCCAGACATGAGCCATCATCAAAGAGATCGTAGCAGAA
ACAAAAGGAGATATCAAGAGGATGGATTGACTTAGACTTGACCTATATTATCAAACATTATTGCTATG
GGATTTCTGCAGAAAGACTGAGGCGTATACAGGAACAATTGATGATGTTAGAAGGTTTGATT
AAAGCATAAAACCATTACAAGATATAACATCTTGTGCTGAAAGACATTATGACACCGCAAATTAAATT
GCAGAGTTGCACAATATCCTTGTGAGGACATAACCCACCCAGCTAGAACATTATCAAACCTTTGTGAA
GATCTTGACCAATGGCTAAGTGAAGATGACAATCATGTTGAGCAATTACTGTAAGCTGGAAAGGGACG
AACTGGTGAATGATATGTGCATATTATACATCGGGCAATTAAAGGACAAAGAGGCCCTAGATT
TCTATGGGAAGTAAGGACCAGAGACAAAAAGGAGTAACATTCCAGTCAGAGCGCTATGTGATT
TATAGCTACCTGTTAAAGAATCATCTGGATTATAGACCAGTGGCACTGTTGTTACAAGATGATGTTGA

AACTATTCCAATGTTAGTGGCGGAACCTGCAATCCTCAGTTGTGGCTGCCAGCTAAGGTGAAGATAT
ATTCCTCCAATTCAAGGACCCACAGACGGGAAGACAAGTCACTGTACTTGTAGTTCCCTCAGCCGTTACCT
GTGTGTGGTATCAAAGTAGAGTCTTCCACAAACAGAACAGATGCTAAAAAAGGACAAAATGTTCA
CTTTTGGGAAATACATTCTCATACCAGGACAGAGGAAACCTCAGAAAAAGTAGAAAATGGAAGTCTAT
GTGATCAAGAAATCGATAGCATTGCAGTATAGAGCGTGCAGATAATGACAAGGAATATCTAGTACTTACT
TTAACAAAAAATGATCTTGACAAAGCAAATAAGACAAAGCCAACCGATACTTTCTCAAATTAAAGGT
GAAGCTGTACTTCACAAAACAGTAGAGGAGCCGTCAAATCCAGAGGCTAGCAGTCAACTCTGTAAACAC
CAGATGTTAGTGAACAATGAACCTGATCATTAGATATTCTGACACCACTGACTCTGATCCAGAGAATGAA
CCTTTGATGAAGATCAGCATACACAAATTACAAAAGTC **TGA** ATT TT TTT ATCAAGAGGGATAAAACAC
CATGAAAATAAAACTGAATAAACTGAAATGGACCTTTTTTTTAATGGCAATAGGACATTGTGTAGA
TTACCAAGTTAGGAACAATTCTCTTCTGACCAATCTGTTTACCCCTATACATCCACAGGGTTTG
CACTTGTGTCCAGTTGAAAAAAGTTGTGTAGCTGTATGTATACCTTTGTGTCAAAGGACAT
TTAAAATTCAATTAGGATTAATAAAGATGGCACTTCCCCTTTATTCAAGTTTATAAAAAGTGGAGACA
GAETGATGTGTACGTAGGAATTTCCTTTGTGTCTGTACCAACTGAAGTGGCTAAAGAGCTTG
TGATATACTGGTTCACATCCTACCCCTTGCACTTGTGGCAACAGATAAGTTGCAGTTGGCTAAGAGAGG
TTTCCGAAGGGTTTGCTACATTCAATGCATGTATTGGGTTAGGGGAATGGAGGGAAATGCTCAGAAAGG
AAATAATTATGCTGGACTCTGGACCATATACCCTCCAGCTATTACACACACCTTCTTAGCATGC
TACAGTTATTAACTGGACATTGAGGAATTGGCCGCTGTCACTGCTTGTGCGCATTTTTTAA
AGCATATTGGTCTAGAAAAGGCAGCTAAAGGAAGTGAATCTGTATTGGGGTACAGGAATGAACCTCTGC
AACATCTTAAGATCCACAAATGAAGGGATAAAATAATGTCATAGGTAAAGAACACAGCAACAATGACT
TAACCATATAATGTGGAGGCATCAACAAAGAATGGCTGAAACATTATAAAAATTGACAATGATTAT
TAAATATGTTTCTCAATTGTAACGACTTCTCCATCTGTGTAACTCAAGGCCAGTGTCAAATTCAAG
GCTGTTAGTACCTACATCAGTCACAAACTTACACTTATTACTAGTTCAATCATAACCTGCTGTGG
ATGCTTCATGTGCTGCTGCAAGCTTCTTTCTCATTAAATATAAAATATTGTAAATGCTGCAGAA
ATTTCATTTGAGATTCTACAGTAAGCGTTTTCTTGAAAGATTATGATGCACTTATTCAATAGCT
GTCAGCCGTTCCACCCCTTGACCTTACACATTCTATTACAATGAATTTCAGTTGCACATTAA
ATGTCATTAACTTAGGGAAATTTCAGTTGAATACTGAATAACATATAATGTTATATTAAAAGGACATT
GTGTTAAAAGGAAATTAGAGTTGCAGTAAACTTCAATGCTGCACACAAAAAGACATTGATTTC
AGTAGAAATTGTCTACATGTGCTTATTGATTGCTATTGAAAGAATAGGGTTTTTTTTTTTTTT
TTTTTTTTAAATGTGCTAGTTGAATCATTCTCATAGTGTCCCCGAGTTGGACTAGGGCTTCAA
TTTCACTCTTAAAAAAATCATCATATATTGATATGCCAGACTGCATACGATTAAAGGGAGTACAA
CTACTATTGAAAGCTAATGTGAAGATATTATAAAAGGTTTTTCCAGAAATTGGTCTTCAA
TTATACCTCACCTTGACATTGAATATCCAGCCATTGTTCTTAATGGTATAAAATTCCATTTCAT
AACTTATTGGTCTGAAATTGTCACTAGCTGTGGCTGACCTAGTTAATTACAAATACAGATTGAATAG
GACCTACTAGAGCAGCATTTATAGAGTTGATGGCAATAGATTAGGCAGAACTTCATCTAAATATTCT
AGTAAATAATGTTGACACGTTCCATACCTGTCAGTTCTTCAACAAATTAAATTAAACAAAGC
TCTTAGGATTACACATTATTTAACATTGATATAGAGTTGATTGATTGCTCATAAGTTAAATT
GGTAAAGTTAGAGACAACATTCTAACACCTCACATTGAAATTATGCCACCTGTCTTCATAAAAG
CTGAAAATTGTTACCTAAATGAAACTTCATGTTGAAGATAAGTTAAATATTGTTCTTGTAA
CAATTTCGGGCACCGCATATTAAACGTAACCTTATTGTTCCAATATGTAACATGGAGGGCCAGGTCAA
ATAATGACATTATAATGGGCTTTGCACTGTTATTATTCTCTTGGAAATGTGAAGGTCTGAATGAGGGT
TTTGATTGAAATGTTCAATGTTTTGAGAAGCCTGCTTACATTATGGTGTAGTCATTGAAATGGA
AAAATGGCATTATATATATATAAAATATATTACATACTCTCCTACTTATTCACTTAC
ATCCCCATAGAATTGACAAGAATTGCTATGACTGAAAGGTTTCAGTCAGTCTAAATTAAACTTATT
GCAGTATTCTACATTAGCCTGAAATGCATTCTGTAGGTAACTCTGAGTTCTGGAATATTCTTAGACT
TTTGATGTGCAGCAGCTTACATGTCTGAAGTTACTTGAAGGCATCATTAAAGAAAGCTTACAGTTGG
GCCCTGTACCCTCCAAGTCCTTGTAGCTCCTCTGAACATGTTGCCATACCTTAAAGGGTAGTTGA
ATAAAATAGCATCACCATTCTTGCTGGCACAGGTTATAAAACTTAAAGTGGAGTTACCGGCAGCATCAA
TGTTCAGCTTAAAAAAATAAAGTAGGGTACAAGTTAATGTTAGTTCTAGAAATTGTCATGCA
TCATAACGATGGCTGTGGTGCACAAAGTGCCTCGTTACCTTAAATACTGTTAATGTGTAC
AGATGGAAGGGTGGAACTGTGCACAAAGTGGGGCTTAACTGTTAGTATTGTCAGTCTGTTAA
CTGCCAGTTCAAAAGTTCAACCTGTTCATATAGAATATATACTAAAAAATTTCAGTCTGTTAA
CCTTACTCTGATTCTGAGCAGTACTCTGCTGTGCAGCAGTGGCTCTGTTAAATGCTATGC
ACTGAGGATAACACAAAAATACCAATATGATGTGTACAGGATAATGCCCATCCAACTCAGATGT
TTATTGTTGTTAACACCCTTATCTTAGTGTATAAAACTCCACTTAAACTGATTAAAGTCTCAT
TCTTGTCAAAAAAAAAAAAAA

LERGGEAAAAAAAAAPGRGSESPVTISRAGNAGELVSPLLLPPTRRRRRHIQGP GPVLNLPSAAAAPP
VARAPEAAGGGSRSEDYSSSPHSAAAARPLAAE EKQAQSLQPSSSRSSHYPAAVQSQAAERGASATAK
SRAISI LQKKPRHQQLLPSLSSFFFSHRLPDMTAIKEIVSRNKRRYQEDGFDLDTYIYPNIIAMGFP AE
RLEGVYRNNDVVRFLLSKHKNHYKIJYNLC AERHYDTAKFNCRVAQYPFEDHNPPQLELIKPF CEDLDQW
LSEDDNHVAIHC KAGKGRGTVMICAYLLHRGKF LKAQEALDFYGEVRTRDKKGVTIPSQRYYVYYSYLL
KNHLDYRPVALLFHKMMFETIPMFSGGT CNPQFVVCQLKVKIYSSNSGPT RREDKF MYFEFPQPLPVCGDI
KVEFFHKQN KMLKDKMFHF WVNTFFIPGPEETSEKVENGLCDQEIDSIC SIERADNDKEYLVLT KND
LDKANKDKANRYFSPNF KVLYFTKTVEEPSNPEASSSTSVTPDVS DNEPDHYRSDTTDSDPENE PFDED
QHTQITKV

AUG initiation = 47.2 kDa
CUG initiation = 64.9 kDa

+++++

#12

NM_139239

CUG initiation conserved in mammals.

NFKBID nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, delta

Synonyms: MGC11314; TA-NFKBH; IkappaBNS; MGC149503; NFKBID

TTCGCCGGAGCGCGACCCGGGACTCCCAGGCCTGTGGCGGGCCCTGCCAGGACTGGCGGTGCCATAA
CCCCTAGTTAAAAACTCGCGGTACCGGACCAAGATCGGGGACCCGGCGCGCCCTCGCGGGGGAAACA
GCGAGGCTGGCGCAGCGCCAGGCCAACGCAAACCGTGAAGAAGCTCTGGAAGAGCAGAGGCGCCAGC
AGCAGCAGCCGACGCTGGGGGTGCAGGGACAATTCTCCCTCCCCAGAGCAGCCCTGACCCCATCT
GTGAATGAGGCTGTGACTGGCACCCCTCCAGCACACTCGGAGACTGTGGTTCTGGACCTAGCAG
CCTGGGCTTCCAGACTGGACCCAAACACGCATCGCTACACTGACAGCCCTACTCTGCCCTGCTT
CTGCTGCCGAAATTTCTGCCCTCTGACTTCTACCCACCCCTCGGACCCAGGGCAGCCGTGCCATTCCC
CAGGGCATGGGCTGGACCTGGAGAGTTCTGCACCCCTCAGGACCCCCACAGTCCCCGCTGTGGT
CCCTGGACC ATCGCTGGAGGTGGCCCGAGCTCACATGCTGGCTTGGGCCACAGCAGCTGCTGGCCAGG
ATGAGGAGGGGACACGCTCCTCACCTGTTGGCTCGGGCTCGGGCTGCGCTGGGGCATATGCTGCGGCT
GAGGTGCTCCAGGTGTACCGCGTCTTGACATTCTGAGCATAAGGGCAAGACCCCTCTCTGGTGGCGC
TGCTCCAACCAGCCCTGATTGTGGAGGATCTGTTGAACCTGGAGCAGAGCCAA TGCGCTGACCATC
AGGGACGTT CGGTCTTGACGTGGCGTACCTACGGGCTCCAGGAGTTCTTGGCTGTGCTTA ACTCT
GGGGTCAGGTTGACCTGGAGGCCAGAGACTTCGAGGGCCTACCCGCTCCACAGGCCAT CCTGGCCCT
TAACGTTGCTATGCCCTTCCGACCTCTGCCCCGGGTGCTGAGCACACAGGCCAGACAGGCTGGATT
GTGTCCACATGTTGCTGCAAATGGGTGCTAATCACACCAGCCAGGAGATCAAGAGCAACAAGACAGTTCTG
CACTTGCCGTGCAGGCTGCAACCCACTCTGGT CAGCTGCTGGAGCTGCCCGGGAGACCTGCG
GACCTTGTCAACATGAAGGCCACGGAACACAGCCCTCACATGGGGCTGCCCTGCCCTGGCCGG
CCCAGGAGGCCATCGTGC GGACCTGTTGGCAGCTGGCGACCCCACTGCGCAACCTGGAGAATGAG
CAGCCC GTTACCTGCTGC GCGCCGGGCCCTGAGGGCTCCGGAGCTGTTGAAGAGGAGCCGTG
GGCGCCGCCAGGCCCTGCTCTTAGACTCAAACCCAGACCC TGGACTGATTTCAAGCTCCACCGTC
GC GGACAGCCAGCGTATGCTAATGTTGCAAACCCATGATAATGTATGGAATATCCTGCCATTGGGGTT
TTACATTAAAACCCAGAATGGCTGCAAGGGGTGAACAGGCCCAATATTGGGGTGTGATACCCCT
CTTCTACCCACAAGGAGCCCTTGTGATGATTCTGTAATCGAGGCCCTTGATTGTTCTGTGAAACAC
CCTGCACCCCTAGTCCCTTCCACTGAGATCTTCCGGTTCTCTCCCTAACTCAGCT

LAQRGRGPGGPQSTPRNPRVSRGERSHCPTQTVKLLEEQRRQQQPDAGGVQQFLPPPEQPLTPSVN
EAVTGHPPFP AHSETVGSGPSSLGFPDWDPNTHAAYTDSPYSCPASAENFLPPDFYPPSDPGQPCPFPQG
MEAGPWRVSAPPGPPQFPVVPGPSLEVARAHMLALGPQQLLAQDEEGDTLLHLFAARGLRW AYAAEV
LQVYRRLDIREHKGKTPLLVAAAANQPLIVEDLLNLGAEPNAADHQGRSVLHVAATYGLPGVLLAVLNSGV
QVDLEARDFEGLTPLHTAILALNVAMR PSDLCPRVLSTQARDRLDCVHMLLQMGANHTS QEIKSNKTVLHL

AVQAANPTLVQLLLELPRGDLRTFVNMAHGNALHMAAALPPGPAQEAIVRHLLAAGADPTLRNLENEQP
VHLLRPGPGPEGRLQLLKRSRVAPPGLSS

AUG initiation = 33.5 kDa
CUG initiation = 48.7 kDa

+++++

#14

NM_001015072

Extension conserved from human to fish but non-AUG initiation is conserved only in mammals. In the second human paralogs, UFSP2, the extension is conserved but is initiated by AUG codon and is part of a larger extension. In-frame AUG codon exists in mouse, Norwegian rat, kangaroo rat and dog upstream of the CUG.

UFSP1 UFM1-specific peptidase 1 (non-functional)

(the “non-functional” designation is apparently based on the fact that it is missing the conserved N-terminal domain of the protein but it is actually there if one starts from the first or second CUG)

Synonyms: UFSP; UFSP1

AGCATTGTGGGAGCTCCTCGTCGGTCCGGTCGGTGGCTGCCTATTGCGGCCTGCGGTGATCAACGAGGC
CCGGGGAGCGCGTCCCAGTCTGCGCCGGCTCTGCGCAGCTGGCCAAGACCAGGCCAAAGGA
TGTTGGAGCCTGAGGTCGCTCCGCGCCCTAGGAGGACGCTGTGCCTGGCTCTGGCTGGACCTCCGCTCCGCC
ACCGCCCTGGAGCCGCTGAGGGACGTCACGTGGCCTGTCCCCGCCGAGCCGCCGCCCCGTCCGCC
GCTGCTCTCGGCCACTACCTCTACTACCACTACGGCTGCGACGGCCTGGACGACCCGGCTGGGCTGG
GCTACCGCACTCTGAGACGCTGTGCTCGTGGCCAGAGGCCAGCCCCGGCGTACCTGGACTGGCC
GTACAGGGGCCCTGGAGGACATGGGACAAGCCCCCGGCTTCCGGGCTCCGGGACTGGATGGCTG
CGTGGAGGCCAGCCTCTGCCCTGCTACTCGGAGGGCCCCAGGGACGCCCTGCCACGTACCCGGGGAG
TGGGGCTGCACGGGAGCTGGAGAGGCTTACTCGCACTTCCGAGGGGGTGGGGGCCAGTCATGGTTGG
GGGGACGCAGATGCCAGGTCAAAGGCCCTGCTGGAGTCTGCGTAGGGTCAGGCACGGAAGCCTATGTC
GGTATTGGACCCCTCACTACTGGGCACTCCAAAAGCCCCAGTGAACACTACAGGCTGCTGGGTGGGTGG
GGCAAGAGGTGAGTGCAGCCTTGACCTGAGGAACTCCTTACAACCTGTGCTTGACGCCCTAGCTCC
CAGCAGCAGCGCACCTGGACTGAGGACGAAGTTACAGAACTGAGATTCTGGGTCCCAGACACGCAC
TGTACCTCCCCTGGTGTCCCTGCAAAGCCTGGCGCTTTGACATCAATAATAAAAGTGGCAGGGCTGAGC
AAAAAAAAAAAAAA

LEPLRDVHVGSLSPSRGPVRLALLSGHYLYYHYGCDGLDDRGWGCYRTLQTLCSWPEGQPAVGVLAAVQ
AALEDMDKPPGFRGSRDWIGCVEASLCLAHFGGPQGRLCHVPRGVGLHGELEYSHFAGGGGPVMVGGD
ADARSKALLGVCVGSGTEAYVLVLDPHYWGTPKSPSELQAAGWVGWQEVSAAFDPNSFYNLCLTSLSQQQ
QRTLD

AUG initiation = 15.0 kDa
ACG initiation = 23.2 kDa

+++++

#15

NM_020153

AUA initiated extension in mammals, AUG initiated in other vertebrates

C11orf60 chromosome 11 open reading frame 60

Synonyms: IFT46; C11orf2; FLJ21827; C11orf60

GTGACTCCTTGATCTTCTCGTGAGAGCCTTCGCGGGTGGAGGCTTAACGCCAGGAGGTCTCACGA
GAGTGGAAAGCAACTCTCGCGAATTAAAATTATCTTTGCCTAGCAGTACAACAGGCTGGTTGCTT
GGCGTGGAAATCTAAAGTGGCTGGCTTGAGACTGGAGACCCAGCCCTAGGCTGGGTTCTTCC
ATTATAGAGGAGACGGATTCAAAGGGTACAGACCAAGGTTGAAACCAAACATAATGATGAGCGTCT
AGAGATTAACGACTCCGAAGAGTTGCAAGTATTATACTCCAACCCAAAGACACCAAGGACTTCCTCGTT
CTGCCCCTTCTAACAAAGGCTATCGCTGATAACAGCAGTGTGAGTATGAAGAGGAAAATAGCAAGGTC
CTAAGGGAGGGCATGCCACAGGCTCCAGGCCACAGAGGCAAAGACATGGACCCCTGTTCCACCTGCCCCG
AAAGTCTCAAGTGCACCCAGACCTTCATGTACTTGAGAGGGTGGATGGTACAGAGAGGCCAGAAGC
ACAGAAAGGAGAAGAAGGAGACCTCACAGTTGACACCTCAACGGGCTTGTGAAAATGAGGATGACGAT
GATGATGATGATGATTCTGAAACTGATTCTGATTCTGATGATGATGAGAGCATGGAGCCCTCT
GGAAGGGGCTATGACCCTGAGACTATGAGCATTGCGAGTTCTGCTGAAATTAGGAACCTTCCAGT
ACATCAGTAGGTACACACCTCAGTTGATTGACCTGGACCACAAACTGAAGCCTTCATTCTGATTTATC
CCAGCTGTCGGGATATTGATGCAATTCTAAAGGCCCACGCTGTGATGAAAGCCTGACAACCTTGGCCT
ATTGGTATTGGATGAACCTTCTACAAAGCAGTCAGACCTACGGTGTCTCACTCTGGTTAACAGAGAATT
CTAAGCAGCACAAACATCACACAACATATGAAAGTAAAAGCCTAGAAGATGCAAGAAAAGAATCCAAAGCC
ATTGACACGTGGATTGAGAGCATCTCTGAATTACACCGTTCTAAGCCCCCTGCGACTGTGCACTACACCAG
GCCCATGCCGACATTGACACGCTGTGAGGAAAGTGTGAGGAAAGCCTGACAACCTTGGCAGGTAA
GCCCTGCCACGGCAGAGATTGATTGACGCTGGCAGAGTACATTGACATGATCTGCCATTCTAGACATC
CCTGCTACAAGAGTCGGATCCAGTCCCTCATCTGCTCTTCCCTACTCAGAATTCAAGAACTCACA
GCATTAAAGCTCGCTGAAGGCAAGAAAGCATTCACTCCTCATCCAATTCCACCTCCAAGCTGGAG
ACATGGAGACATTAACCTTCAGCTGAAGACTTCCAAGCTGCTGTTCAAGGCTGAGCTGGCCCTCTGC
CCCAGCTGAGATGGACAGATCGTGTCACTGACTTGATGTCCTGCCATGCCACAGCTGGCTCAGGGC
AGTGCATGTCCTGCTGCCCTCTGCAAGAGGGCACAGAACATGTTGTTAACGACCTGCCCTGCTCAG
ATTGCTGTCCTGGGGAGTTAACGATCTACACCCTGTGGGATTGAGTTAACGAAATTGGAATTCTG
AGATCCCATGGAGGTTAGATTGGGAGGAAAGCTAAAGATGTCCTTTGTGAGAGGGATGGAATTGTT
TCTTCATTCTGAAAGTTAGTGAGTAAAGATTATAAAATCAAAAAAAAAAAAAAA

IEETDSEGLQTKVENQTYDERLEINDSEEVASIYTPTPRHQGLPRAHLPNKAMADNSDEYEEENSKVL
REGMPQAPGHRGKDMDPVPPAPASLKCQTPSHVLERVGWYRESQKHREKKETSQQLPQRGFSENEEDDD
DDDDSSSETDSDDDDDEEHGAPLEGAYDPADYEHPVSAEIKELFQYISRYTPQLIDLDHKLKPFIPDFIP
AVGDIDAFLKVPRPDGKPDNLGLLVLDEPSTKQSDPTVLSLWTENSQHNITQHMKVKSLEDAKNPKAI
DTWIESISELHRSKPPATVHYTRMPDIDTLMQEWSPEFEELLGVSLPTAEIDCSLAEYIDMICAIDLIP
VYKSRIQLHLLFSLYSEFKNSQHFKALAEKKAFTPSSNSTSQAGDMETLTFS

AUG initiation = 40.0 kDa
AUA initiation = 46.2 kDa

+++++

#16

NM_001005404

ACG initiated extension conserved in vertebrates

YPEL2 yippee-like 2 (*Drosophila*)

Synonyms: FKSG4; DKFZp761C2021; YPEL2

GCCGCGCGGTGGGGAGACTGTGGCTTAAGAGCGTGCCGGAGCCCCAGCCGGCCGCGCTTC
GCCGCTGGCACCCCCAGCGGAGCCAAGCCCCACGCTGGCCGACAGGGCCGCTGCGCCGGCTGCTGAG
AACTAGCCCTAGACCTCTGCGTGAGGGTTCTCTGCCGAAGACATCACCAGTGTGAGCCTGCCACACC
CACCCGCTGCCAACACAGGCCCTTACCTGTGTCTCCGGTTTCCCGTGCAGACCCATCCTGTGGAGTG
CCTCGTGGCTGCCAGAGTTCACCCACACTCAGCAGCACCAATGGTGAAGATGACAAGATCGAAGACT
TTCCAGGCATATCTGCCCTCTGCCACCGGACCTACAGCTGCAATTCACTGCAGAGCTCACTGGCCAATCA
TGATGAACTAATTCCAAGTCATTCCAAGGAAGTCAAGGACGAGCATAACCTCTTAACTCAGTAGTTAATG
TGGGCTGTGGGCTGCAGAAGAGCGAGTGTGCTAACAGGACTGCATGCAGTCAGCAGACATTTACTGTGAA
AACTGCAAAACCACTCTGGGCTGGAAATACGAACATGCTTGAAAGCAGCCAGAAATATAAAGAAGGCAA

ATACATCATTGAACTAGCACACATGATCAAGGACAATGGCTGGGACTGATTGGACAGCATCTACCCAACCC
AGTGTCCACGTGAACGCCATTCAACCGAACATTCTCCCAAGCGTGAGAGAGTGACTGACACTTGGTCCA
TCCATTAGGGCCTTGCATCCGGGCATCCTCCCACCCCTGACGCCATCTTCTGGTGAACGGGCCTCTAA
ATCGCTGTCTCTGTCTTGTATCTGTTGTGAGTTGATCTGGCTCTCTCTGTCTAGT
TTGGCTGAAAACAAAACAACAAAAGGAACAGATCCTTGACCGCATGGGGCAGCCCACCTGGTAAGGGC
CCCAGGGCCCAGTCGAGAGCTGCCTGATGGCCTCTGTCAGGAGAGCAGTGGCACGGGGCGTGAGGAAGA
GGGAAAGGGAAACTCTAAGGGCCTGGCGGGGAAGGGTGGAGGGTAGGAACAAAATTGCG
CCGCTCTGGAGACCTGATAACTTAGGTTGAAATAATTGACTTGTCTAAAGGACAAGAGAAAAAAA
ATACCTCATGACTGCATTCTCTGACTAGAACGTTCTGACACCAAATGTGCCAGGTTAGCAAAT
GAGCACAAGATGTGGCCCTGATTCTAGTTGGTGGGCAAGGGCCTGGTCTCCTGGCTGAGTGGGGAGT
GTCCTGGCAGCAGCGAGTGACCTGGCAGTGGCAGGTGGCGATGACTCTGATGCCACTCAGTC
TGGCAATCATCTTGCCTCTAGCCACCGTAGATAAGGTGTGAAGGGACTGCTGTTGCAATGGGCTT
ACCATCAAATATCCAAAGGCTTGAACAGCAACCAAGTAAAATCAGTAATTGAGGAGAGCAGGGCACAA
AGGGGCTGCAGTTGGAGCTCTGAAGAAATGGCTCAGATATTGAGTCAGAGAAATAAAAGTAGGATCA
GTTAGCAATTCTAAGTGCCTCCTCTGACCCCTCATAAGAGGAGTGTGGTGAGGGAGGGACTGGTAG
GGGTCACTCCAGGAGGAGGGTTACATTGGAACAGTTCAAGGTTGGCATCTTCCTCTGGTTTA
CACTGGCTTCCGGATCGCAATTCTGTTAGAGTTGGGTGTTCTCCAGTCTGTTACTG
TAGACTGTAGAAAGCAGGGCCCCAGGCCTGAGCTTAGTAATAACCTGGCTGGTAGATTCTCATGCC
TAATTGCCCACCTAGGCCTGAATGTCTGCATGGAGAGAAATCTCTGTCAGTGTGGCCAGCAGCAGGG
AGGAGTCTGCCAAATTCCGATATCACCCCTCCCCATCCAAGCATCTTCGATTAGGAAGTGGAGAG
CACATCCTGTAAGGCCATAAGAGAAAGAGGAGTTGTTACATTAACTAACACTGTGAAGTCTGTTCTA
CAGCAATTCCAGCATTACACAGTATATGACTGAAACTCATTTAACTGGGTTAATTCTATTAGACTGA
ATATATTATTGTTAAGATACTGTCGCTGTTAGGTAAATTCTCAGCATCTCCTCAAGTAGGCCACCTCT
CGGAAAATTCAACCTAAAAGCTCACAAAGAATGAGTTCACTGGGAGATTCTGAAAGTGAAGTGA
ATGAAAGCAGCCAACAGCCCAGGAGCTTCAGAATAGCCTGCAAGCAGAACAGTTCCATTAGAGCG
CGTCTGGTGGAAATGCTTTGTGTCTCCACGCGCTGATGGTGAATGGAGGCCAACAGCTGTG
GGCTTAGAAATCAACTTTGTTCCCAAGGCTTCTGTCAGATCTTCAGTGTCTCATAGCCCTGGGA
GATCAAGTTGTTCTCCCCACTTACTGCAAGGTAGACTGAAGTTCAGAAGAAATACTGAATTCTGCTCC
AGAAGAATAGTTCTCTGGCTCACAGGCCAAGTCTCAATGAAATCGTTTTAACTTCAATTCTAA
GCTGGCTTCCGGCACAGAACCCATGGATTCCCTCTCCCTCCCCCTCTCAAGGAAATAGTCTCC
TTTATGGATTTCATTGGACTCTTCCTCAGCGATTGTCCTGGCTGTTATTGATAGTCCTCCATAAGA
AAATGGGTTAACATGGGTTAGGTATTGCTTCAAAACTACAAATGGAATGTGGTGACATAAACTAGA
CATGGGTTGCCCTCAAGTTCAAGGGGACCAATGTGCCACTGTTCTCCTGGGATGAGGCCCTTGACT
GTTGGATGGATCAGAGCAGGCTCCAGTCAGACCCCTGGTCTGAATGTTTTTCGGTAGACTATCCAGT
GAGCCTCAGGGTCAAGGCCATACTTGCTGTGAGAGAGCTGAGTAGAGTGTGGTTCCATAAC
TACAGGGGAAAAAAAGTCATTAGGTTCCCTTGTGTCACTGAAACAAAAGTGTCTTACAACGTT
GCTCTGTTCATGGGTTGTCTATCTAACATTGAGCAGCATTGGAGAGGCCACAGCTGAGCTATGGAGATGCT
AAATTAACTCATGCCCTCAGTCAGTCATTCTTAATTCTCACCAAATTATTGACTTAGAGCATAACCA
AAGACCTCATTCACTCAGGTTGGGTAATTGGAGTTGTGGTAAGTTGGGGGGGGTG
TTGGGAGTAGAGACAGGGTAAGGGGACGTGAGAAAGGAAAGGCATGAAGTTCTATACCTCAGCCAGCAGC
TGCCTCGTTGGAACTGAAGTCCAGCCAGCAGACTCTAGCTCCATCTCCCTGTGCCACCCCTAGGTCA
TATGACCTGGCACCTGGAGTAGACCCAGACCCCTCGGGACCCGGACATTAGTCTCAGGCTGCTGATG
GATTGATTGACATGAACAAACACAGCCAAACTCGATACCCACAAGCTGTCAGCTGAACCTGACTGAGTG
TTCTCTGAGTTACGAGGATAGGCTAGAGTCATTCTGTTACTGGGGATCAGTGTGTGCGAAAGAGATGA
CCCTTATAAAGAGATTTCAGTGGATATATATAAAAGAAACAGTTGTTGAAATATAACTTTGTAA
TAATATTAAATTTTAAATAATATATTGGTCTGTTCTCAGATCCCCTGAGAGCACCTTTATTTC
CTTTAAATTCTATGGTTCTTGCATTCTGAAAGTATATTAAAGGAAACAGTGTACCAACAT
GTTTCAGTTTTTTAAAGGTTCTATCACTTAACTGGATCAAGGTTGAAGCAATGCCCT
CTGCATTCTCTCCAGTGGACAGACTCTGCACTAACATTAGGTTGAGAAATTGAAATATTCTTGC
ATCAGTATTGGCTAGAAAAGAAAATAAAACCAAGTTAATTAGTAGTAACAAACTACAGTGTATTCT
CCTGTTGGAAAGAATTCCAACAAATCAGAATCACGTTTACTGAGTGGGGACACTTAACTACAGTTACACCTCGGGCGCA
AAGCACTTCGATTGTCCTCTGTTCTCGAGTGGGGACACTTAACTACAGTTACACCTCGGGCGCA
TAAAGTTCTCTCTCTCTGTTGAGTGGAGCTGTTGCGGGTTGCTCCTTTCTGAGTGGACCAACAGCAGAACCCACGAGGAT
TTGTTTGAGTAGGAGCTGTTGCGGGTTGCTCCTTTCTGAGTGGACAGTGGACCAACAGCAGAACCCACGAGGAT
TAAAGGAGATGTGGTTGTGAAGATGGAGCAGAGTCAAATCTGTGCTTCAACTGAGATGAGAGTGTAT
TAATCACGTATCGCAGGGCTCCAGCTGTTAGAAGCCACATCATGTTAACATTAACACTGGTTGGATTAA
AAGAACATTAATATTATAACACATATCTTAGGGTAAACAGCTTTTTAAGGTCAAGATTGCT

AGGTTTAGAAAGAGGCTGAGAAATCAAATCTGAACACAATCAACTTACATATTAAAGGAATCTGCCTC
AAATGAGAAAATATGCTAGTTATCTAGATAGAGGAAAGAGATATTACTTTTAAAATTAAAATAGTTA
TGAAATCTGGCAGAAAAGGTAAAGCCTAGAAGAAACTATGAAAGCTATTCTCATGTTACCAATTCTATCT
GCGCATATGTTTGTATAACATTCGGTGCAGTGGGAGTCGGTCCCTTCCAACCTGCAGAGACTAT
CTTCCAATACAGAATCTGCTATTATGCTGTGTTACAAACTGTATTGTTGGGTTGGGTTGGTT
TCTTGGTGGCATTTCAGGTCACTTGCTCTATAACAAAGTAATTGTTTCAAATAATTGTCTTC
CCTTTCTGTATTGTACATAGTGATTCACTTGTGAGAAAAGTGCATTGTCATATTCCAATC
TGTGTTGGTGCCTTGAAGAAAATAAGTTCAAATATTAACTCTAAAAAA

TAFTCVFRCFPDCPSGSASWAAPFPHSAAPMVKMTRSKTFQAYLPSCHRYSIHCRAHLANHDELIS
KSFQGSQGRAYLFNSVNVCGPAAERVLLTGLHAVADYCECKTLGWKYEHAFESSQKYKEGKYIEL
AHMICKDNGWD

PROBABILITY of export to mitochondria: 0.3792 with extension
PROBABILITY of export to mitochondria: 0.4946 w/o extension

AUG initiation = 13.6 kDa
ACG initiation = 17.0 kDa

+++++

#17

NM_020335

AUA initiation conserved in vertebrates

VANGL2 vang-like 2 (van gogh, Drosophila)

Synonyms: LPP1; LTAP; STB1; STBM; STBM1; KIAA1215; MGC119403;
MGC119404; VANGL2

GGATCCGATCTGATTCTGATCCTGATTCTGATCCTGGTCCC GCC ATGGAGCCTGAGCGCCCCCT
ATTCCCCCCTGGCCCCCAGCCCCGGGCCTTGAGGGGAAAGAGGCAGGGCTGGGACGGAGCAGGGGGT
GACCAGACTCAAGAACCCCCCTCAACATCCCCATCGCGCGCTGCCGTCCAGGAGGCCGAGTTCG
GAGCGACCCGGAGCGCTCGGATAAAAGCGACGGGCCAGCGGGGCCGCCGAGGCCACCCGGCAGT
TCGCAGCGCGGGAGCGCTCGCTGGATTCTCTGAGACAAGCCCACCCGCCAGCAAATAGAGTCCCTCA
GGGTGACAGTTGACTTCTGAAAGGTGCCTTGGCTAAAGAACGCCGGTGCTGAAGGAGGTGGCTGTGGG
CCCCCAAGAGGCCAGCTGCGGCCCTGGAGCGCTACAAGGCAGGCCAGACGCCATGGACACCGA
GTCCCAGTACTCGGCTATTCTACAAGTCGGCACTCCCGCAGCTCCCGCAAGCACAGGGACGCCGG
ACCGACACCGCTCAAGAGTCGAGATGGGGCCGAGGGACAAGTCGGTACAATCCAGGCTCCGGGGAG
CCCCTGCTGGACAATGAGTCCACACGAGGGGATGAGCGGGATGACAACCTGGGGAAACGAGCACAGTAGT
AACGGGACCTCAGAGCACAGCATCTCCATGATGACCTCACACGCATGCCAAGGACATGGAGGACAGTG
TCCCTCTGGACTGCTCCCGTCACCTGGGTGTGGCAGCGGGGCCACCCCTGGCACTGCTGTCTTCCTCACG
CCTCTGGCCTTCCTGCTGCTGCCCTCACTGCTGTGGCGGGAGGAGCTGGAGCCTTGGGGACGGCCTGCGA
GGGCCTTCTCATCTGTCGCCCTCAAGCTGCTCATCTGCTACTGGCAGCTGGCTCTGTTCTCCGCC
GGCCAAGGCCCTCGCTGCCCGCTTTGTGCTGCCCTGCTTATGGTGTGGTTTCTGCTCGTG
GTCTCCTACTGGCTTCTATGGTGTGCGCATCTGGATGCTCGGGAGCGCAGCTACCGGGCTGGTGA
GTTCGCCGTGTCGCTGGAGCCCTCTTCTGCACTACCTGCCGTGGCTGCTGGAGCTGCC
AGCTCCAGCCTCAGTTCACGCTCAAGGTGCGCTCCACCGACGGGCCAGCCGCTCTACAACGTTGCC
CATCTCAGCATCCAGCGCTGGCAGTGGAATCTGGAGAAGTATTACCATGACTTCCCTGCTACAACCC
TGCCCTCTCAACCTGCCAAGTCCGCTCTGGCAAGAAAGTGTCTGGCTCAAGGTGATTCCCTCGGAG
AGGAAAACAGCACAACACTCCACTGCCAGTCTCGGGCTGTGATTGAGCGGGCAGCTCGGAGGCCAGG
AACAGTCACAATGAGTACTACTATGAGGAGGCTGAGCATGAGCGAAGGGTGCAGAAGGAGGCCAGG
TGTAGTGGCGGTGGAGGAGGCCCTCACTCACATTAAAGCGGCTGCAGGAAGAGGAGCAGAAAAACCC
AGGTGATGGACCCCCGGAGGCAGCCAAGCCATTTGCATCCATGGCCGTGCCATGCGAGAAGTACCT
CGGACCACCAAGCAGCAGCCCTACCAACCATGGAGAGCATCCTGCAGCACCTGAATTCTGCATCACGCA

TGACATGACGCCAAGGCCTCTGGAGCGATACTTGGCGCTGGACCTACCATCCAGTACCACAAGGAAC
GCTGGCTGCCAACAGTGACATTGGAGCGAGGAGCCGGTACCAACGCCCTAAGGATGGCATCGTT
TTCCTCTAAAAGCCAGGACTTCAGCTGGTGGCAGCACCAAGAAGTCCCATTCTCAAACCTCTCCGA
GGAATTGTGGATCCCAAGTCACACAAGTTGTCATGAGGCTGCAGTCTGAGACCTCAGTG**TG**
ACAGCAGGGGGAGTGGGAAACTCTGGGGGCTCTGAGGGGGTGGGAGGGGGCTTGGTCTCAGGCCAGCC
ACATTCTGCCACCCCTCTTCTTCTGCTCTTTTTACTGAAATTAAACGCACCCCCACCTTCTCTCC
TCGCTTCTCCCTATTTCACCCATGTGAACCTGGAGAGACCATCCTGCTGTCAACAGTACCTGGAGGA
CTCCCACCTCACCAACAACCTTGTATTACTCTAGGCCCTGAGGAATCAGTGCCTCTCCCTCTT
CCCTAGTCTTTCCCAGATTACAGTCTCCTGAAGGGCACAGGGCCCTGCTGATTGTACTTCCCCTCC
TGAGCCCCGACTCACAAATCCAAGTCTAAAACATTCTCTCAGTGGCCAACAGGGTTCTGGGGC
ACATGGACATGACTCCAGAGAGCCACAGTGCACACTCCTCAGGGCAGCAACTGCCCTCCTGCCC
CCCCAGCCACAACAAACCTGGGTTCTAGGGCAGGGATACTCTGCCACACAGGCCAGTTAGAAATCTCC
TTGCTAGGAGCATTGCTTCCACATATTTAGAGCAAAGAAGGATCCATCCTTTCCCAGAAATCTCCA
CCTAATGTTTGTGTTGTATGGTCACGTGACCATAGGCAACCACGTGAAACCCCTCTGTGACC
CAGGGACTTAGGGAAGGTACCTTCTCCAATGTGTCTTCTAGGCAGCCCTGAGGAGGAGGGCTGAAT
AGATCCCTGAGGTTTGGAGAGACCCCCACTGACTCCTGCTCCCTAACCTCACCTCAGTCC
CGCTCTCCCAGTGAAGGATGGTATGTAGACTCCTGTACAGACATAGTGGCTGAGACCTGACCC
CCTGTGGTCTAGACAAATGTTTATTTGTCAACAGCACCCTGCTCTGCCCTCTCGACTCC
AGAGACCTGTTGCCTCATCTTTGGGAAGAGCCGGCAGCTCCTCATCCCTGCCCTAAC
TCTTGCTCAGGGTCTCGTTCTGCCCTCAGGGTCCCCACCCCTTCTCCCTGCTGATTCT
GAGCTCTGGCTCGTGTATTGGGTTGAGGGCAAGGATTACTGCCTTGTAGGTACTC
CCCCATTAGCTCCATAGTCTTGACCAAATCCAATTCTGATAATTAGATCTCATTGAGCAA
ATTGCTGGCCCTTAATAAATATTCAATATACTGAGCCTTGACTCAGACATTGCCAAGGAG
AGTAGAATTAGGAAGTACCCATATACATCCAGCAGGATCCACATGGAGGACCTTCTGATGGCTGCAATG
ACTAGGCCATTCTGAGTAACTCACAGTGTCTTGTAGGCCCTTCTTCCGAAAGACTGGTTGG
TACTTACCTTGAGCAGCACATCCTGGATAAGATCCCAGTGTCTCCCTGGAGGCTCCCC
GCACCAAGCCCTGGGAATGATGGAGCCTAGTGATGGGTTCTCTGCTGCTGCTTCTGCAAAAGTTCACT
TGTACCCACCGCATGCTAGAGAGGAGCTATTGGCAATGCTTACCTTGTCCCCAAGGGTGGTTGT
GGAGCTCACTTAGGCAGGCCCTGGCTGGGAGGGTTATGAGATAGGCCTGTATGAAATATGCTCTGT
TCTGGGGTCTGTCCTTTCTCTCTCAAAAACTTGTGTCAAGAGACTCCCTGTAGTCACATAAATA
CCTCACTATCCTGAAAACAGGGCTGGATGGACTGGGCTATTGCTTGTGGACAGGATGGAGTGTG
GTGTGGTCTGAGGAGCAGGTTGGGTGGGGAGAGGGAAAGGATTGGATCTTAGTGCTGCCCTAGGT
AGGGGCTGGGGAGTGTATTAAAGATCCTGCCATGTTTAATCACTGTGATTGTTCTGCTTCT
TTCCTAAAAAAATTCTCTCAACTCTCTAAGCAACTAAGGCTGTGCTGAGAATGGTAGCATTTG
GTCTTGTCTCAGAACTGTGGTATCTTGCTTTCTCATTATTATTATTATTATTATTAC
TATTGTTTTAAATGTCAGGATGAATTGTCAGACATATGCCATGTTGTCTCTGCTTCT
TGGGAAGTTGTCCTCATGCTGTGAACTGCTGTGGGTGTGAGCTGACTCAGTCCCTGTGAGCAGTTCCC
CACTGTGTCTGCCATCATGCGCTGGATCTGCTATTCTCTGCTGTGGGGTATGCCACCTTACCC
CCTTGACACCATAAGGCTGCTGTGGCTGGCCTCACAGCAGTGTCTTGCTGACTCATGGCATCTCG
TTCATCCCCACCGTGCCTAGCAGGCCCTCCTTCAACCACCTCGGAACGCTTGCTTCTCCCTC
CAGGACGCTGTGCCTCAGTCCTCACCTACCTGCCACTCTGCCACTGTCCCCATTGGCTTCTC
ACTGGTCTTGTGCTCTTTGTTCTATTCCCTTTGTCTCATTCTTCTCCATTCCCC
CCATTTCAGCCCTTAACCTTCTCTTCCATCTCACTCAGTATTCAAATGGCAAACCC
CACCTGCGATGAGACATCGGACTCTCCGAACCTTCTCATCTGACACGTCTTCTCC
CTCCTCCATTGGTCCCAGGCTAACCTCCCTGTGGTGTCTGTCAGTCCTGTCT
CTCTGCCCTTCCCACAGGGCAGTATCTGCTGATGGATTAGCCTGGTGTGATTGTTG
TCCGTGCGAAAAGGAAGAGGGCTTTTGAGTCCCTCAAGTGGAGATTGTAATGAGATT
TTGGATCTAGATTCTTCTTTGGGGGGTGGGGTTACAGAGCTGAGACCTGTG
AGAAAATTGTAATGAAATTCTTCTAATATATAAAAGCTGTTCTACAGTTG
CAGTGGATCTAA
CATTACGGCAATTAGGATTCTTCTAAACATAGGA
ACTAAACTGTACAAATT
TAAAGACATTGACTTTGGGAAAAAAA
IESLRVTVDLKVPLGLKKPVLEVAVGPPKRQPAALERYKARRSDAMTESQYSYKSGHSRSSRKH
RDRDRHRSKSRDGGRGDKSVTIQAPGEPLLDNESTRGDERDDNGETTVVTGTSEHSISHDDLTRIAKD
MEDSVPVLDCSRHLGVAAGATLALLSFLTPLAFLPPLLWREELPCGTACEGLFISVAFKL
LILLLLGSWA
LFFRRPKASLPRVFVLRALLMVLVFLVVSYWLFYGVRILDARERSYQGVVQFA
SLVDALLFVHYLA
VLLRLQLQPQFTLKVVRSTDGASRFYNVGHLSIQRVAVWILEKYYHDFPVYN
PALLNLPKSVLAKKVSGFKV
YSLGEENSTNNSTGQSRAVIAAAARRDN
SHNEYYYEAEHERRVRKRRARLV
VAEEAFTHIKRLQEEQ

KNPREVMDPREAAQAIIFASMARAHQYLRTTKQQPYHTMESILQHLEFCITHDMTPKAFLERYLAAGPTIQ
YHKERWLAKQWTLVSEEPVTNGLKDGIVFLLKRQDFSLVVSTKKVPFFKLSEEFVDPKSHKFVMRLQSETS
V

AUG initiation = 59.7 kDa
AUA initiation = 65.0 kDa

+++++

#18

NM_017457

CUG initiated extension conserved in eutherian mammals

CYTH2 cytohesin 2

Synonyms: ARNO; CTS18; PSCD2; SEC7L; PSCD2L; CTS18.1; Sec7p-L; Sec7p-like;
CYTH2

ATACCTACACCGGCTTTGTACGACTGTTGGCCCTGGAGAACGATCCTTGGTGGCGAGGGCGGGAGGA
CGAAAGCGCCCACGTGGATTGGACAGTGTCAAAAAGAGGGCGGTCCCTACTGAAGGGCGGTGGCGA
CGAAGGGAAGAGTCTTCAAGCGCTGAGGACTGGCGTGAGGAGGCAGGCGGTGGCTCCGGCGTTGAG
CGGGCTACCCGAGCCCAGGGCAACCGGATCCAGGCCACTGGCGGACCGCCCGGATTCCCGCG
GGCCTTCCTAGCCGCATGGAGGACGGCTCATGAACCCCCAGACCTGACTCCGGAGGAGCGGATGGAGC
TGGAGAACATCCGGCGGAGAACGAGGAGCTGCTGGAGATTCAAGGCCCTGGGGAGGAGCTCACTGAA
GCCATGAGCGAGGTGGAGGGCTGGAGGCCAATGAGGGCAGTAAGACCTGCAACGGAACCGGAAGATGGC
AATGGGCAGGAAGAAGTTCAACATGGACCCCAAGAAGGGGATCCAGTTCTGGTGGAGAATGAACGTG
AGAACACACCCGAGGAGATGCCCGCTCCGTACAAGGGCGAGGGGCTGAACAAGACAGCCATGGGAC
TACCTGGGGAGAGGGAAAGAAGTGAACCTGGCAGTGCTCCATGCTTTGTGGATCTGCATGAGTTACCGA
CCTCAATCTGGTGAGGCCCTCAGGCAGTTCTGGGAGGGCTACCCGGAGAGGCCAGAAATTG
ACCGGATGATGGAGGCCTTCGCCCAGCGATACTGCCTGTGCAACCCTGGGTTTCCAGTCCACAGACAG
TGCTATGTGCTGTCCTCGCCGTACATGCTAACACCAGTCTCCACAATCCAATGTCGGGACAAGCC
GGCCTGGAGCGCTTGTGGCATGAACCGGGCATCAACGAGGGCGGGACCTGCCGTAGGGAGCTGCTA
GGAACCTGTACGACAGCATCCGAAATGAGCCCTCAAGATTCTGAGGATGACGGAATGACCTGACCCAC
ACCTTCTCAACCCGGACCGGGAGGGCTGGCTCTGAAGCTGGAGGGGGCCGGGTGAAGACGTGGAAGCG
GCGCTGGTTATCCTCACAGACAACACTGCCCTACTACTTGTAGTACACCAAGGACAAGGAGCCCGAGGAA
TCATCCCCCTGGAGAATCTGAGCATCCGAGAGGTGGACGACCCCGGAAACCGAACCTGTTGAACATTAC
ATCCCCAACACAAGGGCAGCTCATCAAAGCTGAAACTGAGGCGACGGCCAGTGGTGGAGGGAAA
CCACATGGTGTACGGATCTGGCCCCACGCAGGAGGAAGGACGAGTGGATCAAGTCCATCCAGGCC
CTGTGAGTGTGGAGCCCTCTATGAGATGCTGGCAGCGAGAAAGAGGGATTCAGTCAAGAAGAAC
GAGCAGCCCTGACCCCTGCCCTAACCTCCATTATTACCGAGCTGCCCGCTGGTGGCCGGACCC
CTGGGCCCTGGGGCTGTGGATCCTGGTCCCTGTTGGAAAATTCAACCACCTCTAGCTCCTACTGTTCT
TGTAATTAAACACGCTGTTGTAATCTTATTAAACCACTGGCCTGCTGACCCCTCATTTCTG
GGGTTGACAGAGTCGAGGTGCTCCGTGGAGCCAGCCTGTTCCCTGGACAGGGGCCCTGGACCCGCTGTCT
CTGGGTGCTGCCCTGGCTGTCCTGGTGGCTGTTCTGGTTCACCCGAGCCCAGCAGGAGTGGAGTAA
GGGAGAAGGTTAATATGGTGGAGTCTGGAGTTGAATTGGCGGGGACAGAGTTAGAATGCAGGGATT
AGGGTCAAGGTCTAGCATTCTAGAACAGTTACTTGACAGCCATCAGCCAGTACGCTACGCCATCAGC
TAGTTCCCATCGTTGCTTCATGGGCTTGAGGTCTTGAGGGCAGGAGATGAAGCTGGTGGAGTGTGA
GCTGGGCCAGGGCTTGAGGACAACCTGGAGCTGGAAGAACATGCGACCCACTCAGGGAGGGTCAGGGAAAG
GATGAGTGGGGAGGTGGCCCATGTCCTGAAGGGCTTGCTGATGGGATGTCCTGAAGGGCTGGCAGCCT
TAGATCGGGCTAACAGGGCAGGACTGTGGCCAGTGCCAGAGCCAGGCTTGTCTCAAAGGATCAGC
CTCCTTGGAGGACATTGTGTCAGGATGCGCTCATGAAGGTGAAGTCTGGCATGTTGGTACCCCTGAATC
TTGACAGCCAAGAAAATTGAGGATGCGCTCATGAAGGTGAAGTCTGGCATGTTGGTACCCCTGAATC
CTGGGCTAACGGTCTGTGACCTTTGCCCTCCAGTTGGGTATGTTGCTGAGGCTTCAAGGATGGCTCG
TGGATCCGAGCAGGGTGCAGTGGCGGATTGGGAGGTCCCATGTCACTCTCCATGCCCGCTTTGAAGCT
GAGGCGCTTTAGTTAACAAACTACAAGTCCCAGCAGGGACCGGGACCGCGGGTGGGAGAGGCCCTGTGG
CCCCGAGGCGTGCCCGAGTTGTAGTCCTCCCTGCCGCTGTTGCTGAGGCTCTCCGATGGGATG
GGCGCTCGGAATTTCGGGCTTGATCCCTGTCCTGCCACAGGCACCTGCCGGCTGAAGGCC

CCGCGGTGGGGTACCTCGCCCCCTCGCGGAAGGTGGACTACAGTTATCGGCAGGCTGTGCGCGCCA
AAGCCACGGTGACCCAGACCCGAGGTTTCCGGCGTCGAGTTCCGAGACTCCGTGGCGCGTTGT
CTTCTTTCTTAGTCAGATCCCGTACTTTGTGGAGGGTAGAGGAGGCTTGACCGCCGCCCCGGGG
GCTGGTGGAAATGGAGTCCAATGAGAAAATAGAATTCCCACTTCTCTTCCACAGGTGCCAGGGAA
ACGGAGTCATCGACCCAAGAAGTCGTGGAGATGAGGTCCAGGGTAAACAGCGGTCCCACACTATGT
CACCCCTTCCTCGCCCTCCCCGGATGAACCTGCATGCAGGGCGCCGGCTCCGTGGCAGGCAGGGCAGGA
AGAGGCGCGAGCCCGCTGGCGAAGAGGGAGTGCAGGGGGAGGGCAGAGGCAGGCAGGCCGGT
CGCGTGGGGCCTGGCCGCCAGGAGGGCTCTGGCTGGATTCTAGCATGGAAGCCGTCAAGGGCA
GGAGGCAGGGCCTGACGTGTTGGATTGAGGTGCAGGAGGGCCCTGGCTGCTTCAGGGAGAATAATT
TGGAGGCAGCGGGAGGTGAGAGGAAGCTGTGCGGTGCAGGATGGCGTGGCTGGAGGTGGAGAACGTT
GGGCCACGCTAGGAGTCTGAAAGAGGAGCCAAAGGACTTGCTGAGAGCCGCGCAGCTGGTAATGACG
GTGCCATGACTGAGGTGGCTGAAAGGGAGTATTAAGGGAAATCAGGCATTCCGTTGACCATGCCACAGTC
CATCTTCCCCCAGGACCTGAGGATTTGCGTCCGCTCCCTCCTCGCCAGGACCCCCAAGCTCCAGCAC
GCTTCTGATTTTTTGAGTTTTTTTTGTTTGTGTTTGTGTTTGTGTTTGAGAGGGAGTC
TCACTTGTGCCCTAGACTGGAGTGCAATGGGCCATCTGCTTACTGCAACCTCACCTCCAGGTGCA
AGCATTCCCCCTGCCAGCTCCGAGTAGCTGGGATTACAGATGTGAGCCACCGTACCCAGCTAATT
TGTATTTTAGAGACGGGTTTACCATATTGCCAAGCTGGCTCGAACCTCTGACCTCAGGCATC
CATCTGCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCGGCCAGGCCACTTCT
TTGGGGTGGCAATGGTTGGATATCGTTGCTCTACTAAAATTCATGTTGAGATTGAGCCCCAGTGTGG
CAGGTGTTGGGATGTGGGCTCATAAAGGAGCTGTTGGCTTGAGGTGGAGTCCTCATGGATAGATTAA
TGCCTGCCCTAAGGGGTGAGTGAGTGCACACAGCTGCATCCTCTGCCCTCCCCAGAACGAGCAGCAGC
CTTCAGCACCAGTGATCTGTCACAGCTGCATCCTCTGCCCTCCCCAGAACGAGCAGCAGC
AGGGATGAGGGAGGTCTCAGTGAGTGACCTGCCAGTCTGAACCTCCAGGCATCAGAATCTGAGCCA
AATCAACCTCTCCTTATAAAGACCCAGCCTCAGGTCTGTCAGAGCAACACAAAATGGACTCAGCAC
GGATCAAATTGTGCTCCCCCACAAAAACTTATATTAAAATCCTAACCCCCAGCA

LRTGAEAAVAPGAFERAHPSPRANADPGPTGGTAPDSRAFLAAMEDGVYEPPDLTPEERMELENIRRK
QUELLVEIQRLREELSEAMSEVEGLEANEWSKTLQRNRKMAMGRKKFNMDPKKGIQFLVENELLQNTPEEIA
RFLYKGEGLNKTAGDYLGEREELNLAVLHAFVDLHEFTDLNLVQALRQFLWSFRLPGEAQKIDRMMEAFA
QRYCLCNPGVFQSTDTCYVLFAVIMLNTSLHNPVRDKPGLERFVAMNRGINEGGDLPEELLRNLYDSIR
NEPFKIPEDDGNDLHTFFNPDRGWLLKLGGGRVKTWKRWFILTDNCLYYFEYTDKEPRGIIPLENLS
IREVDDPRKPNCFELYIPNNKGQLIKACKTEADGRVVEGNHMVYRISAPTQEEKDEWIKSIAAVSVDPFY
EMLAARKKRISVKKQEQP*

AUG initiation = 46.5 kDa
CUG initiation = 51.0 kDa

+++++

#18

NM_001010908

Non-AUG extension conserved in vertebrates

C1QL3 complement component 1, q subcomponent-like 3

Synonyms: C1ql; K100; C1QTNF13; C1QL3

TCAGAGTAATGCCAAACTCTCTGAGTGGGATGAGCAGAGCAGATGCTGCAATGAGATGCCAAAGCGGCT
CCCCTCCTCTGTGCCCTGGGTGCCTATAAAATTGCTCCGGCGCGCTTGTGAGCCTCTCTGCC
AGGTGGTACCCAGGCAGAATTCTGCCCTCAGTCTCTCTCGCTCCGCTCCGGCGTGAGGCCTCGCC
CTGCTCGCTCGCTCTCCGCCCCAGCTCTGAGCCTCGCCGACCGTGCCGCCGCCGCCGCTGG
GCGCACCCGGGGACGCCGGCCACCGGGGCTTGGGGTGCCTATTCGAGTTGTGGTGGCA
AAGGAGGAAGAGAGAAAGGAGGCAATAAAAAAAAAAAGGCAGGGACGGCGAAGTGGCAGGCAACT
AGGAGGAGGAGGCAAAAAAGGCAGACTTCAGACGGAAAGTTGGTGCAGACAGGCAGTCTGCA
AAAAGTCGATCGCAGGCAGGCCGGCGCATAAAAGTGTGAGCTGCCGGAGCTCAGGAGGCCGGCTGG
CTCTGCCCTCCGGTGGCCGCCGGCTGAGCCACAGGTGCGAAGGAGGCTCGCGGGGGCGAGG

GCGCCCCGCGCACCCCTCCCCGGCCCCACCCGGGCTGGACTTCGGCTCAAGTCACTGGCGCCCGC
 GCTCCCTCCCAGCGCAGCCTCCGCGGGGGAGCAGGAGTCGGCAGCAGCGGCCGCACCGCACGCCA
 GGGGAGTTGGGTTCGCAGGGGTTGTTTCGGCTCTGAAGAGTCCCCGCCAACCTTCAAAATTCTGTC
 CAAAAGCAGACAAGAGGATGCCCGCGCTGAGCGGCTGGTGGGCAGCAGGAGGCGCTGATCGCCGCC
 GGGCGCTGGGGTGGTGATGGTGTGCTGCTGGTATCCTCATCCGGTGTGGTGGCTGAGCTGGCCGGCACG
 TCGGCGCACTACGAGATGCTGGCACCTGCCGATGGTCTGGCAGCCCTACGGGGCACCAAGGCGCCAG
 CACCGCTGCCACGCCGACCGCGGCCATGCAGTCCCTGCCACCTTCATCCAGGGCCCAAAGGCGAGG
 CGCGAGGCCGGAAAGCGGGTCCGCCGGGCCCCCGAGAGCCGGCCACCCGCCATGGGCC
 CCGGGCAGAAGGGCAGCCGGCCAAGGCCCTGCCGGGCCGCCGGCCTGAACGCC
 CGGGCCATCAGGCCGCCACCTACAGCACGGTCCAAAGATCGCCTTACGCCGCCCTAAGCGGCAGC
 ATGAAGGCTACGAGGTGCTCAAGTTCGACGACGGTACCAACCTCGAAACACTACGACCCACC
 GGCAAGTTCACCTGCTCCATCCGGCATCTACTTCCACCTACCGCCTGATGCGGGAGGGAGGG
 CACCAAGCATGTGGCTGATCTGCAAAAACAACCAGGTGCGTGTAGTCAATTGCCAAGATGCTGATC
 AGAATTACGACTATGCCAGTAACAGTGTGGTTCTCATTGGAGCCGGAGATGAAGTCTATATCAAATT
 GATGGGGAAAGCCATGGAGGAAACAACAAACAGCACGTTTCTGGATTATTATTTATGCTGA
 CTGATAATGCAGAAACTAACGTTATTATCTGAGTTGAACACTGGATTGCTATGGCTAACGTCAGTGA
 CAAGGATCCCAGGGATGCCATGGCAGGGCACCTCAGTTGTATATGTGGGAAATCAAATGCTACCTG
 ACTCACATCTGTATCACTCAGAAACATTATGAAAAAAATATCAAAGCAAGATAAGCAGATGTGTGATCCAC
 TACGCCAAAGCAAATACTCCTTATCGTAGTGTCCATGTGAATGAAGTCTATATAGATCACAATT
 ATAGACAAATCTAACAGACATTGAATTATTCTCTATATATGATACTTGGTGTACTGTGATCTGCTGC
 TTTTATCCATATGTCAGCTTGGTTCTGTGAGTTACCTGCTTATTATGATACTGGAGTCATTCTAG
 TGTGGGAAGAATGATTTTGCCTGCAGGAGAAGGTCTAATTGAAATAATGCTGTTGCCCCAAAGAAA
 TTGTTGCCTGTACTCTGTTAACCTAGAGCTAGACCTGGAAATGATTCAACTCAAGCCTAACCTGG
 AATTTCTGGATTGAGGGATTCCCAAGCTATGATCTTTACATTCTTTCTTATATGAATT
 CTTTCTCTCTTCTGGTGTAGTCTTAAAAATAGAGATTGACAGTGTATCATAGGATTACCTCTAAC
 TGTGAAAATGTGTAATTATGACGGTATTAGAAAAAGCCCTATGTGTCCATTGTCACACTGAATACAT
 TTAGATT

ILSKSRQEDRPALSRLVGSRRRLIAAGALGVVMLLLVLIPVLVSSAGTSAHYEMLGTCRMCDPYGGTK
 APSTAATPDRGLMQSLPTFIQGPKEAGRPGKAGRPGPPGEPPGPMPPGEKGEPEGRQGLPGPPGAPGL
 NAAGAISAATYSTVPKIAFYAGLKRQHEGYEVLFDDVVTNLGNHYDPTTGKFTCSIPGIYFFTYHVLMRG
 GDGTSMWADLCKNNQVRASAIAQDADQNYDYASNSVVLHLEPGDEVYIKLDGGKAHGNNNKYSTFSGFII
 YAD

AUG initiation = 26.7 kDa
 AUU initiation = 30.1 kDa

+++++

#20

NM_001008223

Non-AUG extension conserved in vertebrates

C1QL4 complement component 1, q subcomponent-like 4

Synonyms: C1QTNF11; MGC131708; C1QL4

ATACCACCCGCGCCGGAGGGAGGGAGGAAGGTTAGGGAGGGCGGAGAGGGACCGCGCCCGCAGGAGAGA
 GGCAGGGCCAGGGCTCTAGCAGGGACTGGGCCGGCAGGGTAGCAAGGTGAGTCGGTGCTTGCCTAAG
 AGGCAGAGCGCAAACCTACTAGGAGATCGCGCCCGGTGAGCAGCACCCGAGCTCAGAGCCGGACGTC
 CGGAGCGGGAGCAGTCCCTCTCCATCAGGGAGTGGCTATCTGGCAGTCTGGACCCAGGCACCGC
 GCCATCCTGAGAGAGCAGCAGTCTGGAGAGCAGGCATCTCAGATCCCTAAGAAACCAGCCGAGAAG
 CCGCGGATCTCAGGTGCCAGGATCGTTAGGACTGAACGGAGGGTACTAGAGGACCACTGGCTCTGGACC
 GTCGGGAGCTGCCCTGACGTAACCCACGAGGGGCTCCCTGACGGACGGCTGGGAGCGGCCACCGCC
 GCGGCTGGAGCCCGCAGAGGGTAAGGGAGCGGGGGCAGCCGTGGGGAGTGCAGACCCAGGCC
 AAGCGGGTCACCGCCTCTGGCCCGGGAGAGCCCGCCGGCAGCCATTGCGCCAAGAGTGTAGGAA
 GATTTGCTGCCCTGGCAGCGTCGGCTGAGCCCGCAAGAGGGTGGCGGGCGCCGTCGGAGTGGC

CATGGTGTGCTGCTGCTGGCCATCCCGCTGCTGGTGACAGCTCCGCCAGCGACTACGAGA
TGCTGGTCGCTGCCGCATGGTGTGCACCGCATGGGCCCCGTGGCCCTGGTCCCACGGCGCCTGCT
TCCGTCCCCCTTCCCGCCAGCGCAAGGGAGAGTGGGCCGGCGGGAAAGCAGGCTGCGGGGG
CCCTGGACCACCAGGTCAAGAGGGCCCCCAGGAGAACCGGCAGGCCAGGCCGGGGGG
CAGGTCCGGGGGGGGGGGGCTGCCGGCTACGTGCCTCGATTGCTTCTACGCCGGCTGCGGG
CCCCACGAGGGTTACGAGGTGCTGCCTCGACGACGTGGTACCAACGTGGCAACGCC
CAGCGGAAGTAACTTACTTGCCCCATGCCAGCGTCACTTCTCGTTACCGTGCCTACGCC
ACGGCACCGCATGTGGGCCACCTCATGAAGAACGGACAGGTCCGGGCCAGGCCATTGCT
GAGGACGCG
GACCAGAACTACGACTACGCCAGCAACAGCGTATTCTGCACCTGGACGTGGG
GACGAGGTCTTAC
GCTGGACGGGGAAAAGTGCACGGCGAACACCAACAAGTACAGCACCTCTCCGGCTT
CATCTACCC
CCGACTGA
GCCGGCCCCGGGGCTGCCGGCTGCCCTCTCCCGCTCACCACCTCTGCC
CCCACCGAGGCCACCCACCCACCTTGAGAGCCTGGCGTGGGGTGGACCC
CTCGTTCCGGAGGCC
CTAAATGGGCGAACCTTGCTCAAGGGTATAAGTGGCCGGAAAGAGGAGGCC
AGAGCAGGCC
AGAGCGACTTCCGGAGGGATACCCCGCACCAAGTGC
CGCTGGACCC
CATAGGGCAGAGGACTAGGGAGCAGTTTAATAGCGGGACTCAGAGGCC
CCCGCAGCTTGC
GAGGGAAATAACAGAACAGGAGGCC
ATTAGGAAGAGAACATTAAACAGGAGGCC
TAGTGCAGGTTCTCGTCACAAC
TTCTCGCCACCC
CTCGTCC
CTCGTCTCCACTTCAGGCTCAG
GAGGAAACCAGTGAATTCTTCTGGCATT
AAACGCATT
TGTACAGTCCCCATT
CCCCCTATCCGGACTAGGCC
CTGGGCTACAGCTGCTGCTGCC
CTTCTAATAAA
GTGAGG

IAPKSEEDLLALAASRLSRRKRVAGAAVGVAMVLLLVIAIPLLVHSSRGPAHYEMLGRCRMVCDPHGPRGP
GPDGAPASVPPFPPGAKGEVRRKGAGLRGPPGPPGPPGEPRGPPGPPGPGVAPAAGYVPRIA
FYAGLRRPHEGYEVLRFDDVTVN VGNAYEAASGKFTCPMPGVFFAYHLMRGDGTSMWADLMKNGQVRA
SAIAQDADQNYDYASNSVILHLDVGDEVFIKLDGGKVHGGNTNKYSTFSGFIYPD

AUG initiation = 24.9 kDa
AUU initiation = 28.0 kDa

+++++

#21

NM_001002914

Conserved all the way to fish where the extension is initiated by AUG

KCTD11 potassium channel tetramerisation domain containing 11

Synonyms: REN; C17orf36; MGC129844; REN/KCTD11; KCTD11

ATTAGAGGCTCCAGCCCCGCCACTTGCGAGACGTGAGATCGGGCACACCTGAGCGCGGCCGGGGCGGTCTG
GGCCACATCCGGGGCGACGTGCCTGAGTCACCCCGTCCCGCCAGCGTCTGCCAGTCCAGCCAGTCCGCC
GTCTCTCGCGTCCGAGACTCGCCTCCAGCCTCCACCTCCGCCGGCGCGAGCCTCGCGGGGGGGGG
GGCGGGCGCCAAGGGGCGGGCTGCTCTTAAAGGGCCCGGCGCTGCCCTAGGCCACTTCTGGGG
GCGGAGAGGACCTCAGCGGCTCGGCCACACCCAGGGAAAGCGGCCGCCGGTCCGAAACTCCTGGCT
GTTTCCATCAGAGCCCTCGGACACTCCAGCCCAGGCTGAGCACGCATCGCCTCCGGGATACAAG
GGGGCTCCGCCATCGCTCCCGTCAGTCGGCTCCATCTCTGGGACCCGCCGGCAGCCAGGCCAGGC
CTCTGAGTGGCCCCAGGCCCTGGCTGGACTCGTCCACGGCCAGCGATCTGCCGGGTCTGGAGGCC
ATCCCTTCAGAGTCGGCCCTGTGCTGCCACCGTCACCTGCTGGTGGATTCCGAAACCCACTGCTGAA
GACCACAGAGGGGTGCGCTGACCAACCCAAATCGGATACGTCAGACCTCAAGCTCCCTCCCTCTG
GCTGCCCTCTGCTCTTTCATCTCTCAACCTTTGGGATTCTGTGCTCTGACACCACCTCCCA
TCCACCACCAAAGTAGCGGGGTGAGCCCCAACCTTA
CTGGGTGTGCTCCACCTGTGCCCTCCAACCCAGC
GAATCTGACAGCTCGACCCAATTCTGCACACACCCAGGAAGTTCTGCTCTTCTTCTCGGTGCTC
CTGTACTTCCAAAATTCTCTCTCTGTGCCCTCTGCCCTCTTGGGGCCCGTGACCC
AATGTGGGGGGCACACTATATCCACCACTTGGAGACCC
TGACCCGCTTCCAGACTCTATGCTGGGG
CATGTTAGGGCCGGCACCCCATGCCCAACCTCAATTCC
AAGGAGGCCACTACTTCATCGACC
GGGATGGCAAGGCC
CTCCGGCACATCTCAATT
CTGAGGTGGGCCCTGGACCTGCC
CTGGGCTGGACCTGCC
GGAGAGACAGCGCTGCTCAGGGCAGAGGCTGACT
TCTACCA
GAGATCCGCC
CTCTGGACGCC
CTGCCGCTGCCGG

ACTGGAGGCCTCTCAGGGGACCCCTGCACCCACAGCTGCCCTGCTCCACGCAGATGTAGATGTCAGCCCC
GCCTGGTGCACTTCTCTGCTCGCCGGGACCCCCTCACTATGAGCTGAGCTCCCGCAGGTGGACACCTTC
CGAGCCAACCTTTCTGCACCGACTCTGAGTGTCTAGGTGCTTGCGGGCCGATTGGTGTGGCCAGTGG
GGATAGGGCAGAGGGAGCCACATTTCATCTGGAGTGGGGCCAGGAGAGCAGGGAGGTGGTGGCACC
AGTATGGGAGACTGGGCTGCAGCCGCTGTGGACTGGGGCCAGGAGAGCAGGGAGGTGGTGGCACC
CCAAGCTCCTGGAGGAGGTGCTGCGGGTGGCTCTCGAGCACGGCTCCGACTAGACTCTGCTTCCCCGA
CCCCGAAGACCTGCTCAACTCCAGGTCTCGCCTTGTCCGGACTGAGGATGCTGTTCTCAGTTGACT
GTGGGGAGGAGAGAATGGGTACTAGCACCCCTGAAGCCTCTTCAGCTCTGCTTCAAGGAGCTATGAG
AGTCGGACTCTCTGCACCTGACTGGAGCTCAGATGTGGCAGGAATTCCAAACCTGAGGCCACCAAGG
ACTCACAAGTGGTCCAGAAGGTCTCAACCTGTGCTGACCCCTGGAGGGTAGGGAGGTTCTCAGCTTG
TTCTGCCTAAGGCTGAGCACCTCAGTCTCCTTGATTTGGAGCTCAGTGTGTTAAGGGCTTGGAAAAGG
GGGGAAACATCTCTTACCCAGACTAGACCTAGCAAACCTGGAAAGGATATTGAGGTCTGGGAAAAGG
GGACTTTGCATTTCCAATGCGGTCTTGGACCATGGCTCTACTCCTGAAGCTGGTGGCCTGGCCTG
GCCTGACCAATGAGAGGCCAGAACACTCTGGAACATCGGAAGAGGAGTCTTGCTATGTTCCAAGCCATC
TACTGAGGGAGGCAGAAAGGCCACAACCCACCTAGGTTGATGTATGGGAGCTAGGACAGTCCCCATGGCA
ATGGGGCTGGAGCATCCCTCATCTGGAAGAATCCATACTGATGGCAGGGCTGGCAGGGGGAAAGAGGGTA
GTATCTGTGGCTCTGGCCTTCTTCATGTGCGTGCATATCAGCCCGTGTGGCTGACTGATGTATAGGT
CCCTGGCATCCTGGTTCATATCTGTGTTGCTGACTACAGTGTCTGTGATGTCCGATGTCCAGGCCTGTT
GGGGTTGCCTAGCGACTCTTCTGGCACAGGGTGTGCTGTGTTACCTGTGAGGTGGTTGACAATTAGTA
GTTTAATCACAGGGTGTGTTGTTGTTGAGTCACTGGGATCTTCTGGTGTGAGGCAACCAGGACTGTTGCTAGGAG
GTAGCCAGGAGGGCCTGTTGGGTTGAGTCAGTGGGATCTTCTGGTGTGAGGCAACCAGGACTGTTGCTAGGAG
GCTTAGCTGGCCTCTGAGGCTGTATGGAACTCTTGGTGTGAGGCAACCAGGACTGTTGCTAGGAG
ATAGCTGGGAAGGCCAAGCCGCCAGGGCAGAGAGAGGAGACGAAGAGTTGGGACAGTGGGGAGGA
GATGGGAAGGGATGGGATTCTGGGTCCCAGAGCGGGTGGGAACTCACGCACAGCTTCACTGGTGG
GGGTGGGGCACACATTATTCACGGTCAAGAAGAAAATAAACTGCTTTGGAACCAC
AAAAAAAAAAAAAAAAAAAAAA

ISPPPVPSPPSF GGPTLVGGTLYSTTLETLTFPDSMLGAMFRAGTPMPPNLNSQGGGHYFIDRDGKA
FRHILNFLRLGRLDLPRGYGETALLRAEADFYQIRPLLDALRELEASQGTPAPTAALLHADVSPRLVHF
SARRGPHHYELSSVQVDTFRANLFCTDSECLGALARFGVASGDRAEGPSFHLEAPRVPVELPEVEYGR
GLQPLWTGGPGERREVVGTPSFLEEVLRVALEHGRFLDSVFPDPEDLLNSRSLRFVRH

AUG initiation = 25.9 kDa
AUU initiation = 29.9 kDa

+++++

#22

NM_025160

Two isoforms but both have the 5' extension

ACG and CUG extensions conserved from human to fish

WDR26 WD repeat domain 26

Synonyms: CDW2; MIP2; FLJ21016; WDR26

CCCCGCCCCCTCCCCGGCCCCCGCCTCCCCCTCCCCCGCCTGCCCAACGGCTTCCCACCACGGCTCT
CGGCGAGGAAACTCTGGCCTCCGCTTCTCCTCCCTCCGACTCGGACACCGGGAGCCTCCCCGGCCCCCG
GGAAGAAACCCCGAGCCTCGCGGGAGGGAGTAGGAGAGGCCGGGCTCGGCAGGCAGAGCAGGCC
TCCCCCTCGTCCCTCGTCGTCGTCCTCCCTCCCTCCCTCCCTCCGTTGGTAGTGGTGGACTTCC
CCCCGCTGCTGCCCTCCCGCCGCGCTGCTGCCCCACCGAAGTAGCGGCCACAGCCTGGTCAGCGGA
GCATCAGCAGGCAATGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGGGG
GGCGGGGGCCAGGGACAGACCCCGGAACTCGCCTGCTTGTGGCCACAGGGAGTCGTCCCCCTCGTC
GTCGTCGTCGCCGGGGACCTGGCCACGCCAATGGGCTCCCTGCCCTCCGCCGCCAGCAACA
ATAGCAACAGCCTGAATGTCAATAACGGGTTCCCGGGCGGCCGCCATCCTCAGCCACCGTCGCA
GTCGCTCCGCCACCACGCCGCCCTCCCTTGGCCACCCAGAACTGGCAGCAGCCTCAAGAAGAA
GAAGCGGCTCTCCAGTCAGATGAGGATGTCAATTAGGCTAATAGGACAGCACTGAATGGCTTAGGGCTCA

ACCAGACTGTTGATCTCCTCATGCAAGAGTCAGGATGTCGTTAGAACATCCTCTGCTACCAAATTCCGA
AATCATGTATGGAAAGGAGACTGGGATAAGGCAGAAAATGACCTGAATGAACCTAAAGCCTTACTGCATT
TCCTCATGCTATTGTTGGTAAGAGGCGCACTTGAATCTCTCAAACGTTGTTGGGATAATTGTGAGGATGA
AGTTTGCTGCTGCAGCAGAAGTACCTAGAATAACCTGGAGGATGGCAAGGCTGGAGGACTTCAAGTT
CTACGCTGTGAATTGACGCCGTGAAATACAATACAGAGCGATTATGTTCTTAGTGGGATCTGATGT
TAGCCATGCAGAAGACCTACGTGCAAAAGCAGAATGGGAAGGCAAAGGGACAGCTTCCCAGTAAACTAT
TGGATAAACTTCAGACCTATTACCAACCATCAGTGTGCTTCCCCACGGCGTTACAGACTCTCGCG
CAGGCGTGGAACTACAAAGGGATCGGTGCCTATATCACAATACCAAACTGATAATAATCTAGATTCTG
GTCTCTGTTATAGACCATGTTGTTAGTAGGAGGAGCTTCCATGTTATACGCAGCAGATACTTACGGAGC
ATTGTAATGAAGTGTGGTCTGTAATTCTTAATGATGGCACTAAACTAGCAACAGGATCAAAGATA
ACAGTTATCATATGGCAAGTGTGATCCGGATACACACCTGCTAAACATTGATAATAATCTAGATTCTG
TTATGGCGTTCTTATATTGCATGGAGTCCAGATGACAACATCTGTTGTTGTTGGCCAGATGACTGCT
CTGAGCTTGGCTTGGAAATGTACAAACAGGAGAACTAAGGACAAAATGAGCCAGTCTCATGAAGACAGT
TTGACAAGTGTGGCTTGGAAATCCAGATGGGAAGCGCTTGTGACTGGAGGTCAAGCGTGGCAGTTCTATCA
GTGTGACTTAGATGGTAATCTCCTGACTCCTGGAAAGGGGTAAGAGTGCATGCCATTGGTGTGAGTG
ATGGAAAGACTGTTCTGGCATCAGATAACACACCAGCGAATTGGGGCTATAACTTCGAGGACCTTACAGAT
AGGAACATAGTACAAGAAGATCATCCTATTATGTTTACTATTCAAAAATGCCGATTAGCTTGT
AAATGTAGCAACTCAGGGAGTTCATTTATGGGACTTGCAGACAGAGTTTAGTAAGAAAGTATCAAGGTG
TTACACAAGGGTTTATACAATTCTCATGTTGGAGGCCATAATGAAGACTTCATCGCTAGTGGCAGT
GAAGATCACAAGGTTTACATCTGGCACAAACGTAGTGAACGCCATTGCGGAGCTGACAGGGCACACACG
TACAGTAAACTGTGTGAGCTGGACCCACAGATCCATGATGCCAGCGCCTCAGATGATGGCAGT
TTAGAATATGGGACCAGCACCTTATAGACCAACAGAATTGAAAGAGGAATGCAGTAGCATGGATAGT
TGA~~TGG~~GAATTGGAGCAGACGACTCTGTTAACTTAAATTAGTCGTATTTAATGGCTGGATTG
GTGCAACAAACATGATTGATAGCTGGACAGACATGCTCGTATGAAAAAGAACATTCTGAAGCCCGA
TTGGGCCAAACATTACACCTGCTCATAGTAACAGTTGAGATGAAGCACGTCGTTAGAACGTTG
GACACCATGTTGAATTATTCCCCATCGGTGTGAAGAACTGTGCTACATTCCAGGCTTACCCATTGAAC
AGTATATATATTTTCTCCTGTTGTCAGGATACCATTCTGTTGCTCTGTGTAATG
AAGTTAAATGCTGTTGGAAACTTATTAAACAGTTAGAAGGCTTGATAGAAAGAGTGATTAGTCT
GAAGAGTATACATTGGATAGGAAAGAATTCTCTTTGTTCTCAAATCTTCCGCTTATTTAGCTT
GAGATCTTGCAGCTGGTCATGGATTCTAGCCTGCCGTTGCGAGTATATACTGATCCAGATGATA
ACCAGTGAACATATGCAAAAGCACTCTCAATATTACATTGACAAAAGTTGTACTTTCACATAGCTT
GTTGCCCGTAAAGGGTAACAGCACAATTAAAATAAGAAGTATTAGGATTAAAGT
ACTTCATTGATACATTGGAATCTAAACCAGCTTAAACAGCTTCTCAATGACTTAGATACAGTA
TAACGTGCTCTGGAAATACCACATGAGACATGGTCAGAAACAGTGCTGGAGGACATTACACAAGA
AATTCAAGAGTAATGCTTGAAGATTCCCCCTTTGTTTATTCTGAAGGAACATCAGTACCCGATCTT
GAAGAAATTCAAGATTCAAAAGAATTAAATACCAACATGAGACATCAGTAGTCAGTTGGTTTCAG
TAAAGCTTGTCCAAGTTGTTCTCAACTTAGGAAGTAATTGGTGTGATCTAGCAAAGAGTAGGA
GCGATACAACCACTTGGAAAGTTAGTATAATTGAAATTAGAAGAATTCAAGGTTACAGACATA
CTTAAACTGGGATTAAACCTCATAGTCATTCTTAATTGCCCTTAATATTGACATATAAGGATA
TAAATTAAAGAATATTCTCAGTTTCTCAGATTTCAGATATTGCCATACTGAAACCTCATTCTAA
ACTGGTCTTCCCTCTCCCTCTGTTAGTTAAGGAAAGGTTCTCATGGAAAAGCAGTT
ATTTGAGAATTTCACATTAAAGCCAAAGCCTGATGTTCACATTGTCAGTTGCTTTAACTCCGTA
GCTGAATACAACATAACATAGCTCCCTGAAATTCCCTCCCTTCAGTAGAGGATAAATAGGGTACT
CCAGCTGTTGGAAATGAGAATGGGATAACCCCAAATAAGATTCTGGTAAGAAAAGTCTGACTAAGT
TCCCCTTAGGTGTAGCCAAGTCTATAATATAATTGATCATGAATAGACTCCAGAACAAAGAGAA
GAGGGCTTGGATGTCAAAATTCTCCATTGTTGGGAGGGATGGAGAAGGGAAAAGTCAG
GGAAGGGAGAAGAAGAAAACACATTGGCCTTAAATAGCATGCATTCTCCAGAACCCAGATG
GAAAATGACAGATCGTGGTCAGCAGACTCTTATCAAATTGGTATGTAAGGATTTCTCAAGCT
GCAGAGACCACCACTTAGAGAATCAAAATTCTGTTGATACTCTAAGTTAAATATGTTACAGT
TTATCTGGTACTTCATTCTTAACATTACTTTACTTAAAGCTGAAATAAAATCTTCA
AACTGTATGTAATTAAAGTGGCAACTTCTCTTACCTCAGTATAGTTAAACATT
TCATGTATACCTGTAAGGCCGTTAACACACCCAGTCTATGGTAACATCA
TCCACAGCTGTTACAAATTGGTTACAGTGTGCTACGGTCAATTGTAACACT
AGGGCAATTGGCACTCAGATTCTACTCTACCTAACCAACCCCTAGATCTGAG
TCTGAAGTACAGTTAAAACACTTTAAAAGTGGAGTAAAAGTGAGG
CAGACATTTCAGAGTCTGAGTTACAAGAACATAAC
TCCTATTAAACGGAGTAACAACATGCAAGGTTCTATAGCAGCTAGGT
GAGTTGTTCCGGGTCT
TCTTAACGGCAGCTCCTGTACATACTGGTACTTATTGCTGAAACGT
CTGTTCATACATTGCCATG

CAGTGACTGTGCTCGAAAGGTGAAATCAATGGTAAAGAAGCCTAGGATTTATGGTATGAGAAACTGATCG
CAGGATTCTGATCCATAGCTATTAAATTGAAAACCTCTGATAGTGTCTAGGCTCTATAGTAGGCTTAAG
AGTGAGTCTTATACTATGAAGCCAGCCACAAGATACTAACAAATGTATAAATTGAAAAGTACCAACTT
TTGCACATAGATCAAATAGAGTGGAGAATTGGGCTCCAAAATCAGGTTAGTAGGCTCATTCATGTG
CTTAAAAATGTTTGTGTTATGCATTGATTAAGGGCTAGAGCAAATATAAAATTAAATGTCTTTCC
CCAAATAATTGGTTATTAAATGGTAAATGGACTTGGGATTACTGAGATTAGTATGTGGTAGTAAAG
AATTGGAGTGGGAGTTAGTGGGAGGGATGACAAGCTCAGCTTGGTTTGTTAGTCTTC
TAACAGCTACATTCTGTTGCTTATTGCCAGATAATGTGTTATTTAAACCAGGTAAAGTATTAA
CAGAGAACATTATTGATATACCTACAAATTGGCCATTATTGGGTGGGGATAAAAAGGA
TTCCCTAACCTTAAATAATCAGAATATCCTGCAATTGAAAGGAGAAAAAAATCCAATTTCAGTATT
CTTAGTCTGAAAAAAACAGCAGTAGTAATCTGGTAATGTTATTAGGAGTGTCTGCCTCTACAGTGG
CATATTCCATGACATTAAATAGAAAACCTCTTCCCACCCCTCCCTTGACATTTCAGGAA
TGTGGTCTGCAAACAAGAACTGACCTATCACCTCTGCTCACAGTGGTAAATATCATACTC
TTGAAAGGGAAAATATCCATGTCAGTTCTTACAGTAGAAGAATGGTCTTGAATGGGAGGAAAG
AAGACATTTCATAATATCCTTTGCCCCTCTAAATGTATAAATATTGATTGGTTACTTT
AATTTCATTCAATTGCAACTGCTTTGGGTGCCCTCAAACACATTACTCCATTTCAGAAC
CATATCTGATTTAAATAACTGCTAACTTAAGAAGAGGTGAATTGCAATTACACTTAGACACTTCTAA
AAACAGGGTGAACTTAATTGAAATTTCAGTGTGTTAAGCTATTGTTAGTAGCTGCCCTATAGTTAC
AAAAATAAAAACATAAAAACCCATCCCTACTTAATCCAGATAGCCTGGATTAAACAAACCAATA
GTATAAAAATGTGTGTGTGTCTGAGAGAGAGACTTGATTCAATTAAAGTCCTGG
AACAGCTGTTCTAGGATCTGTCATTACTGAAAGTCTCGTCACATGTGAAGTGCCTCTGAGTTA
AGGTTGGTTAATACTGGTATATTGTTATAAGATTGAAATTGTCCTCCATCTTAACCTAAACATTTC
TTATCAGGGTCAATGTGCTAGAACTGAATTGTAACATTAGGCACAGATGGAAAAAAATGTTATTGG
CCTGAAAATGTGTGTGGCGAGGGGAATAGATCCACAAAAGCATGTACTTACAAACCAAGCTG
TAGAGATCAAGAAAAGAACTTAAGTGTGATCTCAAGATTCTAAATTGTCAGATTACATGGCATTGT
GTGGAACTAGTTAACACTTAGAGAGCTTTGGTATGTAATAACTATTGCTATGGACTGATTAAATGTTCAA
AAGATTGTGTTCTCAATTGGTGGGTTGATTGTTTTAACTGCCTCTCAGATTATTTACT
TAGTTAAATTCTTGTCTTATTCAATTAAAGTATAAAACTCAGGTCCTGATATTATTCACTTGT
TTACTAATTATTACAAACCCCTTGTGACTTTATTGTTAAATGTGTAATGTATTAAACGTCTTAA
AATTGTTCAACTGAAACTACATTAACTTGATTGCTTACTGGATTGTTAAAGACACTTTT
CCATGTCAGTGCAGCACTTAACCAGTCGTTGATTCCCTTCTTCATCCAATAACAGAATGAATA
CCTAAATAA

TASLGEETLASASSSSDSDTGGASPPRKKPRASAEGVGEPGASAGRAGLSPPSSSSSSSSSSSSVVVV
VGLPPAAPAAAVPHRSSGHSLVSGSIMQANGAGGGGGGGGGGGGGGGGGGGQQTPELAQCLSAQNGES
SPSSSSAGDLAHANGLPSAPSAASNNSNSLNVNNGVPGGAAAASSATVAAASATTAASSLATPELGSS
LKKKKRLSQSDEDVIRLIGQHLNGLNQTVDLIMQESGCRLEHPSATKFRNHVMEGDWKAENDLNEKL
LVHSPHAIVVRGALEISQTLGIIVRMKFLLLQQKYLEYLEDGVLEALQVRLCELTPLKYNTERIHVLSG
YLMCSHAEDLRRAKEWEKGKTASRSKLLDKLQTYLPPSVMLPPIRLQTLRQAVELQRDRCLYHNTKLDNN
LDSVSLIIDHVCSSRQFPCYTQQILTEHCNEVWFCKFSNDGTLATGSKDTTVIIWQVDPDTHLLKLLKTL
EGHAYGVSYIAWSPDDNYLVACGPDDCSELWLWNVQTGELRTKMSQSHEDSLTSVAWNPDGKRFVTGGQRG
QFYQCDDGNLLDSWEGVRVQCLWCLSDGKTVLASDTHQIRGYNFEDLTDRNIVQEDHPIMSFTISKNGR
LALLNVATQGVHLWDLQDRVLRKYQGVTQGFYTIHSCFGHNEDFIASGSEDHKVYIWHKRSELPIAELT
GHTRTVNCVSWNPQIPSMMASASDDGTVRIWGPAPFIDHQNIEECSSMDS

AUG initiation = 72.1 kDa
ACG initiation = 81.3 kDa

+++++

#23

NM_005078

Conserved in mammals – CUG in good context conserved from opossum to human

TLE3 transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)

Synonyms: ESG; ESG3; GRG3; FLJ39460; HsT18976; KIAA1547; TLE3

AGTCCCTATTTGGAACTGCTCTCGCGGAGTTCAAGACCTCGTCGTCCCCTGCCGTCTGTGTGG
TATCCGTAGGTCCGGGGCACTTTTTGGTGTGTGTGTGTGCGCGCAGGGAGAGGGAGGGGAAGGAAGGCCGTAG
AGCTCCCCGGTCTCGGTCTCCCTCCAGCGCTCTGCCTGTGCGCTCCGACCGCAGCG
GTGGCGCTGAGCGCTCTCGTGGGACCGCGGTGCTCAAGCTGGAGCAGCAGAACCTTGCGGAGGCT
TCCCCTGGAAAGGGCTTGCAGCGCGCGGGCAGGGAGTAGCAGACGAACAGCAGGCTCTCC
TCTCCGAGGGGGGGCGTGAAGCGCCCGCGCCGGCTCCCGCCCTCACCTCCGGCGCTCGGGCCCA
AGCCACACCTCGTAGCCCGCTCCTTCTCTCGCCGAGCCCCAGGCCGACGCCCTAGGATCG
GACCGCGCGCCCGATGCTTGGGCTGCCCGCGCCCGTACAGGCGCAGCCGCTGAGTGACCC
GGAGGGGGCCGAGCTGCCAGGAGGGCCCCCGCTCGGACCCCGCGACCTGCTGGGAGCC
GACGATGTCGCGCAGTCCCAGGCAT **TAA** CGCGCTCGACGCAGCCCAGGGAGGCCGGCGCTGCCG
CCGAGCCCTAGGCCACAAAGCCCGCCGCCGCCGCCGGCCGCCGCCGACGACTTTGCCGCTGC
TCCGCGCTTTGTCTCCACTGGGCGGGCGCCGACT **CTGGGATTTCGCTGCGAGAACGAGCTGGGG**
GGCGGGGGCGAGCTCTCGGTTCCCGCCCGCCCGCTCGGCTCGCTCCCCCTCCCCCGACCTCC
CCGGCCCGGCTCGCGCTCCACGCTCTCGGAATCACGACCCCTCCCTGCCATGTATCCGAGGGCAG
ACATCCGGCTCCCCATCAACCCGGCAGCCGGATTAAATTACGGTGGCTGAGTCTTGTGACAGGATCA
AAGACGAATTCCAGTTCTGCAAGCTCAGTATCACAGCCTCAAAGTGGAGTACGACAAGCTGGCAAACGAG
AAGACGGAGATGCGCAGGCCATTATGTGATGTTAGATGAGATGCTCATGGCTGAACATTGAAATGCACAA
GCAGACAGAGATTGCGAAGAGACTGAACACAATTAGCACAGATCATGCCATTCTGTACAAGAGCACC
AGCAGCAGGTGGCGCAGGCAGTGGAGCGCCAAGCAGGTACCATGACGGAGCTAACGCCATCATCGGG
CAGCAGCAGCTCCAGGCCAGCACCTCTCCATGCCACACAGGCCCGGTCCAGTTGCCACCCACCC
GTCAGGTCTCCAGCCTCCAGGAATCCCCCAGTGACAGGGAGCAGCTCCGGCTGTCGACTGGCGCC
TGGGAGCCAGGCCATCTGACGGTGAAGGATGAGAAGAACCCATGAACACTCGATCACAGAGAGAGAA
TCCAGTGCATAACTCTGTGTCACCCTCGGAAAGCCTCCGGCCAGTGAGAAGCACCGGGCTCTCGGA
CTACAGCATGGAAGCCAAGAAGCGGAAGGCCAGGAGAAGGACAGCTTGAGCCGATACGACAGTGAG
ACAAGAGTGTGATCTGGTGGATGTTCCAATGAGGACCCGCAACGCCCGGTACGCCGGCACAC
TCCCCTCTGAAAATGGCTGGACAAGGCCGTAGCCTGAAAAAAAGATGCCCAACAGCCCTGCCCTGGT
GGCCTCTCCAGTAGCACACCTCCTCCAAGACCAAAGACCTGGTCATAACGACAAATCTCCACCCCTG
GGCTCAAGTCAAACACACCAACCCAAAGGAACGACGCCAACTCCAGGCACAGACCCAGGGCTC
AGGTCGATGCCGGTAAACCTCCGGCATGGACCCGATAGGTATAATGCCCTCGGCTTGCACGCCAT
CTCCATCACCAGCTCCTATGCCGCCCTCGCATGAGCCACCATGAGATGAACGGCTCCCTCACCA
GTCCTGGCGCCTACGCCGGCTCCACACATCCCACCCAGTGAGGCCGCCGCTGTCAGCCGCT
GCCCTGGCGATGCCAATGGTGAAGCTTGGAGCTGTTGACCCCTACCCCCCGATGCCGGCAC
AGGCCTCCCTCAAGCCTGGCTCCATTCCCTGGAGGAAACAGCGTACTCATTCCATGTGAGTGTGATG
GGCAGATGCGCCGTGCCCTCCCCACGACGCCCTGGCAGGCCGGCATCCGAGGCACGCCGGCAG
ATCAACACACTGCCACGGGAGGTGGTGTGCCGTGACCATCAGCAACCCACGAGGCACGTCTACAC
AGGTGGCAAGGGCTCGTGAGATCTGGACATGCCAGGCCAGCAAGAGCCCCATCTCCAGCTGG
ACTGCCTGAACAGGGACAATTACATGCCCTCTGCAAGCTGCTCCCTGATGGCGCACGCTCATGTGGC
GGCGAGGCCAGCAGCTCACCATCTGGACCTGGCTGCCACGCCCGCATCAAGGCCAGCTGACGTC
CTCGCTCCGCCGTTATGCCCTGCCATTAGCCCTGACGCCAAAGTGTGCTCTCTGTCAGCGATG
GGAACATTGCTGCTGGGACCTGCACAACCGACCCCTGGTCAGGCAGTCCAGGGCACACAGATGGGCC
AGCTGCATAGACATCTCCATGATGGCACCAAAGCTGTGGACAGGGGCCCTGGACAACACGGTGCCTCG
GGACCTGGGGAGGGCCGACAGCTACAGCAGCATGACTCATTCCCAAGATCTCTCGCTGGCTACTGCC
CCACTGGGGAGTGGCTGGCTGGCATGGAGAGCAGCAACGGTGGAGGTGCTGCACCAAGCCTGAC
AAGTACCGAGCTGCACCTGCAGAGAGCTCGTGCTCTCCCTCAAGTCCGCTACTCGGGCAAGTGGTCTG
GAGCACTGGGAAAGATAACCTCTCAACGCCCTGGAGGACGCCATTGGAGCCAGCATATTCCAGTCTAAAG
AATCCTCGTCTGTTGAGTTGACATTGAGGATGACAAATACATTGAAACAGGCTCTGGTACAAG
AAGGCCACAGTTATGAGGTATCTA **TAA** ACAAGAACCCAGGCCAGGCCAGGCCAGGCC
ACTCGGCTCTGACAGGGAGACCCCGAGGCCAGGGGCCAGGGATGGCGAGGATGGGCCGAGGCC
AGCGTTAGGGCTCGCCTCCGCCGGCTGAGAGGGCACGTGCCCGTCACAGTCTGGACTCCGGCTGG
ATTGATGTGTCACAGACTCGGAAGGGTCTGCTCCTCTCCCTGCCATGAGTCTGTTGCT
CAAATAGATTATGGAGGCTTATGGCTCCGGTCCCCACAGACCCGCTCATGAGTCTGTTGCT
CCCTTTCTTTGCCCTGTCCTCACCTGGTGGGGTGTGGAGTGGACCAATGTTGCTGGGG
ATGGGGGGTCTCTTGGCCGATTGTGCAAGATTGAAAAATGTAATAACAGACTCCTAT

TGCGGACTGATCAGTGGGAGAGGAGGCCCTCCACCGAAACTCTGAGTGTGTATTCGCCTGCTGTAT
TTGTAATCCACTCGTGGTGGCTTTTTTTTTTTAAATAAACAGATGCTCTCACCTGGAA
GAGGAGACAGGGAGGGGAACCAATTGAAGAAGAGGAGAAAAGTCTTAGAGTGTGAAAAGGAAACCAGGT
TGGCCGTAAAGGTGCCTGCTGGAATGCGTGTGCCTCACACGGGTCTGGCATCCGACTGATAACCAGCCG
GCCAGACTGAGGGATGGAAGGCACTGAGATGGGGGCCGTCAGGCGACACCCGAGAAATGGAGCTTC
TGTGGTCTCTGCACTCTGGCTGCCTCTGGCTCTGTGTCTCTCTTCTGGTCTCTCCCTCTCCCT
CCTCAGCCTGGTCTTCTCTTGGTGACACTTAGTTATTGGTGAGCAATGGAAGTTCAAAGGAACCTCC
CTCTCCAGCTCTCTGAATCTTGGACACAGCCTAAAAGGACAAAAGTTAGAAGACAGCATAGCAACTC
AGCTCAGGGAGCTACCAGAAAAATAGCAACTGATGTGGGTGCTTTTTTTTTAATTGAATAAA
AAGAATTAGAAGTGTGCTTTATAAAATGCCTCTCCCCCTCCGCCTACAGTCTCTCCCTCCCC
TTAGAGGGGGAAAGTGTATAACCTACAGGGTTGTGAGTCGAAAAGAGGATCCCCCTACCCCCACCC
GGGAGAGCAGTGGGGTTGGGGGGAGAGGGGACACAGATCTGCACACTGTGGATATTCTTG
AGATTGAGTCTCTGTGGCCAAACAGGTTAGGTAGACTATGCCTCTGGCAGGTGCCACCTTTGGTAC
CAACATGTTCTGAGGTGTTAGGATTGGGTTGGGTTTTTTGTTGTTGTTTTTTCCCTTTGGTCTTT
TTTTTTCTCCTTTAAAGAAAAGCTAAAGGCCGCTGTGAGTCCTGGCAGGCTCTCCATGGATGTAGC
ATATGAAGATAATTTTATGCTTTATGATTATTGATTATTGTAATGTGTGATTCCGTCTGCTGAGGAG
GTGGGAGGGGCTCAGGAAAGCCACCCACCTTCAGTGAGGTTGCTCCCCAGCTGAGCGCACCGGCATGG
GATGTGGAGGCTGGCAGCACACCCCTGTGCCTCTCCAAGGCTGGCGCTGGGCGTCCAGAGTCTCTGG
GTCTCAGATGTCCATCTGCCACCTCTGTTAAGGCTCTAGCCAGAAGGGAGGGTGAGGGTAGAAAGAAAGTT
ATTCCCGAAGAAAAAAAGAATGAAAAGTCAATTGACTGAACGTGTTTATATTAAAGTTACTATTAA
AAGGTTAAAAAAAAAAAAAA

LGFRCENELGPGASSRFPGRPLGLSPLPRTSPARALGASTLSESRLPAMYPQGRHPAPHQPGQPGFK
FTVAESCDRIKDEFQFLQAQYHSLKVEYDKLANEKTEMQRHYVMMYEMSYGLNIEMHQTEIAKRLNTILA
QIMPFLSQEHQQQVAQAAVERAKQVTMTELNAIIGQQQLQAQHLSHATHGPPVQLPPHSGLQPPGIPPVVG
SSSGLLALGALGSQAHLTVKDEKNHHELDHRERESSANNSVSPSESLRASEKHRGSAADYSMEAKKRKAEEK
DSLSRYDSDGDKSSDLVVVDVSNEPDATPRVSPAHSPPENGLDKARSLKDAPTSPASVASSSTPSSKTD
LGHNDKSSTPGLKSNTPTPRNDAPTPGTSTTPGLRSMPGKPPGMDPIGIMASALRTPISITSSYAAPFAMM
SHHEMNGLTSPGAYAGLHNIPPMSAAAAAAAYGRSPMVSFGAVGFDPHPPMRATGLPSSLASIPGGK
PAYSFHVSAQDMQPVPFPHDALAGPGIPRHARQINTLHGEVVCATISNPTRHVTGGKGCVKIWDISO
PGSKSPISQLDCLNRDNYIRSKLLPDGRTLIVGGEASTLTIWDLASPTPRIKAELTSSAPACYALAIspd
AKVCFSCCSDGNIAVWDLNQTLVRQFQGHTDGASCIDISHDTKLWTGGLDNTVRSWDLREGRQLQQHDF
TSQIFSLGYCPTGEWLAvgMESSNVEVLHHTKPDKYQLHLHESCVLSKFAYCGKWFVSTGKDNLNAWRT
PYGASIFQSKESSVLSADIADDKYIVTSGDKKATVYEVIY

AUG initiation = 83.4 kDa
CUG initiation = 88.7 kDa

+++++

#24

NM_002250

GUG/CUG initiated extension conserved in mammals (AUG in *Pteropus vampyrus*)

KCNN4 potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4

Synonyms: IK1; SK4; KCA4; hSK4; IKCA1; hKC4; KCa3.1; hIKCa1; KCNN4
GTCCTTCGGTGTCTGGGTGTGGTGAAGTAGAGGTGTGTGACAAAGTACAGACCATTGTGTGACAAAGC
CCATCGTGTCTGTGTGTCTTATCCACGTGGATGGACGTCTCTTCTTGCTCTGCCCAAGACACAC
CCTAGCCCCCTCCTATTCTCAAAAGGGGAGCTGGGAGCCTCCCCCTACCCCTGGGCGTCCCTGCCCT
CCCCGCCCTGCCCTGGCGTACCAACTCCCCAGAGGGCACAGGGCTCTGCTGTGCCTCAGAGCAAAGTCCC
AGAGCCAGCAGAGCAGGCTGACGACCTGCAAGCCACAGTGGCTGCCCTGTGCGTGTGCGAGGTGGGGAC
CCTGGGAGGAAGCTGGCTGAGCCCCAAGACCCGGGGCATGGCGGGATCTGGTCTGGCCTGGG
GCCTTGAGACGCCAAAGCGCTTGCTGGAGCAGGAGAAGTCTGGCCGGCTGGCAGTGGCTGGCAGG

AACTGGCATTGGACTCATGGTGTGCATGCAGAGATGCTGGTTCGGGGGTGCTCGTGGCGCTCTACCTGTTAACATGCACGATCAGCATTCCACCTCTTAACCTCCTCTGCCTCATCGTGGCCTTCATGCCAAAGAGGTCCAGCTGTTCATGACCGACAACGGGCTGCGGGACTGGCGCGTGGCGCTGACCAGGGCGGCAGGGCGCAGATCGTGCTGGAGCTGGTGTGGCGTGCACCCGGCGCCGTGCGGGGCCCGCCGTGCGTGAAGGATTAGGGCGCCGTGACCTCCCCCAGCCCTGGCGGGATTCCCTGGGCAAGGGAAAGCGCTGCTGTCCTGGCCATGCTGCTCTACCTGGTGCCTCGCAGCGCGTCCCTGCGCAGCGCGTCCCTGCTCAA CGCTTCTACCGCAGCATCGCGCTCTCAATCAAGTCCGCTCCGCCACTGGTTCGTGGCCAAGCTTACA TGAACACGCACCCCTGGCGCTGCTCGGCCCTACGCTTGGCCTCTGGCTGACCAACGCCCTGGGTGCTGTCCTGGCCAGAGGGCAGGGCTGAGGCTGACACACTTTGGCTGATCCCCATCAC ATTCCGTGACCATCGGCTATGGTGACGTGGTGCCTGGGACCATGTGGGAAGATCGTCTGCCTGTGCACTGGAGTCATGGGTGCTGTCACAGCCCTGCTGGTGGCCGTGGTGGCCGGAAAGCTGGAGTTAACAGGAGAAGAACGACGTGACAACATTCATGATGGATATCCAGTATAACAAAGAGATGAAGGAGTCCGCTGCCGAGT GCTACAAGAACGCTGGATGTTCTACAAACATACTCGCAGGAAGGAGTCTCATGCTGCCCGCAGGCATCAGCGCAAGCTGCTGGCCCATCAACCGCTCCGCCAGGTGCGGCTGAAACACCCGGAAAGCTCCGGGAACAAGTGAACCTCCATGGTGGACATCTCCAAGATGCACATGATCCCTGTATGACCTGAGCAGAATCTGAGCAGCTCACACCGGGCCCTGGAGAACAGATTGACACGCTGGCGGGAAAGCTGGATGCCCTGACTGAGCTGCTTAGCACTGCCCTGGGGCGAGGCAGCTTCCAGAACCCAGCCAGCTTACAGCTGGACCCACGAGGAGGAACCAGGCTACTTCCCCAGTACTGAGGTGGGACATCGTCTCTGCCACTCCTGACCCAGCCCTGAACAAAGCACCTCAAGTGCAGGACAAAGGGGCCCTGGCTTGGAGTGGGTTGGCTGCTGATGGCTGCTGGAGGGGACGCTGGCTAAAGTGGTAGGCCTGGGCCACCTGAGGCCAGGTGGGAACATGGTCACCCCCACTCTGCATACCCTCATCAAAACACTCTCACTATGCTGTTGGACGACCTCAGCTCTCAGTTACAAGTGCAGGCGACTGGAGGCAGACTCCTGGGTCCCTGGAAAGAGGGTACTAGGGGCCGGATCCAGGATTCTGGGAGGGCTTCAGTTACCGCTGGCGAGCTGAAGAACTGGGTATGAGGCTGGGGGGCTGGAGGTGGCGCCCCCTGGTGGGAC AACAAAGAGGACACCATTTCAGAGCTGCAGAGGACCTGGTGGGAGGAAGAAGTGTAACTCACCAGCCTCTGCTTTATCTTGTAAATAATGTTAAAGCCAGAA

VAAALCVLRGGGPWAGSWLSPKTPGAMGGDLVGLGALRRRKRLLEQEKSAGWALVLAGTGIGLMVLHAEMLWFGGCSWALYLFLVKCTISISTFLLLCLIVAFHAKEVQLFMDNGLRDWRVALTRQAAQIVLELVVCGLHPAPVRGPPCVQDLGAPLTSPQPWPFGFLQGEALLSLAMLLRLYLVPRAVLLRSGVLLNASYRSIGALNQVRFRHWFVAKLYMNTHPGRLLLGLTLGLWLTTAWVLSVAERQAVNATGHLSDLWLIPITFLTIGYGDVVPGBTMWGKIVCLCTGVGMGVCCCTALLVAVVARKLEFNKAEKHVNFMDIQYTKEKESARVLQEAWMFYKHTRRKESHAARRHQRKLLAIAINAFRQVRLKHRKLREQVNSMVDISKMHMILYDLQNLSSSHRALEKQIDTLAGKLDALTELLSTALGPRQLPEPSQQSK

Probability of export to mitochondria = 0.4881

AUG initiation = 47.7 kDa
GUG initiation = 50.1 kDa

++++++

#25

NM_004494

GUG initiated extension conserved in mammals

HDGF hepatoma-derived growth factor (high-mobility group protein 1-like)

Synonyms: HMG1L2; FLJ96580; DKFZp686J1764; HDGF
GAGGGAGGAGGAGGTGGGACCGGGCGGGGGTGGAGGAAGAGGCCTCGCGCAGAGGAGGGAGCAATTG AATTTCAAACACAAACTGCACGAGCGCGCACCCACCGCGCCGGAGCCTGCCCCGATCCCGCGCCCGCC CCGTCCGTGCGCGCGCGGGAGACGCCGTGGCCGCGCCGGAGCTGGGCCGGGGCCACCATCGAGGC GGGGCCGCGCGAGGGCCGGAGCGCGAGCGGCCACCGCGCACCGCAGCGCAAACCTGGCTCGCTTCC CGGCCCGCGCGAGGCCGGCGCCGGAGCCCCGCCATGTCGCGATCCAACCGCAGAAGGAGTACAAA TGCGGGGACCTGGTGTGCAAGATGAAGGGCTACCCACACTGGCCGCCGGATTGACGAGATGCCTGA GGCTGCCGTGAAATCAACAGCCAACAAATACCAAGTCTTTTTCGGGACCCACGGAGACGGCATTCTGG GCCCCAAAGACCTTCCCTACGAGGAATCCAAGGAGAAGTTGGCAAGCCAAACAAGAGGAAAGGGTTC

AGCGAGGGGCTGTGGAGATCGAGAACAAACCTACTGTCAAGGCTCCGGTATCAGTCCTCCCAGAAAAA
 GAGCTGTGTGGAAGAGCCTGAACCAGAGCCCGAAGCTGCAGAGGGTGACGGTATAAGAAGGGAAATGCAG
 AGGGCAGCAGCAGGAGGAAGGGACTGGTCAATTGATGAGCCAGCAAGGAGAAGAACGAGAAAGGAGCG
 TTGAAGAGGAGAGCAGGGACTGCTGGAGGACTCTCTAAACGTCCAAGGAGGAGAAAACCTGAAGG
 AGAGGAGAAGGAGGCAGCCACCTGGAGGTTGAGAGGCCCTCTATGGAGGTGAAAAGAATAGCACCC
 CCTCTGAGCCGGCTCTGGCCGGGCTCCCCAAGAGGAAGAAGAAGAGGAGGATGAAGAGGAAGAGGCT
 ACCAAGGAAGATGCTGAGGCCAGGCATCAGAGATCATGAGAGCCTGAGCCACCAATGTTCAAGAGGA
 GCCCCCACCCCTGTCTGCTGCTGGTACTGGGAAACTGGGCATGGCTGCAAACGGAAAC
 CCTTCCCACCCAACTGCTCTCTACTCACTTTCCACTCCAAGGCCAGCCATGGAGATTGA
 CCTGGATGGGGCAGGCCACCTGGCTCACCTCTAGGTCCCCTACTCTATGATCTGAGTCAGAGCCATG
 TCTTCTCCCTGGAATGAGTTGAGGCCACTGTGTTCTCGTGGAGCTATTTCCAGGCTCTGCTGG
 GCCTGGGACAACGTCTCCACCTCCTGACACCCTCTCCACTCTCTAGGCATTGGAACCTCTGGGTTG
 GGATCAGGGGTAGGAATGAAAGGATGGAGCATCAACAGCAGGGTGGGTTGAGGAGGGCTGGGAGGGCAA
 TCCTCAAATGCGGGGTGGGGCAGCACAGGAGGGCGCCTCTGAGCTCCTGAGCTCCTGAGGCTACACCTA
 TTATCCAGCTGCTAGATTAGGGAAAGTGGGACAGCTTAGGGAGGGCTCTTCCATAATCCTT
 GATGATTGACAACACCCATTTCCTTGCAGCCACCCAAAGAGTTGGGAGTTGAGTTAACATCAAGA
 GAATTGGGCTTCAAGTTGTCGGCCAAGGACCTGAGACCTGAGAGGGTGTACTTACCCATTGGGTG
 GGAGTGGTGAACATCTGCCCCCTTAGATCTCTGAAGCCACAAATAGGATGCTGGGAAGACTCCTAGCT
 GTCCTTTCTCCACACAGTGTCAAGGCCAGCTATAGTCATATATACCCAGACATAAGGAAA
 AGACACATTAGGAAATGTTTAATAAAAGAAAATTACAAAAAAATTAAAGACCCCTAACCC
 TTGTGTGCTCTCATTCTGCTCCTCCCCATCGTTGCCCCCATTCTGAGGTGACTGGAGGCTCCCC
 CTATTTGGGCTTGATGACTTCTTTGATGCTGGGCTTGTAGTGTCTCCAGTGTCAATTCTCATCC
 ACATACCTGACCTGGCCCCCTAGTGTGACCAAGATCTGATTTGTAACCCACTGAGAGGACAGAGAGA
 AATAAGTGCCTCTCCACCCCTTCTGACTGGTCTCTATGCTCTACAGTCTGCTCTTTACCC
 TGGCCCTCTCCCTGGCTCTGATGAAAATTGCTGACTGTAGCTTGGAAAGTTAGCTCTGAGAACCGT
 AGATGATTTCAGTTCTAGGAAAATAAACCGTTGATTACTATAAAAAAAAAAA

REEEEWPGGGWRKRPRAEAGAIEFQTQTTARARTHAGALPRSAPAPSVERAGDAVAAPELPGPATIEA
 GAARGPERSGAATAARANLGSRFPARRGARGARSPAMSRSNRQKEYKCGDLVFAKMKGYPHWPARIDEMPE
 AAVKSTANKYQVFFFGTHETAFLGPKDLPYEESEKEKFGKPNRKGFSEGLWEIENNPTVKASGYQSSQKK
 SCVEEPEPEPEAAEGDGDKKGNAGSSDEEGKLVIDEPAKEKNEKGALKRRAGDLEDSPKRKEAENPEG
 EKEAATLEVERPLPMEVKNSTPSEPGSGRPPQEEEEEDEEEATKEDAEAPGIRDHESL

AUG initiation = 26.8 kDa
 GUG initiation = 31.6 kDa

+++++

#26

NM_013313

ACG initiated extension conserved in vertebrates – in humans three of the four paralogs have the ACG extension

YPEL1 yippee-like 1 (Drosophila)

Synonyms: FKSG3; MGC64992; YPEL1

CAGGCAGCCCGCGGGCGGGAGGATGCCCTCGCGCCCTGCCCTAGGCCTAACGGCCTCAGCGCGTC
 CCAGGGCGCGCCGGAAACGCTCTGAGAGCCGAGGCCCGCGCTGACCGGGGCCGGATGGCGCTGC
 GCCGGGGCGCGAACCGCGAGCGGCCGTTACGTTCTCGTCCCAGCTGTGGACAGTGCCACACGC
 CCTCCTGGACAACGACCGCTCTGCGTCATTCCAGTGTGCTTTGTCAGAGAGGCCAGCAGAGCCAC
 GGTTCTCCAGAACCCAGCCCTGAGCTGAGTGAGGAGGCCAGGAGATGGTAAAATGACAAAGTCAA
 AACTTTCCAAGCGTATCTGCGAACACTGTCACCGAACGTACAGCTGTATCCACTGCAAGAGCACAC
 ATCATGACGAGCTCATCTCCAAGTCCTTCAGGGGAGCCAGGGACGCCCTACCTCTCAATTCCGTGG
 AACGTGGGCTGCGGCCCTGCAGAGGAGAGGGCTCTCACCCGGCTGCATGCGGTTGCCGACATCTACT
 CGAGAACTGCAAGACCACGCTGGGTGAAATACGAGCATGCCTTGAGAGCAGTCAGAAATATAAGGAAG

GAAAATTCACTATTGAGCTTGCATATGATCAAAGACAATGGCTGGAGTAATGTGCGAACTTCCCTTC
TCCTTGAATGCTTTGTAAAGAACTGTGAATGTAATGGAAACGTAGGAGCATCTGGTACAGCCTT
TCTTGCCTCTGACCTCAAAGGCTAGCTGCGCATAGCTCTGACACTCTCGGCCATCTGTGGGTAAGGT
GTCCCCTGGATCTGCCTCTCGTGTACACAGTTGTTCTGAAAATTTCAATGAGCTTTCTAACCTCT
CAAGTTCTAGAGAAAGAATTAACCAACTGTGACTTACCTGCCTAGTTAATATCTCCCTTCACCTTGTC
TTCAATATAGTTGGGCTCTGCTTTTAAGGTTCAGTTGAAAACCAAACGGGGCCGGGTGCGGTGGCTCA
CGCCTGTAATCCCAGCACTTGGGAGGCCAAGATGGGTGGATCACCTGAGGTAGGAGTTCTAGATCAGCC
TGGCCAACATGGTAAACCCATCTACTAAAAATACGAAAATTAGCCGGCATGGTGGCGAGTGCCTGT
AATCTTAGCTACTCAGGAGGTAAGCAGGAGAACACTTGAAACCTGGACACGGAGGTGCACTGAGCTA
AGATCATGCCATCGCACTCCAGCCTGGGGACAAAAGTGAGACATCGTCTAAAAAAAAAAAAAGCTG
GGTATGGTGGCGCATGCCTTAATCCAGCTACTCGGGAGGCTGAGGCAGAGAACACTTGAAACCCAGGA
GGCGGAGGTTGCAGTGAGCCAAGATCGCGCACTGCACTCCAGCCTGGCAATAGGGCAGACTCCGTCTCA
ATTTAAAACAAAAGAGAACCAACTGAGTCTCTGAAGAACCCACAGGGACAGGGCTCTTAGATAGCAAGTC
TCACCATTCCCTTTTAGAGAAAAGGTATTGTAGCCCACCTCCACCCGCTGTTTCTTAAATTGCA
GAACTTCAAATTGCTATTCCCTTGCAAATGAACCTTAAAGTACAGTGTATTAAAGAATCTTCAGAG
GCAGTCACAGACTTACACACTAAGGGATTGGTTAGCTGTCAAAACAGAGGCCAGCACAGA
TGACATTAGATACACTCTAAATTGAGAATGGTGTCTAGTGGAACATGTTATTAAAGCCAGTAGATTCC
TTATCTAGAAAGCAGGTGAGCTAGCCCTAGAGAAGGCTGTCCCGGGCCCGCAGAGGTGCCCTACTGAG
GTGACAGCCTCACAGGGCTGGTACCGGGGTTGTGCCCTCAGCAGTGACAGCAGCTAGGTGTCAGGCAG
TTGCTGAGTGGCTGGTCCATGTCTATAGAGTAACACACTGGACCGAGGAAAAGTCAGATTCATTCTAC
CCTGGATGTACTTGAAGAAAAGAATTATTTGCATATGAAAGAGGCCAGAACCCACAGGAAAACCTCA
AAACTTGACATTGCCAGAATGTTAAATTGTCAGAAAAGGTTAACAGAACAGTTAGCCTTTGTG
ATGAAGACGCCTGCCCTGCTAGACGCGTTGCCCTCGTGTGCTGCCCCATGTCACTTGAACACTG
AGAGGGCCTGTGCAATCTCTAACGGCCTGTGTTCTGCCATATTTATTATAAAATTACAATCCACTCA
TCCACCTGCCCTCACCAGGAGTGGCACCCATAAGGTTAGGCCACTTGCAAGGGATGGAGGTCAA
AACCACTCCCAGATAAGTTGGTTTCAACATTAGTAACCTGTCTCAGGGCAGAGGGCAGGCCAGGGGAC
CGAGGGCAGCAGATAGGAGAGCACTGAGCCGGATAGTCTCAGCCTGGCAAGTGGCTCTGAAGCTGCC
TCAGACAAGGCTAGTCTAGGGCAAGAGTGCAGCTGGTGCACAAATAAGAACGTGCCACCTGCCAGCT
CACACCTCCCCGACTTCAGCCCTCTAAACCCAGACCTGCCCTCAGGCACTGGCTGTGCCAC
TCGAGCTCACTGCCACACACAGCATGCCCTGGTGCATCTTGCCTAACGCTGGCCATGAAGTTAGCA
GTGGAAATGCCGTGCCCTGGCATGGCACTGAGATGCATCCACTCAGCAGGAGTGACAGAGGCAGA
AGTTCTTAAAGCACATCTTCACTTAGAAAGGAAGGAATCTTGACTGTCTTGAAGCCTCCACAT
CCGGCTATGCCCTGCAAGCTGCTTATCCCTGCCCTAGTCTCCCCCGAGGGTTAGGCTGCCAGCACA
TCCTGCTCCTGAGCTCGCGTGCAGCCACCCAGAGCGCAGGGTCACTGCACGCTGCAGGGCTTGCTG
CCATGGTCTCAAGCCTGAAGAGGCTCCGCCACAAGCTGGCCATGAAGTTAGCAATGCCCTGGCTTCAG
TCAATTGCTTGAGACTGTGAAGAGGCTGAAAGACACCTTCCGGTGAAGAAGGAGTTACTGAA
TATCTTAAACTGACCCCTCCCTTGAGTGAGTCTCATTCTCTCCATGTGGGAACCCAGCTCCGATGC
CCCAGGGACTAGGGGAAACAGTTGGAGGTTGTGCCCTCCAGCCTGCCACGGTGCAGGACAGCCAAG
TCCTGAGTGAACATGCTTCACTACATGAAAGAAACTCTAAACTCTACCGAGTGGTTTTGTATA
TACTAAAGTTCTATTAGAGCTTCTGTTGGCAAGTTGCTGCTCCTCTATTGGGCACTTGGTT
TTTGTACTGTCTTTGTGACGGCATTGATTGAACATTTTACTAGTAGTACTGCTTATGACTTTGTATT
TTTTTTTTGTAAATTATACCAACAACACTTTTATCACTTTTTTGTGTTGGCTTCTGCAAAATACAAGC
TCATTTAAACCAAATGAACAGACCATGAGCTGCCCTCAGGGAAAGTGTATTACAGGACCATATCCAC
CACCTCTTAAATTCTAAACAATATCATCTAGGACTCTATTAAAGTTATTAAAATATTCCCTTG
GAGCCTGGGAGGTGATGTCAAGGGTTATAATGGCACAGTGCAATTGCTGTAGGAATGTGGTTGGCATTG
TTTATACACACAGTATTTTATACCTTAATGCTTATTCTGATGGCATCTGTCAAGATATTAGAATTGAA
AATAAGAATCTTCCAAAATCTTAAATTACCTGATGCCCTCATCAGTGTAAAAATTCAAATGGTT
TAATAGCTAAAAACTACAAATTAAAGCTCTAAACAAACAAACTACAGAAAATGTAACCTCATTGCAA
AGGTCCTTGGTGGCTGTCCCTGCCCTGGAGCAGATGGCCTGAAGCCCTCCCTACTGTGAGGCC
CCGGGTGAGGCTGGACGGTACCCATGGTGGCTTCACTGCAAGGAGCAGGACTGCCAGCTCAAGCACGG
GCCCTCAGCTCCCTGCTCTGCCACACGCCAGCCCTGGCATTCTGTGAGGTTACAAATA
AAGCTTCTGATGTCAAATGTTAAAAAAAAAAAAAA

TALACIFQCCFCQRQQSHSSRTSPELSEECPGEMVKMTKSRTFQAYLPNCHRTYSCIHCRAHLANHDEL
ISKSFQGSQGRAYLFNSVVNVCGPAEERVLLTGLHAVADIVCENCKTTLGWKYEHAFESSQKYKEGKFII
ELAHMIKDNGWE

PROBABILITY of export to mitochondria: 0.5853 with extension
PROBABILITY of export to mitochondria: 0.5239 w/o extension

13.6 kDa from the AUG codon
17.3 kDa from the AUU codon

+++++

#27

NM_022106

AUA initiation conserved in mammals

C20orf177 chromosome 20 open reading frame 177

Synonyms: dj551D2.5; C20orf177

TCAGCTCAGCCAGGGAGCTCAGCGGAGCTGCAGCGCCCTCCGCCCTCAGCTCCCCTGCCGCAGCGCGCCGCAG
CCGGCGTCCCCCGCGGGCGCCGGAGAGGAAGGAAGGGCTGGCAGCCTCGTCACGTGTCCGCTGCAGTCGC
GAAACAGTTCCGGTGGTGAGGAGACCTTCAAATATAAGAGGAATAAGAAGTCACCTCCCCAGCTGTCA
TCATCTCCAGCAGATTGAGAAGAATTTGAGCACTACAGGAAAGACAGTCCATCAAACCCGAGATGA
TGATCAGCCACGTGATTTTCAGAAGAGGAATAGGGTGAATGAATCTCATCAGAAAAGCAGCAATATGA
ATGCTGCCCATCTTGAATAAGTCAACATTCAAAGAATTCTTCAGGAAAAGGCAGAGTAAATCCAA
GTACCCACGCTTCTCCCAGCCGAGAACGAGCCTCACAGCTGTCACTGAAGAAAAACTTAA
AGAAAGCATTCGGAGAACAGCAGAAAGGAATCCACTCGGTTCCAGGTGTCAAGGGGCCCTCAGGGAA
ATAAAATGTTCTGATTTCACTGAAAGAATCCCATTCTCCTTCTCCCTCACACCTCAGATCTCAATCTCGAGCTGA
AGAAAATTGATCCAGTTACTTTGATCTCACCTGGTCAGGCCATACAAACCTGAATACTATTATCTTA
ATTTCTTCATCCCTTCAGCTCTGGACCTACGAGATATGGCCCTGCTCTGAACGCAGAGAACAAA
ACGGAAAGCCGTGCCCCGAGTGGGAGGACTTCTTGGAAAGTATATCGATAGACTTATTCACTGAGTGGCT
GCAAGTCCAGACTGTACAGTGTGAAAAGCAAAGGGGGCAAGCAAGGCCACTGCCCTGGGACCT
CAGGGGCACTGAAAAGCCCTGGGAGAAGTAAGCTAATTGCTAGTGCTCTGCCAGGCCACTACCTCACCA
GAAGGGGCTTCAAAGTCAGGCCATTGATGTGCTGGTACCAGGTTTGTCTCAGAGGCAAACCC
TGCATTGAGACTTCCCTAGACCCATTGATGTGCTGGTACCAGGTTTGTCTCAGAGGCAAACCC
TTGAAATGAGGACAGAAGAAAAGAAAATCAAGTAAGAGTACGAAGCTGCAGCGCTGGATCTGTCC
GGCAGTGGAAAGCAGCTAAGGTGAAACCAGCGTCACATTGAGTTCCAAACAGGCAGCTGTGATTCT
GGACTCAGCAGATTCTGTAAGGCTCCAAAACACAAGCACATGCACATCCTAGGAAAAGGAAAGGCAG
AGAGCTGTGGTCATGCCACTGTATCGAGTGAAGAAAAACTGAAAACAAACGGAGTAAAGCAAACACATAT
AAAATGAAATATCTAAAATGCTGAATTGCCAAGACCTGCAGGTACCTCAATGTTAGAGCGCTTCAA
AAGTCAAAATACTGTGAATTAAAGGAATTTCACAAATACTGACATTAAAGTAGTTGACTGGCATTGTTGT
CCACCTTATTCTACCCCTGAGTGGGTTATTCAAAAGGAAGTGTCTTCAATAAGCCTTCTTGTAT
TGTCACTCTAGGAAATGAGAGCCATTAGATAAAATTATGTAATGTCATATAAGGAATAAA
ATGGCACCTCTCCAGGGAAAGTGTCACTGAAACCTCAGCTACAGTAGCCGTCTGTGAGAGCAGCTAGTG
GTGTTACCTCCCCATTTCACATGCACGTAAGTATATGAAATAGTGCAGACTGTTCAAATGGTGTGGAAT
CCTAAATGTTAAAATAAGGTCTTCTGCCCACCCCTCGCTTACTTTTATAAACTCTCAAGCAAAA
TTCTGTTATTACCCCTAGGAGAAGCTTACTGAGAGACAGTGAATCTAGTACTTTAATACATTTC
TCTGACATGGTTTTCTTTGAGGGCATTAAACTTAGAGGTGGTGGTAAACCTACTTTG
AGTTCTCGAACTGAGGTTAAAATAACTTGCAAGAATTTCAGAATGGCTTAGCATGATTACTGCT
GTTGGTGGGCTGAGAATGAAATATTGACATTCTGGAATTGCTGGCATGTAAGCTTCTCCAGAGAGGC
ACCCCAAGGGAAATCACTCTTACAATTGTAAGGAAGGGCTGAAAAGGATCAAACACATGGACCTAC
ATTCAGTGAATAGTTACAAGTTACTGATTGGGTTCCACACCCCTGTGGCTTACTGCAAAATAATGAT
CTGTTCACTGCAAGAGCAGGATTATTATTGCTTGGGAGGGGGAGAGTGGAAATATGAGT
AAGGTTGCTGAATGAAATTCTAAACTCGCTTACTGGTCTTCAGGCTTCCAACACTCTCTCAAAGCCTTCA
TTTCACTGCACTTAAATAACATCTTGTGCTTCTGTTAGTTGAGTTCTGTTGATTTGTTCAATTGCGCA
GTGTATTAAATACGGCCATGTCATTAGTTGAGTTATCCCTTAAACAATTACTGTATTGTTGTTG
ACGTAGAGGTTCAATTTCACCTGGGGCAAATGAAAACCTGGCATTTCATTTGGAACATATA

ATAGCTTGTAAACTTTCAGACAGCAGTAAATGTCGAAAAAATATCAAAAACAGCATAAAGACAAGATTATGTAGCTCTAATTATACTGATATAATTATAAAAAACAATGTCAAGGGTTATATTTAAGGTCTTTAAAAA
TCTGATTTGATCATACCAAATGACATAATATTTTATGGTAGCCTTTACTTTCAAGACTTAATTTCAGACTGGAGAAAGCATAACTTTATGCTAAGATCTTACTTAAGCTTTATGTGAACAAAAGATGTACATATAGTAAGTATTACTCCGTAGTCCTC
AAATTACTATAACTTTGACTTAGTATATGTTTATATTGGAAAACAGCACTACGCTTAGTTCCCTG
TAGTTCTGAGTGATGTCGTGTCCTGCCCTTTTGAGCACAGATTAGTCTGTTATCCATGGCTGCACACTTCTGCTAGATTTTATGCAGCTCTATGAAGTTCATGGCCA
TAGATATTCAAAAGCAAGATATTCTATACATATGTGTATATGTATATACCTCTATGTTAATACTAAAGTGTTATGCTGAGTTGCTGCTTCCCCGTATGTATCCATGTGATGCTCTAGAGACCTTGAATGGTTG
AGGGTAAAGTGAATTATTAGTAATTCTACTTGCTGTGAGCTGAAACAACTGTTGATTAAGAATTTAGGCTGGCGTGGCTCACGCCGTAACTCCAGCACTTGGGAGGCCGAGGCCGA
TCACCTGAGGTGGCTGGCGTGGCTAACAGACCAAGCAGCCTGACCAACATGGAGAAACCCCTGCTCCACTAAAAATACAAATTAGCGGGTGTGGTGCATGCCGTAACTCCAGCTACTCGGAAGTTGAGACGAGAATCTCTGAA
CCGGGAGGCCGAGGTTGTGGTGCATTGCCACTCCAGCTGGCAACAAGAGTGTGGTGCATTGCCACTCCAGCTGGCAACAAGAGGGAAACTCCGTTCAAAAAAAAAAAAAAA

IRGIKKSPQLSSSSRLSKNILSTTGKTVHQTRDDQPRDFKKRNVRVNEHQKSNMAGPSWNKVQHS
KNSSGKRQSKSQVPHASSQPRSSLTAVTQPTEEKLKESISPEARRKRNLGSRCQGASGNKLFLDFQSMKI
IENADEDSASDLSDLSERIPIPPSPLTPPDNLRAEEIDPVYFDLHPGQGHTKPEYYYPNFLPSPFSSWDL
RDMALLLNNAENKTEAVPRVGGLGKYIDRLIQLEWLQVQCEKAKGGKARPPTAPGTSGALKSPGRSKL
IASALSKPLPHQEASKSGPSRKKAHFHHEIHPSHYAFETSPRPIDVLLGTRFCQRQTLEMRETEKKKS
SKSTKLQRWDLSGSGSSSKVETSGHIRVPKQAAVILDSADSKASKTQAHAPRKKKAESCGHATVSSEK
KLKTNGVKQNTYKLK

AUG initiation = 42.1 kDa
AUA initiation = 48.7 kDa

+++++

#28

NM_006079

CUG initiation conserved in mammals

CITED2 Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2

Synonyms: MRG1; P35SRJ; CITED2

ACAGCTATTGTTGCAGCTGCCGGCGGTCTGCCAGCTGTGAGGGCAACGGAGGGAAATAAAAGGAA
ACGGCTCGAATCTGCCCGAGCGGCCGCTGCAGACCTCGCGCCGACATCGCAGCGAAGCGCTTGC
ACGCCAGGAAGGTCCCCTCATGTGCTGCTGAGCCGGTCCTGGACGCGACAGGCCCTCGGTCTTCGG
AGCAGAAATCGAAAAACGGAAGGACTGGAAATGGCAGACCATATGATGGCATGAACCACGGCGCTTCC
CCGACGGCACCAATGGGCTGCACCACATCACCCTGCCACCGCATGGGATGGGAGCTTCCGAGCCCCAT
CACCACCAAGCAGCAGCAGCCCCAGCACGCCCTCAACGCCCTAATGGGAGCACATACACTACGGCGCGGG
CAACATGAATGCCAGAGCGGCATCAGGCATGCGATGGGCCGGGACTGTGAACGGAGGGACCCCCCGA
GCGCCTGGCCCCCGGCCAGGTTAACAACTCCCAGTTCATGGTCCCCGGTGGCAGCCAGGGAGGC
TCCCTGCCGGCCAGCATGCGACTGCAAGCTAACAAACAGTATTTCACCATCACCCCTACCCCCACAA
CCACTACATGCCGGATTGCAACCTGCTGCAGGCCACCAAGATGAACGGACAAACCCAGCACTCCGAGATT
GCAACCCCAAGCACAGCGCCGGCAGCAGCACCCCGGCCGCTGGCGGCAGCAGCACCCCCGGCGCTCT
GGCAGCAGCTGGCGGGCGGGCAGCAGCACAGCAGCGGGCAGCGGAGCGGCAACATGCCCG
CTCCGTGGCCCAGCTCCCGCTGAATGCTGCCGCCAATGTCATAGACACTGATTTCATCGACGAGGAAG
TTCTTATGTCCTGGTGTAGAAATGGGTTGGACCGCATCAAGGAGCTGCCGAACTCTGGCTGGGCAA
AACGAGTTGATTATGACGGACTTCGTGTGAAACAGCAGCCCAGCAGAGTGAGCTGTGACTCGATCG
AAACCCGGCGAAAGAAATCAAACCCCAACTTCTCGCGTGAATTAAAGAAACATTCCCTAGACACA
GTATCTCACTTTCAAGATCTGAAAGGTTGAGAACTTGGAAACAAAGTAAACTATAAACTTGTACA
GGTTTAAAAAAATTGCTGCCACTTTTCTGTTTGTAGCCTGACATTCAAC

ACCTCCCTTATGTAGTTGAAATATCTAGCTAACTGGTCTTTCGTTGTTTTACTCCTTCCTC
ACTTTCTCCAGTGCTCAACTGTTAGATATTAACTCTGGCAAAGTGCCTAATCTTGGAATTGTTAGATGG
TTTCAAATGACTGAAGTGCATTAGTGCAGATTACAGAGTGAAGGAAAATTGCATTAGTGGTGCATGAACCT
CGAAGGGCAGATATTACTGCACAAACTGCCATCTCGCTCATTTTTAACTATGCATTTGAGTACAGACT
AATTTTAAAATATGCTAAACTGGAAGATTAACAGATGTGGGCCAAACTGTTCTGGATCAGGAAAGTCAT
ACTGTTCACTTCAGTGGCTGTCCCCCCCAGCCCCCACCCCATATGTACAGATGATAATAGGG
TGTGGAATGTCGTCACTGGCAAACATTTCACAGATTATTTGTTCTGCTTCACATTGACACT
GTGCTAATAGTTATTCAGTACATGAAAAGATACTACTGTGTTGAAAGCTTTAGGAAATTGACAGT
ATTTTGTCACAAACATTGAAAGAAACTTGTAAATTCTATTCTATTAAATTGCCAATGTCAATA
AAAAGTTAAGAAA

LDATSPSVFGAEIAKTEGLEMADHMMAMNHGRFPDGTNGLHHPAHRMGMQFPSPHHQQQQPQHAFNA
LMGEHIHYGAGNMATSGIRHAMPGTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQKLNN
QYFNHHPPYPhNHYMPDLHPAAGHQMNQHFRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSSGGAGSSNS
GGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLSLVIEMGLDRIKELPELWLQNEFDFTDFVCKQ
QPSRVSC

AUG initiation = 28.5 kDa
CUG initiation = 30.6 kDa

+++++

#29

NM_182603

CUG initiated extension conserved in mammals but maybe not in Arfotheria (mouse, dog, orangutan, cattle, etc, have delimiting upstream in-frame stop codon).

ANKRD42 ankyrin repeat domain 42

Synonyms: SARP; FLJ37874; ANKRD42

GCGACGCCCTGCTGCCCTCCAGCCAAGTGGCTGGAGTCGGGAGGCTCGAAAGAGACTCCGAGAAAGTAC
CAGCGGAAGGCGGGCGCCGCTACGGCGATTGCGAGGGAGTAGCAGACGAAGACGGTGGCCGCGCACTAGC
CACCACGTGTGGAGGATAAACGGCTACACGGCATTCCGGCGCGAGTCTAGGAAAGAGTTAGCGACGA
CGGGGAAAGAAAATGTGAAGAGAGCGACCGCCGCTCCAGGGTCGCTGCAGGAAGCTAAGTGCAGACGCCG
GCTTCTCCCGCAGTGACTTGAGAAGGGTCAGTGAACCTCGGCCACTGCCGCAGCGTCTCTAGGGAGAGA
GTTAGGGGAGATACTGGCACAGTCACAGCTGCTTGGGAGAGAGTTAGGGGAGACAGCACCTTCTGCAG
CAGCAGCTGAATTAGTGAAGTTGGAGGCCACAAACTACCGACTCAGGGGAACAGCCAGAGAACCC
GAGGCCTCCGCCTCAGTGGCTCTGGGAGGGAGTCAGTGACATTGGGACCCGCAACTAGTGACTTCGGG
GATAGAGTCAGTGACGATCGCAGTCGCCCTCAGTGGCTCTGGGAGGGAGGTGTCGAAGGCGGCCA
CAGCGTTGGTAGTTCTGGGAGGAAGTAAGTGGAGACCGCGGCTACCGCAGCCAGCGACTCCTCTGGTGA
GCGCAGTGAAAGACGCCAGCTACCGCTCAGTGGCTTGGGAGAGAGAAAGTGAAGACGAAGGTTCCGC
TGCAGCTCTGGGAGAGCAAGAGAGGACCTTGGGCCCTAGTGACGACGGAGAAGAGGGCGCTG
CCGCTGCAGTGGCTCGTGGGTGAGAGCAAGTGAAAGACCGCCGAGCATAGGGGCTGGACTCAACTCCTC
CCCAGAGTCGGAGGTGTCGCCATGCCGGGGTGGCAATTCAAGGCCCTCCACTTCCTCTAGGGAGACT
GCAAACCCCTGTTCCAGGAAGAGCAAGTGCATTGGCAGCATACATGATGCACTACGGAGCTGGAGATGTAA
GCAGCTTCAGAAATAGTGTGCTTCATTGGCTGCTGGCATGGAGCTGATATCACACACGTAACAACGA
GAGGTGGACAGCATCTCACATAGCTGCAATCAGGGTCAGGATGCTGTACAGGCTTATAATGAAT
GGAGCAAATCTGACAGCCCAGGATGACGGGGATGCACTCCTTACATCTGCTGCAACTCATGGACATT
TTCACTTACAATAATGCTCCGAAGTGGAGTGATCCCAGTGTGACTGATAAGAGAGAATGGAGACCTG
TGCATTATGCACTTTCATGGCGGCTTGGCTGCTGCAACTTCTGTTAAATGGGTTGAGCATAGAA
GATGTGGACTACAATGGAAACCTCCAGTTCACTTAGCAGGCATGGAAGGCCACCTCACTGTTCAAATT
CCTAGTCAGTAGAATGAGCAGTGCAGCAGTTAAAAGCTTCAATGATAATGGAGAAAATGACTGG
ATTGGGCCAGAGGTTCTCAAGCAGAACATTACAGTTATCCAGGGGCTGAGTATGAAGGAAAAGAC
CTAGAGGATCAGGAAACTTACAGTCACTTCCAGGTCACTGTGGCTGCCTTAAAGGGTGATTGGGATGCTTAA
GAAATTAGTGGAAAGATGGAGTAATCAATATTAATGAGCGTGTGATAATGGATCAACTCCTATGCATAAAG

CTGCTGGACAAGGCCACATAGAGTGTTCAGTGGTTAATTAAAATGGGAGCAGACAGTAATATTACCAAC
AAAGCAGGGGAGAGACCCAGTGATGTGCAAAGAGGTTGCCATTGGCAGCAGTGAAGCTTAGAGGA
GCTACAGAAATATGATATAGATGACGAAAATGAAATTGATGAAATGATGTGAAATATTTATAAGACATG
GTGTTGAGGAAGCACTGATGCCAAGGATGATTTATGCTGAGTGACTGGATAAAACAGATGCCAGAAGA
CCATCAAAGAACTGCAGGGCCAGCTGGAGTGAACGACTACGTAGAGAAAAATAGAATGTCAGCTTGAT
GAATATCGAGCAGAAGTTGATCAACTCAGGGAAACACTGGAAAAAATTCAAGTCCCAAACCTTGTGGCTAT
GGTTGGTGCCTTTAATACATTATTTCTCAAGAAGTATATACAAGAGTGGCCAAGAGTACAAGCTT
TGGGCTGGGCTTTGGACTTGAATCCTGGCTATAGTCTTTTGATTGGCTACTCATTAAACCTTT
TGTGCCTCAACTATAAACTGGAGATGATAACTTATACCTTCTAGAGCTGATGACAGGATTAAGGAATACA
CACACACACACACACACACACACACACACTCTATATGTAACTATAACTTCTAGATACATACA
CACATCTGTCTATCTATCTCAAACAGCAGCCTAGTCATAAGTATTGTTAACATTGTACCATAGAAG
GAGAAAATGTTTCTATAAGTAATCAGAATTCTAAGACAGCTAGAGGATATGTATATTTAATATTCT
AAATAACCAAAATAAAAGTGTGTTGAATGGAATCCATATTCTTCCATAGGGAAAGTTCTCATCAATCA
TCATGGATGAATTAATTCTGTTGAGGCTTGTGTCACCTCAAATCTGATCTATTAATATTAGAATCTC
CTGGCTCTTTTATTTCACAAACTAGTCATTGGCTCTTAAATAACCCTTAAATTTAAATTTAA
TAAAACATGTTGTTGTATAAGTGTAAAACAAGATAGATGTTCTCTGACACTAAGTCCACTCTCCAG
AGGAAGCTATTGTTAACATTAGTACATACCATATTAGACAAAATTCTATTATCTGCAAACATGTATGT
GTATTCTATGCATATGCAAATATAACCACCTCCTCAGTTCTCTGGGTTGGTGTCTCACAGTGGTTG
ATTTCCAAGAAATATAATTAACTATAATAATT

LERDSEKVPAGEGRRYQDSQVVADEDGRRRTSHVWRINGLHGHSGAESRERVSDDGERKCEESDRRSRA
AGSLSDAGFSRSDLRRVSENLGHCRSVSREVRGDGHSHCSWERVRGDSTFCSDVNFSVGHQTTD
SRGTAREDRLRLSGPWEGVSDIRDPRTSDFGDRVSDDRSRRFSGSWEGGSVEGGHSVGSSWEVSGDRGY
AASDSSGVSGSEDASYRFSGWERESEDEGFRCFSWEREADLGPRPSDDGEGRCRCSGSWVRASEDRRS
IRGLDSTPPQSRRCCAMPVANSGPSTSRETANPCSRKKVIFGSIHDRAVAGDVVKQLSEIVCLHWLLWHG
ADITHVTTRGWTASHIAAIRGQDACVQALIMNGANLTAQDDRGCPLHLAATHGHHSFTLQIMLRSGVDPNV
TDKREWVPHYAAFHGRGLGCLQLLVKGCSIEDVDYNGNLPVHLAAMEGHLHCFKFLVSRMSSATQVLKAF
NDNGENVLDLAQRFFKQNILQFIQGAEYEKGLEDQETLAFPGHVAAFKGDLGMLKKLVEDGVININERAD
NGSTPMHKAAQGQHIECLQWLKMGADSNTNKAGERPSDVAKRFAHLAAVKLEELQKYDIDENEIDEN
DVKYFIRHGVEGSTDAKDDLCLSLDKTDARRPSKNCRASWSMNDYVEKN

AUG initiation = 43.1 kDa
CUG initiation = 76.1 kDa

+++++

#30

NM_014310

CUG initiated extension conserved in mammals.

RASD2 RASD family, member 2

Synonyms: Rhes; TEM2; MGC:4834; RASD2

TAGGAGCTGCTCCTCATCACTGGAGTCGCCCTACCTCTGCCCCAGCCGAGAGCCCCAGGCAGGG
CCCCGGATCGGACGTCCCCAACGCCCTCGGGCACCTGGCTCAGCAGGAGGCCCGCTCGGGCAGGGCA
GGGCCGGCGGGCGAGCCGAGGCCGCCCTGCCCGGCCAGGCCCTCGGAGGCCACCCATGG
GCACCTGCCCTTGCCTCCTGCCCCGCCGCCAGCCGGCTCCGAGCAGCGCAGGGAGGATCC
CCGCGCAGTGACCCGGAGCACCACAGACTCTGGAGGCTGGCGCTGGAGCAGCAGGCAGCTCCCGC
AGCTCCGGCGCTCCAGGCAGCTCTGAGCGTCCCAGAGGCCGCCATTCCCAGCCCCGAGCC
ATGATGAAGACTTGTCCAGCGGAACCTGCACGCTCAGTGTGCCGCCAAAAACTCATACCGCATGGTGGT
GCTGGGTGCCTCTGGGTGGCAAGAGCTCCATCGTGTCTCGCTTCTCAATGGCCGTTGAGGACAGT
ACACACCCACCATCGAGGACTTCCACCGTAAGGTATAACACATCCGCGGACATGTACAGCTCGACATC
CTGGATACCTCTGGCAACCACCCCTTCCCGCCATGCGCAGGCTGTCCATCCTCACAGGGATGTCTTCA
CCTGGTGTTCAGCCTGGATAACCGGGAGTCCTCGATGAGGTCAAGCGCCTTCAGAAGCAGATCCTGGAGG
TCAAGTCTGCCTGAAGAACAAAGACCAAGGAGGGCGGAGCTGCCATGGTATCTGTGGCAACAAGAAC
GACCACGGCGAGCTGTGCCAGGTGCCACCACCGAGGCCAGCTGCTGGTGTGGCGACGAGAAC

CGCCTACTTCGAGGTGTCGCCAAGAAGAACACCAACGTGGACGAGATGTTCTACGTGCTTCA
CCAAGCTGCCACACGAGATGAGCCCCGCCCTGCATCGCAAGATCTCCGTGCAGTACGGTGACGCC
CCCAGGCCCTCTGCATGCCCGCGTCAAGGAGATGGACGCCATGGCATGGTCTGCCCTCGCCCG
CCCCAGCGTCAACAGTGACCTCAAGTACATCAAGGCCAAGGTCTCGGAAGGCCAGGCC
ACAAGTGCACCATCCAGTGAGCGAGGGATGCTGGGCGGGCTGGCAGTGCTTCAGGGAGGTGG
AGATGCCCACTGTGCATCTCCCCACCGAGGCCGGCAGCAGTCTTGCACAGACCTTAGGCACC
CTGGAGGCCCGGGCGCTGGCCTCCGCACATTGGTCTGCCTCTCACAGCTTCTGAGTCCGTT
ACAGCTCTGGTGGTTCATCTCCTCTGTGGGAGGACACATCTGCAGCCTCAAGAGTTAGGCAG
TCAAGTTACACCTCCTCTGGGTTGAAAGAAATGTTGATGCCAGAGGGTGAGGATTGCTGC
ATGGAGCCTCTGGACAAGCTCAGGATGAAAAGGACACAGAAGGCCAGTGAGAAAGGTCT
CTGGCATAACACCCAGCTGGTTGGCAGCTGGAGAACTTCTCCCAGCCTGCAACTTAC
TCTGGTCAGCTGCCCTGCACCCCTCCACCCAGCACACACACAAGTTGCCAGCTGC
CATTGAGGCCAGTGACTCTGTCTGAAGGGGCGTGGCACACCTCTAGACCA
CACGCCACCTCCTGACCGCGTCCCTCAGCCTCCCTAGGTCCCTCGGCCAG
GTGGTTGCAGCTGGTGTATGTTAGTAGAAATGGAAATCATTGTA
CTCCCTCTGGCAGGCCACCCAGTCAGATCCACGCCCTCACCGGG
CTCCCTCTGGCAGGCCACCCAGTCAGATCCACGCCCTCACCGGG
CCCAAACAGGGTTCGCTGGCTGTTGAGCTAGACATTGACCTCCGCCATTGAG
ACAATTGACAAGCGTGGGGGGCAGGCCAGGACTGCTTTTTA
CACCGAGACTCGGAGGGGACACGATGAGCACCAGGCC
TCCACCTAGAACCAAGGCTGCCCTACTGTGCTCG
GGTTCACTGTGCTCTAGGAAACCAAGAGGGTCCCAGTC
CTTCTGCACACACAGCGTGGGGAGGCGGGAGGAGCAG
GACAGAAAAGCTCAGGGCGGGCTCTCTGT
AGGCATGGTGGGTTTGAA
TTAACATAAACCAAGGCC
CCATGTATATA
AGGTTCA
CTGGGCATTCA
GGACATCTG
CCATG
GCACCTG
CCCT
CTCAC
GGAC
AKVLREGQARERDKCTIQ

AUG initiation = 30.4 kDa
CUG initiation = 40.6 kDa

+++++

#31

NM_002506

UUG/CUG initiation in mammals and most tetrapods

NGF nerve growth factor (beta polypeptide)

Synonyms: NGFB; HSAN5; Beta-NGF; MGC161426; MGC161428; NGF

CGGACCGTGCGCTGGGAGCCGGAGGGAGCGCAGCGAGTTTG
CTGGATGGCAGCTGGACCCAAGCTCAGCTCAGCGTCCGGACCA
TTCTATCCTGGCCACACTGAGGTGCATAGCGTAATG
TCAGTCCATGTTGTTACACTCTGATCACAGCTTCT

GATCGGCATACAGCGGAACCACACTCAGAGAGCAATGTCCCTGCAGGACACACCATCCCCAAGTCCACT
GGACTAAACTCAGCATTCCCTGACACTGCCCTCGCAGAGCCCGAGCGCCCCGGCAGCGCGATAGCT
GCACCGCTGGCGGGGAGACCGCAACATTACTGTGGACCCCAGGCTGTTAAAAAGCGGCAGCTCCGTT
ACCCCGTGTGCTGTTAGCACCCAGCCTCCCGTAAGCTGAGACACTCAGGATCTGGACTTCGAGGTG
GTGGTGCTGCCCTCAACAGGACTCACAGGAGCAAGCGGTACATCCCACATCCATCTTCCACAGGG
GAATTCTCGGTGTGACAGTGTGAGCGTGTGGGTTGGGATAAGACCACGCCACAGACATCAAGGGCAA
GGAGGTGATGGTGGTAGGTGAACATTAAACAACAGTGTATTCAAACAGTACTTTTGAGACCAAG
TGGCGGGGACCCAAATCCCGTTGACAGCGGTGCGGGGATTGACTCAAAGCACTGGAACATATTG
CCACGACTCACACCTTGTCAAGGCCTTACCATGGGATGCAAGCAGGGATAACCTGGGGTTATCC
CGGATGAGATACTGGCCGGGAAAAGTGGGCTCAACAAGGGAGGATGAAGAGAAGGGCTAAAGGCTGG
CGGAAGGACATCCCTCCCCCTGGGGCATTCAAAAACCTCTGCGGGGGCCCTCCCTAACCGGAAAC
CTGGTAAAATGGATTTTAAGGGTGGTAA

LASGRAVQGAGWHAGPKLSSASGPNSFTKGAAFYPGHTEVHSVMSMLFYTLITAFLIGIQAEPHSESNVP
AGHTIPQVHWTKLQHSLDTALRRARSAPAAAIAARVAGQTRNITVDPLRFKKRRLSPRVLFSTQPPREAA
DTQDLDFEVGGAAPFNRTHRKRSSSHPIFHGEFSVCDSVWVGDKTTATDIKKEVMVGRSEH

AUG initiation = 18.3 kDa
UUG initiation = 22.7 kDa

++++++

#32

NM_152283

GUG (or ACG) initiated extension conserved in mammals – AUG in several species;
GUG->AUG *Bos taurus*, *Sus scrofa*, *Mus musculus* (two transcriptional isoforms in mice
one with AUG and one with GUG)

In human two 5' variants – one adding in-frame AUG just in front of the GUG

ZFP62 zinc finger protein 62 homolog

Synonyms: ZET; ZNF755; FLJ11344; FLJ34057; FLJ34231; FLJ58781; FLJ59694;
MGC176438; DKFZp667F2013; ZFP62
GTGAGTGTGATGATAAGCTCCGCTGAGGTGAAATAGAGACCCGGAGGCGCGTCTAGCCCTCATCT
GGGGAAAGCGCACCTGCATACAGACGGGTGACCGGGGAGGAGGCGATCTGCCGCGTGTCTGCAAGCAGA
AAAGGAGTTAACTAGTGTACATTGAAAGACGAGCACTGAGGATGAGGAACCAACTGAAGAATATGAAAAT
GTTGAAATGAGCATCTAAGTGGCAAAAGTGGAGGATCCTATCCCTGAATCTAAGGTTGGTACACATG
TGTGTTGGGATAGCAAGGTAGAGAACTAACAGAAAAAGCCTGTTGAAAACAGGTGAAGGAGGACAAAGCA
GCATCAGGGAAGCAATCAGCAAAGCAAGAGTACAGCAAATATAAGACAGAACAGGAAGGTGAGGCATCT
GAGAAGAGCTTGCATCTGAGCCCACAGCATATCACACACCAGACTATGCCATAGGACAGAGAGGCAGTGA
GCAAGGCAAACGTGTGGAGAACATTAATGGAACCTCCTACCCTAGTCTACAGCAGAAAACCAATGCTGTTA
AGAAATTACATAAAATGTGATGAATGTGGAAATCCTCAAATATAATTCCCGCTTGTCAACATAAAATT
ATGCACACTGGGAAAAGCGCTATGAATGTGATGACTGTGGAGGGACTTCCGGAGCAGCTCGAGCCTTCG
GGTCCACAAACGGATCCACACTGGGAGAAGCCGTACAAGTGTGAGGAATGTGGAAAAGCCTACATGCTCCT
ACTCCAGCCTTATAAACCAACAAAGCACCCATTCTGGGAGAAGAACTGTAATGTGATGAATGTGGAAA
TCCTTCATTATAAGCTCTGTTCTGGACCAGCATAAAAGGATCCACACTGGGAGAAGCCCTATGAATGTGG
TGAGTGTTGGAGGCCTTCAGGAACAGCTCTGGGCTCAGAGTCCACAAAGGATCCACACGGGGAGAAGC
CCTATGAATGCGACATCTGTGGAAAACCTTCAGTAACAGCTCTGGCCTTAGGTCATAAAAGGATCCAC
ACAGGTGAGAAACCTTACGAATGTGATGAGTGTGGGAGGCCCTCATTACTGTAGAACACTTCTCAACCA
TAAAAGCATCCACTTGGAGATAAAACCTATAAAATGTGATGAGTGTGAGAAATCTTTAATTATAGCTCTC
TTCTCATTAGCATAAAGTCACTCCACACTGGAGAGAAACCTTATGAATGTGATGAATGTGGAGGCTTTC
AGGAACAGCTCAGGCCTCATAGTCATAAAAGGATCCACACAGGAGAGAAACCTTACAAGTGTGATGTCTG
TGGCAAAGCATTCACTAGCTATAGCTCAGGCCCGCAGTCCATAAAAGCATTCAACCTGGAGAAGAAAGCCCATG
AATGTAAGGAGTGTGGAAATCCTTAGTTAGTTAACTCACTACTTCTCAACACAGAACTATTCAACCGGA
GAGAGACCTTATGTATGTGATGTGTGGAAAACGTTAGAAACATGCAGGCCCTAAAGTCCACAGGAG

GCTCCATACTGGGAAAAACCATATAAGTGTGATGTGTGGAAAGCCTATATCTCACGCTCTAGCCTTA
AAAATCACAAAGGAATCCACCTGGGGAGAACGCCTATAAATGTAGCTATTGTGAGAAATCCTCAACTAC
AGCTCTGCCCTTGAAACAGCATAAAAGGATTACACCAGGGAAAAACCCCTTGGGTGTGAGTGTGGTAA
AGCTTCAGAAATAATTCTGCCCTTAAAGTACATAAACGAATCCACACTGGGGAACGACCTTACAATGTG
AAGAATGTGGGAAAGCATACATCTCTCTCGAGCCTTATAAATCATAAAAGTGTACACCTGGGAGAAG
CCCTTAAAGTGTGACGAGTGTGAGAAGCCTTCATCACATACCGAACCCCTACAAACCACAAAAAGTTCA
TCTTGGGGAGAAGCCTACAAATGTGATGTGTGAGAAATCTTTAATTACACATCGCTCCTTCTCAGC
ACAGAAGGGTCCACACTAGAGAGAACCCATGAATGTGACAGGTGTGAGAAGGTCTTCAGAAACAACCTCA
AGCCTTAAAGTCTATAAAAGAATCCATACCTGGGAGAGGCCCTATGAATGTGATGTGTGGAAAAGCCTA
CATCTCACACTCAAGCCTTATTAACCATAAGAGTACCCACCCCTGGCAAGACACCCCACATGTGATGAAT
GTGGAAAAGCTTTCTCAAGCAGAACCTTATAAGCCATAAAAGAGTCCATCTGGGAGAAACCCCTC
AAAGTGTGTTGAGTGTGGGAAATCTTCAGTTACAGCTCTCTCTCAGCACAAGAGGATCCACACAGG
GGAGAAACCCATATGTGTGATAGGTGTGGGAAGGCCTCAGGAACAGCTCAGGCCACAGTCATAAAA
GGATCCACACAGGTGAGAAACCCATGAATGTGATGAGTGTGGGAAGGCATACATCTCACACTCAAGTCTT
ATCAATCATAAAAGTGTCCACCAGGGGAAGCAGCCTATAATTGTGAGTGTGGGAAATCCTCAATTATAG
ATCAGTCCTTGACCAGCACAAAAGGATCCACACTGGAAAGAAGCATAACCGATGTAATGAGTGTGGTAAGG
CTTTAATATCAGATCAAATCTACCAAGCATAAAAGAACCCATACTGGAGAGGAATCTTAAATGTGATA
TATGTGGGAAGTTATAGTGGCACATCCAGAAGAGAACCTATGAGGGAGGGAAATGCCCTGGATGGGGCAG
GATGAGGATGCCTCTGAGCAGGAGCTTACCAAGTCTCCGAACCTCAAATGGAAGAAATACCTTATG
AATGTAAGAATGTAGGGGTATGGCTGTAAATTACACAGTGTAAATGAAACCATCCTAGAGGATTATGA
GGAATCTTCTATGTGATTTCAATCATAGCAAGCAAGAACGGCTCAGTGTCAAGTAGTTCAAGCTCTT
ACAGGATATAAAACAGTCATACTTGAGAGAAAACCTAGATCTGAGTGTGGAATGTGAAGCAAATCTT
AAAATCAGTAGACATTCTGACATAAAACACAGATGAGGAAGGGCTCAATTAGAAGTTACGTAATCAC
CATCAGAAAGTTCATGTTGTAATTCTGTTACTAGAAATGTAGGAATTTCAGGTATAGCTTGAATCCC
AATTACACATTGGTCAGTGGAAAACTAAGGGCTCCAACAGGAAATTCAAGGGAGGATAGGTTCAAGGG
ATATAAATTATTAAATTAGTGGCTTTAAGTATAAAACTTGATGTAATTGGTTGGAGGGGGCAGTGA
TGATGACTTCTGAAACAAAATTGGATTTCCTTTAGGAAAGTAGAAAGCATAGACTTACAAGTCTAAC
GGAGATAGGAGAGACTCATAAAAATGCAAATTGATGAACGTACTATTGTGATACATTAGTTGAATG
GATGAAACTTTTAAAGTTCAAGATGAACCTCCATAATGAATGATGAATTGTGATGAGGGATAACCTGG
AAAGTGGTATTACACATTATGCTACAATAAAAGTTCTACCGTGGAGAGGATTTGACACATTCACT
AATGGAACACACCGTCAACATGAATTCCACCTTACATGACAGAAGTGTGATTCAAGGATTCTATGAATAGA
AATGCTGAGAAGGAACGCATTATTGAGCAAGCTAAAAGCTAAAGTACCGTCACTCTAGAGAGAAGGAA
ATTAATGTTCTTAATAATCCTGTTAAATGTTGATTGTTGGATGTGTTATTGAAAGATGTCA
AGGACATGTATATGTTGCTGTTAAAATGTTAACGAAACTTGTCAAGGCTCACTCTCTTGTCA
TGAAAGCCAGCTCTGTGGCAGGTAAAGTGGATTCCAATAAGAAATTCTTAAATCAAAA

VSHLKTSSTEDEEPEEYENVGNAASKWPKVEDPIPESKVGDTCVWDSKVENQQKKPVENRMKEDKSSIREA
ISKAKSTANIKTEQESEKSLHLSPQHITHQTMPIGQRGSEQQKRVENINGTYPQLQQKTNALKLHK
CDEC GKSF KYN SRLV QH KIM HTGE KRYEC DDCGGTFRSSSSLRVH KRI HTGE KPYK CEE CGKAYMSYSSLI
NHKSTHSGEKNC KDEC GKSF NYSS VLDQH KRI HTGE KPYEC GECG KAFR NSSL RVH KRI HTGE KPYEC
ICGKTFSN SSG LRVH KRI HTGE KPYEC DECG KAFR NSSL GLVH KRI HTGE KPYK CDV CGKAF SYSS GLAV HKS
IHPG KKAHE CKE C
GKSFSYNSLLLQHRTIHTGERPYVCDVCGKTFRNINAGLKVRHRLHTGEKPYKCDVCGKAYISRSSLKNHKG
IHLGEKPYKCSYCEKSFNYSSALEQH KRI HTREK PFG CDECG KAFR NN SGLVH KRI HTGE RPYK CEE CGK
AYISLSSLINHKS VHPGEKPKF KDEC EKAF I T YR TL NHK VHL GEK PYK CDV CEK SF NYTS LLSQH RR
TREK PYEC DRCEK VFRNN SLL KVH KRI HTGE RPYECDV CGKAYI SHS S L INH KSTH PGK TP HT CDEC GKA F
FSSRTLISHKRVHLGEKPKFCKVECGKFSYSSLLSQH KRI HTGEK PYVCDRCG KAFR NSSL TVH KRI HTG
EKPYEC DECG KAYI SHS S L INH KS VHQ KQPY NC ECG KSF NYRS VLDQH KRI HTG KK PYRC NECG KAFN IR
SNLTKH KRHTGEESLNVIYVGYSQTSQKRTYE EGGN ALD GGR MRP L

AUG initiation = 95.6 kDa
GUG initiation = 102.4 kDa

+++++

#33

NM_001102654

UUG/CUG initiated extension conserved in mammals and likely most vertebrates

NTF3 neurotrophin 3

Synonyms: NT3; HDNF; NGF2; NGF-2; MGC129711; NTF3

```

AGTTGAAGCTCCTCTCCCTCCGAACAGCTCCGCCACCGCCCCGCGACGCAGCCGGCGCAACTACTTTC
TTCTCTCTCCTTCTTCTTCTCCTCTCCTTTCCCTGCTGGTAGTGCTGCGGGGGTGGGGAGACT
TGAAATGACCAGCTCGCTCCACCTTCTTCATGTCGACGTCCCTGAAACGCCACACGGATGCCAT
GGTTACTTTGCCACGATCTTACAGGTGAACAAGGTATGTCATCTGTTTATGTATATTCTCGCTT
ATCTCCGTGGCATCCAAGGTAACAACATGGATCAAAGGAGTTGCCAGAAGACTCGCTCAATTCCCTCATT
ATTAAGCTGATCCAGGAGATATTTGAAAAACAAGCTCTCCAAGCAGATGGTAGCTTAAGGAAAATTA
CCAGAGCACCCCTGCCAAAGCTGAGGCTCCCCGAGAGCCGGAGGGGGAGGCCAAGTCAGCATTCC
AGCCGGTGATTGCAATGGACACCGAACCTGCTGCGACAACAGAGACGCTACAACCTCACCAGGGTCTGCTG
AGCGACAGCACCCCCCTGGAGCCCCGCCCTGTATCTCATGGAGGATTACGTGGCAGCCCCGTGGTGGC
GAACAGAACATCACGGCGAACGGTACCGGAGCATAGAGTCACCGAGGGAGTACTCGGTATGTGACA
GTGAGAGTCTGTGGGTGACCGACAAGTCATCGGCATCGACATTGGGACACCAGGTACGGTGTGGG
GAGATCAAAACGGCAACTCTCCGTCAAACAATATTTTATGAAACCGCATGTAAGGAAGCCAGGCCGGT
CAAAACGGTTGCAAGGGTATTGATGATAAACACTGAAACTCTCACTGCAAACATCCAAACCTACGTCC
GAGCACTGACTTCAGAGAACATAAAACTCGTGGGCTGGCGTGGATACGGATAGACACGTCTGTGTGT
GCCCTGTCGAGAAAATCGGAAGAACATTGAATTGGCATCTCTCCCATAATAAATTATTACTTTAAATT
TATGATATGCATGTAGCATATAATGTTATATTGTTTATATATAAGTGACCTTATTATTAA
CTTCAGCAACCCTACAGTATATAAGCTTTCTCAATAAAATCAGTGTGCTGCCCTCAGGCCTCT
CCCATCTGTTAAACTTGTGTTGTGATCCGGCTCTCAGGAGTCACTCTGAAATCTGTGTACACCAGT
ATTGCAATTGATATTGTCAAGGCCATGACTGTTTTAGTAAACTTGTAAATCAAAAAAAAAAAAAA
LNDRARVHSLHVDPGNHDTAMVTFAILQVNKVMISLFYVIFLAYLRGIQGNMQRSLPEDSLSI
IKLIQADILKNKLSKQMVDVKENYQSTLPKAEPREPERGGPAKSFAQPVIAMDTELLRQQRYYNSPRVLL
SDSTPLEPPPLYLMEDYVGSPVANRTSRRKRYAEHKSHRGEYSVCDSESLWVTDKSSAIDIRGHQVTVLG
EIKTGNSPVKQFYETRCKEARPVKNGCRGIDDKHWNSQCKTSQTYVRLTSENNKLGVWRWIRIDTSCVC
ALSRKIGRT

```

AUG initiation = 30.8 kDa
UUG initiation = 33.3 kDa

+++++

#34

NM_003252

Possible GUG initiation. Strong conservation (in mammals) at least at nucleotide level.

TIAL1 TIA1 cytotoxic granule-associated RNA binding protein-like 1

Synonyms: TCBP; TIAR; MGC33401; TIAL1

```

TCTTCACGTCCCAGCGCGGGGGCGCCGGCTCCCTTAACCACAGGTTCCAGAAGTCCTCTGCAGAA
GTGCTTCCCTCTCATTTCAGGACCAATTCCAGAGACTTCGGCTTCAGACGTTCTCTTTTGCC
CGATCTCTCCGGAGCTGGCTGGGCTCGGCCGCCAGAGGCCACAGCGACGTGATCCGCGTGAGC
GGGTCCCAGGGTTCCCTCGGCCGCCCTTTCTCTCCCTCGTGTCCCCCTTGCAAGGCTGTCGGCT
GGCGCTGAGCGGTGACCTGGGGTCGCCCTGCGCTCTGCCCTGTTCTGCCTGGCTGGCGGGCC
ATTGGTTCATCTCCCTCCCTGCTCCTGGTTGGAGCGCAGTGTCCGGAGCGGGCTGGGGAG
AGAGCCCGAGAGCAGGGTCGGTGCCTTCTCTGTCCTCCAGCCGGTCCCAGAGCCCCCTCCCTCC
TCCCCACCCCCCTCCCTCCCCAACCTGCCCTCCCCCTGTCCTGGGATCGCTCCGTCGACACCACCATGA
TGGAAAGACGAGGGCAGCCCCGGACTCTATACGTAGGTAACCTTCCAGAGATGTGACAGAAGTCCTATA
CTTCAGTTGTCAGTCAGATTGGACCTGTAAAAGCTGTAATGATAACAGAGCATAAGCAATGACCC
ATATTGCTTGTGGAATTATGAACACAGAGATGCAGCTGCTGCATTAGCTGCTATGAATGGGAGAAAAA
TTTGGAAAGGAGGTCAAAGTAAACTGGCAACCACCAAGTAGCCAGAAAAAGATACTTCAATCAC

```

TTCCATGTGTTGGGATTGAGTCCAGAAATTACAACAGAAGATATCAAATCAGCATTGCCCTT
TGGTAAAATATCGATGCCGGTAGTTAAAGACATGGCAACTGGAAATCCAAGGCTATGGTTTGAT
CTTTTATAACAAACTGGATGCAGAAAATGCAGATTGTCATATGGCGTCAGTGGTGGGTGCGTCAA
ATCCGAACCAATTGGGCCACTCGAACACCACCTGCACCTAAAAGTACACAAGAAAACAACACTAAGCAGTT
GAGATTGAGATGTAGTAACACAGTCAAGTCCAAAAATTGTAAGTGTACTGTGAGGAATTGCGTCTG
GGTTAACAGATCAGCTTATGAGACAGACATTCTCACCATTTGGACAAATTATGGAAAATAAGAGTTTCCA
GAAAAGGGCTATTCTATTGTCAGATTCAACCCATGAAAGTGCAGGCCATGCCATTGTTGGTGAACGG
TACTACGATTGAGGACATGTGGTAAATGCTATTGGGTAAGAATCTCCTGATATGACTAAAACCTCC
AACAGGGTGACTATAGTCATGGGCCATGGAGCAAGTGTATGGAACCCCACAAACAGTATGGACAGTAT
ATGGCAATGGTGGCAAGTACCGCCTATGGAGTATACGGCAACCATGGAATCAACAAGGATTGGAGT
AGATCAATCACCTCTGCTGGATGGTGGATTGGTGCAGCCTCCCCAAGGACAAGCTCCTCCCC
CTGTAATACCTCCCTAACCAAGCCGATATGGTATGGCAAGTTACAAACACAGTGAGGGGACTCTA
AAAAAAATTGTAATTCATGATAGGCTCGATTTCTGTGACACTCTGAAGACATGAAAGTAGACATCGGA
AAATGAAAATTATTAAATTGAAATGTTGAAACCTTACAGCAGATTGCTTTGGTGAAGGACA
CGTGTCTCTAGTTCTGCCTTTAAGTTTGTTCATGATGGATATGAAACATGATTTCATGTACA
AAAACAAAATAAGTCATAAAAGACAATTCTGACTACAAATTGATATAAGGAAAATGGCTAATAC
ATTTGATTCTAGATACTATTCCATTTTATCTGCTGTTAGTATTAACTCACTGTGTTTAAAG
AGCAAAAAAGGGAGGATCGTAAAACCTGGGAATCACATATAAGTCATCCTGAATCCTGATACTCCCCTC
CCCTCCCTGAGGTGGACCACATTGAAGTCAGCAGAGAAAAGTGTGATATTGAGAAATGCGTGATT
TTGGAGTCGTTGGAGGAATATTTCTTCTATGCCTAACAGAAACTGAAGCCAGACTGAAGTTTGC
ACCCTAAAAAGGAACAGCATTGTTGAGTTACTGAGCAAATGTTGGTCCACGTTAACAGACATATT
TAAAACCTCCAAAAGTGTGATTATAAAATTGAGTATTTCATTTGATTTGGGGAAATCCAAGTA
TGGTGTGTTGATTGAAGTCAGACAGTCATACTTGCTTTACATGAAGTTAACATGATACATATTGAA
TATTCAATAACTACAGTTAAAAGCATGCTCAACATAGAAGTAGCAGAACATGTAATTATTGAA
ACACTAACACACTCCGCTGATTGAATGCAGTGGATTGATCAGAATGTTAAGACTGACATTCCAAGGTT
GGCTACTATGTAATTAAATTACACAAATTGTCAGAAAAGCCTTAATTAAATTTCATACAATCTT
GATGCATTAGTATGTTCTAAATGTCATTGGAAATTAGTTTGTGTTTTGTTTTTTTTTTTT
TGCTTACATTCTGGTATGTAATAACCTGATTAAAACCTTGTAAACCAATTCAAGGTTACTATAAGT
TGTATAGTACAAGTGTGTTTAAAATCTTGGGTTTTAAAATTAAAGATATAATTGCCAAGAATT
TTTTAACAGATTGCTAAAACATCTTATTAGACACTTCATGTACCAATTATAATTGGATATTCA
TTAAATAGTACACAGAGTTGGCTTTATTTCATTAATTTTCTTGTGGCAGTGTGATGGTAT
AATAAGCCTGAGCAGAGGCTTAAGTTGATGTCAGAGTTGAAAGGAATCAATTGGAAGATGCAGAA
GACCGAGGTTGCTTCAAGGTATTTTCAGGCTGTGGTAAAATTGCTCAAATTCTATCAAACAG
GAATGTAATTAGATAAAATCCTATGTATTGAATTGTCAGAGCTAGGGAGTGCAGATTGTTGGCAATGT
ATTCAAAATGCTGGCTGGCACCAAGAGAAAATAGCCTTTACAGTTACATAGTAAGATGCGATTAGTA
CCCACAAATTACTGTTCTAAACATTGAAGTTACGATTAGCTTTAAAATAATGATTATAATTGG
TGGTCACAATAATTGGTATTACTTCCCTTCCACTAGCAATATAGCAGATTGTTACCTAACATA
AAAATTCAAGGGCTGAAATTCAAGCTGGCCAAATTGTTATGGCACCTAGTTACCTAACATA
ATCTATTACACTCTCTGTTATAAAATACCTTATTCTTTGTTATGTTACCTGGAGGAATTGTTT
TGGAAAAGTTAATTATTTATAGGGAGAACACTCAATAAAATTATGTTACCTGCCCCGAGTTAAA
ATTTTATGAGTATATGTAACACTGAACAACGTAAAGACTTTTTAATTGATAAAAATGCTTAGTATGCCT
GTTTGGTCTGCCAGTAATTAAAGTAGCTTATTGAGATAACTAACAGCTAAATAGCTGAGTGT
GACTGTATTCTATGATTAAATAAAATTATCCAGACTAGTTATTGCCACAGTAAAAA

V
TWRVAPALCPVSCLAGGGHFVHPPPPPAPPWSAVSGAGWGERAREQGSVPFPLSPAGAQSPPLPPHP
LPSPLSPLSRDRSVAPTMEDDGQPRTLVGNLSDRVTEVLILQLFSQIGPCKSKMITEHTSNDPYCF
VEFYEHRAAAALAAMNGRKILKEVKVNWATTSSQKKDTSNHFHVFGDLSPEITTEDIKSAFAPFGKI
SDARVVKDMATGKSKGYGFVSYNKLDENAIVHMGGQWLGRQIRTNWATRKPAPKSTQENNTKQLRFE
DVVNQSSPKNCTVCGGIASGLTDQLMRQTFSFGQIMEIRVFPEKGSFVRFSTHESAAHAIVSVNGTTI
EGHVVKCYWGKESPDMTKNFQQVDYSQWGQWSQVYGNPQQYQYMGANGWQVPPYGVYQWPWNQQGFGVDQS
PSAAWMGGFGAQPPQGAPPVIPPPNQAGYGMASYQTQ

AUG initiation = 41.6 KDA
CUG initiation = 50.6 KDA

+++++

#35

NM_024794

ACG initiation conserved in mammals

EPHX3 epoxide hydrolase 3

Synonyms: ABHD9; FLJ22408; MGC131519; EPHX3

CCATTTATTCCCCCTTCCAACCTCTCCTCCCCAGCCTCGCACCTACCCTGGTTCCTGTCCCTGT
CGCGCCCAGGTGTTACCTGGCACTCAGGTGAGTGGTGCCTCTGGCTGTTCTGTCGGAGCCGCCGCC
TCTTCCCTCAGCGCGTCCCACAAATCCGACGGCAGGGAGGGCCAGGCCAAGGGCGATGGGCCCTGA
GCCCTGACACCGCTTCGCCGCTGCTGAGGTGCCCTGGCCGGCAGGCCGCCGGTGGTCCCAGGCC
GACATGCCGGAGCTGGTGGTACCGCGCTGCTGGCGCCGTGGCGCTGTGCTGAAGCTGCTGCCGC
CATGTGAGCCTGGTGGTCTCGGTGGCGCTGGTGCCGCGGCTACGGCTGCATAGCGCTACGCAC
TGCTGTGCCGGCCCGGGCGCGCTGCTGCGGGCGCCGTGGAGCGCTCCCCCGCCTGCCTGAGCGAC
TCGCTGGGTGAGCACGGTTCTGAACCTCAAGAGCTGGGCTGCGCTGCACTATGTCCTGGCTGGAC
AGGTAACGGACCCCTCATGCTGTTCTGACGGCTCCCTGAGAACACTGTTCTGGCGTTACAGCTCC
GGGAGTTCCAGAGCCGCTTCATGTTGGCTGGACTTGCGAGGCTATGGCCCTCGGATGCACCTCGG
GATGTGGAACGCTACACAATCGACCTGCTGGTGGACATCAAAGATGTCATCCTAGGCCTGGGTTACTC
GAAGTGACATCCTTGCGGCCATGACTGGGTGCCCTCCTGCGCTGGCATTTCTCCATCTACTACCCATCCC
TGGTCGAGCGGATGGTGTGGTCAGTGGTGCCCATGTCGGTGTACCAAGAACTATTCCCTGACCAACATC
AGCCAGTTCTCCGTTCCACTACATGTTCTGTCAGCTGCCCTGGCTGCCGAGAACAGGATCCATGCTGAC
GTCTGACTTCAGATTCTGAAGACCACCTCACCCACCGCAAGACAGGATCCATGCTGACCCCCAGCG
AGCTGAGGCCTTCTTATAACTCTCACAGCCTGGTGGCCTCACTGGGCCCTCAACTACTACCGAAAC
CTCTTCAGGAACCTCCCCCTGGAACCCCAGGAGCTGACCACACCCACATTGCTGCTGGGGGAGAACAGGA
CACTTACTGGAGCTGGGCTGGTGGAGCCATCGGCAGGCCCTTGTGCGGGCCGCTGGAGGCCCACA
TCCGCAAGGCATAGGGCATTGGATCCACAGAGCAACCCCCAGGAGATGCAACAGTACATGTCCTTC
TTGCAAGACCTGCTGGAC **TAG** TGGTCCTTGCTGGCTGCCAGGAGCATGGATACTCAGGAAGGAACACA
CACCCATTATCTGTTGCTGGAGTCATACATGGTGAACCTTGAATCGCTCATA
GGCATGGACTCTGGATCCACACAAGCGCACCTATGGGTGCGCTGGACACACCAACCCCTATACTCAC
ACACAGGCATGGATGCACTGTTGTAACAAACACTTGACCCCTGGAACTGGGTATACTCTCTTCCAGT
GGAGCCAGATGCCAGACCGAGTGTCCACCTCCCTGGGCTCACTCTGCGCTTGCCAAAGT
CGCTTCTCTGCCATAGCTGCACAGACCTTAAACCTGACCTTCTGCTGCCCTGCACCTCCTGCCTGG
GTCTTCAGCTCAGTGCTACTCTGAATAACAATGCTGACCACATTGGCTACTTCAGCTAAATGTATGT
TAATTGAAATGAAATACAATGAAAATGCAGTGCTCAGTCCCA

**TEGPQAKGDGPLSPDTASPLLQVPLAGSAAVPERGDMPELVVTALLAPSRLSLKLLRAFMWSLVFSVALV
AAAVYGCIALTHVLCRP芮GCCRRSASPACLSDSLGEHGLNLKSSGLRLHYVSAGRGNGPLMLFLHG
FPENWFWSWRYQLREFQSRFHVVAVDLRGYGPSDAPRDVDCYTIIDLLLVDIKDVILGLGYSKCILVAHDWGA
LLAWHFSIYPSLVERMVVSGAPMSVYQDYSLHHISQFFRSHYMFLFQLPWLPKLLSMSDFQILKTTLT
HRKTGIPCLTPSELEAFLYNFSQPGLTGPLNYRNLFRNFPLEPQELETTPTLLLWGEKDITLELGLVEAI
GSRFVPGRLEAHILPGIGHWIHQSNPQEMHQYMWAFLQDLLD**

AUG initiation = 40.9 kDa
ACG initiation = 44.5 kDa

+++++

#36

NM_018646

ACG initiated extension conserved in mammals. Extension initiated by AUG in bats:
Myotis lucifugus and *Pteropus vampyrus*.

Synonyms: **TRPV6 transient receptor potential cation channel, subfamily V, member 6**

CAT1; CATL; ZFAB; ECAC2; ABP/ZF; LP6728; HSA277909; TRPV6

AGAGTCCTGGCTGGCTCTGCCAAGTGTAAACAAACTCACAGCCCTCTCCAAACTGGCTGGGCTGCTGGGAG
ACTCCCAAGGAACCTCGTCAGGAAGGAGAGACAGGAGACGGGAGACGGTGGGCCGGCC
CTTGGGGGGCTGATGTTGGCCCCAAGGCTGAGTCCCCTCAGGGTCTGGCTCGGCCTCAGGCCCAAGGA
GCCGCCCTACACCCCCATCGGTTGTCACTGCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGT
TCTGCAGATGGTCCAGAGACGGGAGTCTGGGCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAG
AGGATCTGGGAGTCTCTCTCTAGCTGCCAAGATAATGATGTCAGGCCCTGAACAAGTTGCTCAA
GTATGAGGATTGCAAGGTGCACCAGAGAGGGCCATGGGGAAACAGCGCTACACATAGCAGCCCTATG
ACAACCTGGAGGCCCATGGTGTATGGAGGCTGCCCGAGCTGGCTTTGAGCCCATGACATCTGAG
CTCTATGAGGGTCAAGACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCAGCCCTGCT
TGCCCGAGGGCCAGTGTCTGCTGCCAGAGCCACAGGCACTGCCCTCCGGTAGTCCCTGCAACCTCATCT
ACTTTGGGAGCACCCCTTGTCTGCTGCTGTAACAGTGTGAGGAGATCGTGGCTGCTCATTGAG
CATGGAGCTGACATCCGGCCCAGGACTCCCTGGAAACACAGTGTACACATCCTCATCCTCAGCCAA
CAAAACCTTGCCTGCCAGATGTACAACCTGTTGCTGCTCACGACAGACATGGGACCACCTGCAGCC
TGGACCTCGTGCCTGCCAATCACCAGGGTCTCACCCCTTCAAGCTGGCTGGAGTGGAGGGTAACACTGTGATG
TTTCAGCACCTGATGCAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGACTCTATGA
CCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACCTATCATCACCACCAAGAAGCGGG
AGGCTGCCAGATCCTGGACCAGACGCCGTGAAGGAGCTGGTGGCTCAAGTGGAAAGCGGTACGGGCGG
CCGTACTCTGCATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTCACCATGTGCTGCATCTACCG
CCCCCTCAAGCCCAGGACCAATAACCGCACGAGCCCCGGACAACACCCCTTACAGCAGAACGCTACTTC
AGGAAGCCTACATGACCCCTAAGGACGATATCCGGCTGGTGGGGAGCTGGTACTGTCATTGGGCTATC
ATCATCCTGCTGGTAGAGGTTCCAGACATCTCAGAATGGGGTCACTCGCTCTTGACAGACCACCT
TGGGGGCCATTCCATGTCCTCATCATCACCTATGCCCTCATGGTGTGGTACCATGGTATGCGGCTCA
TCAGTGCCAGCGGGAGGTGGTACCCATGTCCTTGCACTCGTGTGGCTGGTGAACGTATGTACTTC
GCCGAGGATTCCAGATGCTAGGCCCCCTCACCATCATGATTGAGAACATGTTTGGCGACCTGATGCG
ATTCTGCTGGCTGATGGCTGGTACCTGGCTGGCTTGCCTCAGCCTTCTATATCATCTTCCAGACAGAGG
ACCCCGAGGAGCTAGGCCACTCTACGACTACCCATGCCCTGTTCACGACACTCGAGCTGGCTTAC
ATCATCGATGGCCCAGCCAACACAGTGGACCTGCCCTCATGTACAGCATCACCTATGTCGCTTGC
CATCATCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGCGACACTCACTGGCGAGTGGCCC
ATGAGCGGGATGAGCTGTGGAGGGCCAGATTGTGGCCACCGGTGATGCTGGAGCGGAAGCTGCCTCGC
TGCCTGTGGCCTCGCTCCGGATCTGGGACGGAGTATGGCTGGAGACCGCTGGTCTGCGGGTGG
AGACAGGCAAGATCTCAACCGGAGCGGATCCAAACGCTACGACAGGCCCTCCACACCCGGGCTCTGAGG
ATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGCTGTCCTCAGCCCCACCTGTCCTTCTATG
CCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTCCAATTGGGAAAGGCTGGCAAGGGACCTGAGGAG
AGACCTGCGTGGATAATCAACAGGGGCTGGAGGACGGGAGAGCTGGAAATATCAGATCTGCGTG
TTCTCACTCGCTCCTGGAACTTGCTCTCATTTCTGGTGATCAAACAAAACAAAACACCC
AGAGGCTCATCTCCAGGCCAGGGAGAAAGAGGAGTAGCATGAACGCCAGGAATGTACGTTGAGAA
TCACTGCTCCAGGCCGTGCTTACTCCTCAGCTGGGAGAGGAAGGCCAGGCCAAGCACGGGCTGGC
AGGGCGTGAGGAACCTCTCTGTGGCTGCTCATCACCTTCCGACAGGAGCACTGCATGTCAGAGCACTT
AAAAACAGGCCAGGCTGCTGGCGCTGGCTCAGGCTATAAGTGGGAGAGAGGCCCTCCCA
GGCACCCAGGAGGTGAGGAACTGCAAGAGCTGGAAAGCGTGTGAGTGGAGGAGACAGGAACGGCT
CTGGGGGTGGAGTGGGCTAGGTCTGCCAATCAAAGTCGTTTCCGATCCCTGAAAA
AAAAAAAAAAAAAA

TGPLQGDGGPALGGADVAPRLSPRVWPRPQAPKEPALHPMGLSLPKELKLCLWSKFCRWFQRRESWAQ
SRDEQNLLQQKRRIWESPLLAAKDNDVQALNKLKYEDCKVHQRGAMGETALHIAALYDNLEAMVLMEAA
PELVFEPMTSELYEGQTALHIAVVNNMNLVRALLARRASVSARATGTAFRRSPCNLIYFGEHPLSFAACV
NSEEIVRLLIEHGADIRAQDSLGNVLHILILQPQNKTFACQMYNLLLSYDRHGDHLQPLDLVPNHQGLTPF
KLAGVEGNTVMFQHLMQKRKHTQWTYGPLTSTLYDLTEIDSSGDEQSLELIITTKKREARQILDQTPVKE
LVSLWKRYGRPYFCMLGAIYLIIICFTMCCIYRPLKPRTNNRTSPRDNTLLQQKLLQEAYMTPKDDIRL
VGELVTIGAIILLVEVPDIFRMGVTRFFGQTLGGPFHVLITYAFMVLVTMVMRLISASGEVVPMSFA
LVLGWCNVMYFARGFQMLGPFTIMIQKMIFGDLMRFCWLMAVVILGFASAFYIIFQTEDPEELGHFYDYP
ALFSTFELFLTIIDGPANYNVDLPFMYSITYAAFAIIATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVA

TTVMLERKLPRCLWPRSGICGREYGLGDRWFLRVEDRQDLNRQRIQRYAQAFHTRGSEDLDKDSVEKLELG
CPFSPHLSPMPSVRSRSTSRRSANWERLRQTLRRDLRGIIINRGLEDGESWEYQI

AUG initiation = 83.2 kDa
ACG initiation = 87.3 kDa

++++++

#37

NM_033315

non-AUG extension in mammals - multiple potential initiation start sites

Synonyms: RASL10B RAS-like, family 10, member B

RRP17; MGC47540; VTS58635; RASL10B

GGCCTCCGGAGGGAGAGCTGGGCTTCTACCCCCCTCGGCGGGCCGCATCTGCCCGCGCGCCC
GCCCTGAAGCCCGCCCCGACTGGCAGGCGGGGAGCCCTACTTCTCTCCCCCGGGCGGGGAGCCGGGG
GGCAGCGCCGGAGCCGGGGAGCTCAGCCCCGCCGACCGGCCGGCAGGGCAGGGGAGCTAGGACGG
CCCCGGTCCAGGTGGAGGCCAGAGGGCCCAGGGCAAGCAGAGGCAGACAATGGTTGGTCTGACGGTGGC
TGAGCCCCCAGCCCTGGAATATGCAAGCCGGGGAGCCCTAGACAGGGCAAGGACAGGGTGGCGGAGTG
GGCGGGAGGCATGGTCTCCACCTACGGGTGGCGTCTGGGGCGCAGGGTGTGGCAAGAGTGCCATC
GTGCGCCAGTTCTGTACAACGAGTTCAAGCAGGTCTGCGTCCCCACCACCGCCCGCCTTACCTGCC
TGCTGTCGTATGAACGGCACGTGCAAGCACCTCAGATCTCGACTTCCACCCATCAGGCCCTCCCTG
TCAATACGCTCCAGGAGTGGCAGACACCTGCTGCAAGGGACTCCGGAGTGTCCACGCCATACCTGGTC
TACGACATCTGCTGTTGACAGCTTGAGTACGTCAGACCATCCGCCAGCAGATCTGGAGACGAGGGT
GATCGGAACCTCAGAGACGCCATCATCATCGTGGCAACAAAGCGGACCTGCAGCGGGACCGTGTACCC
CGCGCTGGAACGTGTCGCACCTGGTAGCAAGACCTGGAGTGTCCACGTGGATGCTCGCCAAGTAC
AACTGGCACATCTGCTGCTCTCAGCGAGCTGCTCAAGAGCGTCGGCTGCGCCCGTTGCAAGCACGTGCA
CGCTGCCCTGCGCTTCCAGGGCGCGTGCACCGCTGCGCCATCATGTGAAGCCTGCGGCCCTC
GGGCTGACCGGCACTGGCGAGCGAGGGCGGGCGTACTGCGGGCTGGGGCGGGAGCGGGCGGGAA
ATGGAACGTGACGGTCCCCTGAGGCCCTGAGCCACCGCACCTCCCGTGAGAACAGAGCGCGAGA
GGGAGCCCTCGTAACCTGCCAGCCCTGCCCCCTGCCCCCTGGCTTCTGGACAGCCGCTTCAGTGCT
GTATTAGTGCAGTGCCTGGGGGGGAGCCCGCGGGGGTGCCACAGCCTTGGGATGGGGTGAGCGTGAAT
GGAGGCTGGGGTGGCGAGGTGCGCCTTGGCGGGCCCCACGTGTCTTCTCAGAATGTGTCTGTCTT
GCCTGGTGTCTCTTCCCGTGTCCGCCACCCAGCGTCTGTTGGTACTTACCTGTCTACCTACCCCTC
CAGTCCCTCCAGCTCGCTCACAGGCTCTCATTCGTCCATCCCTGTCAGATCTGGCAGCTTC
TTTGTGAGGCCAGGCTCTGACTGTCAAGCACCAGGCAAGGGCAGAGATGCGGGTGGCCAAGGACCA
CGATCAAGGGTCCGGGGACCGAGGTCCCAGATCAGTGAGGGAGAAGGTTGAGCTCTCGGCTTCCAGG
GAGACCTCCCCGCCAGCAGCCCCCAGAGACACAACACTACCTCCAGCCTTAACTCGATGGTCCGTCC
CTGCCAGGTGCCCTCACTCTCAGGCCAGATCACCCCTGGGTTAAAACTTTTCTTT
TTTTTTTGACAGAGTGTGAAAGGGAGCCCCCAAAGGATAGCTTCTTTCATGATGCCAGGCTCCAG
TCCTTATTCCCTCTGCATACTGCAATCTGATCTGCAGACTGGGAATGTTGGTCTGGGTCTGGTC
GTGGCAGGATGGTGCCTAGGGGGTTAGGTTGCTCCAGTGAAATTCTGTTGCCGTCTCAACCCCA
TCTGACTACCCCAGACTCTGCCCTGCTCAGATCTCAGACTATCCTGATTAATCTGGGAAGAACAGAGCCA
GGGAAAGAATGGTGGGGACCCCTGTACTTGGGGAGACACACCTGCATCTCCTCTGCCACAGATGGAGG
CCCTCAGGATCTGACACCCCTTGTCCCAACACCAGTCAGCCCTACCCCTAACACTCACTCCACCCCATTT
CTCCGGCTGCCCTGGCGGGTTTCTACCTCTCGTCACCGGAGCTGATCACTGTCAGTTGTACCGATTTAG
AAATAACAATAATAATGAAGATTCTAGGAATGGCATGAGGGATTGATGGGGACTGGAGGGAGGGACAAG
TGGTGCCCTGTCCCTGCTCCCTGGCAAAGAAAGCTGTCTTGAGGCTGAGCCCTCAGCCCTGGCTGG
TGGGGGGACAGCAAGGTCCCTGTTATAAGAGGGCAGAGAGGGACAACCTCCGCTTGGCCAACCTAGCCAA
GGCTGCAGCATATAGACCAGGAAATCAGGTAGCCAGACTGGTGTGGAGCAGAGTCTGGGGGAAGGGTCG
TGGGTGGGAATTATCACCAACATCCATTGTAGGGGAATCTATGATTCTGCTTCCCAGCGGATTC
CTCTGTCACCAACTGGGGGTAGCACAGCCTCACAGCAACCCCTGACCTGGCAGTCTAGTGTCT
GCATTCTAGTCCCTGCTGTGCAGGACTTGGCAAGTGACCTGCCCTGTGAGCCTCCCTGACAC
AGAGGAGGTGGCTCCCTCCCCACACCTTAGAGTGGCTGGAGGGTAACAAAGAGGGCTGCCCTTTAG
TCTCCTGACCCCTGCCCTGGTTACCCAGAGGGAGCGATGAAGGATGGCAGCATCTCACATGCCCAT

CACCAACTCTGAGGCACCTGGGTGGGGGGCGGAGCCAGGCCTCTGGCTGCTCCCCTGTGGAGCCATT
GGAATGTATCCCCTGACAGGCCCTCCGCCTCACCTAACCCAGGTCTGGATTTCAGGCCCTCCAC
CCCCATTCTGAGTCTCTGTCTCCTCTCCTCACCGCTCCAGGGTTCCCACCACAGGGCTGGAAGTGT
GTGTGACGCCATTGAGCTTACCGAAGTCAGATAAAATCAGGGAGTGTTCCTCGTTCTGTAA
AAAAAAAAAAAAAAA

LVLTVAEPPAPGICSPGEPEQTAARTRWRSGAGGMVSTYRVAVLGARGVGKSAIVRQFLYNEFSEVCVPTTA
RRLYLPAVVMNGHVHDLQILDFFPISAFPVNTLQEADTCCRGLRSVHAYILVYDICCDFSFEYVKTIQQ
ILETRVIIGTSETPIIIVGNKRDLQRGRVIPRWNVSHLVRKTWKCGYVECSAKYNWHILLLFSELLKSVGCA
RCKHVHAALRFQGALRRNRCAIM

AUG initiation = 23.2 kDa
UUG initiation = 26.5 kDa

+++++

#38

NM_001080510

CUG initiation in most mammals. The extension is AUG initiated in mouse, rat and opossum. In addition, the downstream in-frame AUG is immediately preceded by out-of-frame AUG in perfect context which means in the downstream in-frame AUG is usually inaccessible. Delimiting upstream in-frame stop codons in *Macaca fascicularis*, *Procarvia capensis*, *Rattus norvegicus* and *Sus scrofa*.

C17orf95 chromosome 17 open reading frame 95

(homology to AdoMet-dependent methyltransferases)

Synonyms: C17orf95

AGTTCTGCGCGTGTGAGTCTTTCGCCTGCTCGGCTTCTCGCTCGCAGCGCGGCAGGGTTATCAC
CAGATCTGGCTTCCCTTCTGCCGTACGGTCTACGCCACGTGGCCGGCTCCGCTCGCGCAG
TCTGGCAGCCGGAGCCTCCCGGGTCCCCACCCGCCGGGCCAACGACGCCACTGGCGAGCACG
ATTTCCGAGGACAGGGGGTCGGGCCAGCGCTTCGATTCTCGGAGGAGCCGGTCCGGGGCGACGG
GCTGCTCTGGAGGTCCACGTCCCGCAGGTCTGCATCTCACTATGGAATGTATGTTGGCCCTGTGCTGT
GGTCCTGGCCCAGTACCTTGGTTACAGAAGATCTCTGCCAGGCAAGGCCATCTAGAGATTGGAGCTG
GAGTGAGCCTCCAGGAATTGGCTGCCAATGTGGTGAGAAGTAATACTGTCAGACAGCTCAGAACTG
CCTCACTGTCTGAAGTCTGCGCAAAGCTGCCAATGAATAACCTGCCACATCTGCAGGTGGTAGGACT
AACATGGGTCATATATCTGGATCTCTGGCTTACCAACACAAGATATTATCCTGCATCTGATGTGT
TCTTTGAACCAGAAGATTTGAAGACATTTGGCTACAATATATTTTGATGCACAAGAATCCAAGGTC
CAATTGGTCTACTTACAGTTAGGAGTGCTGGCTGGTCACTTGAAGCTTACTCTACAAATGGGATAT
GAAATGTGTCACATTCTCTTGAGTCTTGTGAGCAGACAAAGAAGATATAGCAGAAATCTACCCCTCCAG
GAAGACATACAGTGAAATGCTGGTCATTCCTTGCAAAAGGACAGTCTCTGAATTATACCTACAACCTGT
TCTGGGACAGTATCAATACTGATGAGCAACCTGGCACACAAACTATGAGCAGACCCACTTCAGCTTGAGAAT
GCAGTGGGTCTGAAGATGGTCAAGTCTGTTGCCTTAGATTTGATGTCACCTAGACAACACTAAACTCA
TATGAAACAAAAATTAAACGTATTACAAGTAAAAAAA

LAARSLPRSPTRPGPNDALEHDFRGQGVRAQRFRFSEEPGPAGDAGAVLEVHPQLHLQYGMYVWPCAV
VLAQYLWFHRRSLPGKAILEIGAGVSLPGILAACKGAEVILSDSSELPHCLEVCRQSCQMNNLPHLVVGL
TWGHISWDLLALPPQDIILASDVFFEPEDFEDILATIYFLMHKNPKVQLWSTYQVRSAGWSLEALLYKWD
KCVHIPLESFDADKEDIAESTLPGRHVTVELVISFAKDSL

PROBABILITY of export to mitochondria: 0.9710 with extension
PROBABILITY of export to mitochondria: 0.4276 w/o extension

AUG initiation = 21.4 kDa

CUG initiation = 28.2 kDa

+++++

#39

NM_023110

Likely ACG initiated extension in mammals

FGFR1 fibroblast growth factor receptor 1

Synonyms: CEK; FLG; OGD; FLT2; KAL2; BFGFR; CD331; FGFBR; HBGFR; N-SAM; FLJ99988; FGFR1

AGATGCAGGGCGCAAACGCCAAGGAGACCAGGCTGTAGGAAGAGAAAGGCAGAGGCCGGACAGCTCGG
CCCGCTCCCCGTCTTGGGCCGCGCTGGGAACTACAAGGCCAGCAGGCACTGCAGGGGGCGGAGG
CGGAGGAGGGACCAGCGCGGTGGAGTGAGAGAGCGAGCCCTCGCAGCCCGCAGCAGCTCGG
GCGCTCTGCGCACAGGCCGGCTCGCCGGGGAGCTGGAGGAGCTGGCGGCCACGTCCGGACGGGAGACCC
TCGTAGCGATTGCGCGACCTCGCCTCCCCGGCGAGCGCGCCGCTGCT**TGA**AAAGCCGCGGAACCC
AAGGACTTTCTCCGGTCCGAGCTCGGGCGCCCGAGGGCGACGGTACCGTGCTGCAGTCGGCACG
CCGGCGCCGGGCCTCCGAGGGCGATGGAGCCGGTCTGCAAGGAAAGTGAGGCGCCGCGCTGCGT
CTGGAGGAGGGGGCACAAGGTCTGGAGACCCGGGTGGCGAGGGAGCCCTCCCCCGCCCGCTCCG
GGCACAGCTCCGGCTCCATTGTCGGCCGGCTGGAGGCGAGCACCGAGCGCCGGAGTCG
AGCAGCCGGCGAGCTTGTGCAACCCGAGAACAGAGCCGGGGCGGGAGCCG
GGGACGCGGGCACACGCCGCTCGCACAGCC**ACG**CGGACTCTCCGAGGCGGAACCTCACGCCAGCG
AGGGTCAGTTGAAAAGGAGGATCGAGCTACTGAGTGGAGTCC**ATG**GAG**ATG**TGAGCCTTGTCAAC
CTCTAACTGCAGAACTGG**ATG**TGAGCTGGAGCTGAACAAGCCAGGCCCTGGGAGCCCTGTGAGTC
TCTGACCGCTAGGCCGCTCCGACCTTGCTGAACAAGCCAGGCCCTGGGAGCCCTGTGAGTC
TCCCTCTGGTCCACCCCGGTGACCTGCTGCAGCTTCGCTGCGCTGCCAGTGTGCAAGAGCATCAA
CTGGCTGCGGGACGGGTGCAAGCTGGCGAAAGCAACCGCACCCGATCACAGGGAGGAGGTGGAGGTG
AGGACTCCGTGCCCGAGACTCCGGCTCTATGCTTGCTAACAGCAGCCCTCGGGCAGTGACACC
TACTTCTCCGTCAATGTTCAAGATGCTCTCCCTCTCGAGGGATGATGATGATGACTCCTCT
AGAGGAGAAAGAACAGATAACACCAAAACCAACCGTATGCCGTAGCTCCATATTGGACATCCCAGAAA
AGATGGAAAAGAAATTGCAATGCACTGAGCTGGCGCTGCAAGACAGTGAAGTCAAATGCCCTCCAGTGG
CCAAACCCACACTGCGCTGGTAAAAATGGCAAAGAATTCAAACCTGACCACAGAATTGGAGGCTACAA
GGTCCGTTATGCCACCTGGAGCATCATAATGGACTCTGTTGCCCTCTGACAAGGGCAACTACAC
TTGTGGAGAATGAGTACGGCAGCATCAACCACACATACCAGCTGGATGTCGTGGAGCGGT
CCCATCTGCAAGCAGGGTTGCCGCAACAAAACAGTGGCCTGGTAGCAACGTGGAGTT
GGTGTACAGTGACCCGAGCCGCACATCCAGTGGCTAAAGCACATCGAGGTGAATGGAG
CAGACAACTGCCCTATGTCAGATCTGAAAGACTGCTGGAGTTAATACCACCGACAAGAGATGGAGG
CTTCACCTAAAGAAATGTCCTCTTGAGGACGCAGGGAGTACGTGCTGGCCTGGGTAACTCT
CTCCCATCACTGCACTGGTGAACGTTCTGGAAAGCCCTGGAGAGAGGCCGGCAGTGATGAC
TGTACCTGGAGATCATCATCTATTGACAGGGCCTCCTCATCTCTGCATGGGGGTCGGTC
TACAAGATGAAGAGTGGTACCAAGAAGAGTGA
CTTCCACAGCCAGATGGCTGTGCAACAGCTGGCA
CATCCCTCTGCGCAGACAGGTAAACAGTGTCTGCTGACTCCAGTG
ACTCCATGAACACTGGGGTTCT
TTCGGCCATCACGGCTCTCCAGTGGACTCC
GACCGCTGGAGCTGCCCTGGGACAGACTGGCTT
AGGCAAACCCCTGGAGAGGGCTGCC
GGTGGTGTGGCAGAGGCTATCGGGCTGGACAAGGACA
ACCCAGCCACAACCCAGAGGAGCAGCT
GGCATGGAGTATCTGGCCTCCAAGAAGTG
GGACAATGTGATGAAGATAGCAGACT
CAACCAACGGCCGACTGCC
AGTGTGAGTGGATGG
GCCTGAGGAACTTTCAAGCTG
GGCAGGGTACCGC
GACAGGCCAGTA
ACTGCACCAAC

AGCTGTACATGATGATGCGGGACTGCTGGCATGCAGTGCCCTCACAGAGACCCACCTCAAGCAGCTGGTG
GAAGACCTGGACCGCATCGTGGCTTGACCTCCAACCAGGAGTACCTGACCTGTCATGCCCTGGACCA
GTACTCCCCAGCTTCCCACACCCGGAGCTACGTGCTCTCAGGGGAGGATTCCGTCTCTCATG
AGCGCTGCCGAGGAGCCCTGCCGCCACACCCAGCCAGCTTCCAATGGGGACTCAAACGCC
TGA CTGCCACCCACGCCCTCCCCAGACTCCACCGTCAGCTGAACCCACCCACAGGCCCTGCTGGG
CCACCCACTGTCGGCTCCCTGCCCCCTTCCCTGCTGGCAGGAGCCGGTGCCTACCCAGGGCCTTCTGTG
GGCCTGCCCTCACCCACTCAGCTCACCTCTCCACCTCCTCCACCTGCTGGTGGAGAGGTGCAAAG
AGGCAGATCTTGCTGCCAGCCACTTCATCCCTCCAGATGTTGGACCAACACCCCTGCCACCAGG
CACTGCTGGAGGGCAGGGAGTGGGAGGAATGAACAGGCATGCAAGTGAGAGCTTCTGAGCTTCTCCT
GTCGGTTGGTCTGTTGCCCTACCCATAAGCCCCTGCACCTGGTGGCAGGTGCTTGTGCTCAGGG
CTACAGCAGTAGGGAGGTCACTGCTCGCTCGATTGAAGGTGACCTCTGCCAGATAGGTGGTGCCA
GTGGCTATTAACTCGATACTAGTTGCTTGACCAAATGCCTGGTACCAAGGAGGTGGAGGCGAA
GGCCAGGTTGGGGCAGTGTGTTGGCCCTGGGGCCAGGCCAAACTGGGGCTCTGTATATAGCTATGAA
GAAAACACAAAGTGTATAAATCTGAGTATATATTACATGCTTTTAAAAGGGTCGTTACCAAGAGATTAA
CCCATCGGTAAGATGCTCTGGTGGCTGGAGGCATCAGTTGCTATATATTAAAACAAAAAGAAAAAA
AAGGAAAATGTTTAAAAGGTATATATTGCTACTTTGCTGTTTATTTTAAATTATGTT
TAAACCTATTTCAAGTTAGTCCCTCAATAAAATTGCTGCTGCTTCATTATCTATGGGCTGTATGAA
AGGGTGGAAATGTCACGGAAAGAAGGGACACCCACGGGCCTGGGCTAGGTCTGCTCGAGGGCACCG
CATGCTCCCGGCCAGGTTCTGTAACCTCTTCTAGGTCTGCACCCAGACCTCACGACGCACCTC
CTGCCTCTCGCTGCTTTGAAAGTCAGAAAAAGAAGATGTCAGCTCGAGGGCAGGAACCCATCCATG
CAGTAGAGGCCTGGCAGAGAGTCAGGCCACGGCCATCGACCATGGATGGTTCTCCAAGGAAACC
GGTGGGGTTGGCTGGGAGGGGCACCTACCTAGGAATGCCACGGGTAGAGCTACAGTGATTAAGAGG
AAAGCAAGGGCGCGGTTGCTCACGCCCTGAATCCAGCACTTGGGACACCGAGGTGGCAGATCACTTC
GGTCAAGGAGTTGAGACCAGCCTGGCAAACCTAGTGAACACCCATCTACTAAAATGCAAAATTATCC
AGGCATGGTGGCACACGCCCTGAATCCAGCTCCACAGGAGGCTGAGGAGAACATCCCTGAAGCTGGGAGG
CGGAGGGTGCAGTGAGCCAGATTGCCACTCCAGGCTGGCAACAGAGAAAACAAAAAGGAAA
CAAATGATGAAGGTCTGCAGAAACTGAAACCCAGACATGTGCTGCCCTCTATGTGGCATGGTTTGC
CAGTGCTCTAAGTCAGGAGAACATGTCACCTGAGGCTAGTTGCACTCAGGTCCTGGCTCGTTCT
TGTTGGTATGCCTCCCCAGATCGCTCTGTATCCATGTGACCAGACTGTATTGTTGGACTGTCGCA
GATCTGGCTTCTACAGTTCTCTGCAAACCTCCATCCTGCTCCCTAGGAACGGGGAAAATTCTCC
GAATGTTTGGTTTTGGCTGCTGAAATTACTCTGCCACCTGCTGGTCACTACTGTCCTCACTAAG
TGGATTCTGGCTCCCCGTACCTCATGGCTAAACTACCACTCCTCAGTCGCTATATTAAAGCTTATATT
TGCTGGATTACTGCTAAATACAAAAGAAAAGTTCAATATGTTTCAATTCTGTAGGGAAAATGGATTGCTG
CTTTAAATTCTGAGCTAGGGATTGGCAGCTGCAGTGTGGCAGTATTGTAATTCTCTTGT
CTCTGTAAATAGCACCTGTAACATTACAATTGTATTGTTAAAGAAGGCATCATTGGTGAACA
GAACTAGGAAATGAATTAGCTCTAAAGCATTGCTTGGAGACCCACAGGAGTGTCTTCTGT
AAACAGTGATGATAATTCTGCTTGGGAGAGGTCTAGAAAATATAAGGCACATCATAATTACAGTGATGTCCTG
CTGTTACTACTCAAATCACCCACAAATTCCCAAAGACTGCGTAGCTGCAAATAAGACAGTGAAAT
TGACCTGAAAAAAAAAAAAAA

TADSPAEPPRARVSLKRIELTVEYPWRCGALSPTSNCRTGMWSWKCLLFWAVLVTATLCTARPSTLP
EQAQPWGAPVEVESFLVHPGDLQLRCLRDDVQSIWL RDGVQLAESNRTRITGEEVEVQDSVPADSGLY
ACVTSSPSGSDDTYFSVNVDALPSEEDDDDDDSSEEKETDNTKPNRMPVAPYWTSPPEKMEKKLHAPPA
AKTVKFKPSSGTPNPTLRLKNGKEFKPDHRIGGYKVRYATWSIIMDSVPSDKGNYTCIVENEYGSINH
TYQLDVVERSPHRPILOAGLPANKTVALGSNEFMCKVYSDPQPHIQWLKHIEVNGSKIGPDNL PYVQILK
TAGVN TTDKEMEVHLRNVSFEDAGEYTCLAGNSIGLSHHSAWLTGLEAERPAVMTSPLYLEIIIYCTG
AFLISCMVGSIVYKMKSGTKSDFHSQMAVHKLAKSIPLRRQVTVSADSSASMNSGVLLVRPSRLSSGT
PMLAGVSEYELPEDPRWELPRDRVLGKPLGEFCFGQVVAEAIGLDKDKPNRVTKAVKMLKSDATEKDL
SDLI SEMMEMMKGHKNIINLLGACTQDGPLYVIVEYASKGNLREYLQARRPPGLEYCYNPSHNPEEQLS
SKDLVSCAYQVARGMEYLASKKCIHDLAARNVLTEDNVMKIADFGLARDIHIDYYKKTTNGRLPVKWM
APEALFDRIYTHQSDVWSFGVLLWEIFTLGGSPYPGPVVEELFKLLKEGHMDKPSNCTNELYMMRDCH
AVPSQRPTFKQLVEDLDRIVALTSNQEYLDLSMPLDQYSPSFPDTRSSTCSSGEDSVSHEPLPEEPCLPR
HPAQLANGGLKRR

AUG initiation = 91.9 kDa

ACG initiation = 96.7 kDa

+++++

#40

NM_153369

CUG initiation conserved in mammals. Extension initiated by AUG in *Pteropus vampyrus* and *Loxodonta africana*. Curiously in mouse the first downstream in-frame AUG has become GUG. The next available in-frame AUG there is 48 codons further downstream.

KIAA1919

Synonyms: NaGLT1; MGC33953

GATAGCGGTTGGCAGAGGCCACATTCCGCCACGTGACCCGCATGCGCTGCTTGAGAGCGAGC
GTCCTTTGCCACGAACACCTGCGCGTGCAGAAGTCTCTCTGCCCCGGTCCGGCGTCCGCTGG
GAGGTGGCGCCCGAGGGACACCGTGCCTGGGTGCGGCGAGTGCCTCTCGAGCTCCAGAAGTCT
GCTGAAGAGCGCGCGCAGCAAGACGACTTCTCCGGAGCCCGAGCTGGAGTTAGAGGTGGAGCTCCG
GGGGCCGGGCCCCGGCTGGGGCAGGGCTCTGCAGGGGAGGCCCGGGAGAATGAGCCGGAGCC
GGAGGTGGTGGTGGCTCTGGCAGAGCGGTGGACCGGGAGCAAGCTGGTGGTCAACACCTTGATGC
TGTGTGCCTCCCTCCTGGGCTGGATTGAGTGTGCTATAAGTGGACCCACGTTCAAGATTGGCAACA
AACGTGAACCGAAATATCAGTAGTCTGCTTTCATTTTGAGGTGCGCTGGATATTGAGTGGCTC
TGTGATTGGTGGATTCTTGTGATGTCAATTATTTTACTTTGGAAATCTCAATGTCGGCTACCA
CCGTTGGCTTTATCTTGTCCCTTTGCAAGACAGCAATATTACTCACTGTATGATGTCATCTCGGT
GTTTCAATTGGCATTCTGGATACAGGTGGTAACGTCCTTATCTGGCTATTGGGGGACAAAGGAGCCC
ACATATGCAGGCCTTACACTCTCTTGCCTGGGTGCCTTTGGCTCCACTGCTAGCTAAACTGGCTT
GGGTCCGACAGCGTCTGCTGAAAACCACACAGAGTCTGACTTCCATCTGCACTAACCAATCATCTGAT
GCTGACTCAGAAGCTCTGTTGGAGTACCTAATGATAAGAATTACTGTTGGCTTATGTTATCGGTAC
TTACATGTTCTAGTTCTGTCAATTGGCTTGTGTTAAAGAATAGCTCAAAGCAAGAAAAGCAA
GAGCATCTGCTGAGACATTGCAAGAGAAAATATCACAAACGCCCTTCTTGTCTCCTTTCTGTTCTC
TTTTTTATGTTGGAGCTGAGGTAAACATATGGCTTCTATGTTCTCATTTGCAACCACCCATGCTGGCAT
GAAAGAAAAGTGAAGCTGCTGGTTGAACCTCCATCTGGGGACATTGCAAGCCTGAGGGCCTGGCAA
TCTTTCTGCTACCTGTTACAGCCTGGAACCATGATTGTTGAGCAACATTGGCAGCCTGACTTCATCT
TTATTTCTGGTGTCTTTGACAAGAACCAATTGTCCTGAGTAACTGTTGAGTGGTATGGGCTTCAAT
GGCAACCACATTCCGAGTGGTCTTGGATTGAGCAGTACACGACCATCCATGGAAATCTGAGCAT
TTTTGTAATTGGTGTCTCCCTGGAGAAATGGCTATTCTGCAAGTCACTGGATTCTCAAGGAAATAC
CCTGATTGCTGTAGTTCTGTATACCTCTTGGAGCATCAATAGCTACTGGTATTGTTATTCTGTGCT
ATATAAATTAGCCACTTCACCTCTGATGCCAGCGAAAAGAGACAGAAAGAGTGGAGGACAGAAAGCTC
TGCTCTAGCTCCGGCTAAATGAATATGAGGAAGAGAATGAAGAGGAGGATGCAGAAAATGGAATGAA
ATGGATTGAAATGATTGAAACGAATGATAACATGAGGCATTCTATAATAGAGACATCTAGAAAGTAGTCT
GACGGAGCCCACAGCTGAAGTCTATAATCAATACCCATCAAATGCACTGGTGGTGGAGTCTTCCTTTA
ATACTGGCAGTGCCATGTGAAGCACTGCCAGAAAACCAGGACAAAGGGACTAACGTTAGAGAAGATGG
ATTACTCACTGACATCTTGAAATAACTGCCACTTCTAAGGACGCCATTAGGCAGAGCATTAGCAGAATT
TCACCATGCTTGGGATTAAAATTGAGTCATATTATAGCAAATGCTAAAATTGGTAAATGTTCT
GTAATTCCAGAGCTTCCATAAATAACCAATGGTCTCATACACGTACAATAGGATCTGTTATAACGCT
ATCATTGACATGTGACTGATTAGGAATATTTATGGTCTCCACCCCTCAGAGCACCCAAAGAAGGCATT
GTTTTGAGAAGGGGGTATTCTCAGAATGAGGACCATGAGGAGTTGAATTGGTGAACAAATTATC
AGGTATCAATATTTCCAGGAAAGTTGAGGAGAATAGTTCATAAAAGAGAAATGGCCGGGTGCAGTGG
CTAAAGCCTGTAATCCCAGCACTTGGGAGGCCAGACGGCAGATCACGAGGTCAAGGAGATCGAGACCAT
CCTGGCTAACACAGTGAACACCTGCTCTACTAAAAATACAAAAATTAGCCAGGCGTGGCAGGCACC
TGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGTGGAGCTGAGTGG
CCGAGATCATGCCACACCGTACTCCAGCCTGGCGACATCTAGACTCCGCTCAAAAAGAAATGGAGAGG
TATTTTATTAAATTGCTTACCTTATCTGTAGGGTTGGATCAAAGAATGTAATGACTCATTTGATTCTAA
TTTCAATGAGAGAAACACAGAAACAGGGCCAAGGGGTGGCTCAAGCCTTGACTTCTAGCCCTGCA
CTCTGTCACCTAACGTAGGCCTGTTATCTGCCCTCTGATCTGTGATTCTAGAAATGAAATGAGATGA
TTTCCAATTGGGAATCTGTAAGTTAAAGTTGCAAAATGTCTGCTAGAAATTAGTACAAGTGACAG

CATGTTCTTCATTCAAATGAGAAGTGAAACAAAATGATAAATAGCTTAAGAATGTGCTAATGATAAAT
GATTACATGTCAATTAAATGTTACTTAATGTTAACCTTATTGAATAATTACCTGAAGAATATATTTT
TAGTACTGCATTTCATTGATTCTAAGTGCACCTTTACCCCCATACTGTTAACATATCTGAAATCAGAAT
GTGTCTTACAATCAGTGATCGTTAACATTGTGACAAAGTTAATGGACAGTTTCCCATATGTATATA
TAAAATAATGTGTTTACAATCAGTGGCTTAGATTCAAGTAAACAGTAATTCAATTATGATAGTA
TCTTTACAGACATTAAAAATAAGTTATTTTATATGCTAATATTCTATGTTCAAGTGGAAATTGGAGAC
ACTATGTTCTAATTAAAGAACTGTTCACTGTTATTAGATTAGTAGTAGGGCTGTAGATGGGTCTTCG
AGGTCTTGCCTACCTCAGTAACAGAGGCTGGGCTGAGCCTGCTCGAGATTAGCTCTCACCAAGGTCT
GTGTGTTAATCATATCGTAGTTACCGTAGGTACCGCTGCAAGTGTAAATTACAGTGAGGCATAGTCTTTC
CACTGATTGGAATACTTGTGTCACCGATCCAGGGTTCTCTTCTGAGTCTGAAATACCTGGGAAACAG
TTAGAAGGTGACTGGTAGCCTAGGAGTCAGAAAGGGTACTTCATATATGTGGCAGAATTGG
TGCAGGTGATCCTTGAATAAGACAAGAATTGATGAAATTGTTACTCCTCCAATTGGCTGAAT
AACATCTCCTTCTACCTGATTGGAGTAGGCTTGCTTACTTAATTCTATAATTGCCATTGTT
TTGTCTCAAGTAAAAAAAAAAAAAA

LELEVELRGAGPPAAGQRLLQAEAPAENEPEPEVVVWSQSGGTGSKLRWFITLMLCASFLGLGLSVAIVG
PTFQDLATNVNRNISLSFIFVGRALGYLSGSVIGGFLVDVMNYFLLGISMSATTVGLYLVPFCKTAILL
TVMMMSIFGVSIGILDGGNVILAIWGDKGAPHMQALHFSFALGAFLAPLLAKLALGPTASAENHTESDFH
PALNQSSDADSEALFGVPNDKNLLWAYAVIGTYMFLVSIVFFCLFLKNSSKQEKA
RASAETFRRAKYHNAL
LCLLFLFFFYVGAEVTVGSYVFSFATTHAGMK
ESEAAGLNSIFWGTFAACRGLAIF
FATCLQP
GTIVLS
NIGSLTSSLFLVLFDKNPICLWIATSVYGASMA
TFPSGVSWIEQYTTIHGKSAAFFVIGASLGEMA
IPAV
IGILQGKYPDLPVLYTSLGASIATGILFPVLYKLATSPLDRQRKEDRK
SEDQKALLSSGLNEYEE
EE
EDA
EKWNEMDFEMIETNDTMRHSIIETSRSSLTEPTAEVYNQPSNALVFESSPFNTGSAHV
KHL
PETRTK
GTNV

AUG initiation = 56.2 kDa
CUG initiation = 62.0 kDa

+++++

#41

NM_001144886

CUG initiation conserved in mammals

CITED1 Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 1

Synonyms: MSG1; CITED1

AGAAGGGGCACCGCTGAGCTCCGAGAGAGGAAGCTCGCTCTGCCGGCTGCCCTCTTG
TAGTCCGCCGGCGA
GGGGCAGTTCTCGGTGAGGAGGAAGAGAGCAGCGACGGCACAGCACCCGCGGGCCCTCCACAACAGC
TCCAG
CTGGCAGCATCACTCCGCCAATTATCCAACCTCTGCCAAGGCTCTGAA
ATG
CCAACAACGTCG
AGGCCTGCACCTGATGTCAGGGTGGCACCTCACCTGCGAAGGAGGATGCCAACCAAGAGATGAGCTCCG
GGCCTACTCCAACCTGCGGTGAAAGATCGCAAAGCAGTGGCATTCTGCACTACCCCTGGGGTAGCCTCAA
ATGGAACCAAGGCCAGTGGGCTCCACTAGTCTCTGGGATCTCCAATAGGCTCTCCTACAACCACCC
CCCAC
AAACCCCCATCCTCAACCTGCA
ACCCGCCCTCACTGCTGGCTAGTATGCACCTGCA
GAAACT
TAATAGCCAGTATCAGGGATGGCTGCTGCCACTCCAGGCCAACCCGGGGAGGCAGGACCCCTGCAA
AAACT
GGGACTTTGGGCCCAGGCCAGGGCAGAATCACTCTCCTCTGCTGGT
GCCAGAGCCCTGCTATC
ATCGATTGGACCCAGTGGATGAGGAAGTGTGATGTCGCTGGTGGTGA
ACTGGGTTGGACCGAGCCA
TGAGCTCCGGAGCTGTGGCTGGGAGAATGAGTTGACTCACTGCC
ACTTTCCATCTAGCTGCT
TAA
GCCAAGTGTCCCTAAAGATGGAGGAATAAGCCACCAATTCTGTTGAA
AATAAAATAAGTTACTTACAA
AGAGACGGCCAAAAAA

LAASLPANLSNFQGSEMPTTSRPALDVKG
GTSPA
KEDANQEMSSVAYSNLAVKDRKAVAILHYPGVASNG
TKASGAPTSSSGSPIGSPTTPPTKPPSFNLHPA
HLLASMLQKLNSQYQGMAAATPGQP
GEAGPLQNWDF
FGAQAGGAESLSPSAGAQS
PAIIDSDPVDEEVLM
SLV
ELGLD
RANEL
PELWL
GQNE
FDFTAD
FPSSC

AUG initiation = 19.9 kDa
CUG initiation = 21.6 kDa

+++++

#42

NM_006645

Possible GUG initiation conserved in mammals

STARD10 StAR-related lipid transfer (START) domain containing 10

Synonyms: PCTP2; CGI-52; MGC14401; NY-CO-28; SDCCAG28; STARD10

CAAGCCGAGAGTCGGGATCGACACGTGGGGGCCCTGAGCGAAGATAACCGTAATAAATAGTAACCTAA
CGGTCCAGTCATCGTTCTGTGGTCCTTCTTTATGATTACAAGGAATGACCCTTCATGCCCTCTCCT
AATTCACTCACAACAGTCCTTTACAATGGACAACAGGTTAGAGGAAGTCAGGCAGATTCCAGCA
TCATAGAGAGTAAAGGACCAGGGAAAGGATCAGGATTCAAGGACTGCACCCAGGCTCTGCTCCAGCTTGCT
GTGTGACTTGGGTAAATTGTTCCCTAGGAACTGAGCTTCTCATTTGTAATGCAAACAGGCTGTTG
GGAGGATCAAATGAGATCCAGGGGTGAAAACAGCTTAGTTACTTCAGGAATTACCCACGCGGTATATA
AAGGAAAATATTATTAGTCAGGTGATTGAGATTGAGGAACCCATTCCCTCATTCTGCAAATTGCAA
CCTGAGGGCCCAAAGAGGGACAGGGGCTTGCCTCAGGCTCTCAGCAGGCTGTGAGCAAGAGCTAAAGCTAA
TCCTCCTGCCTTGGGCTGGAGCCCTCCTGTACCCCAGGGTCAGTGTCTTTGTTGGATACAGGCTTA
GATTGACTGACTGTACCCCTGAGAACCTAGGGAGTCCCTGTCCCAATTCTCTACCCCCACCTTGGC
CTGATGGAGGAAGACCCCTGCTGTGAGATGAGCACAGAGCAAGAAGCTGAGGAGGATCTGGAGAATT
CTGGAGGAAGAGGAGAGTGTGCTGGAGCTGTACAGACCCCTGCTTCAGGTCCCAGGAAGSTGGCGTCAG
CATCTGCAGCCGCGTCGACGTTGTCGGAGCCTCCCGGAGGACCCAGGAGAGCCGACTAGGACCAAGGGC
CTGGGCTCCCCAACACTCCCCATGAGAACGCTGGGGCTCTACAGAGCCCCAACGGGCTCGGCCGGTCT
GGCCGTGAGAGTGTCCAGGTGCCGATGACCAAGACTTCCGAGCTCCGGTCAGAGTGTGAGGCTGAGG
TGGGCTGGAACCTGACCTATAGCAGGGCTGGGTGTCGTGCTGGGTGCAGGCTGTGGAGATGGATCGGACG
CTGCACAAGATCAAGTGCCGGATGGAGTGCTGTGATGTGCCAGCGAGACACTCTACGACGTCTACACGA
CATTGAGTACCGCAAGAAATGGGACAGCAACGTATTGAGACTTTGACATCGCCGCTGACAGTCAACG
CTGACGTGGCTATTACTCCTGGAGGTGTCCTAACGCCCCCTGAAGAACCGTGTGTCATCACCCCTCGCTCC
TGGCTCCCATGGCGCTGATTACATCATTATGAACTACTCAGTCAAACATCCAAATACCCACCTCGGAA
AGACTTGGTCCGAGCTGTGTCCTACCTCAGACGGCTACCTCATCCAGAGACAGGGCCAAGAGCTGCGTC
TCACCTACCTGGCCCAAGGTGGACCCCAAAGGCTCTTACCCAAGTGGGGTGAATAAATCTTCTCAGTTC
CTGGCTCCCAAGGCCATGAAGAACGATGTACAAGGGCTGCCCTCAAGTACCCGAGTGGAAACAGAACCT
GCCTCACTCAAGCGTGGCTGCACCCGGAGCAGAGCCGTTGCCAGCCTGGCGCTGTCGGAGCTGTCGG
TGCAGCATCGGACTCACTGGAGAACATCGACGAGAGCGCGGTGGCCAGAGCAGAGAGGAGGGATGGC
GGCGCGGGCGGGCAGGGCAGCGACGACACCTCGCTCACCTGAGGCCGACCGCTTCAGGGACGGAGA
CAGGACCGGGCGAGCCCTGGGGCGGCCGCTCTGCACTTTCTCCCCCACC CGCACCTGGTGG
CACCGGGCCAGGCCAGCGGGTGCTGCAGCCTGGCTGGACAGAGCCCCAATAACGATCCCACAGCCTCA

VASASAAASTLSEPPRRTQESRTRTRALGLPTLPMEKLAASSTEPQGPRPVLGRESVQPDDQDFRSFRSEC
EEAEVGWNLTYSRAGVSVWVQAVEMDRTLHKIKCRMECCDVPAETLYDVLDIEYRKWDNSVIETFDIARL
TVNADVGYYSWRCPKPLKNRDVITLRSWLPMGADYIIMNYSVKHPKYPKRDLVRAVSIQTGYLIQSTGPK
SCVITYLAQVDPKGSLPKWVNKSQFLAPKAMKKMYKACLKYPEWKQKHLPHFKPWLHPEQSPLPSLALS
ELSVQHADSLNIDESAVAESREERMGGAGGEGSDDDTSLT

AUG initiation = 33.0 kDa
GUG initiation = 36.6 kDa

B)

#1

NM_002097

Extension well conserved in vertebrates. CUG initiation conserved only in mammals.

GTF3A general transcription factor IIIA

Synonyms: AP2; TFIIIA; GTF3A

TGCGCGATCTCCGGAGC ATGC CGCAGCGCGGCCGACGCGGGCGGTGCCTGGTACCGCGCGCGCTCC
CGGAAGTGTGCCGGCGTCGCCGAAGGTT CAGCAGGGAGCCGTGGCCGGCGCCGGTCCCGCACGT
GTCTCGCACGTGCCAGCGCCCTGGCCTGGCCTGGAGGCGCCGGCGCC CTG GATCCGCCGGCGTGGT
CGCGAGTCGGTGTGCTCCTGACCATGCCGACCGTTATTGAGCCGGCAGAGACTCAGCTCCGACCC
CGCCGCCCGCCGCTTCCCAGGAGGTT CATCTGCTCCTCCCTGACTGCAGCGCCAATTACAGCAAAGCC
TGGAAAGCTTGACCGCACCTGTCAAGCACACGGGGAGAGACCATTGTTGTG ACTATGAAGGGTGTGG
CAAGGCCTTCATCAGGGACTACCCTGAGCCGCACATTGACTCACACAGGGAGAAAAGCCGTTGTT
GTGCAGCCA ATGCTGTGATCAAAATTAACACAAAATCAAACATTGAGAAACATTGAAACGCAAAC AT
GAAAATCAACAAAAACAATATATGCA GTTTGAAGACTGTAAGAAGACCTTAAGAAACATCAGCAGCT
GAAAATCCATCAGTGCCAGCATACCA ATGAAACCTTATTCAAGTGTACCCAGGAAGGATG TGGGAAACACT
TTGCATCAGCAAGCTGAAACGACATGCCAAGGCCACGGGGCT ATGTATGCAAAAGGGATGTTCC
TTTGTGGCAAAACATGGAGCGAACCTCTGAAAC ATGTGAGAGAAACCCATAAGAGGAAATACTATGTA
AGTATGCCGAAAACATTAAACGCAAAGATTACCTTAAGCAACACATGAAAACCTCATGCCAGAAAGGG
ATGTATGTCGCTGTCCAAGAGAAGGCTGTGGAAGAACCTATACAACCTGTGTTAATCTCCAAAGCCATATC
CTCTCCTTCATGAGGAAGGCCCTTTGTGTGTAACATGCTGGCTGTGGCAAACATTGCAATGAA
ACAAAGTCTCACTAGGCATGCTGTTGACATGATCCTGACAAGAAGAAATGAAGCTCAAAGTCAAAAAT
CTCGTAAAAACGGAGTTGCCCTCTCATCTCAGGGATATATCCCTCCAAAAGGAAACAAGGGCAAGGC
TTATCTTGTGTCAAAACGGAGAGTCACCCAACTGTTGGAAGACAAGATGCTCGACAGTTGCAGTACT
TACCCCTGGC TAAGAACTGCACTGCTTGTAAAGGACTGCAGACCAAGGAGCGAGCTTCTCAGAGC
ATGCTTTCTTATTAAAATTACTGATGCAGAACATTGATTGATTCTTATCATTTC

LDPPAVVAESVSSLTIADAFIAAGESSAPTPPRPALP RRFICSF PDCSANYSKAWLDAHLCKHTGERPFW
CDYEGCGKAFIRDYHLSRILHTGEKP FVCAANGCDQKFNTKSNLKKHERKHENQQKQY ICSFEDCKKT
FKKHQQLKIHQCCQHTNEPLFKCTQEGCGKHASP SKLKRHAKAHEGYVCQKGCSFVAKTWELLKH VRETH
KEEILCEVCRKTFRKDYLKQHMKTAPERDVCRCPREGCGRTYTTVFNLQSHILSFHEESRPVCEHAGC
GKTFAMKQSLTRHAVVHDPKKKMKLKVKKSSREKRSLASHL SGYIPPKRKQGQGLSLCQNGE SPNCVEDKM
LSTVAVLTLG

+++++

#2

NM_001418

Region conserved from insects to human.

EIF4G2 eukaryotic translation initiation factor 4 gamma

Synonyms: P97; AAG1; DAP5; NAT1; FLJ41344; EIF4G2

GCAGACGGCAACGGGGCGCTGATTGGCGCGAAGGAGCCATTGCGGGAGACTCTGGTGGGTTGGCTGC
CCCAAGAGTGATAAGTTCGGCTTCAGACACGCCATTAGCGCCAGCAGTCGGAGCTCT
AGCGGGTACCGAGTGGCGCTGCAGCAGCGACTCCTCTGAGCTGAGTTGAGGCCGCCCCGACTCCTTC
TCCCCCTCCCTCCCCCTTTTTGTGTTCCGTTCCCTCCCTCCATCCCCGACGACCGG
ATCCTGAGGAGGCAGCTGCCGTGGCAGCTGCTGAGTTCTCGGTGAAGGTATTCATTCTCCTGTCCCCCTC

CCCTCCCCACCCCATCTATTAATATTATTCTTGAAGATTCTCGTTGTCAAGCCGCAAAGTGGAGAGT
GCGATTGCAGAAGGGGTGCTCTCGTTCACTGCTCTCGGGCGAGGAGGAAGTAGGGGTGCACCTCA
GCACTATCCCAGACTGCTGCCAACAGCGAGTCCTGGGAAAACCCAGGGCAAACGCTCAGAAATGGA
TTCCTGCACGAAGCACTAGACGAGATCACAACCTCCGCAGCAAACAACCTCGCAAACGAAAAGAACGACAT
GATGCAATCTTCAGGAAAGTAAGAGGCATACTAAATAAGCTACTCCTGAAAAGTTGACAAGCTATGCT
TGAGCTCTCAATGTTGGGTGTAGAGTCTAAACTCATCCTAAAGGGGTCTACTGCTGATTGTGGACAAAG
CCCTAGAAGAGCAGAAAAGTATAGCTCACTGTATGCTCAGCTATGTCAGGATTGGCAGAGATGACCAAAC
TTTGTAGGGCCAGCAGCAGAGGGTCAACCAGGACAGAACGAAAGCACACATTAGCAGCCTCTAATTTC
CAAATTACAAGATGAAATTGAAAACGAACTAGAAATGTTGATGCTATGATAAGCGTAAAATCCCCTCC
TCCCCGAGGAGGAGGAACAGAGAGGCCATTGCTAAGATCAAGATGTTGGAAACATCAAATTCTGGAGAG
CTTGGCAAGCTTGTATTCAGAATCTATCCTCATAAGTGCATCAAACACTTTGGAAAAGAGAA
GAGAGTCCAACCTAAAGATATGGGAGAGGATTGGAGTGCCTCTGTCAAGATAATGAGGACAGTGGACCTA
GATTAGACCATGAACGAGCCAAGTCCTTAATGGATCAGTACTTGTGCCGAATGTGCTCTGATGTTAAGT
AAGGAATTGCCAGCAAGGATTGTTCTGCTGCAGGATACCGTAGAGTTGCGAGAACACCATTGGTTCC
TCGCAAGGCTTTCTGACAATGGACCAAAGACGATCAATCAAATTGCTCAAGATGCACTAAAGATCTAG
GGGTGTTATTCTGCTCTATGGCTCAAGGGATGAGAAGTGACTTCTTGTGGAGGGACCGTTCATGCCA
CCAGGATGAAAATGGATAGGGACCCACTGGAGGACTTGCTGATATGTTGGACAAATGCCAGGTAGGG
AATTGGTACTGGTCCAGGAGTTATCCAGGATAGATTTCACCCACCATGGGACGTCATGTTCAAATCAC
TCTTCAATGCCATGGGGACACATCATGCCCTCACACAATCGCAGTTGGAGAGATGGGAGGCAAGTT
ATGAAAAGCCAGATTAGCCTGAGGCCTGCTCAGTCGTTCTAATGAATAAAATCAAGTGCAAAGCTTC
GCCAGATAACTATGATTCTCTAGTGCACAACCACCCAGCACTAAACACCACCTCTGGACAGACAC
CTCAGCTTGGTCTAAAACATAATCCACCGTTATCCAGGAAAGCCTGCCAAGACAGCAAAGCCACCA
CCGTCAAAGGAAGAACTCCTAAACTACTGAAACTGTTGACTGAATATCTAAATAGTGGAAATGCAA
TGAGGCTGTCATGGTAAAGAGAAATGAGGGCTCTAAACACTTTCTCTGAGATGTTAGCAAAGTAA
TCATCCTGTCACTAGATAGAAGCGATGAAAGATAAGAAAAGCAAGTTCTTGATCAGTTACTCAAACAG
GAAGGGATAGCCACAAGTGACAACCTCATGCAGGTTCTGAATGTTAGGACCATGTCCTAAACTGGA
GGTTGACATCCCTTGTGAAATCCTATTAGCACAGTTGAGCTCGCCATCTTCAGAGCTGGTGA
GCATTTCAGAACTAGCTCAACCCTAGAAAGTGGACCCATTTCCTCTTCTACTTGTCTCAGCAG
TTAGCTAAATTACAAGATCGAGAATGGTTAACAGAAACTTTCTAACAAAGCAAGGTCATATGAGAAAAT
GCTCCCAGAAATTGATCAGAATAAGGACCGCATGTTGGAGATTGGAGGAAAGGACTGAGTTCTTAT
TCCCACCTCAAATTGGAGAAGGAACTGTTGAAGCAAATAAGTGGATCCATCCCCTCAAACCATATAT
AAATGGATTAAAGATAACATCTCCCAAACCTCATGTTAGATAAAAGGATTGTAACATCTTAAATGACTAG
CTTCTTACAGTACATTCTAGTGAAGTAAACCCCCCAGCGATGAAACAGATTCTCTGCTCTTCCA
AAGAACAGTTAGAGCAGGAAAACAACACTACTACTATCTTCAAGCCAGTAATGCAGAAATTCTTCTCATGAT
CACGTTGATCTAACAGTCAGTGCCTGTATGCTCTCCAGGTGCACTGCTATAACAGCAACTTCCAAAAGG
CATGTTACTTCGCTTTTGTCACCTCTATGACATGGAAATTATTGAAGAAGAAGCTTCTGGCTTGG
AAGAAGATATAACCAAGAGTTCCGGAAAAGGAAGGCTTGTCCAGGTGAATCAGTGGTAACCTGG
TTAGAAACTGCTGAAGAAGAAGAATCAGAGGAAGAAGCTGACTAAAGAACAGCCAAAGCCTTAAATTGT
CAAAACATACTGTTGCTATGATGTAACGCACTTGACCTAACCAACTGCAGAAATTCTTCCGCTGTAATGT
TTTCACAATATTAAAGCAGAACGACGTCAGTTAGGATTCTCTGCTACAGGTTAGTGTAAATGTTAGTGT
GTCTTAATCATAGTCTACCATCAAATATTAGGAGTATCTTAAATGTTAGATAGTATATTAGCAGCATG
CAATAATTACATCATAAGTTCTCAAGCAGAGGAGTCTATTGCAAGGACCTCTTGCTGCCAGTTATCAT
AGGCTGTTTAAGAAACTGAATAGCAACACTGAATAACTGTAGAAATGCACTTGCTCAGTAATACT
TGAGTTGTTGCAATATTGATTATCCATTGGTTTACAGAAAATTCTTAACGTAAATTGATGGTTGTT
GCCGTAATAGTATATTGCTGTATTCACCTCTAGTAATGGCTTATGTGCTAGATTAAATATCCTG
AGCCTGGGCAAGTGCACAAGTCTTTAAAAGAAACATGGTTACTTGCAACAAACTGATCAGTTGAGA
GATCGTTAATGCCCTGAGTGGTTTGTGGGTGAAACAAATGGTGAGAATTGAAATTGGTCCCTCCT
ATTATAGTATTGAAATTAAAGTCTACTTAATTATCAAGTCATGTTCATGCCCTGATTATATACTGTAT
CTATCAATAAACATTGTGATCTGATGTAGTGA

VESAIAEKGASRF SASSGGGSRGAPQHYPKTAGNSEFLGKTPGQNAQKWI PARSTRDDNSAANNSANEK
ERHDAIFRKVRGILNKLTPEKFDKLCELLNVGVESKLILKGVILLIVDKALEEPKYSSLYAQLCLRLAED
APNFDGPAAEQPGQKQSTTFRLLLISKLQDEFENRTRNVDYDKRENPLPEEEEQRAIAKIKMLGNIKF
IGELGKLDLIHESILHKCIKTLLEKKKRVQLKDMGEDLECLCQIMRTVGPRLDHERAKSLMDQYFARMCSL
MLSKELPARIRFLLQDTVELREHHWPRKAFLDNGPKTINQIRQDAVKDLGVFIPAPMAQGMRSDFLEG
FMPPRMKMDRPLGLADMFGQMPGSGIGTGPVIQDRFSPTMGRHRSNQLFNGHGGHIMPPTQSQFEGMG

GKFMKSQISLRPAQSFLMNKNQVPKLQPQITMIPPSAQPPRTQTPPLGQTPQLGLKTNPPLIQEKP
KPPPSKEELLKLTEVVTEYLNSGNANEAVNGVREMRAPKHFLPEMLSKVIILSLDRSDEDKEKASSLISL
LKQEGIATSDNFMQAFLNVLDQCPKLEVDIPLVKSYLAQFAARAIISELVSISELAQPLESGTHFPLFLC
LQQLAKLQDREWLTLELFQQSKVNMQKMLPEIDQNKRMLIEILEGKGLSFLFPPLLKLEELLKQIKLDPSPQ
TIYKWIKDNIISPKLHVVDKGFVNILMTSFLOYISSEVNPPSDETDSSSAPSKEQLEQEKOULLSFKPVMQF
LHDHVDLQVSALYALQVHCYNNSNFPKGMLLRFFVHFYDMEIEEEAFLAWKEDITQEFPKGKALFQVNQW
LTWLETAEEESEEAD

+++++

#3

NM_001017371

Region highly conserved from zebrafish to human - the non-AUG initiation goes at least as far back as chicken.

SP3 Sp3 transcription factor

Synonyms: SPR2; DKFZp686O1631; SP3

CCTCGCCCGCCTGCCGCCTTTTGTCGGCGTGTGAGTGTGGCCCCAGCGTGCCCTCCC GGGGTGGGTT
CGGGCGGAAGGCGGAGGCCGGCGCGCAGCCGCCGCGCCCTGCCCGGGACCGGGAGCGGGGGTGCTT
GGAGCGGGGGACGCCAGGCGTGGCTGGCGGGACCAGGAGGAGGAGGAGGAGGAGGAGCGCG
GCTGGCGCTTGCCCGGGCGCAGTCGGCGGGACCGAGTCGTACTTCTGTGCGAAAGGCCCGCACCTA
ACCGCCACCCCTCCCCCTGTCTCCCTCTGAACCCGCCATTGGGGTAGGACACTCAGCCGTACCGC
TCGCTCTGCTGGCCGCTACCTGCAGCAAGATAGGCCGACAGGTGATTGGCTCTGCACAGTTAGGAGGAGCACC
GCCGCAGCCGGGGCCCCCGCCGCCGGAGCGACAGGTGATTGGCTCTGCACAGTTAGGAGGAGCACC
AAACCGATGGAGGTTTGTCAAGGCCACACTACAATATAAAAGATGAAGCTGGTAATCTAGTCCAGATT
CAAGTGCTGCTACTTCAAGTGGCAAGTATGTTCTCCCTCAGAATTGCAAGAACAAATATTCC
GTTGCACCAGGATCAGATTCAAATGGTACAGTGTCCAGTGTCAATATCAAGTGTACACAGATCCA
GTCAGCAGATGGTCAAGGGTCAAATTGGTTCACAGGCTTCAAGATAATGGGGTATAATCAAGAAA
GCAGTCACATTCAAGATCATTCTGGCTTAATCAAACCTTAATGCCCTGGAACACCTCTGCTAACATC
CAGAACATCTCATACCACAGACTGGTCAAGTCCAGGTTCAAGGGAGTTCAATTGGTGGTTCATCTTCTGG
TCAAACCCAAGTAGTTGCTAATGTCAGCTTCTGGCTCTGGCTGCAAGGAAATATTACGTTGTACCAATCAATAGTG
TCGATCTAGATTCTGGACTCTCGGGCAGTTCAAGACAATGACTGCAGGCATTAAATGCCGACGGACAT
TTGATAAACACAGGACAAGCTATGGATAGTTCAAGACAATTCAAGAAAGGACTGGTGAGCGGGTTCTCTGA
TATTAATGAAACTAAATCTGATACAGATTATTGTGCAACATCCTCTCATCACAGTTGCTGTACGA
TAGATAGTACAGGTATATTACAACAAACACAAATAGCTTCAACTACATCTAGTGGCAGGTTATTCTCA
GATCTTCAGGGAAATTATATCCAGTCGCTGTTCTGAAGAGACACAGGCACAGAAATTCAAGGTTCTAC
AGCACAGCCTGGTACAGCATCTACAACCTCAAGAGTCCTCAGCAGCCAACCAGTCAGCCAAATTGTGC
AAGGTATTACACCACAGACAATCCATGGTGTGCAAGGCCAGGGTCAAAATATATCACAACAGGCTTGCAA
AACTTCAGTTGCAAGTCAACTCTGGAACCTTTAATTCAAGGCACAGACAGTGACCCCTCTGGACAGGT
AACTTGCAAACGTTCAAGTACAAGGGTCCAGAACTTGCAAGATTGCAAATACAGAAACTGCTGCC
AACAAATAACTTGAACGCCCTTCAACACTTGGTCAAGTTGCGGCAGGTTGAGCCTTCACTTC
ACTCCAGTTAGTCTAAGCACTGGTCAGTTGCAAATCTACAAACAGTTACAGTGAACACTATAGATTCTGC
TGGTATACAGCTACATCCAGGAGAGAATGCTGACAGTCCTGCAGATATTAGGATCAAGGAAGAACCTG
ATCCTGAAGAGTGGCAGCTCAGTGGTGAATTCTACCTGAATACCAATGACCTAACACACTTAAGAGTACAG
GTGGTAGATGAAGAAGGGGACCAACAAACATCAAGAAGGAAAAAGACTTCGGAGGGTAGCTTGACCTGTCC
CAACTGTAAAGAAGGTGGAGAGGTTACCAATCTGGGAAAAAGAAGCAACACATTTGTATACCCAG
GATGTGGTAAAGTCTATGGGAGACCTCACATCTGAGAGCTCATCTGGTGTGGCATTCTGGAGAACGCCCT
TTTGTGTAAGTGGTACTGTGGTAAAGATTACTCGAAGTGTGAATTACAGAGGACAGAACAGAAG
ACATACAGGTGAGAAGAAATTGTGTCAGAATGTTCAAACAGCTTATGAGAAGTGACCAACCTTGCA
AACATATTAAAACACACCAGAAATAAAAGGTATTCACTCTAGCAGTACAGTGCTGGCATCTGTGAAAGCT
GCGCGAGATGATACTTGATTACTGCAGGAGGAACAACGCTTATCCTTGCAAATATTCAACAAGGTTCTGT
TTCAGGGATAGGAACTGTTAATACTTCCGCCACCAGCAATCAAGATATCCTTACCAACACTGAAATACCTT
TACAGCTTGTACAGTTCTGGAAATGAGACAATGGAGTAATATTACACAAATACTTATTCAATTGTGGTT
ATTTTTATACAGTAGTGAGAAGAAATTGTTCTAAGTTCTAGATATCTTTATTGATGTGCAAAATT

TTTGGATTGACAGTAAC TTGTTATACATGACACTGAAATGCC TACTTGATATTCCATAGTATA
 TAAAAATGGTAAAATTGATGGTTAGGTACTTGAAATCTAGAAGAAATGAAATTACCAAGTT
 ATATAAAGAGAAAATTGAATTAAACAATGCGAATGGTAGTCTAACCAATGCATCAATCCTGTTGGGTTA
 GTGTAAGAAATGAGAACATGTTGGTATTATCTATTGTAAGATAAAAAGCTGGTGGGTGAAAGAAATCATG
 TTATGATAAAAATTGTAATTTCATGACTGGAATTTTATTATGCATAACTGACAAATCAAGTT
 CCAAGCAAATGTTACATAGTGTAGGTTACTTAGCTTATCAATTGTCATTGAAAGCTAATTATTTAA
 TTAGGTTAACTATGTACAATATTAAAGCATTACTCTGTAAGATTGAAAACACATTAAACATGGAA
 CTCTAGGGATAGTCACCTTTAAATCCTGTTGAAAGCCATGTTAAGATTAAATTGCCAAATAATGTC
 TTGTTAATATTCTCAATAACGAAGTGGCAATATAACCAATGTTAAAAGTTAAAATGTATAAGT
 TGAGGCATTGGGGTAAGAGAATGTTATAGTGAATTATCCTTCTGACTATTGGAGGACCAAAAAA
 ATAAGGTGATTGCGTCTTAGCAGTGTATTATCCAATCTGTTCCAAAACCATGTCCTCCAGGGCCT
 TAAAGCCATCATGTAATTACAGTAAGTGTACATATGCAAACATAACAAAATCACTCCATAGTGC
 GATACTCCAACCATA GGATATTAGTCATAGAAGAACTAGAGGTTTATGATATTTTAAAGTCTTTT
 TTTTTGTCTAGGTAGTCAGTCTGCACCTAAATATCAATCATTCCCTTTGCTTCTCCCTTAAATT
 TATATGTATCCAGTACATTAAATTGAGAACCGTATGTTTTATTATGCTGTATTCTTTATTTTA
 ATTATTGTTATATTCAATTCAAAAATGTACAAAATAAGTTACATTGCTGGTCTGTAAGAGCTATAC
 AGTTTCTAAATGTATACCTGTAACTGCAGCAGTCACCTATTCAAAAATTGAAATTCTGTTCATTTG
 TTATTCTTAAGACCACCTCAAATTAAAGGCTACCTATTGACGTTAAAGTGTATTATAACAGTGTGGT
 AGTTAATAAAACACTATTTTTTCTTTGAAAAA AAAAAA AAAAAA AAAAAA AAAAAA
 AAAA

IGPPSPGDDEEEAAAAGAPAAAGATGDLASAQLGGAPNRWEVLSATPTTIKDEAGNLVQIPSAATSSGQY
 VLPLQNLQNQQIFS VAPGSDSSNGTVSSVQYQVI PQI QPQSADGQQVQIGFTGSSDNNGINQESSQIQIIPGS
 NQTLLASGTPSANTQNLIPQTGQVQVQGVAIGGSFPQQTQVVANVPLGLPGNITFVPINSVDLDSLGLSG
 SSQTMAGINADGHLINTGQAMDSSDNERTGERVSPDINETNTDTDLFVPTSSSQLPVTIDSTGILQQN
 TNSLTTSSGQVHSSDLQGNYIQSPVSEETQAQNIQVSTAQPVVQHLQLQESQQPTSQAQIVQGIFTQPTIHG
 VQASGQNISSQALQNLQLQLPNGTFLIQAQTVPSPQVWQTFQVQGVQNLQNLQIQNTAAQQTILTPVQT
 LT LGQVAAGGAFTSTPVSLSTGQLPNLQTVTVNSIDSAGIQLHPGENADSPADIRIKEEEDPEEWQLSGD
 STLNTNDLTHLRVQVVEDEGDQQHQEGKRLRRVACTCPNCKEGGGRGTNLGKKQHICHI PGCGKVYGKTS
 HLRAHLRWHSGERPFVCNWMYCGKRFTRSDELQRHRRHTGEKKFVCPECSKRFMRSDHLAKHIKTHQNKK
 GIHSSTVLASVEAARDDTLITAGGTTLILANIQGSVSGIGTVNTSATSNQDILTNTIEPLQLTVSGNE
 TME

+++++

#3

NM_175886

Highly conserved throughout vertebrates and beyond but it looks like in all other species, including chimps, there is an AUG in place of the ACG – there is actually a human paralog on a different chromosome, much more common in the EST database, with AUG too.

PRPS1L1 phosphoribosyl pyrophosphate synthetase 1-like 1

Synonyms: PRPS1; PRPS3; PRPSL; PRS-III; PRPS1L1

ATTAGAGTCTGTGCTTCACTCCCGTCCAGCCTCAGCGCAGCTGGATCGCTCGACGGAGTGCTCTGGTA
 GTTGGCCAAGACGCCGAATATCAAATCTTCAGCGCAGCTCCACCAGGACTTATCCCAGAAAATTGCTG
 ACCGCCTGGGCCTGGAGCTAGGCAAGGGTGTGACTAAGAAATTAGCAACCAGGAGACCTGCGTGGAAATT
 GATGAGAGTGTGCGTGGAGAGGATGCTCACATCGTTCAGAGTGGTTGGCGAAATCAACGACAGTCTAAT
 GGAGTTTGATCATGATTAATGCCTGCAAGATTGCTTCAGCTAGCCGAGTTACTGCAGTCATCCCAGTGC
 TCCCTTATGCCCGACAGGATAAGAAGGATAAGAGCCGGTCCCCATCTGCCAAGCTTGTGCAAATATG
 CTCTCTATAGCAGGTGCGGATCATATCATCACCATGGACCTACATGCTCTCAAATTAGGGCTTTTGAA
 TATCCCAGTAGACAACTTGTATGCAGAGCCAATGCTCTGAAGTGGATAAGGGAGAATACCTGAGTGGAA
 AGAACTGCATTATTGTCTGCCAGATGCTGGTGGAGCTAAAGAGTGACCTCCATTGAGACCAAGTGAAT

GTGGACTTTGCTTGATTATAAGAACCGAAGAAGGCCAATGAAGTGGACTGCATAGTGTAGTGGAGA
TGTGAATGATCGTGTGGCTATCCTGTAGATGACATGGCAGACACTTGTGTTACAATCTGCCTCGCAGCTG
ACAAACTCTCTCAGCTGGAGCAACCAGAGTTATGCTATCTGACTCATGGAATCTTCTGGCCCAGCC
ATTTCTCGCATCACACTGCATGCTTGAAGCAGTGGTAGTCACCAATACCATACCTCAAGATGAGAAGAT
GAAGCATTGCTCCAAAATACGAGTAATTGACATCTCCATGATCCTGCAGAAGCCATAAGGAGAACTCATA
ATGGGGATCTGTTCTACCTGTTAGCCATGTTCTTAAACAGAATAACTCTAGGTTATGCTATT
TAAAATAAAATAAGATTAATAAAAAAA

TPNIKIFSGSSHQDLSQKIADRLGLELGKVVTKKFSNQETCVEIDESVRGEDVYIVQSGCGEINDSLMELL
IMINACKIASASRVTAVIPCFPYARQDKDKDSRSPISAKLVAANMLSIAGADHIITMDLHASQIQGFFDIPV
DNLYAEPTVLKWIENIPEWKNCIIVSPDAGGAKRVTSIADQLNVDFALIHKERKKANEVDCIVLVGDVND
RVAILVDDMADTCVTICLAADKLLSAGATRVYAILTHGIFSGPAISRINTACFEAVVTNTIPQDEKMKH
SKIRVIDISMILAEAIRRTHNGESVSYLFSHVPL

+++++

#5

NM_003213

Region and non-AUG initiation conserved from zebrafish to human.

TEAD4 TEA domain family member 4

Synonyms: TEF3; RTEF1; TEF-3; EFTR-2; TEFR-1; MGC9014; TCF13L1; hRTEF-1B; TEAD4

GTGGCCCCGCCCACTCCCTCCGGCTCCCTCCCTCCGCCGGCGCGCAGCTCATCTCATTCCAGCCCTCATTC
CGCGCATTCAGCGTCTCCTCGCACACTCGAGGCCAGGGGGCGGGAGGGCCGCAGCTCCGGCGCCGCC
GTCGGCCAGAACGATGCCGGCGGAAGAGTTGGCGCTGGCGACTCCTTGAACCTGGCTTAGCG
CACCCATCCCACCTTCCGACCCCTGGGACCGGTCAACGAGCGCTCTCCAAGCGGAGCCTTGGAGGGCA
CGGCCGGCACCATACCTCCAACGAGTTGGAGCTCTCCACCTCCCTGAGGGGAGCACC GCCCTCTGGGGGC
AGTCAGGCAGTGGACAAGCCATCGACAATGACGAGGCCAGGGCGTGTGGAGGCCGGATATTGAGCAGAGTT
CCAGGAGGCCCTGCCATCTACCCGCCCTGTGGCAGCGCAAATCATCTGTGGACGAGGGCAAGATGT
ATGGTCGAACGAGCTGATTGCCCGTACATCAAGCTCCGGACAGGGAGACCCGACCAGGAAGCAGGTC
TCCAGCCACATCCAGGTGCTGGCTCGCAAGCTCGAGATCCAGGCCAAGCTAAAGGACCAGGCAGC
TAAGGACAAGGCCCTGAGAGCATGGCTGCCATGTCGCTGACAGATCATCTCCGCCACGCCCTCCACA
GTAGCATGGCCCTGCCGGGGCCCGGCCAGCAGTCTCAGGGTTTGGCAAGGAGCTTGCCAGGC
CAAGCCGAACGCCATGATGTGAAGCTTCTCTCAGCAAACCTATGCTGTCCAGCCTCCGTGCCT
GCCAGGGTTGAGTCTCTGAGGGCCGCCCATGCCCTCTGCCCGGCCAGGGCCATGGCAGGG
GCAGCGTGGCCAGCTCCAAGCTCTGGATGTTGGAGTTCTGCTCTGCCCTCTGGAGCAGCAGCAGGCC
ACGTACAACAAGCACCTGTCGTGACATTGGCAGTCCAGGCCAAGCTACAGCGACCCCTACCTCGAAC
CGTGGACATGCCAAATCATGACAATTCCCGAGAAAAAGGGTGGACTCAAGGATCTCTCGAACGG
GACCCCTCAATGCCCTTTCTGTGAAGTTCTGGCAGACCTCAACACCAACATCGAGGATGAAGGAGC
TCCTCTATGGGGCTCCAGGAGTATGAGAGCCCGAGAACATGATCATCACCTGCTCCACCAAGGTCTG
CTCTTCCGGCAAGCAGGTGGAGAAAGTTGAGACAGAGTATGCTCGTATGAGAATGGACACTACTCT
ACCGCATCCACCGTCCCCGCTCTGTGAGTACATGACAATTCCACAAGCTCAAGCACCTCCCTGAG
AAAGTACATGATGAACAGCGTGTGGAGAACCTCACCATCTGAGGTGGTCAGCCAGTGGAGCAG
GACCTTGCTGTGCATTGCCATGTCAGGGTGGAGAACATGATCATCACCTGCTCCACCAAGGTCTG
GGCTGGTGAAGAATGAGAGACTGGGAGCAGGGAGGGGGAGAGAGACGTGTGAGCAGGAAACGGGAGC
TGGGGAGGGGACCTGCAGGGCAGCCCCCTGAAGTGCAGAGAGAGCTGAGAGGAGCAGTTGTGACTCTACC
CAGGAACAAACTGTGCCTGAACCTGAGGTGCCAACCCAAATAACCCAAGATGCTGTATTTCAGAG
AAAAA

LEGTAGTITSNEWSPTSPEGSTASGGSQALDKPIDNDAEGVWSPDIEQSFOEALAIYPPCGRRIIILSDE
GKMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLARRKAREIQAKLKQDAKDKALQSMAAMSSAQIISAT
AFHSSMALRGPGRAVSGFWQGALPGQAGTSHDVKPFSSQTYAVQPPLPLPGFESPAGPAPSPSAPPAPP

WQGRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDKFPEKKGLKDL
FERGPSNAFLVFWADLNTNIEDEGSSFYGVSSQYESPENMIITCSTKVCNSFGKQVVEKVETEYARYENG
HYSYRIHSRPLCEYMINFIHKLKLPEKYMMNSVLENFTILQVVTNRDTQETLLCIAYVFEVSASEHGAQH
HIYRLVKE

++++++

#6

NM_003214

Region and non-AUG initiation conserved from zebrafish to human.

TEAD3 TEA domain family member 3

Synonyms: TEF5; TEAD5; TEF-5; DTEF-1; ETFR-1; TEAD3

TCCTCAACACAAACTTCCGTCCCCTCGCTCCCTCCGCGCTCGGCCTCCCGCTCCAGCCCCGGCTC
ATTCCGCACATTCCGGCCAGCCCCCTCCCCACGACCCCCCTCCCCGGCCCCCTTGCAGCTCCCTCGGGC
CCGGCGAGCGGCCGGCCGGAGCGCCCCCGAGCTCGGACCAGGCTCAGCCGCCAGTGGCTCAGGGC
CAGAGCCCAGAGCAACCAGCACAA~~TAGCGTCAAACAGCTGGAACGCCAGCAGCAGCCCCGGGAGGCCCCG~~
GAGGATGGGCCGAGGGCTGGACAAGGGCTGGACAACG~~TAGCGGAGGGCTGTGGAGCCCCGACATCGA~~
GCAGAGCTCCAGGAGGCCCTGGCCATCTACCCGCCCTGCAGCCGGCGAAGATCATCCTGTAGACGAGG
GCAAGATGTACGGCGAAATGAGTTGATTGCACGCTATATTAAACTGAGGACGGGAAGACTCGGACGAGA
AAACAGGTGTCCAGCCACATACAGTTCTAGCTCGGAAGAAGGTGCGGGAGTACCAAGGTGGCATCAAGGC
CATGAACCTGGACCAGGTCTCAAGGACAAAGCCCTCAGAGCATGGCGTCCATGTCCTCTGCCAGATCG
TCTCTGCCAGTGTCTGCAGAACAAAGTTCAAGCCCACCTTCCCTCTGCCAGGCCGTCTCTCCACTTCC
TCGCGGTTCTGGAGCAGCCCCCTCTCTGGGACAGCAGCCTGGACCCCTCAGGACATCAAGCCCTTGC
ACAGCCAGCCTACCCCATCCAGCCGCCCTGCCGCCAGCCTCAGCAGTTATGAGCCCTGGCCCCCTCC
CCTCAGCTGCTCTGTGGCTGTGGCAGGACCGTACCGATTGCTCTCCCTCCGGCTGCCGCTCTGGAG
TATTCAAGCCTCATGGAGGTGCAGCGAGACCTGACACGTACAGCAAACACCTGTTGTGCACATCGGCCA
GACGAACCCCGCCTCTCAGACCCACCCCTGGAGGCAGTAGATGTGCCAGATCTATGACAAATTCCCCG
AGAAAAAGGGAGGATTGAAGGAGCTATGAGAAGGGGCCCTAATGCCCTCTCCTGTCAAGTTCTGG
GCCGACCTCAACAGCACCATCCAGGAGGGCCGGAGCCTCTATGGGGTCAGCTCTCAGTACAGCTCTGC
TGATAGCATGACCACATCAGCGTCTCCACCAAGGTGTGCTCTTGGCAAACAGGTGGTAGAGAAGGTGGAGA
CTGAGTATGCCAGGCTGGAGAACGGGCCTTGTGACCGTATCCACCGCTGCCATGTGCCAGTACATG
ATCAACTTCATCCACAAGCTGAAGCACCTGCCAGAAGTACATGATGAACAGCGTGTGGAGAACTTCAC
CATCCTGCAGGTGGTCAGGAGCCGGACTCCCAAGGAGACCCCTGCTGTCAATTGCTTTGTCTCGAAGTCT
CCACCAAGTGAGCACGGGCCAGCACCATGTCTACAAGCTCGTCAAAGACTAGGGTGCCTCTGCCCTCC
TTAAGGATGCAGGGTGAGCATCTCTCCACACCTGCCCTGGCACCCCTGGGGGGGTCCAGGATTGAGGAT
TCATCTACCTGCCAGGCCCTCAGGCCAGGACCCAGGAGGCCCTCCACCTACCCAGCACACACTCCCT
GCCACTGTTCTGCCCTTAATTGTGGAGAAGAGAGGAGAGGGCTAGCGGTGGGGCAGCCTGTCCGG
GGCGCTGACCCACCATCACCCCTGCTGCCAGCCTGCCGTGACCTCAGAGAGGTGGGGATAGGGGACACC
TTCAGCCTCCAGCATGTGTGCCACTGTGACCCACCCACCCCTGGGGAGCATGATGGGCAGGTGAGGGC
AGGATGGAGACCAAGGGAGTCAGTGAGCAGAGGCCCTGGAGGTGTCCGGTTGGGGACTGAGGACAGA
GGGGCCACACTTCCTGCCCTTGTGCTGCCAGGCCCTGGTGCAGACTCCTTGATGGCTTGTTGGC
CTCAGACTCCGCACAGCGAGCGTAGGTCTCTGGGTTAGATGAAGTGCCAGGCTCCAGGAAGTTGAGGG
ACCCACAGGAGAGGTGGCAGAGCTGGAGTCTCATCCAGGCTGCTTGTCCCCAGAGCCCAGGGTTATAC
TACCTCCCTGGGGCGGGGGCTGGCCAGGGTAGGGAGAGGCTCTGCAGTGTGGAGTGGAGCCTCATCGA
GGGGCGCTGGGTTAGGGAGCACCTGTTAGACTGGGATGAAGAAGGGAGCACAGCAGCTACTAGACCC
CATTAGCACCTCATTAGGCCACAAGCCAGCCAGGGGCCAGGAAGATGGGCACCCCCCAGCACCCCTCCA
GATTGAGAGCAAGGTAGAGGAAGGGAGTCCCAGCCTCTGGCAGACAGAGGCCAGAGGGAGAGAGTAGCA
GAAGGCTTTGATTTCTCTGCCCTGAGGCTTGAAATCTGACAAACCCCTGGGGACTGCTCCCTTAGG
TTCTTCCCCACCTCAATCTACCTGCCAGGCTAGAGTAGCAGCTCCAGACCCAGTTCTGGGACTGAAGGTTAAC
CTTCACCTGCTGCCCTCTTAACACCCAGGCCAGAGCAGCTGGCCTGTCAGCAGGCCACCTGTGG
GTATTATGAGTTCATATGAAGTACTGTGCCCTCCCTCATCCGACCCCTGCCAGCTTCTGA
AGGTCTCACTGTTGCATATCGCTCAGGCCACCTCAAACCCACCTAGGTTATAATGTATATTAT
ATTTTTTGTGTATTTAAAATCCAGCTGTATGGGTTATATCATAAATGCAGCTGGGGTTGGAGCAGG

GGCCCTCAAAGGCCAGCTCCTGCTAAAAAAAAAAAAAAATTAAAGTTATTGTTGTGGGTCA
GTCATGAAAAAAAAAAAAAA

IASNSWNASSSPGEAREDGPGLDKLDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIIILSDEGKMYGRNE
LIARYIKLRTGKTRTRKVSSHIQVLARKVREYQVGIKAMNLDQVSKDKALQSMASMSAQIVSAVLQN
KFSPPSPLPQAVFSTSSRFWSSPPLLQQPGPSQDIKPFAQPAYPIQPPPLPPTLSSYEPLAPLPSAAASVP
VWQDRTIASSRLRLEYSAFMEVQRDPDTYSKHLFVHIGQTNPADPPLAEDVRQIYDKFPEKKGLKE
LYEKGPNAFFLVFKWADLNSTIQEGPGAFYGVSSQYSSADSMTISVSTKVCSPFGKQVVEKVELEYARLEN
GRFVYRIHRSPMCYMINFIHKLKHLPEKYMMSVLENFTILQVVTSRDSQETLLVIAFVFESTSEHGAQ
HHVYKLVKD

+++++

#6

NM_031895

Conserved in vertebrates though in some fish initiated at AUU while others at AUG,
though the AUG initiated one might be another paralog.

CACNG8 calcium channel, voltage-dependent, gamma subunit 8

Synonyms: CACNG8

CCCCGCTTCTGCCTGCGCTGTGAACCCCCCCCAGCCGCCGCACGGCCCCGCCCCGCTGCCCCGGTGGT
GGCCCACGGCCCCCGCTGCCGTGGTCAAACCTGGAGTCGCTGAAGCGCTGGAACGAAGAGCGGGGCCTC
TGGTGCAGAAGGGGGTGCAGGTGCTGCTGACGACGGTGGCGCCTCGCCGCTTCGGCCTCATGACCAT
CGCCATCAGCACTGACTACTGGCTCTACACGCGCCCTCATCTGCAACACCACCAACCTCACGGCCGGG
GCGACGACGGACCCCCCACCGGGGGCGCGCCCTCGAGAAGAAGGACCCGGCGGCCCTACGCAC
TCGGGCCTCTGGAGGATCTGCTGCCTGGAAGGGTTGAAAAGAGGCCTGCGTGAAGATCAATCATTCC
GGAGGACACGGACTACGACCACGACAGCGGGAGTATCTACTCCGAGTTGTCGGCCTCCAGCATTTCC
CCATCCTTAGCGCCATCCTGCTGCTGCTCGGGGGTGTGTGCGTGGCGGCCCTCCCGCTACAAGTCCAAG
AGGAACATCATTCTGGCGCAGGGATCTGTTCTGCTGGCAGCAGGCCTGAGCAACATCATGGCGTATCGT
GTACATCTCCGCCAACCGGGCGAGCGGGCCGAAGCGGGACGAGGAGAAGAAAACACTACTCGTACG
GCTGGCCTTCTACTCGGGGGCTGTCGTTCATCTGGCCAGGTGATAGGCGTGTGGCCGTAACATC
TACATCGAGCGCAGCCCGAGGGCGCACTGCCAGTCTGCTCGGACCTGCTCAAGGCCGGGGCGCG
CGGCAGTGGCGGGAGGGCGCCCTCGGCATCCTCCGTCTGCCAGTACCGCTTCCGCTACGCCGCCGCT
CCCGCTCTAGCTCCGCTCCAGCGAGCCGTCGCCGCGCGCTCCCGCCGGCCCCGGGGCCCG
GGCTTGCCCTCACGGACATCTCATGTACACGCTCAGCCCGACCCCTCCAAGGGCAGCGTGGCCGCG
GCTGGGGGGCCGGCGCGCGGGCGCCGCGCGCGCGGGTGGGGCGTCCGGCGCGCCGGGGCGCCGGG
GCGCGGGAGGGCGGGCGGGGGCGGGTGCAGCGGGACCGCGGGGGCGTCCGGCTTCCCTACGCTG
CACAACGCCCTCCCAAGGAGGGCGGGCGCGCGTACCGTCACGGTCACCGGTACCCGGCCGCCCG
GCCCGGCCACCCGCCCTGCGCCGCCGGGACCCCTGGCCAAGGAGGGCGCCGCTCCAACACCA
ACACGCTAACAGGAAAACACGCCGTGTAGGGCGGGGGAGCCAGGGCGTGTCCGGGG

LESLKRWNEERGLWCEKGTVQVLLTTVGAFAAFLGLMTIAISTDYWLWYTRALICNTTNLTAGGDGTPHRG
GASEKKDPGLTHGLWRICCLEGLKRGVCVKINHPEDTDYDHSAEYLLRVVRASSIFPILSAI
LLLG
GVCVAASRVYKSERNIILGAGILFVAAGLSNIIGVIVYISANAGEPGPKRDEEKKNHYSYGS
WSFYFGGLSF
ILAEVIGVLAVNIYIERSREAHCQSRS DLLKAGGGAGGSGGGPSAILRLPSYRFRRRS
RSSRSSEPS
PSRDASPGPGPGPFASTDISMYTLSRDP SKGSVAAGLAGAGGGGGAGVAGFGAAGGAGGGGGGGAGA
ERDRGGASGF LTLHNAFPKEAGGGVTVTGPPAPPAPAPSAPAPGTLAKEAAASNTNLNRKTPV

+++++

#8

NM_016178

CUG initiation well conserved in mammals.

OAZ3 ornithine decarboxylase antizyme 3

Synonyms: AZ3; OAZ-t; TISP15; OAZ3

GTTGCCAACCTCTGCCACCCACCTGTGAACCTCACTTGCACAGGGAGGCAC **TGA**ACTGAAAAA**CTGC**
CTTGTAAAGAGGTGTGCCCTCTGTCTACTCCCTTCTTATCAAGAGGGAAAACACGTAACTACCTC
TACCCGATCTGGTCACCATAACGCCATTACCTTACTGTTACAAGTACCGGATCACTCTCCGGAGAACAT
GCTGCCCTCGTTATAAAAGCATCACTTAA **GA**AGAGGAGGACTTGACACTCCAGCCCCGTTCTGCC
TCCAGTGCTCC **TGA**GTCCTAGTAGGCCTCCAGGAGGGCAAAGCACCGAGCAGGTAACCACGACCAGCT
TAAAGAACTGTATTGGCTGGGAACTTGACGGTGTGGCTACTGACCCCTGCTCACCAGGACCCAGTAC
AGTTAGACTTTCACTCCGCCTTACCTCCCAGACCTCTGCCATTGGCACGGCTTCTGTGACCGTCGA
CTCTTCTGGATATCCCATATCAGGCCCTGGATCAAGGCAACCGGGAAAGTTGACTGCAACCCCTGGAGTA
CGTGGAAAGAGAAGACAATGTGGACTCTGTGTTGTGACTTCCAGAATGATCGAACGACAGAGGTGCC
TGCTGGGGCCTTCAGCTACATGGCTTGAGGTGGTCAGACAGATCACCTGCCCTCCCTGGAC
AATGTCATCTTATGGGTATCCCCTGAAAGGGATGTTGGCACCTGCCAGTGAGCCTCT **TGA**ACATG
CTTATTCCAACGCTTGAGGGCTGGAAGCCTGACACATGAAATCAGGGCCCGGGATGTGATTCAAGGAC
ACTTTCCATCCTAGGAATAAAGGTAGTGAATCAAAAAAAAAAAAAAA

LPCKRKRPSVYSLSYIKRGKTRNYLYPIWSPYAYLYCYKYRITLREK**M**LPRCYKSITYKEEEDLTQPRS
CLQCSESLVGLQEGKSTEQGNHDQLKELYSAgnLTvLATDPLLHQDPVQLDFHFRLTSQTSAHWHGLLCDR
RLFLDIPIYQALDQGNRESLTATLEYVEEKTNVDVFNFQNDRNDRGALLRAFSYMGFEVVRPDHPALPPL
DNVIFMVYPLERDVGHLPSEPP

+++++

#9

NM_021961

Region and non-AUG initiation conserved from zebrafish to human.

TEAD1 TEA domain family member 1 (SV40 transcriptional enhancer factor)

Synonyms: AA; REF1; TCF13; TEF-1; TEAD1

TTCCGAACATTCTAGCATCGCTCGCGCCGCCGCCGCGCCGCTGAGCCGAGCCGAGCCTCTGCTGCCGCCG
CCCGGGCCCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC
CCCGCGCGCGGGGGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC
TCCACCTTGCCCCGGAGCGCGGGCAGCAGCCCAGCGCGCCAGCCGGCCCCGGGGCAGGAGCAGCGGTGCTAGGC
AGGGTTGGGTGGCCGGGCCAGGGACCGGGAGCAGGGAGCAGGGAGGGAGCCGGGACCCGAGCAGAGGGGGGA
AGCGGCCCGAAGTTGCCTCGGACTCGCCGGCGCTGCGGTGGCTCCCTGGCCAGGACTGTTGCTGCC
GCTGCCGCCGCCCTTCATTGCACATTCAAGTGGAAATTTCAGGAGTCAGCAGAAACATTGTGTCCAA
AAAGACT**TGA**GTCGAGTTACACCAAAACCCAGGAGGAGACTCCCTGAAACTCCCTCCCTTCGGT
TTATTTCTGAAAGGCTCAGGCTTCGGC **TGA**AAATCCACGCCAAATTGAGCCAGCAGCTGG
GCCGCAGTGAAGGCCCTGCCAAAACATGGAAAGGATGAGTGACTCTGCAGATAAGCCAATTGACAATGAT
GCAGAAGGGGTCTGGAGCCCCGACATCGAGCAAAGCTTTCAGGAGGCCCTGGCTATCTATCCACCATGTGG
GAGGAGGAAAATCATCTTATCAGACGAAGGCAAATGTATGGTAGGAATGAATTGATAGCCAGATACTCA
AACTCAGGACAGGAAGACGAGGACCAGAAAACAGGTGTCTAGTCACATTCAAGGTTCTGCCAGAAGGAA
TCTCGTATTTCATTCCAAGCTAAAGGATCAGACTGCAAAGGATAAGGCCCTGCAGCACATGGCGGCCAT
GTCCTCAGCCCAGATCGTCTGGCCACTGCCATTATAACAAGCTGGGCTGCCCTGGATTCCACGCCGA
CCTTCCCAGGGGCCGGGGTCTGCCGGGAATGATTCAAACAGGGCAGCCAGGATCCTACAAGACGTC
AAGCCTTGTGCAGCAGGCCCTACCCCATCCAGGCCAGCGGTACAGGCCAGGCCAGGTTGAGCCTGC
ATCGGCCCCAGCTCCCTCAGTCCCTGCCCTGGCAAGGTGCTCCATTGGCACAAACCAAGCTTCCGCTGGTGG
AATTTTCAGCTTCTCGAGCAGCAGCGAGACCCAGACTCGTACAACAAACACCTCTCGTGCACATTGGG

CATGCCAACCATCTTACAGTGACCCATTGCTTGAATCAGTGGACATTGTCAGATTATGACAAATTCC
TGAAAAGAAAGGTGGCTTAAGGAACTGTTGGAAAGGGCCCTAAAATGCCTTCTCGTAAAATTCT
GGGCTGATTAAACTGCAATATTCAAGATGATGCTGGGCTTTATGGTGTAAACCAGTCAGTAGAGAGT
TCTGAAAATATGACAGTCACCTGTCACCAAGTTGCTCCTTGGGAAGCAAGTAGTAGAAAAAGTAGA
GACGGAGTATGCAAGGTTGAGAATGGCGATTTGTATACCGAATAAACCGCTCCCAATGTGTGAATATA
TGATCAACTTCATCCACAAGCTCAAACACTTACCAAGAGAAATATGATGAACAGTGTTGGAAAACCTC
ACAATTATTGGTGGTAACAAACAGGGATACACAAGAAACTCTACTCTGCATGGCCTGTGTGTTGAAGT
TTCAAATAGTGAACACGGAGCACACATCATATTACAGGCTGTAAAGGACTGAACATGGTATTATAT
ATATAGATATCTGTATATACACACACATATGTCACACACACTCTCTCCATTATCGAACGACTGA
CTGTAACCTCACACACAGGGTGGTGCCTGGCCCCGAGGTACCCCGACTTTCTAAATCTGTTGAG
TGAAGTCATTTCATGTGTCATACTATCATTGAGCTGTGAAAGTCTGGTACAGTTGTAAGGAGAA
ATTGAGTTGTTCTATGTCAGATGTGAGCCCACAATTCTGGAAAGGTGAACCTGAACAACC
CAAGTCCTCTGCAGAGCCTGTTCTAAATTGAGTAGAAAATATTGAGACAGAGCATTGCCATGGGA
CATTACAGCCTTATACAAATGTATTAGTCTCTTTCCAACATAAAATTCTGTTAAGATACAA
GTAAAATTAACTTTAAATATAATGTAATTAGTACACAAAACATAAGAATCTTAGACTTATCTTGTAA
CTAATTAGGTGGAAAGTTATGAAAGAATGTAATTCACTAAATTATTTAAATGAAACCTTTTTCT
TTTGAAACCAAATGTTAAACTATAGCCTTAAGAAATGCTGGTAGAAGTGTCTAAATGAGACAAATTG
ACTTTATCCTCAAGGTTAACACTAATCTCTAACTTAACTCTGAACAGGTATTACAAGGAAGAA
AACTCACCCTTATCCTAACATATATAGTATATTAAAAAAATATAAAATTGTATTGACTAATGTGATG
ATGGATTATTTAATGAAAAAGAAAAATGGCTCTTTGCAATAAGTAGATACTGAAAAAAATCTAAA
CTTACAATGTTATAGTCTGTGTCAGTTATTTATATGGACGACCAATTCTTATTAAGATGAG
TAAATATTGAACCACTGAATTAAATAACAAAATTAAATTGGCATGAATACGAATACTGCACTGTG
AGATGCAAAGTATACAGAATCTGTGGCTGGAGAAAATTCTACAAATAGACAAGTAAAAGGCTCATCAGT
TTAGCATCTGCTCCCCAGAAAATTGTAAGCATCCTCACAGCCTGTGGATACATTCTTATTTCTAGT
GACCAATATGCATATTAAACCTGCTATAACTAGGGCTATATGTGAGGTATGTGTACATACATACACAAAT
GCACATATAGAGTTAACACATTAGTGAACACTTGTGTTAGTGCACTCAGTTGCTAGGTGCTGATATGTA
CGTATATCTCAATGTGTCGTAGACTTAGATACTCCTTGAAGCACATCCATTCTTAGCGTCTCTCA
GTAAGTTACAGTACTGTTGACTTAGTTAAGAGGCCAGCTACCTATCTGACCTTCAAATAGGC
TCATTTGGAGATTCTTGTGCCAGGAGATTCAACTTCCAATCTAAGTATTCCAGAGCATTGCCAGGC
AGAGTTGGTTGATGTGGCCAGATGTTGAGTTATTCCCTTAAGTGTGTTCACTGGGAGAGAACAGGG
GTGCTCTCCAGCTCCAAAGAAATATGTTTGTAAAGTGTAGGAACATGTGACACAATAGAACATGA
AATAAGTTTTAACTTGTAAAACATGTCAAGATTTCCACCAAGCTAGAAAATAAAAACTAGTTCTA
CCACATCCAATTAAACTTACACACCCCCCTCCCTGTCACACACTGCTTGCCTGCTTTCTATTATTA
CATCAGTCAGCATCTGTGGTCCCTAACATGAGGATGTGGCTGGCTGGAAACAGCAAAACACTAAGC
CTGACCTCTCCAAATTGGGAAGACCAAGAGGAGAAAGTGCAAAAGTCTGCCCCATTGGAATGCCATTCT
TCTAGAAACCAGTGGACAGTGCTCTGCCCCCTTCACGGTTGTCACATTACAGTGAATTCTGTTGAACAC
TGGCCTGTGAAGGATTGCCCCAGGTGCTCTTACGGTTGTCACATTCTCTTAAGGAAAGCCCCATCCTCTCCAG
CCCTCTAGGGATGTTCTTGTCTTACGGTCTCTTACGGTTGTCACATTCTCTTAAGGAAAGCCCCATCCTCTCCAG
GACCAGGAGTTATGACCAGGGAGCACAAATGGCTAAAGCCAAGCTGCTTAGAACATTCACTGAGGGAGAG
CTGTCTGGTCATATTCTACCCAGGAATGGTACTTTCACTGAGCTGCAGCCAGGAGGGCTTGGGATTCCCTT
CCAAAGCACAAAATCTGGGACCCAAGAAGAACAGCTAGAGGACAACCTGTGTCACAGAGACGGGGAC
AGCCCAGTCTGCTGACCTCACAGGGTCAGCTGGGCCCCCTGGTGTCTACCACCTGCATCCTCTGCTCA
GAATGCCTTGCAGTTGAGTTCTGGTTCTATGATTGACCTTGAGGTTACTCCTGCTTACAACA
TTCTAAGGATTAAAAGTTACTCTTGTCTGTTCTCTAAAGCTTCTCAGGACAGATAATTCC
CTGTCTAACCACTGGTCCAGTCATCCAGTGGGCTCTCTTGTCTCTCCAGATTAGACCTTGGGTGA
GATTGGCATCACAAACATCTAATCTGAGTCTGTCTTGTCTTGCCTTCATTCTGTATGGCAGTCTCCCTTGTAA
TAAAAGCTTCTAAAGCATACTAAAGAAGCCTTCCAGAGGCCGCTTGCTCTCTCCAGGTGCTCT
CCCCTCGAGACCCCTCTGGTGCCAGGCTGCTTCACGGCCATCTGTGTTGTCAGAGTTGGAGGCC
AGTTTCCACAGCTAAACAGGGAGGAGCTGCAGAATGGGCTCTGGTCTGGCATTCTCCCTCAT
AGAGGCTGAGAATAAAACAAAGGACTTATTACACATGTTCTAGAACCCAGAATGCCCAAGTTACCTGAG
ACCAGGGTTCTCAACCTGACACCATTGACATTGGACTGGTAATTCTTGTCTGCAGAGCTGCTCT
TTGCACTGTAGGAGATTACTAATATCCTGGCCTACCCAGTAGTACCAACTAGCACCTATTCCCCACCC
AGCGTGTCTCCAGATATTGTCAAATATCCCAGTGGGTGCAAATGATCCCTGGTCAAGATCTGTTGCCAA
GATGTTACAGGTACAATGACCCACATTGAAATTGTTCCCTTCAATTACCCCTGTGAAAGCATCTC
CTAGAGCCTTGCAAGAGGCAGGTGACATTGTGTCATATTCTCTGTTCTGAGAACTTCTGTTCAAC
AATTCTCTCGTACAAGTATTCTTCACTCAGCACTGGGAAGTTGGAACAGCTGGTCACCATCATC
CCTTAATCAACTCACACCTGTTAAAGAGTGTGTTGACCTTCATCCCTAGTTACTGGCGTTA

AAAAAAAGTCTCAGCAATTTCATTATTCTCGTGGTCTCATTATCAAACCTTACTTATTCGGCATATT
TCCTCTGGGCTTCTCTAGTTCTGCCTTACAAGCAATGCTTCTGTAAATTATTGAAACCTCTGGAAC
ATTTCACCTTAGAGATGGAGGATGGAAGGATTGGTACAGAGAAGGGCTAAGATACGTTCTGTCTG
GCTGAAAGCACAGTCTACTCTCCTCGTTGTCATGAGAAAGTTGAGGCCAGAGGGGAGGTGACATGTT
TAGAGTCACCCAGCTGGTTAGTGCAGAGAAAAGCGTGAGAGTTGTCAGGATTGCCACTTGGTCC
GGCCTCTCCTGGGGAGGCTGTCAGTCTTAGGTGCTCTAAGCTTAATCCCTCAGAATGTTGACAGGTCA
GCTTAGAGAGATGGGAGATTCAAGGATCCCCCTGTCAGACAGCAGCCTCACCGGATGTCCTCCCAC
ACTGAAGTGTCTGTCCGACCATTGCTATCTGAGGCATCCACAAGCAGTAGGAAAGCTGGGAGGCCATT
TACTTCTGAGGACAATTCCCAGCCACAGGCTCTGAGTCAGTAAATTCTATTGTAAGCATCCTAGCAGCA
AAGTCCTGCACTCAGACCAGCAAAAAACAGCCCCATTCAAGTACTGGTGTCAAAAGTCCCCGAACGA
CTTTAAACCAAGTCTTAAGGTTCAAGTACTGAGTCAGTAACTTACAACCAAGGTACATATTAAATTGTATA
ATTATTAAATCATCTGAGATGACAGTCAATTTCAGGAAACTTGTAACTTGTATAATTGTAT
ATGCTCTGGTACACTACCTGAACTAACGAAGGGTAGAAGCTAATTCTGTTGTCAGTGTTCACACCTGTAAC
ATTAGGAGGATATGTCATGCTTATTCTTATGTTGGTTCTGTCAGGAAAGCTGCAACTGGACATGGCA
TTTCTGAAAAGCCTTAAATCTTAAGATGTTGATGTAGGGTATGCAAGTCAAAAGGCTGCCAGAAC
TGAGCCCTTTGTAAGCTGGAAAGCATTCTACTACTGTTACTTTGTTAGGAAGTTCAATTCAAGAGC
TGCAAAGTGTCCCGTAAGCAGTGCCTTAGTAATACCTTAGTCATGCCAGCCTTCTACACCAAT
TCCTAATGTTCATTTACGAATTGGCCAATTGGAAACAAAACAAGCAAAATTGTCCTCATTTGTT
TGTAAGGCCATTCTCCAGTTCTATAGGAAACTGACTGCTGGGTGAAAATCCGAAACTGGACACAAG
TCAGTTCTTCACCACACTCAAATGTATATACAAAACAAAAGGTTGCAACTCATAGTTACTATGAAA
GCAAATTGTACTTTAATGTTGCCATTAAATTGACCAAAACTTAGCTATTGTAATCTGCA
CTCTAGCATGAAAGTGCCTTGGTTGAGATTCCAGCTTAGAAAAGTGTGCCATAAACGATAATTG
AGAGAGACCAAAATATTGAGATCACCGTAATGCCATTGGTTACGGGATGAGTAACCAACCACAGGC
CTCTGTTACAAGAGCACGACGTGGTCCCCGCTGCTAGTCTGCCACTGGGGCCTCCAAACAT
CCATAGCACACTTCAGCGGAAGGACCCAGAAACTGTTGTTGTCAGTGTGCTGATGACCTAGTGT
TTCACCTCGTCAACCAGCCCTGCGTCCGGATGAGGGGACTCTGCACAAATGACAGAAATCTGGCTGG
ACAGATACTACAGCTTCTCCTCCTGTGTTGTCAGTCTGTGGAGACTTCTTCCATTCA
AATGACAGTGCACCTATCTGGTTACACAATGATACCAATTGAAAGTGGAAAGCCTAAACTGAGACG
ACAGTGCAGAACAAAACAAAAGTGAGTTAGGGCTTAAAATTGAAAGTGTCTTCTTAGGGCAAACATGTT
GACTCCGAGTATTGTGATGAATGTCAGAGAAACTTCAAAGAGCACCATTACAATTGGCATTTC
AAAGAATGTTCCAGCCCTCAAAGGGGCAACTCTTAAAGCTTGTGCTTATCCAAACCTTGTAGAA
ATTGGGAAAGCTGATAGAGGTAAGGAAGACGAGTGGAAAGGACAAGAAGGCCAAACACCAGC
AAAGGAAACTAGGAAAGGCAACACCAGC
ACTAGGAAAAAAAGATTCTTGCTAATATAGATGTTAAACATCAGACATCTTGAAATTAGCCT
CTAAACTCTTAATACATACGTTCTGTGTCCTACCTGGCGTCTTAAGAAATACCTCTGGGCTCTGA
AATTAGGAGTGTGATTCTTACCACTCAAGTTGTAAGTATTGAGAAATTCTCAACTCTCCCTGCC
AGCACTATACACAGTGTGAAAGAATTAGTCATGAACTGCTTCTCTCTCTCTCTCTCTCT
TGCTTGTGAAAGTAGTTCTCTCTCAAAGCCGTGCTTATATGTTAAGAAATGAAAGGTTGTGTTA
AAATTATTGCAATTGCAAAGGGTAGTTCACTGAAGTCATGCCATTAAATAAGATGAAATTGTT
TATTGTCCTACTCTCTCAAAGCCGTAACTTCTCTCTGTGAATTGCAATTGAGTCAC
TCTGCTACACT
CATCGCTTCTGTTGAGATGGCATTATGTTCTCTCGTTATCATGAAATGGGTCAGATTCC
GATTCCACCTCTGTCAGGTGACTCTGTCTGCCATTGATGAGATTCTCTCTCTCCCTTCT
TAAGAGAGGCTGACAGATCTAGGTGTCATCAATTGAAACCAGTCTGATTCTTCTGATT
TCTATCATTAGTTCACTGTTGAAATTAGATATCAACTGCACTCTTAAACATCTCC
ACCTCCTTGAAGATTACTCTGTAGGCCTTCTCAATAGGCTCATGACTGCA
GAGACAAGGAAAAAG
TAAAAACAAAACAGTATGTGCTGAAATGACAAAAAAATTGTAACATTAAAAAGAAACCTGAA
TAGCCTTAATTCTTAATAACACTTAAATTGTAATCGGTTCTGCCACGTGTTGTCACA
TTCTAAATGACTTAATGGGATTCTCACGGTCTGTCAGTGTATAAAATGGGCTTGTGATGT
AAGCGTTTCACTGGTCAGGGTCTTGTGATATTGACTGCTGCTGGAGTGGGCTGGAACCTGCC
CGGGTAACCTATTGCAAGAACCTTCAAAGGCATTGATTAAACCTCTGGCAGTACAGTATTCTG
TCTTAACCTATTGCAAGAACCTTCAAAGGCATTGATTAAACCTCTGGCAGTACAGTATTCTG
TGTAAACGTCTGTGTTAGGTAAGTGTACCTTTGTTAAAATGTTCAAGTGTGTTGCTTAAAGTGA
TTTATCTTGTGATGAGTATTGAAATTAGGATTGTCAGGAGAATTGTTGTTGTTGTTG
TGTAAAAAAAGATTGATTCTAGCTTGCAGGTAATGTCAGGAGAATTGTTGTTGTTGTTG
AAGTGCAATTATTCACTCAAAATGCTATTGTTGTCAGTCTGCCAGGTGTTGTTG
CAAAGTGACACATTGATGCCATTGATAAAGGGTAGACATTGTTGAGCATTGTTG
TAGCTTCTAGAAACTTTGTATT

CATACGGTATCAATGAAAAATAAGAAAATGAAAGTGTGGTCACCTTTTATCTGAAAAAAAAAA
AAAAA

LENPTAKIEPSSWSGSESPAENMERMSDSADKPIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEG
KMYGRNELIARYIKLRTGKTRTRKQVSHIQVLARRKSDFHSKLKDQTAKDKALQHMAAMSSAQIVSATA
IHNLGLPGIPRPTFPGAPGFWMQQTGQPGSSQDVKPFVQQAYPIQPAVTAPIGFEPASAPAPSVPAW
QGRSIGTTKLRLVEFSAFLEQRDPDSYNKHLFVIGHANHSYSDPLLESVDIRQIYDKFPEKKGLKELF
GKGPQNAAFFLVKFADLNCFIQQDAGAFYGVTSQYESSENMTVTCSTKCSFGKVVEKETEYARFENG
FVYRINRSPMCEYMINFIHKLKHLPEKYMMNSVLENFTILLVVTNRDTQETLLCMACVFEVSNSEHGAQHH
IYRLVKD

+++++

#10

NM_001098504

Apparently conserved in mammals though in-frame stop codon in mice, rats and elephants. Likely initiated at a “ACGGUG” tandem downstream)

DDX17 DEAD (Asp-Glu-Ala-Asp) box polypeptide 17

Synonyms: P72; RH70; DKFZp761H2016; DDX17

GTAAAGTTGGAGCCGACTCAGCGCGGGGCCATTGCGAGTCGCTGGAAAGGAAGGAGACGCCCTAAC
CGCGGCACTGCCGGTTGAGCGTAGCCAAACCTGCCACGGGCTTGAGCCCCGATTCTCTGTGTTTG
CTCCCGTCTCCGACGAGAGAGGGCGCGAGCGTGGCTCGCACGGAGACAGCGCTGGAGCAGAG
CGCTGCCCTGCCGCCAACAGCGGAGGCCGCGCATCGTCGTCACCAAGACGGAGCCGAGG
CCCTCCCGAGCCGGCCATCCGTGCCCGCTCCAGATCTCTATCCTTTGGGACCATGCGGGAGGAGG
TTTGGGGACCGGGACCGGGATCGTGACCGTGGAGGATTGGAGCAAGAGGTGGTGGCCTCCCCGAA
GAAATTGGTAATCCTGGGAGCGTTGCGTAAAAAAAGTGGGATTGAGTGGAGCTCCAAAGTTGAGA
AAAAATTATGTTATGTTGAAACATCCGGAAGTAGCAAGGCTGACACCATATGAGGTTGATGAGCTACGCCGAAAG
AAGGAGATTACAGTGAGGGGGGAGATGTTGCTCTAAACCGTGTGGCCTCCATCATGCTAACTTCCC
ACAATATGTAATGGATGTGTTGATGGATCAGCACTTACAGAACCAACTCCAATTGAGCTGCCAGGGATTTC
CGTTGGCTTCTAGTGGCCGGATATGGGGCATGGCTCAGACTGGCTCTGGAAAGACGTTGGCGTATCTC
CTGCCTGCAATTGTCATATTAAACCACAGCCACTTGGAAAGGGGAGATGGCCAATCTGTCTAGTTCT
GGCTCTTACAGAGAGCTTGGCCAGCAAGTACAGCAGGTGGCGATGACTATGGCAAATGTTAGATTGA
AGAGTACTTGTATTGAGGTGCTCTAAAGGTCCCCAGATTGAGACTTGGAAAGAGGTGGAGATC
TGCATAGCCACTCTGGACGCTGTGATAGATTCTGGAGTCAGGAAAGACAATCTCGCCGATGTA
CCTTGATGGACGAAAGCTGACAGAATGCTGATATGGGTTGAACCCCAGATCCGTAATTGTTGACC
AAATCAGGCCGCTGATAGGCAGACACTGATGTGGAGTCACCTGGCAAAGAAGTAAGACAGCTTGAG
GATTCCTCGTGTATTACACCCAGATCAACGTAGGAATCTGGAGTGAGTGCAACCACAAACATCCTCCA
GATAGTGGATGTCTGCATGGAAAGTGGAAAAGACACAAGTGTGATCCAACATGGAAGAAATAATGGCTG
AAAAGGAAAACAAACAATAATTGAGGAGACAAAGAGACGCTGTGATGACTGAGCTGAAGGATGCC
AGAGATGGTGGCCAGCTATGTGTATCCATGGAGACAAGAGTCACCAACAGAGAAAGAGATTGGGTACTTAATGA
GTTCCGTTCTGGAAAGGCACCCATCCTATTGCTACAGATGTAGCCTCCGTGGCTAGATGTGGAAAGATG
TCAAGTTGTGATCAACTATGACTATCCAAACAGCTCAGAGGATTATGTCACCGTATTGGCGAACAGCC
CGTAGCACCAACAAGGTACCGCTATACCTCTCACCCAGGAACCTAAACAGGCCAGAGAGCTTAT
CAAAGTGTGAAAGAGGCCAATCAGGCTATCAATCCAAACTGATGCACTGAGCTGTGGACCACAGAGGAGGCG
GCGGAGGCGGGGTTAAGGGTGGCTCGTCTCGTACCGGACCACTCTCAGCCAACAATCCAACTGTGATG
TATCAGGATGAGTGTGACCGAAGGGCTCGAGGAGTCAGGATGGTGGCCGGAGAGACTCTGCAAGCTATCG
GGATCGTAGTGAACCGATAGAGCTGGTTATGCTAATGGCAGTGGCTATGGAAGTCCAATTCTGCTTGT
GAGCACAAGCAGGCCAATACACCTATGGTCAAGGCACCTATGGGAGCTGTGTTATGGCACCAGTAGCTAT
ACAGCTAAGAATATGGTGTGGCACTTATGGAGCTAGTAGCACCACCTCAACTGGAGAAGTTCACAGAG
CTCTAGCCAGCAGTTAGTGGGATAGGCCGGTCTGGCAGCAGCCACAGCCACTGATGTCACAACAGTTG
CACAGCCTCCGGGAGCTACCAATATGATAGGTTACATGGGAGACTGCTACCAATACCTCCTCCTCCT
CCCCCTCCTCCTCACGTAAATGAAACCACTCAAGTGGTAGTGACTCCAGCAGACTTAATTACATT

AAGGAACACTGTCTTCCTTTTTCTCTCGCCTTCTTTCTTTTCTTTCTTTTTTCTTTTTTT
 TTAATTTCCCCCAACCATCGTATTGTCATGCAGATTAGTAGAATTCACTGCCAGGTTCT
 TCTGCCACAAATGATCCAGTCTGAAATAACATTGTAAAAAAAATATATATATATATAT
 AGCTGACTGGAAGAGATTAATTCTTCCCCAACCTCTGCATGTTGAGATATTGAGCTATTTCATC
 TAAAAGAGTAAGGTATTAGGCCCTTGTGGAGCCCAGTTGTTCTGAGTTGGTGGGAGGGAG
 GGAGGGGAGGGCTGAATTGTTGCAGAGGAAGATGGCATCTGTGCTTAAATTCTCATTACTGGGTTA
 GAAAACAAAGAGGGATTGCCCTGCACATTCTTGTGCTTAAATGTTCTAAGTTGAAACAGGTT
 CCTCGGCCTGTTGACTGATTGCTGGAGTGATGTAAAGTTACTAATTGGTTTATTCCC
 TTCACACTCTGCCTCCCCACTCTCCCCCGTACTGAAAATAACCATTAGTGTAGGCTAGAAATTG
 AATTGCTGAGTTGTGTATCCTTAAATTAAAACCACAAGTGTATTGTTAGGGTAAACTGTAGCAT
 CTCAGCATCTGGTGGAGCTGCCTATATTCTCCAGTTAACTGGGACATCTGTGAAATTAAATT
 CCATCCAGACAGCTGAGCAATGAACATAATGCTGGAAATTACTAACCAAGTTTATATT
 GACCTGAGTGTAAAAGCACATTAAATTATAAAACATATATTCAAATGGGCAAATTTTATTTCAAATG
 CAGTGTAGAGCTAGATTAAAGCAACTCTTGCCACCTACTCTGCCCTTGGCAAAGTTACCTTAACAGC
 AGAATCTTAAGGGTTTATTAAGAACTCTTATTCTCATACCCGTTCTGCACTGCTTCTAACAGC
 TTCTGGTGCAGATTCTCGGCATCCTTGCACTCAGCTATTACAGGTAGGTAGTGTCTAAGAAAAG
 TCATGGAGGACTAAAGCCTAACGTCTTCACTTTCCATCTGAAGGTAGGTAGTTACCTCTTCA
 TGGTAATGCTGTTTACCAAGACTTATAGCAGATGGACCCAGAAAGAATTCTGCTATTGTTCA
 CAACAGGATAGGGACATCAGACAGCCCCAGAAACCCCTCAGATCTGATATGGACTATTAATT
 CTGTTAATTGGTATTCTACATTACAATGCAAGTGAAGGGGGAGGCTTCACTGCATTGGCTAAGGCCTG
 AATGCTGCTCATCTGTAAGATCTACTCGAGGTTGTTCTTAAATTCTTAGGGAGAGAGGG
 ATGGTTCTGAGGGTTCTGAAAGTATGATTCAATGCAACATACAGGTAGGTCTCAGCATAAGCTGAA
 ATATATGCATGAAAAACTTGACATCTTTTAAATTCCACTTCTTAACTTACTCTCTT
 TTGTCCTCCCCCATCTTACAGAAGTTGAGGCCAGGGAGAATGGTAGGCACAGAAGAAACATGGCAA
 ACTGCTGTGCTTCAAACCAAAGTGTCCCCCAACCCCAAATTGTCTAAAGCAGTGGCAGTGTG
 GCATTGTTCTACAACCAAATTCTGGTTTCTTCTTAAACATAGGGTACCCACACAAGGG
 ATGCCCTACTCTCGCAGCTTGAAAGCATCTGTTGAGGGAAAGGTCTGGCAAGCAAGTGGTTAT
 TTGGATTGCTGCTCCCTTCCACTGGACATTGTAATCATAAAATAACAGTAAATTCCAAACCTCA
 AAAACTATTATGGCTGAGCACAGCTGAAATCTAGCAGAGTTAACCTCTGCCTCCATGTC
 ATAATTCAAGGTTCTGCTGGCTTCAGAACATGAGCAGAAGAATGTTTATGCTAGTTATTG
 CATTCAATGGTAAACTTAGGAAAGGTCCAATGTATTAAGCAATGGCTGCTCTCCCA
 ACAATTCTGTTGTGGACTTCTCATCTAAAGGTTAGTGGCTTTGCTGGATCAGTGTCTCTATT
 GATGTTCTGCTGGTCTCCAGACACATTCTGTTGCATTAAGACTTGAAAGACTTGTAGATGT
 GATGTTCAAGCAGGATGCTGAAAGCTATGTTACTATTCTTAGTTGAAATTGTCTTT
 GATACCATCATCTGTTCTTTGTAGGTATAAAACACTGTTGACAATAA
 AAAAAAAAAAAAAAA

TVASATGDSASERESAAPAAAPTAEAPPSVTRPEPQALPSPAIRAPLPDLYPFGTMRGGGFGDRDRD
 RGGFGARGGGGLPPKKFGNPGERLRKKWDLSELPKFEKNFYVEHPEVARLTPYEVD
 ERLRKKEITVRGGD
 VCPKPVFAFHANFPQYVMDVLMDQHFTEPTPIQCQGFPLALSGRD
 MVGIAQTGS
 GKTLAYLLPAIVHINH
 QPYLERGDGPICLVLAPL
 TREL
 AQQVQQVADDYGKCSRLK
 STCIYGGAPKG
 PQI RDLER
 GVEICIA
 TPGRLID
 DFLES
 GKTNLRRCTYL
 VL
 DEADRM
 LDMG
 FEPQIRK
 IV
 DQIRP
 DRQ
 TLMWS
 ATWP
 KEV
 RQ
 LAEDFLR
 DYT
 QI
 NVGN
 LE
 SANHN
 IL
 Q
 IVD
 VC
 ME
 SE
 KDH
 KL
 I
 QLM
 MEE
 I
 MAE
 KNT
 II
 IF
 VET
 KRR
 CDD
 LRR
 MRR
 DG
 WP
 AMCI
 HGDKS
 SQ
 PER
 DW
 VLN
 EFR
 SG
 KAP
 I
 LI
 AT
 DV
 ASR
 GL
 D
 ED
 V
 K
 F
 V
 IN
 YD
 P
 NS
 S
 E
 D
 VHR
 IR
 GT
 TAR
 ST
 NKG
 TAY
 T
 FFT
 PG
 NL
 K
 Q
 ARE
 LI
 KV
 L
 E
 E
 A
 Q
 A
 N
 P
 K
 L
 M
 Q
 L
 D
 R
 S
 A
 Y
 R
 D
 R
 S
 E
 T
 DR
 AG
 Y
 A
 N
 G
 S
 G
 Y
 G
 S
 P
 N
 S
 A
 F
 G
 A
 Q
 A
 G
 Q
 Y
 T
 Y
 G
 Q
 G
 T
 Y
 G
 A
 A
 Y
 G
 T
 S
 S
 Y
 T
 A
 Q
 E
 Y
 G
 A
 G
 T
 Y
 G
 A
 S
 T
 T
 S
 T
 G
 R
 S
 Q
 S
 S
 S
 Q
 Q
 F
 S
 G
 I
 R
 S
 G
 Q
 Q
 P
 Q
 P
 L
 M
 S
 Q
 Q
 F
 A
 Q
 P
 P
 G
 A
 T
 N
 M
 I
 G
 M
 Q
 T
 A
 Y
 Q
 P
 P
 P
 P
 P
 P
 S
 R
 K
 E

+++++

#11

NM_001025366

Region is conserved in mammals. In-frame stop codons and broken frames in: *Callithrix jacchus*, *Felis catus*, and *Echinops telfairi*.

VEGFA vascular endothelial growth factor A

Synonyms: VPF; VEGF; MVCD1; MGC70609; VEGFA

GGCTGGGCAGCGGGTAGCTGGAGTCGTGGCCTGGGGTAGCACAGCGCTCTGTCGGAGGCG
AGCGTTAGGTGGACCGGTAGCGGACTCACCGGCCAGGGCGCTGGTCTGGAATTGATATTCAATTGAT
CCGGTTTATCCCTTCTTTCTAAACATTTTTAAAAGTGATTGTTCTGTTAATTATTT
TTTTGCTGCCATTCCCCACTGAAATCGGGCGACGGCTGGGAGATTGCTCTACTTCCCAAATCACT
GTGGATTGGAAACCAGCAGAAAGAGGAAAGAGGTAGCAAGAGCTCCAGAGAGAAGTCGAGGAAGAGAGA
GACGGGTCAGAGAGAGCGCGCGGTGCGAGCAGCAGAACAGCAGAGGGCAAAGTGAGTCGCTT
TGGGGTGACCGCCGGAGCGCGGTGAGCCCTCCCCCTGGGATCCCGAGCTGACAGTCGCGCTGACCG
GACAGACAGACAGACACGCCCGAGCCCAGCTACCACCTCTCCCCGCCGGCGGAGACAGTGGACGC
GGCGCGAGCGCGGGCAGGGCGAGCCCGCCGGAGCGGGTGGAGGGGTCGGGCTCGCGCG
TCGCACGTAAACTTCTGTCACCTCTGGCTGTTCTGCTCGCTCGAGGAGGCCGTTCCCGCGGGGAA
GCCGAGCCGAGCGGAGCGGAGAAGTGCTAGCTGGCCGGAGGAGCCGAGCCGGAGGAGGGGAGGA
GGAAGAAGAGAAGGAAGAGGAGAGGGGCGCAGTGGCAGACTGGCGCTCGGAAGCCGGCTCATGGACGG
GTGAGCGCGGTGCGAGACAGTGCTCCAGCCGCGCTCCCCAGGCCCTGGCCCGGGCTCGGCG
CGGGAGGAAGAGTAGCTGCCGAGGCCGAGGAGAGCGGGCCACAGCCGAGCCGGAGAGGGAG
CGCGAGCCGCGCCGCCCGTGGCCTCCGAAACATGAACTTCTGCTGCTGGGTGATTGGAGCC
TTGCCTGCTGCTCACCTCACCAGCCAAGTGGATGTCTACAGCGCAGCTACTGCCATCCAATCGAGACCCCTGGT
AAATCATCACGAAGTGGTAAGTTCATGGATGTCTACAGCGCAGCTACTGCCATCCAATCGAGACCCCTGGT
GGACATCTCCAGGAGTACCTGATGAGATCGAGTACATCTCAAGCCATCTGTCAGGAGTCCAAACATCACCAGCAGATT
ATGCGGATCAAACCTCACCAAGGCCAGCACATAGGAGAGATGAGCTTCCACAGCACAACAAATGTGAATG
CAGACAAAGAAAGATAGAGCAAGACAAGAAAAAAATCAGTCAGGAAAGGGAAAGGGGAAACAGAA
AGCGCAAGAAATCCGGTATAAGCTGGAGCGTAGCTGGTGCCTGCTTAATGCCCTGGAGC
CTCCCTGGCCCCCATCCCTGTCAGGAGGAAAGCATTGTTGACAAGATCCGAGAC
GTGTAATGTCCTGCAAAACACAGACTCGCTTGCAAGGGAGGGAGCTTGAGTTAACGAACGTACTT
GCAGATGTGACAAGCCGAGGGGGTGAGCCGGCAGGAGGAAGGGAGCCTCCCTCAGGGTTCGGGAACCAGA
TCTCTCACCAGGAAAGACTGATAAGACAGATCGATAAGAAACACGCTGCCACCACACCACCCA
TCGACAGAACAGTCTTAATCCAGAAACCTGAAATGAAGGAAGAGGAGACTCTGCGCAGAGCACTTGGT
CCGGAGGGCGAGACTCCGGCGAACGATTCCCGGGGTGACCCAGCACGGTCCCTTGGAAATTGGATT
CGCCATTATTCTTCTGCTAAATCACCGAGCCGGAGATTAGAGAGTTTATTCTGGGATTCT
GTAGACACACCCACACATACATACATTATATATATATATATATATATATATATATATATATATCT
ATTTATATATATAAAATATATATATTCTTTTAAATTAAACAGTGCTAATGTTATTGGTCTTCACTG
GATGTATTGACTGCTGGACTTGAGTTGGAGGGGAATTGGCCACTCAGATCCTGACAGGGAAAGAGGA
GGAGATGAGAGACTCTGGCATGATTTTTGTCCCCTGGTGGGGCCAGGGTCTCTCCCTGCC
GGAATGTGCAAGGCCAGGGCATGGGGAAATATGACCCAGTTGGAACACCGACAAACCCAGCCCTGG
CGCTGAGCCTCTCACCCAGGTAGACGGACAGAAAGACAGATCACAGGTACAGGGATGAGGACACCGGC
TCTGACCAGGAGTTGGGAGCTTCAGGACATTGCTGTGCTTGGGATTCCCTCACATGCTGCACGCGC
ATCTGCCCTCAGGGCACTGCCCTGGAGATTAGGAGCTGGCCTGGCGCTCGCTACTCTCACCTGCTTC
TGAGTTGCCAGGGAGACACTGGCAGATGTCCCAGCGAAGAGAAGAGACACATTGTTGAAAGACAGGCC
ATGACAGCTCCCTCCTGGACTGCCCTCATCCTCTGCTCCCTTGGGTGAGCCTAAAG
GACCTATGTCTCACACATTGAAACCAACTAGTTCTGCTCCCCCAGGAGACCTGGTTGTTGAGT
GGTTGACCTTCTCCATCCCTGGCTCTCCCTCCCTGGAGGACAGAGAGACAGGGCAGGATCCAC
GTGCCATTGTGGAGGCAGAGAAAAGAGAAAGTGTATTATACGGTACTTATTAAATCCCTTTAATT
TAGAAATTAAACAGTTAATTAAATTAAAGAGTAGGGTTTTTCAGTATTCTGGTAATATTAAATT
CAACTATTATGAGATGTATTTGCTCTCTGCTCTTATTGTAACGGTTTTGTATATAAAATT
CATGTTCCAATCTCTCTCCCTGATCGGTGACAGTCAGCTTGTAAACAGATATTAAATTG
TAACACTCAGCTGCCCTCCCCGATCCCTGGCTCCCCAGCACACATTCTTGAATAAGGTTCAATA
TACATCTACACTATATATATTGGCAACTTGTTGTTGTTGATATATATATATGTTATGTA
TATATGTGATTGATAAAATAGACATTGCTATTGTTTATATGAAAACAAAACAAGAAAAATA
GAGAATTCTACATACTAAATCTCTCTCTTAAATTAAATTGTTATCATTATTATTGGTCT
CTGTTATCCGTAATAATTGTTGGGAAAGATATTAAACATCACGTCTTGTCTAGTGCAGTTTCGAG
ATATTCCGTAATAATTGTTGGGAAAGATATTAAACATCACGTCTTGTCTAGTGCAGTTTCGAG
ATATTGTATTAAAGAATTAAATTCTGATCTAAAAAAAAAAAAAA

LTDRTDTAPSPSYHLLPGRRTVDAAASRGQGPAPGGVEVGARGVALKLFVQLLGCSRGGAVRA
GEAEPSGAARSASSGREEPQPEEGEEEEKEERGPQWRLGARKPGSWTGEAAVCADSAPAARAPQALARA
SGRRGVARGAESGPPHSRGSASRAGPRASETMNFLLSWVHSLALLLYLHHAKWSQAAPMAEGG
GQNHEVVKFMDVYQRSYCHPIETLVDIFQEYPDEIEYIFKPSCVPLMRCGGCCNDEGLECVPTESNITM
QIMRIKPHQGQHIGEMSFLQHNKCECRPKDRARQEKKSVRGKGKGQKRKRKKSRYKWSVYVGARCLMP
WSLPGPHPCGPCSERRKHLFVQDPQTCKCSCKNTDSRCKARQLELNERTCRCDKPRR

+++++

#11

NM_022002

Region conserved from insects to human. Non-AUG initiation conserved in mammals; in non-mammals the region is initiated by AUG.

NR1I2 nuclear receptor subfamily 1, group I, member 2

Synonyms: BXR; PAR; PRR; PXR; SAR; SXR; ONR1; PAR1; PAR2; PARq; NR1I2
TTCTTAACCCCTTCCAGCTTCCCACCCCTTTGGCTTAGCCATGCCCTCTGATCTGTGTTCTCAGGG
GACCTGCAGGCCAGATATAGCCCCATGCTGTCTCCTACCCAGAGCACACTGTCAGGCTACTTCCAC
TGGTACTGAAATCCAGTATTCACTTACTCTTTCTTCCAATATCCTCATGACATTCAATATTCACCT
ACTCTAGGTCCCTGCCTAAGGCCAAGTCAACTTCTGTCAGTGGGATTGTAATCCAATACCTCCT
AGCCCTAGCAGAACATCCATGTGGATAATCAGAAATGTGACTGGAAAAAGGACAGAGCTCTATGGCTGTGGG
TCCCAGCCCCACTGCTGGCAGTAAGTCCCCAGCAGTGAGCTGTGAAGCACCTACATTCTCGCCTTGGG
TGAAAACAGCAAGGAAGCATCCACTTGAGAAATGTCAACCCCTAGGAAATCCCAGCCTCAAGTCTTCTC
ATCCCCTGGGAAGTGCAAATTGGATAGAGAAGAACCAATTAAAAACAAAACAATCATACTTAGAT
ATTCTGGCTTCTCACCAGGGCTGGATTAAAGCATGTACTCAAATAACAACCTAACGCAATAAAT
AAATGTAAGGAAGTCAAATGTTACCTGAAGACAACGTGGTCACTTTTGGCAATCCCAGGTTCTCTT
TCTACCTGTTGCTCAATCGGGTCTCCCTCTCCCTCTTGTGGGGCCATGCCCTGCTTACTGTTG
CCAGAGGCTTGTACTGTTGCCTTTAGTAGGAGCAGTACTTCACTCCCTCACCTGCCATAAGCA
TCTTATAAACAAAGCAAGTAGAAGAACACATCTGGTATCCACCACATTGGCTTTGTTGATTCTGTT
CACTTGGGAGCACCTGCTAGGGATAAGAAGGTTGAGGCTGAAGAGTGAGGACTCTCAGCTCCCTC
TGGCAGGACCCGGAGAGGAAAGAGCCCTCAGCTGGTCCATCTCCCCACTCCTGGTCAGCCTCTGTTCT
GAGATCAAAGTGGTGGGGTACATTCTGAGAACACTGTGCTCAGCCCCCTCATCTCACACCCCTTCCCTCTC
CCTGTGCGCTGCCCTCTTACATAACCAGCTGGTGTGGACCTGGCCATAAATCAATACTTGTCA
CTTTCACATCAAGTAACACTATCCAGGGAGGTGGTTCAACAAAGGAGGAAGTATAAGGAGATCTAGGTT
AAATTAATGTTGCCCTAGTGGTAAAGGACAGAGACCTCAGACTGATGAAATGCACTCAGAATTACTTAG
ACAAAGGGATATTGCCACTCTTCCCCTTCTGTGTTTGAGTGAAGAGACCTGAAAGAAAAAA
GTAGGGAGAACATAATGAGAACAAATACGGAATCTCTTGTAGTTCAAGTGCTGGACTTGGGACT
TAGGAGGGCAATGGAGGCCCTAGTGCCTACATCTGACTTGGACTGAAATATAGGTGAGAGACAGATTG
TCTCATATCCGGGAAATCATAACCTATGACTAGGACGGGAAGAGGAAGCAGTGCCTTACTCAGTGGGA
ATCTGGCCTCAGCTGCAAGCCAAGTGTACAGTGAGAAAAGCAAGAGATAAGCTAATACTCCTGTCC
TGAACAAGGCAGGGCTCCTGGTAAAGCTACTCCTTGATCGATCCTTGCACCGGATTGTTCAAAGTGG
CCCCAGGGAGAGTCGGAGAACAACTTACCAAGCAGTCCAAGAGGCCAGAACCTGGAG
TGAGACCAAAGAACAGCTGGAAACCTGACTTTGAGTGAAGGACACAGAGCTGTTCTGGAAAG
CCCAGTGTCAACCGAGATGAGGAAGTGGAGGCTCCCAAATGCCGTGATGTTGGGACAAGGCCACTGG
CTATCACTCAATGTCATGACATGTGAAGGATGCAAGGGCTTTTCAGGAGGCCATGAAACGCAACGCC
GGCTGAGGTGCCCTTCCGGAAGGGCGCTCGAGATCACCGGAAGACCCGGCGACAGTGCCAGGCC
CGCCTGCGCAAGTGCCTGGAGAGCGGCATGAAGAAGGAGATGATCATGTCGACGAGGCCGTGGAGGAG
GCCGGCCTGATCAAGCGGAAGAACAGTGAACGGACAGGGACTCAGCCACTGGAGTGCAGGGCTGACAG
AGGAGCAGCGGATGATGATCAGGGAGCTGATGGACGCTCAGATGAAACCTTGACACTACCTCTCC
TTCAAGAATTCCGGCTGCCAGGGTGTAGCAGTGGCTGCGAGTGTGCCAGAGTCTCTGCAGCTGC
GAGGGAGAACAGCTGCAAGTGGAGGCCAGGTCCGAAAGATCTGTGCTCTTGAGGTCTCTGCAGCTGC
GGGGGAGGATGGCAGTGTCTGGAACATACAAACCCCCAGCCGACAGTGGCGGGAAAGAGATCTCTCC
CTGCCACATGGCTGACATGTCAACCTACATGTTCAAAGGCATCATCAGTTGCCAAAGTCATCTCCTA

CTTCAGGGACTTGCCCATCGAGGACCAGATCTCCCTGCTGAAGGGGCCGTTCGAGCTGTGTCACAGA
GATTCAACACAGTGTCAACGGAGACTGGAACCTGGAGTGTGGCCGCTGCTACTGCTGAAAGAC
ACTGCAGGTGGCTCCAGCAACTTCACTGGAGCCATGCTGAAATTCCACTACATGCTGAAGAAGCTGA
GCTGCATGAGGAGGAGTATGTGCTGATGCAGGCCATCTCCCTCTCCTCCCCAGACCGCCAGGTGTGCTG
AGCACCGCGTGGTGACCAGCTGCAGGAGCAATTGCCATTACTCTGAAGTCTACATTGAATGCAATCG
CCCCAGCCTGCTCATAGGTTCTGTTCTGAAGATCATGGCTATGCTCACCGAGCTCCGCAGCATCAATGC
TCAGCACACCCAGCGCTGCGCATCCAGGACATAACACCCCTTGCTACGCCCTCATGCAGGAGTTG
TCGGCATCACAGGTAGCTGAGCGGCTGCCCTGGGTGACACCTCGAGAGGCAGCCAGACCCAGAGCCCT
TGAGGCCACTCCGGCCAAGACAGATGGACACTGCCAAGAGCCGACAATGCCCTGCTGGCTGTCTCC
CTAGGAAATTCTGCTATGACAGCTGGTAGCATCCTCAGGAAGGACATGGGTGCCCCCAGGGCAGT
CAGTCTGTAGGGAGTGAAGCACAGACTCTACGTGGAGAGTGCACGACTGTAGGTGAGGACATCAGA
GAGGCAAGGTTGCCCTTAAAGGCCCTGTTGGGGAGAAATCCCTCAGATCCCACAAAGT
GTCAAGGTGGAAGGGACCAAGCGACCAAGGATGGGCATCTGGGTCTATGCCACATACCCACGTTG
TTCGCTCCTGAGTCTTCTATTGCTACCTCTAAAGTCTGTCTCCACTTCCACTCGTCCCCCTCCTC
TTCCGAGCTGCTTGTGGGCTCCAGGCCTGACTCATCGCAGCGCATGAGTATCTGTGGAGTCCTCTA
GAGAGATGAGAAGCCAGGAGGCCTGCACCAAATGTCAGAACGTTGGCATGACCTCATTCCGGCACATCAT
TCTGTCTCTGCATCCATTGAAACACATTATAAGCACCATAAGTGTAGCTGCTGTGGGTATACAG
CATTGACTCAGATATAGATCTGAGCTCACAGAGTTATAGTTAAAAAAACAAACAGAAACACAAACAAATT
TGGATCAAAGGAGAAATGATAAGTACAGAACAAAAGCAGCACAGGAATTTCCCTGTTGGAGTCAGTGTG
ATGGCGGGCACTGGTACCAAGTGAAGGTTCCCAGGGACATGAGTCTGTAGGAGCAAGGGCACAAACTGC
AGCTGTGAGTGCCTGTGTTGAGGTCTGTTGCCACTTGATGGGCCTGGTTGTC
CTGGGCTGGAATGCTGGGTATGCTCTGTGACAAGGCTACGCTGACAATCAGTTAACACACCCGGAGAAGA
ACCATTACATGCACCTTATATTCTGTGACACATCTATTCTCAAAGCTAAAGGGTATGAAAGTGCCTGC
CTTGTATAGCCACTGTGAGTAAAAAATTTTGCATTTCAAACATTACTTATATAAGGCATTCC
ACACCTAAGAAACTAGTTGGAAATGTAGCCCTGGGTTAATGTCAAATCAAGGAAAAGGAATTAAATA
ATGTACTTTGGCTAAAAAAAAAAAAAAAAAAAAAA

LEVRPKESWNHADFVHEDTESVPGKPSVNADEEVGGPQICRVCGDKATGYHFNVITCEGCKGFFRRAMKR
NARLRCPRKGACEITRKTRQCQACRLRKCLESGMKEMIMSDEAVEERRALIKRKKSERTGTQPLGVQG
LTEEQRMMIRELMDAQMKTFDTFSHKFNFRLPGVLSGGCELPESLQAPSREEAAKWSQVRKDLCLSKVSL
QLRGEDGSVWNYKPPADSGGKEIFSLPHMADMSTYMFKGIIISFAKVISYFRDLPIEDQISLLKGAAFEILC
QLRFNTVFNAETGTWECGRLSYCLEDTAGGFQQLLLEPMLKFHMLKKQLHEEYVLMQAISLFSPDRPG
VLQHRVVDQLQEFAQITLKSYIECNRPQPAHRLFPLKIMAMLTELRSINAQHTQRLRIQDIHPFATPLMQ
ELFGITGS

+++++

#11

NM_001172131

Region conserved in mammals.

HCK hemopoietic cell kinase

Synonyms: JTK9; HCK

GTCAGCTGGGAGCACATCAGAGGCTTAGAGGGAGTGGGAAGGGACTCAGACAGTGCAGGAGGAGAAA
CGCCCGGGCACCAAAGCCCTCAGAGCGTCGCCCTAGTTAGAAAGTCAGTTCCGGCACT
GGCACCCGGAACCTCAGGGCTGCCAGCTGGGGCGCTCAAGCTGCGAGGATCCGGCTGCCCGGA
GACGAGGAGCGGGCGCCAGGATGGGTGCATGAAGTCCAAGTCTCCAGGTGGAGGAATACATTCTC
AAAAACTGAAACCAGCGCCAGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGG
GGCCTAATAGCCACAACAGAACACACCAGGAATCAGGGAGGCAGGCTTGAGGACATCATCGTGGTTGCC
CTGTATGATTACGAGGCCATTACCCAGGAAGACCTCAGCTTCCAGAAGGGGACAGATGGGGCTTCTAGA
GGAATCCGGGAGTGGTGGAAAGGCTCGATCCCTGCCACCCGAAGGGAGGCTACATCCAAGCAACTATG
TCGCCCGCTTGACTCTCTGGAGACAGAGGAGTGGTTTCAAGGGCATCAGCCGGAAAGGAGCGAGAGCGC
CAACTGCTGGCTCCGGAACATGCTGGCTCTCATGATCCGGATAGCGAGACCCTAAAGGAAGCTA

CTCTTTGTCCGTGCGAGACTACGACCCTGGCAGGGAGATACCGTAAACATTACAAGATCCGGACCCTGG
ACAACGGGGCTTCTACATATCCCCCGAAGCACCTCAGCACTCTGCAGGAGCTGGTGGACCACTACAAG
AAGGGGAACGACGGCTCTGCCAGAAACTGTCGGTGCCTGCATGTCTCCAAGCCCCAGAACGCTGGG
GAAAGATGCCTGGGAGATCCCTGGGAATCCCTCAAGCTGGAGAAGAAACTTGGAGCTGGCAGTTGGG
AAAGTCTGGATGGCACCTACAACAAGCACACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTC
GTGGAGGCCTTCCTGGCAGAGGCCAACGTGATGAAAACCTCTGCAGCATGACAAGCTGGTCAAACCTTCATGC
GGTGGTCACCAAGGAGCCCACATCTACATCATCAGGAGTTCATGCCAAAGGAAGCTTGACTTCTGA
AAAGTGTGATGAGGGCAGCAAGCAGCCATTGCCAAACTCATTGACTTCTCAGCCCAGATTGAGAACGGCATG
GCCCTCATCGAGCAGAGGAACATACATCCACCGAGACCTCCGAGCTGCCAACATCTGGTCTCTGCATCC
GGTGTGTAAGATTGCTGACTTTGGCCTGGCCGGGTCAATTGAGGACAACGAGTACACGGCTGGGAAGGG
CCAAGTCCCCATCAAGTGGACAGCTCTGAAGCCATCAACTTGGCTCTTACCCATCAAGTCAGACGTC
TGGCCTTGGTATCCTGCTGATGGAGATCGTACCTACGGCCGGATCCCTACCCAGGGATGCAAACCC
TGAAGTGTGATCCGAGCTCTGGAGCGTGGGATACCGGATGCCCTCGCCAGAGAACTGCCAGAGGAGCTTACA
ACATCATGATGCGTGTGGAAAAACCGTCCGGAGGGAGCGGGCACCTCGAATACATCCAGAGTGTGCTG
GATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCAGGGCA
GGGGGTGCCAGGTGGTGGCTGCAAGGTGGCTCAGCACCACCGCCAGGGCCCACACCCCTTACTC
CCAGACACCCACCCCTCGCTCAGCCACAGTTCTCATCTGCCAGTGGTAGGTTGGACTGAAAATCTC
TTTTGACTCTTGAATCCACAATCTGACATTCTCAGGAAGCCCCAAGTTGATATTCTATTTCCTGGAA
TGGTTGGATTTAGTTACAGCTGTGATTGAAAGGGAAACTTCAAAATAGTGAATGAATTTAAATAA
AAGATATAATGCAAAGTCTTACCAAAAAAAAAAAAAAA

LGGRSSCEDPGPCRDEERAPRGCMKSKFLQVGGNTFSKTESTSAPHCPVYVPDPTSTIKPGPNHSNSNTP
GIREAGSEDIIIVVALYDYEAIHHEDLSFQKGDQMVLLEESGEWWKARSLATRKEGYIPSNYVARVDSLETE
EWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYSLSVRDYDPRQGDTVKHYKIRTLNDNGFYISPR
STFSTLQELVDHYKKGNDGLCQKLSVPCMSSKPQPKWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKH
TKVAVKTMKPGSMSVEAFLAEANVMKTLQHDKLVLHAVVTKEPIYIITEFMAKGSLLDLKSDEGSKQPL
PKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFLGLARVIEDNEYTAREGAKFPIKWTAP
EAINFGSFTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALERGYRMPRPENCPEELYNIMMRCWKNR
PEERPTFEYIQSVLDDFYTATESQYQQQP

+++++

#14

NM_000378

Region conserved in mammals. Broken frame in some: elephants, *Myotis lucifugus*, etc.)
The CUG site designated as initiation site in humans (shown underlined below) is not
conserved at all but a CTG close by in much better context is.

WT1 Wilms tumor 1

Synonyms: GUD; AWT1; WAGR; WT33; WIT-2; WT1

CCAGGCAGCTGGGTAAGGAGTTCAAGGCAGCGCCCACACCCGGGGCTCTCCGCAACCCGACCGCCTGTC
CGCTCCCCACTTCCGCCCTCCCTCCACCTACTCATTCACCCACCCACCCAGAGCCGGGACGGCA
GCCAGGCAGGGCCCGCCGTCTCCCTGCCGATCCTGACTTCTCTTGCTGCAGGACCCGGCTTC
CACGTGTGTCGGAGCCGGCTCTCACGACACGCTCCGGGCTGGGTGCTACAGCAGCCAGAGC
AGCAGGGAGTCCGGACCCGGCGCATCTGGGCAAGTTAGGCGCCGCCAGGCCAGCGCTAACGTCTC
CAGGGCCGGAGGAGCCGCCGGGCGTCCGGGTCTGAGCCGAGCAATGGCTCGACGTGCGGGACCTGAA
CGCGCTGCTGCCGCCGTCCCTCCCTGGGTGGCCGGCGCTGTGCCCTGCCCTGTGAGCGCGCGCC
AGTGGGCGCCGGTGCTGGACTTTGCGCCCCGGCGCTCGGCTTACGGGTGTTGGCGGGCCCGCG
CCACCGCTCCGCCACCCCGCCGCCCTACACTCCATCAAACAGGAGCCAGCTGGGCG
CGCGAGCCGCAGGAGCAGTGCCTGAGGCCCTCACTGTCCACTTTCCGGCAGTTCACTGGCACAG
CCGGAGCCTGTCGCTACGGGCCCTCGTCCGCCAGGCCAGCGTACATCCGCCAGGCCAGGATG
TTTCCTAACCGGCCCTACCTGCCAGCTGCCAGAGCCAGGCCGCTATTGCAATCAGGGTTACAGCAC

GGTACACCTTCGACGGGACGCCAGCTACGGTCACACGCCCTCGCACCATGCAGCGCAGTTCCCCAACCAACT
 CATTCAAGCATGAGGATCCCATGGGCCAGCAGGGCTCGTGGGTGAGCAGCAGTACTCGGTGCCCCCCG
 GTCTATGGCTGCCACACCCCCACCGACAGCTGCACCGGCAGCCAGGCTTGCTGCTGAGGACGCCCTACAG
 CAGTACAATTATACCAAATGACATCCCAGCTGAATGCATGACCTGAATCAGATGAACCTAGGAGCCA
 CCTTAAAGGCCACAGCACAGGGTACGAGAGCGATAACCACACACAGCCCACCTCTGCAGGCCAATAC
 AGAATACACACGCACGGTCTTCAGAGGCATTCAAGGATGTGCGACGTGTCCTGGAGTAGCCCCGACTCT
 TGTACGGTCGGCATCTGAGACCAGTGAGAAACGCCCTTCATGTGCTTACCCAGGCTGCAATAAGAGAT
 ATTTTAAGCTGTCCCACCTACAGATGCACAGCAGGAAGCACACTGGTGAGAAACCATACAGTGTACTTC
 AAGGACTGTGAACGAAGGTTTCTCGTTAGACCCAGCTCAAAAGACACCAAAAGGAGACATACAGGTGTGAA
 ACCATTCCAGTGTAAAACCTTGTCAAGCGAAAGTCTCCCGTCCGACCACCTGAAGACCCACACCAGGACTC
 ATACAGGTAAAACAAGTGAACGCCCTCAGCTGCGTGGCAAGTTGTCAGAAAAAGTTGCCGGTCA
 GATGAATTAGTCCGCATCACAAACATGCATCAGAGAAACATGACCAACTCCAGCTGGCGCTTGAGGGT
 CTCCCTGGGGACCGTCAGTGTCCCAGGCAGCACAGTGTGAACTGCTTCAAGTCTGACTCTCCACTC
 CTCCTCACTAAAAGGAAACTTCAGTTGATCTTCTTCATCCAACCTCAAGACAAGATAACCGGTGCTTCTG
 GAAACTACCAGGTGTGCTGAAAGAGTTGGTCTGCCCTGCCTACTTTAGTTGACTCACAGGCCCTGGA
 GAAGCAGCTAACATGCTGTTAGTTAAAGCCATTGCCATTGGTGTGGATTTCTACTGTAAGAAGA
 GCCATAGCTGATCATGCCCCCTGACCCTCCCTTTTTATGCTCGTTTCGCTGGGGATGGAATTA
 TTGTACCATTTCTATCATGAAATATTATAGGCCAGGGATGTGATGTGCTGTAATGTAACATTGTT
 CATGGTTTCCATTACTAACAGCAACAGCAAGAAATAATCAGAGAGCAAGGCATGGGGTGAATCTTGT
 CTAACATTCCGAGGTCAAGCAGGCTGCTAACCTGAAAGCAGGATGTAGTTCTGCCAGGCAACTTTAAA
 GCTCATGCATTCAGCTAACGAGAAAAATCAGAAACTAACAGTACCTCTGTATAGAAATCTAAAGAA
 TTTTACATTCACTGTTAAATTCAATGTGAACACTGGCACACTGCTCTTAAGAAACTATGAGATCTGAGATT
 TTTTGTGTATGTTTGACTCTTGAGTGGTAATCATATGTGCTTTATAGATGTACATACCTCCTGCA
 CAAATGGAGGGATTCACTGGAGTGTCTTACTGTTAGTGTATAAAAACATGCTGGTATATGGCT
 TCAAGTTGAAAAATGAAAGTGACTIONTAAAGAAAATAGGGATGGTCAGGATCTCAACTGATAAGACTG
 TTTTAAGTAACTTAAGGACCTTGGGCTACAAGTATATGTGAAAAAAATGAGACTACTGGGTGAGGAA
 ATCCATTGTTAAAGATGGCGT
 TTTTAAGGGAGGGATTATTATTACCGTTGCTGAAATTACTGTGAAATATATGTCTGATAATGATT
 GCTCTTGACAACAAATTAGGACTGTATAAGTACTAGATGCATCACTGGGTGTTGATCTAACAGATAT
 TGATGATAACACTAAAATTGTAACCTGCATTTGCTCTAACAGTCTATTCAAAAGTCTATTCAAAAG

LDFLLQLDPASTCVPEPASQHTLRSGPGLQQPEQQGVDRPGGIWAKLGAAEASAERLQGRSRGASGSEP
QQMGSVDRLNALLPAVPSLGGGGCALPVSGAAQWAPVLDAPPAGSAYGSLGGPAPPPAPPPPPPPPH
SFIKQEWSWGAEPHEEQCLSAFTVHFSGQFTGTAGACRYGPFGPPPPSQASSGQARMFPNAPYLPSCLES
QPAIRNQGYSTVTDFGTPSYGHTPSHAAQFPNHSFKHEDPMGQQGSLGEQQYSVPPPVYGCHPTDSCCTG
SQALLRTPYSSDNLYQMTSQUECMTNQMNGLATLKGHSTGYESDNHTTPILCGAQYRIHTHGVRGIQD
VRRVPGVAPTLVRSASETSEKRPFMCAYPGCNKRYFKLSHLQMHSRKHTGEKPYQCDFKDCERRFSRSDQL
KRHQRRTGVKPFQCKTCQRKFSRSDHLKTHTRHTGKTSEKPFSCRWPSCQKKFARSDELVRHHNMHQRN
MTKLQLAL

THE UNDERLAINED CUG IS THE ONE MARKED IN GENBANK

+++++

#15

NM_001172415

CUG initiated extension conserved in mammals.

BAG1 BCL2-associated athanogene

Synonyms: HAP; RAP46; BAG1

GCAGGCCGGGGCGGGCTGGGAAGTAGTCGGCGGGTTGTGAGACGCCGCTCAGCTTCCATCGCTGGG
 CGGTCAACAAGTGC~~GG~~CTAGCGCGGGGGCGGGAGACCGCGAGGCAGCGACCGGGAGCGGCTGGG
 TTCCCGCTGCGGCCCTTCCGGCAGGCCGGAGCCGCCAGTCGGAGCCCCGCCAGCGTGGTCCCG

CTCCCTCTGGCGTCCACCTGCCGGAGTACTGCCAGCGGGCATGACCACCCACCAGGGGCCGCC
GGCGCTCGCAGGCCGCGGATCAGAAAGAAAACCCGGCGCCGTCGACCCGGAGCGAGGAGTTGACCCGGAG
CGAGGAGTTGACCTGAGTGAGGAAGCGACCTGGAGTGAAGAGGCGACCCAGAGTGAGGAGGCACCCAGG
GCGAAGAGATGAATCGGAGCCAGGAGGTGACCCGGACGAGGAGTCGACCCGGAGCGAGGAGGTGACCA
GAGGAAATGGCGGCAGCTGGCTCACCGTACTGTCACCCACAGCAATGAGAACGACGACCTTCATGTTAC
CTCCCAGCAGGGCAGCAGTGAAACCAGTTGTCACAGACCTGGCCAGGTTGAGGAGGTGACCCGGAG
CACAGTCTTTCAAGAAACTCATATTAAAGGAAAATCTCTGAAGGAAATGGAAACACCGTTGTCAGCA
GGAATACAAGATGGTGCAGGGTCACTGTTAATTGGGAAAAGAACAGTCCACAGGAAGAGGTGAACTAA
GAAGTTGAAACATTGGAGAAGTCTGGAGAAGATAGCTGACCAGCTGGAGAGGTTGAAATAAGAGCTTA
CTGGAATCCAGCAGGGTTCTGCCAAGGATTGCAAGCTGAAGCTCTGCAAACCTGATAGGAGAGTA
AAAGCCACAATAGAGCAGTTATGAAGATCTGGAGGAGATTGACACACTGATCTGCCAGAAAATTCAA
AGACAGTAGATTGAAAAGGAAAGGCTGGTAAAAAGGTTCAAGCATTCTAGCCAGTGTGACACAGTGG
AGCAGAACATCTGCCAGGAGACTGAGCGGCTGCAGTACAACATTGCCCTGGCCAGTGAGGTGAG
GAAAAAGGCTGTGCTGCCCTGAAGAATGGGCCACCAGCTCTGCCGCTCTGGAGCGGAATTACCTGATT
TCTTCAGGGCTGCTGGGGCAACTGCCATTGCAATTTCCTACTCTCACACTGGTCTCAATGAAAAA
TAGTGTCTTGTGATTTGAGTAAAGCTCCTATCTGTTTCTCCTCTGTCTGTGGTTGACTGTCAG
CAATCCACCTTCTGGAGAGGGCACCTCTGCCAAATTTCAGCTGTTGGACCTCTGGGTGCTTC
TTTGGGCTGGTGGAGAGCTCTAATTGCTTGGCAGTTCAAGGTTAGGCCCCCTCAGTCTCAGATA
CATGAGGGCTTCTTGTGATCGTAGTCCCATAGCTGTAAACCCAGAACATCACCAGGAGGTGCA
CCTAGTCAGGAATATTGGGAATGGCTAGAACAAAGGTGTTGGCACATAAGTAGAACCAACTTACCTCATT
GTGACCTAATTCCAGAGCATGGCTGGGTTGTTGGGTTCTAGACTTGTCTCACCTCCCAGTGACCC
ACTAGCCACAGGCATGAGATACCAGGGGCCCTGGATGGAGCCTGTTGATGCAAGGCTCT
TGTCCCCAAGCAAGTCTCAGAAGGTTAGAACCCAGTGTGACTGAGTCTGTGCTGAAACCCAGGCCAG
CCATGGATTAGGAAGGGCAAAGAGAACGGACCAGAACATGAGTAAAGCAGGAGGTGAG
ACTTCTCAGGAGTGCATGTGCTTCTCAAAGGATTTCAGGTTAACCATATCCTCTGAGTTCTATGTT
CCTTCAGCTGTTCTATCCATTGACTGCCCCACCCCCACCCCATATTGTTTAAAAAATT
AGGCCTGGCGCAGCAGCTCATGCCTATAATCCCAGCACTTGGGAGGCTGAGGCCGGGATCACTGAGG
CCAGGAGTTGAGACCAGGCCAGGCAACATAGCAAACCCATTCTGCTTAAAAAAAAAAAAAAA
TTAGCTGGCTAGTGGCATGTGCTATAATCCAGCTACTGGGAGGTGAGGCACAAGAACATTGAA
CCTGGGAGGTAGAGGTTGCTGTGAGCCAGATTACGCCCTGCACTCCAGCCTGGTCACAGAGTGAGACT
CCATCTCAGAAAAAAATTGAGTCAGGTGCACTAGCTCTCCTGTAGTCCAGCTACTTGGGAGG
TGAGGCTAGAGGATCACTTGAGGCCAGGAGTTGAGTCTAGTCTGGCAACATAGCAAGACCCATCTCTA
AAATTAAAGTAAGTAAAGTAGATAAAATAAAAAGAAAAAAACTGTTATGTGCTCATCATAAAAGTAGAA
GAGTGGTTGCTTTTTTTTTGGATTAATGAGGAATCATTGCTGGCTCTAGTCATAATTATG
CTTAATAACATTGATAGTAGGCCCTTGCGCTATAACTCTACCTAAAGACTCACATCATTGGCAGAGAG
AGTCGGTGAAGTCCAGGAATTCAAGACTGGCAGGTTAACAGCTCAGAACAGTAGTAGAGGTTAGACTTG
TGGACAAGGCTGGTCCCAGCCCACCCCACTTTAACATCAGAGTGGTCACTATTGATCTATT
GTGTGATAGCTGTGCGTGGCCACACATTAAATGAGAAGTTACTGTGCAACAAACTGCCAACACCA
TTCTAAACTATTCATATATATTAGTCATTAAATTCTACATAACTTGAGAGGTAGACAGATACTCTTATT
TAGAGATGAGGAAACCAAGAGAACCTAGGTCAATTAGCGAACAGGTTGAGAGTAAGCGAACAGACA
CAAAGCTGGGTGGTTGGCTCAGAGCCAGTGTGCTTTCCCTACTGTACTGCCTCTCAACCAACACAG
GTTGCACAGGCCATTCTGCTGTTCTCTGCTACTACCTCTGCTGAACAGCTCCTCACCTCCTCCAACC
TGCTCTAGTCATTCTTAGAGCAGCCAGTGATCATGAAGTGCAATCTGCCATGTCAGTCCCCTG
CTTAGAACCCCTCAAATGGCTCACTTCTCTTTAGGCAAAAGTCTTACCCATGCCCTCTCCATCTCATC
TCAACCCCTCATTTGTTGGCTGTGTCAGCCACTCTCTCAGGTCTCAGATGCACTGCACCC
CTCCTGCCCTGGGGCTTGTCTGCTACTACCTCTGCTGAACAGCTCCTCACCTCCTCCAACC
CTACCCCTGTATAGGTGACTTTGTTCATCTTCAGAACATTCAACTCACATGTCCTGCATGGAGAACCC
CACCTACTGTGTTGAGACCCCTGTCAGGCCACCCAGGTGGATCTCTGACTTCCACATTTCTTCA
CAGCATTACATAGTCAGTAAAGTTACTGTGGGATTATTGGTTAATCTTGCTTAAACACCAGGG
TCCTTGGGTGAAGGAGCTTCTTATCTGTAACAGCATTATTCAAGCATAACTGTAATATAGTTAT
TACATATATAACATATATATAACATAACATATATAACATATAACAAGCATAACTGTTATATAGT
CTTGTATATAGTAAGACCTCAATAATATTGGAGAACAAAAAA

LAQRGARRPRGDRERLGSRLRALRPGREPRQSEPPAQRGPPPSRPPARSTASGHDRPTRGAAAGARRPR
MKKKTRRRSTRSEELTRSEELTSEEATWSEEATQSEEATQGEEMNRSQEVTDEESTRSEEVREEMAA
GLTVTVTHSNEKHDLHVTSQLQGSSEPVVQDLAQVVEEVIGVPQSFQLIFKGKSLKEMETPLSALGIQDG

RVMLIGKKNSPQEVELKKLKHEKSVEKIADQLEELNKELTGIQQGFLPKDLQAEALCKLDRRVKATIEQ
FMKILEEIDTLILPENFKDSRLKRKGLVKKVQAFLAECDTVEQNICQETERLQSTNFALAE

+++++

#16

NM_001099456

Region conserved from zebrafish to human – in non-mammals and even some mammals initiated by AUG.

NPW neuropeptide W

Synonyms: L8; L8C; PPL8; PPNPW; NPW

GTACCCACTCCCTGGCACTTCACTCCTAGAGGGAGGCTGAGCAGGCAGAGAATGGGACGTGTCCCCCT
CAGAGGAGCCTGAGGCCAGTTCCAGCCAGCGGCCACTCAG **TGA**GGTCTCAAGTACCCACGTCCCCCGC
CAGCTGCCAGGGTTCCTCTCCTCCCTCCGCTCCCTCCCCCATCTGGGAGGCCAGCGGTACTGAGGGGGC
GGAACGAGGCGGGGCCACCGAGCGTTATAGCTGGGCTGCAGGGACCACGGCTCGCCTCAGCCTCCT
GCGCTCCGGTACCTGGCGTCCAACCTCACTGCGCGCCAAACCCAGCCAGGCCAGGGTTCTGGGCCCGCC
CGCCGGCGGCCGTCGACGCGAGCGCC **CTGGCGTGGCGCCCAGGGAGCGGGGGCTCCCGCAGCCGGCC**
GCGGCTGGCACTGCTGCTTCTGCTCTGCTGCGCTGCCCTCCGGCGCTGGTACAAGCACGTGGCGA
GTCCCCGCTACACACGGTGGCCGCGCCTGCTATGGGCTGCGTCACCCCTATCTGTGG
CGCCGCGCGCTGCCGCCGGCCGGCCGGCCCTGGCCAGGGACACCCCTCCCCGAACCCGAGCCCGGA
GGCTCTCTCTGCTGCCCTCGTGGGTCAGGAGCTGTGGGAGACGCAGCAGGAGCTCCAGGCAGGA
TCCCCGTCGTGCCCGAGCCGGAGCCGCGCCAGAGCCTGCGCTGGAACCGGAGTCCCTGGAATTCA
GGAGCTGGCCAGAGACTTCGGAGAGACGTCTCCGCCAGCGGTGGACCCGCAGCAAACCGCCTGGCCT
GCCCTGCGTGGCCCCCGGACCGTCTGACAGCGTCCCCCGCCCGCCGTGGCGCCTCCGCGCCTGACCCAG
GAGGAGTGGCCGCGCCTCAGGAGCCGCTCATAGACCCCGCTGCCGTCCGGTCAATAAAATCCGCCTG
ACTCCTGCGCCCCCGCATGCGAAAAAAAAAAAAAAA

LAWRPGERGAPASRPRLLALLLLLLPLPSGAWYKHVASPRYHTVGRAAGLLMGLRRSPYLWRRALRAAAG
PLARDTLSPEPAAREAPLLLPSWVQELWETRRRSSQAGIPVRAPRSRPAPEPESLDFSGAGQRLRRD
VSRPAVDPAANRLGLPCLAPGPF

+++++

#17

NM_002467

CUG initiation conserved in some mammals but the sequence of the extension is not
MYC v-myc myelocytomatosis viral oncogene homolog (avian)

Synonyms: MRTL; c-Myc; bHLHe39; MYC

ACCCCGAGCTGTGCTCGCGGCCACGCCGGCCGGCCGGCCGTCCTGGCTCCCTCCCTGCCTCGA
GAAGGGCAGGGCTCTCAGAGGCTTGGGGAAAAGAACGGAGGGAGGGATCGCCTGAGTATAAAAGCC
GGTTTTCGGGGTTTATCTAACTCGCTGTAGTAATTCCAGCGAGAGGCAGAGGGAGCGAGCAGGGCGCCGG
CTAGGGTGGAGAGCCGGCGAGCAGAGCTGCGCTGCCGCTCTGGAGGGAGATCCGGAGCGAATAG
GGGGCTCGCCTCTGGCCAGCCCTCCCGCTGATCCCCCAGCCAGCGTCCGCAACCCCTGCCGCATCCAC
GAAACTTGGCCATAGCAGCGGGCGGGACTTGCACTGGAACCTACAACACCCGAGCAAGGACGCGACTC
TCCCGACGCGGGAGGCTATTCTGCCATTGGGACACTTCCCGCCGCTGCCAGGACCCGCTCTCTGA
AAGGCTCTCTTGAGCTGCTAGACGCTGGATTTCGGTAGTGAAAACCAGCAGCCTCCCGCGAC
GATGCCCTCAACGTTAGCTCACCAACAGGAACATGACCTCGACTACGACTCGGTGAGCCGTATTTCT
ACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCGGCGCCAGCGAG

GATATCTGGAAGAAATTGAGCTGCTGCCACCCGCCCTGTCCCCTAGCCGCCCTCGGGCTCTGCTC
GCCCTCTACGTTGGGTACACCCCTCTCCCTCGGGAGACAACGACGGCGTGGCGGGAGCTCTCCA
CGGCCGACAGCTGGAGATGGTGACCGAGCTGCTGGAGGAGACATGGTAACCAGAGTTCATCTGCAC
CCGGACGACGAGACCTTCATAAAAACATCATCATCCAGGACTGTATGTTGAGCGCTCTCGGCCGCC
CAAGCTCGTCTAGAGAAGCTGGCCTTACCCAGGCTGCCCAAAGACAGCGGCCAGCCGAACCCGCC
GCCGCCACAGCGTCTGCCACCTCCAGCTTGTACCTGCAGGATCTGAGCGCCGCCCTCAGAGTCATC
GACCCCTCGGTGGTCTTCCCCCTACCCCTCAACGACAGCAGCTGCCAAAGTCCTGCGCTCGCAAGACTC
CAGGCCTCTCTCCGTCCTCGGATTCTCTGCTCTCGACGGACTCTCCCGCAGGGCAGCCCCGAGC
CCCTGGTGCTCCATGAGGAGACACGCCACCAGCAGCAGCAGCTGAGGAGGAACAAGAAAGATGAGGAA
GAAATCGATGTTCTGAGGAAAAGAGGCAGGCTCTGGAAAAGGTCAAGAGTCAGGATCTGGATCACCTCTG
TGGAGGCCACAGCAAACCTCTCACAGCCCAGTGGCTCTCAAGAGGTGCCACGTCACACATCAGCACA
ACTACGGCAGGCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCA
GTCCTGAGACAGATCAGCAACAACGAAAATGCACCCAGGCTCTCGGACACCGAGGAGAATGTCAA
GAGGCGAACACACAAACGTTGGAGGCCAGAGGAGGAACGAGCTAAACGGAGCTTTTGCCTGCGTG
ACCAGATCCCAGGTTGGAAAACAATGAAAAGGCCCAAGGTAGTTACCTTAAAAAGCCACAGCATA
ATCCTGTCGTCAGCAGAGGAGCAAAGCTATTCTGAAGAGGACTTGTGCGGAAACGACGAGAAC
GTTGAAACACAAACTGAACAGCTACGAACTCTGTGCTAAAGAAAAGTAAGGAAACGATTCTCT
ACAGAAATGTCCTGAGCAATCACCTATGAACTTGTCTCAAATGATGATCAAATGAAACCTCACAACCTTG
GCTGAGTCTTGAGACTGAAAGATTAGCATAATGTAAGACTGCTCAAATTGGACTTGGGATAAAAGAA
CTTTTTATGCTTACCATTTTTTTCTTTAACAGATTGTATTTAAGAATTGTTTAAAAAATTT
AAGATTACACAATGTTCTGTAAATATTGCCATTAAATGTAATAACTTTAATAAAACGTTTATAGCA
GTTACACAGAATTCAATCTAGTATATAGTACCTAGTATTAGGTACTATAACCTTAATTTTTT
TTAAGTACATTTGCTTTAAAGTTGATTCTATTGTTTAGAAAAAATAACTGGCAAAT
ATATCATTGAGCaaaaaaaaaaaaaaaaaaaaaa

LDFFRVVENQOPPATMPLNSFTNRNYLDYDSVQPYFYCDEEENFYQQQQSELQPPAPSEDIWKKFELL
PTPPLSPSRRSGLCSPSYVAVTPFSLRGDNDGGGSFSTADQLEMVTELLGGDMVNQSFICDPDDETFIKN
IIIQDCMWSGFSAAAKLVSEKLASYQAARKDSGSPNPARGHNSVCSTSSLYLQDLSAAASECIDPSVVFPYP
LNDSSSPKSCASQDSSAFSPSSDSLLSSTESSPQGSPEPLVHEETPPTTSDSEEQEDEEEIDVVSVEK
RQAPGKRSESGSPSAGGHSKPPHSPLVKRCHVSTHQHNYAAPPSTRKDYPAAKRVKLDVRVLRQISNNR
KCTSPRSSDTEENVKRRTHNVLERQRRELKRSFFALRDQIPELENNEKAPKVVILKKATAYILSVQAEQ
KLISEEDLLRKREQLKHKEQLRNSCA

C)

#1

NM_002006

Does not appear conserved beyond humans.

FGF2 fibroblast growth factor 2 (basic)

Synonyms: BFGF; FGFB; HBGF-2; FGF2

CGGCCCCAGAAAACCCGAGCGAGTACGGGGCGGCCGCAGGAGGGAGGAGAACTGGGGCGGGAGGCTG
GTGGGTGTGGGGGGTGGAGATGTTAGAAGATGTGACGCCGCCGGCTGCAGCTTGGAGGCCCTCTCCCC
GACACCCATCCGTGAACCCCAGGTCCCCGGCCGGCTCGCCGCCGACCCAGGGGGCGGAGAAGAG
CGGCCGAGCGGCTCGAGGCTGGGGACCGCGGGCGCGCGCTGCCGGCGGGAGGCTGGGGGGCG
GGGCCGGGGCGTCCCCGGAGCGGGTGGAGGCCGGGGCGGGGAGGCTCCCCGGCG
CTCCAGCGGCTCGGGATCCGGCGGGCCCGCAGGGACCATGGCAGCCGGAGGATCACCACGCTGCC
GCCTTGGCCGAGGATGGCGGAGCGGGCTTCCGCCGACTTCAGGACCCAAAGCGGCTGTACTG
CAAAACGGGGCTTCTTCCTGCGCATCCACCCGACGGCCAGTTGACGGGGTCCGGAGAACAGCGACC
CTCACATCAAGCTACAACCTCAAGCAGAAGAGAGAGGAGTTGTGCTATCAAAGGAGTGTGCTAACCGT

TACCTGGCTATGAAGGAAGATGGAAGATTACTGGCTTCTAAATGTGTTACGGATGAGTGTTCCTTTTGAA
ACGATTGGAATCTAATAACTACAATACTTACCGGTCAAGGAAATACACCAGTGGTATGTGGACTGAAAC
GAACGGGCAGTATAAACCTGGATCCAAAACAGGACCTGGCAGAAAGCTATACTTTCTCCAATGTCT
GCTAAGAGCTGAATTAAATGCCACATCTAATCTCATTCACATGAAAGAAGAAGTATATTAGAAATT
GTTAATGAGAGTAAAGAAAATAATGTGTATAGCTCAGTTGGATAATTGGTCAAACAATTGTTATC
GTAGTAAAATATGTAACCATTGCCCAGTAAAGAAAAATAACAAAAGTTGTAAGAAGTATATTCTCCCTT
TATATTGCATCTGCTGTTACCCAGTGAAGCTTACCTAGAGCAATGATCTTTCACGCATTGCTTTATT
GAAAAGAGGCTTTAAATGTGCATGTTAGAAACAAAATTCTCATGAAATCATATACTACATTAGAAAAT
CACAGTCAGATGTTAATCAATCCAAAATGTCCACTATTCTTATGTCATTGTTAGTCTACATGTTCTA
AACATATAAAATGTGAATTAAATCAATTCTTCATAGTTTATAATTCTCTGGCAGTTCTTATGATAGAG
TTTATAAAACAGTCTGTGAAACTGCTGGAAGTCTTCCACAGTCAGGTCAATTGTCAGTCTGCCAGGTCA
TGTACCCATACAGCAGCAGCTAGCAACTCTGCTGGTGTAGGGAGTTGTCAGTCTGCCAGGTCA
TTGAGATCCATCCACTCACATCTAACGATTCTCCTGGCAAAATTATGGTGAATGAATATGGCTTAG
GCGGCAGATGATATACATATCTGACTTCCAAAAGCTCCAGGATTGTCAGTCTGCTGGCAAAACCTCTC
CGGACCTGAATTCTGATTTTATACCACTCTCAGTCAGGCTGTCTCTACGTAAAAA
AAGAGATGTACAAATCAATAATAATTACACTTTAGAAACTGTATCATCAAAGATTTCAGTTAAAGTAGC
ATTATGTAAGGCTCAAAACATTACCCATAACAAAGTAAAGTTCAATACAAATTCTTGCCTGTGGATA
TCAAGAAATCCAAAATATTCTTACCACTGTAATTCAAGAAGCTTTGAAATGCTGAATATTCTTG
GCTGCTACTTGGAGGCTTACCTACCTGTACATTGTCAGCTTTGGGTCAGCTCTTTAACCTCTGCTGCT
TCCAAAAGGTTAAAATATAGATTGAAAGTTAAAACATTGTCATGGCTGCAGTCAGTCTGTTCTGAG
ATAAGATTCCAAAGAACTTAGATTCAATTCTAACACCGAAATGCTGGAGGTGTTGATCAGTTCAAG
AAACTTGAATATAATAATTCAACAAAGGTTTACATTGTCAGTGTGTCAGTCTCCAAAGTATTAGGAGA
AATGCAATTGTCAGGATTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTAA
TTCTACTGGTGAACATGCAAAGAAGAGGAAGTCACAGAAACATGTCATGTCAGTCTGCTGACT
GTAGACTGTCCTACCATAGACTGTCCTACCCATCCCCTGGATATGCTCTGTTTCCCTCTAATAGCTA
TGGAAAGATGCATAGAAAGAGTATAATGTTTAAACATAAGGCATTGCTGCCTTCAATTACATG
CTGACTCCCTACAATTGAGATTGCCCAGTGGTAAACATGGTTAGAAACAACTGAAAGCATAAAAGAA
AAATCTAGGCCGGTGCAGTGGCTCATGCCTATATTCCCTGACTTGGGAGGCCAAGCAGGAGGATCGC
TTGAGCCCAGGAGTTCAAGACCAACCTGGTAAACCCCGTCTACAAAAAAACACAAAAATAGCCAGGC
ATGGTGGCGTGTACATGTTGCTCAGATACTGGGAGGCTGAGGTGGGAGGGTTGATCAGTCTGAG
AGGTCAAGGGTGCAGTGGCATAATCGTGCCTAGCAGTCAGCCTAGGCAACAGAGTGGACTTTGCT
CAAAAAAGAGAAATTCTTAAATAAGAAAAGTAAATTACTCTGATGTGCAATACATTTGTTATTAA
TTTATTATTAAAGATGGTAGCACTAGTCTAAATTGATAAAATCCCTAACATGTTAAATGTCATT
TTTATTCTTAAATTGTTGAAAGATAATTGTCAGTGTGTTAACATGTTGTTATTAAATTATTAAAGATA
GTAGCACTAGTCTAAATTGATATAACATCTCTAACATTGTTAACATGTCATTGTTAACATGTT
GAAAATAAAATTATGGGATCCTATTAGCTCTTAGTACCAACTAACAGTGGCATGTCAGTCATGAT
CTATGCTGTTCTATGTCGTTGAAAGCACCAGGATGGGGTAGTGAGCAATCTGCTGCTCAGCAGTCACC
ATAGCAGCTGACTGAAATCAGCACTGCTGAGTAGTTGATCAGTTAACATGAAACTAAGTGA
AAAATTGAATGGCAAATAAGTGTGTTCTCAGAGTATGCGGGAGACCCCTCACCTCAAGATGGATA
TTTCTCCCCAAGGATTCAAGATGAATTGAAATTAAATCAAGATACTGTCAGTGTGTTATTGTTATT
TTTATTATTAAATATACTGTAAGCCAAACTGAAATAACATTGCTGTTTATAGGTTGAAGAACATAGG
AAAAACTAAGAGGTTGTTTATTTGCTGATGAAAGAGATATGTTAAATATGTTGATTGTTGTT
TAGTTACAGGACAATAATGAAATGGAGTTATTTGTTATTCTATTGTTATTTAAATAAGAATT
AGATTGAAATAAAATATAATGGGAAATAATCTGCAGAATGTCAGGTTTCCCTGGTGTGTTCCCTGACTCTA
GTGCACTGATGATCTGATAAGGCTCAGCTGTTTATAGTCTCTGGCTAATGCAGCAGATACTCTCCT
GCCAGTGGTAATACGATTAAAGAAGGAGTTGTCATTAAATCTGTTATGACCCCTTGTGATATTAAAAAA
ACACTATGGATAACAATTCTCATTACCTAGTATTGAAAGAATGAAGGAGTTCAAACAAATGTGTT
CCAGTAACTAGGGTTACTGTTGAGCAATATAATGTTAACTGTTGATGGCAGTATTCTCAAAG
TACATTGCATGTTCTCAAATACAGAGTTAAATAATTCTAGTAATTCTAGATGATTGACTCAGCT
TAAGAATATCTTTGTTTATGTTGAGTTAGAAATGCTCATATAAGACATAGTCAGGCTACTG
TCAGTTTCATTCTAGCTGCTTCAGGGTTTATGAAATTTCAGGCAAAGCTTAAATTACTAAGCT
AGGAAGTATGGCTAATGCCAACGGCAGTTTTCTCTTAATTCCACATGACTGAGGCATATGATCTC
TGGTAGGTGAGTTGTTGACAACCACAAGCACTTTTTTTAAAGAAAAAGGTAGTGAATT

TAATCATCTGGACTTAAGAAGGATTCTGGAGTACTTAGGCCTGAAATTATATATTTGGCTTGAAATGTGTTTCTTCATTACATCTACAAGTAAGTACAGCTGAAATTAGCAGAGGACCCATAAGAGTTCACATGA
AAAAAAATCAATTATTTGAAAAGGCAAGATGCAGGAGAGAGAAGCCTGCAAACCTGCAGACTGCTTTTGCCTATAGATTAGCTCACATGCTGCTCACGTGGACCACTGAGTAGTGTGAGAGAATTAGGCTGTAGAACAAATGGCTTCTCTTCACTGAGTGGCTAATAGCTTAAAGCTTAATTAGCTCACATGCTGCTCACGTGGCA
CCAGTGGATAGTGTGAGAGAATTAGGCTGTAGAACAAATGGCTTCTCTTCACTGAGTGGCTAATAGCTTAAAGCTTAATTAGCTCACATGCTGCTCACGTGGCA
AATCATCTTTATATCAACAGAAGAATAAGCATAAACTAACAGAAAAGGTCAATAAGTACCTGAAACCAAGA
TTGGCTAGAGATATCTTAATGCAATCCATTCTGATGGATTGTTACGAGTTGCTATATAATGTATGT
ATGGTATTTGATTGTGAAAAGTTAAAATCAAGCTTAAAGTACATGGACATTAAATAAAATAT
TTAAAGACAATTAGAAAATTGCCCTAATATCATTTGTTGGCTAAATAGAATAGGGACATGCATATTAAAGG
AAAAGGTCACTGGAGAAATAATATTGGTATCAAACAAATACATTGATTGTCATGATACACATTGAATTG
TCCAATAGTTAAGGAATAGGTAGGAAATTGGTTCTATTTCGATTTCTGAAATCAGTGACATAA
ATAATTCTTAGCTTATTTATATTCCCTGCTTAAATACTGAGCTCAGTAAGTTGTTAGGGATTATT
TCTCAGTTGAGACTTCTTATATGACATTAACTATGTTGACTTCCCTGACTATTAAAATAAATAGTAG
ATACAATTTCATAAAGTGAAGAATTATATAATCACTGCTTATAACTGACTTTATTATTTTCAAA
GTTCAATTAAAGGCTACTATTCTCATCCTGTGATGGAATGGCAGGAATTGTTCTCATAGTTAATT
CAACACAATATTAGCTGATCCAAAATAACCTTAATGCTAAACTTACTGATGATGATATCCAAAGCTTCT
CATTTCAGACAGATTAATCCAGAAGCAGTCATAAACAGAAGAATAGTGGTATGTTCTAATGATATT
TTCTACTAATGGAATAAACTGTAATATTAGAAATTGCTGTAATTATATCAGCTCTGAGGTAAATTCTG
AAATGTTCACTCAGTCGGAACAAATTGAAAATTAAATTCTTAGCTATAAAGCAAGAAAGTA
AACACATTAATTCCCTCAACATTAAAGCCAATTAAAATAAAAGATAACACACCAATATCTTCTCAG
GCTCTGACAGGCCCTGGAAACTTCCACATATTTCACACTGAGTAAAGTCAGAAATAAGTTAAC
ATAACTTCACTAACACACATATGAGATTTCACAAATCCACCTATAATTGTCAAAGTGGTTGAGAA
TATATTTTTAGTAATTGATGCAAAATTTCAGCTTCCATCTCCCTCGTTCTTTGGTT
GGGGAGCTGGTAACTGATGAAATCTTCCCACCTTCTCAGGAAATATAAGTGGTTCTTGGTT
AACGTGATACATTCTGTATGAAATGAAACATTGGAGGGAAACATCTACTGAATTCTGTAATTAAAATATT
TTGCTGCTAGTTAACTATGAAACAGATAGAAGAATCTTACAGATGCTGCTATAAATAAGTAGAAAATATAAA
TTTCATCACTAAAATATGCTATTAAATCTATTCCCTATATTGTATTCTAATCAGATGTATTACTCTT
ATTATTCTATTGTATGTTAATGATTATGTAATTGCTTTCATGAGTAGTATGAATAAA
ATTGATTAGTTGTGTTTCTTGTCTCCC

LVGVGGGDVEDVTPRPGCQISGRGARGCNGIPGAAWEAALPRRPRRHPSVNPRSRAAGSPRTRGRRE
ERPSGSRLGDRGRRALPGGLGGRRGRAPERVGGRGRGTAAPRAAPAARGSRPGPAGTMAGSITTL
PALPEDGGSGAFPPGHFKDPKRLYCKNNGFFLRIHPDGRVDVREKSDPHIKLQLQAEERGVVSIGVCAN
RYLAMKEDGRLLASKCVTDECFFFERLESNNYNTYRSRKYTSWYVALKRTGQYKLGSKTGPGQKAILFLPM
SAKS

+++++

#2

NM_014293

No conservation found.

NPTXR neuronal pentraxin receptor

Synonyms: NPR; NPTXR

CGGCCGCGCGACAGCTCCAGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCGCTGCCCCGCTC
GGCCCAGCGCGCCGGCTCCCGCGCCCCGACCCCGCCGCCGCGCTGCCGGGGGCTCGGGCGCCCCCGCC
GCCGCCTCACGCTGAAGTTCTGGCGTGTGCTGGCCGGGCATGCTGGCGTCCCTCGGTGCCGT
CTGCATCATGCCAGCGTCCCCCTGGCGCCAGCCGGCGGGCGCTGCCGGCGGCCGACAATGCTT
CGTCGCTCGGGCGCCGCCGTCCCCGGGCCAGCGGAGCCTGAGCGCGCTGCACGGCGGGCGGT
TCAGCCGGGCCCCCGCGCTGCCGGGGCACCCGGGCCAGCGCGCACCGCTGCCGCCGGGCGCTGTT
CAGCCGCTCCTGTGCACGCCGCTGGCTGCTGCCGTGGGGGCCAGCAGGGGACGCCGAGCTGCCAGACGGCGCTG
CTGCAGCGGCCGAGCGCGAAGAGCTGCTGCTGAGAGCACGCCGAGCAGCTGCCAGACGGCGCTG
CAGCAGGAGGCCGCGCATCCGCCGACCAGGACACCATTGGCTGAGCTCACGGCAAGCTGGCCGCTGCC
GAGCGGCCCTGCCGCCGCCCTCAGGGGCCGGCGACACCATTGGCCAGGGCCCTGGGACT

CGCCTGCGCTCATTCTGGAGCTGGAGGACGCCGTGCGGCCCTGCGGGACCGCATCGACCGCCTGGAGCAG
GAGCTTCCAGCCCCGTGAACCTCTCAGCTGCCCAAGCCCCAGTCTCTGCTGTGCCACCGCCTACACTC
CAAGATGGACCAGCTGGAGGGCAGCTGCTGCCAGGTGCTGGCACTGGAGAAGGAGCGTGTGGCCTCA
GCCACAGCAGCCGCCGGCAGAGGCAGGAAGTGGAAAAGGAGTTGGACGTCTGCAGGGTCGTGTGGCTGAG
CTGGAGCACGGGTCCTCAGCCTACAGTCTCCAGATGCCTCAAGATCAGCATCCCCATCCGTAACAAC
CATGTACGCCCGCGTGCAGGCTCTGCCAGAGCTACGCATTACCGCCTGCATGTGGCTGCCGTCCA
GGTCCAGCGGCACCGGCAGGGCACCCCTCTCTACTCAGTGCCGGCAGGCCAACGAGATTGTACTG
CTAGAGGGGCCATGAGCCATGGAGCTGCTGATCAACGACAAGGTGGCCAGCTGCCCTGAGCCTGAA
GGACAATGGCTGGCACCACTGCATGCCCTGGACCACAAGGGATGGCTATGGTCTGCCCTACAGGACG
GGGAGCTGCAGGGCTCCGGTGAGAACCTGGCTGCTGCCACCCATCAAGCCTCATGGGATCCTTATCTG
GGCCAGGAGCAGGATAACCTGGGTGGCCGGTTGATGCCACCCAGGCCTTGTGCGTGACATTGCCAGTT
TAACCTGTGGGACACGCCCTGACACCAGCCCAGGTCTGGCATTGCAAACGTGACTGCCACTGCTGG
GCAACGTCCCTCCCTGGGAAGACAAGTGGTGAGGCCTTGGGGGTGCAACAAAGGCTGCCCTCGATGTC
TGCAAGGGGAGGGCCAAGGCATGAGGGGCCACCTCATCCAGGGCCCTCCCTGCCACTTGGGGAC
TTGAGGGGGGTATTCCTCCCTCAGCCTGCCACGCACTGGCCTCCCTGCCACTCCTGGCTGT
GCCTCCCATTCCTCAGCTGACCCACACCTCAGAATGCCCTGCCCTGCGAGTGTGCCCCCTGCCCC
ACCTGAGTGGGAGGAGCCTCAAGTGAACAGTGGGAGCCTGCCACCTGGCACTGCACTGGAGTTGTCT
CTTACCCACCCCTCCCTGCCCATCAACTGTATCTGATTTCACTAATTTGACAGCACCCAGTAGGGTAG
GATTGTGTATGAGGGGACCCACTATCTCAGTGGTGGGGTGGCCGCCGCCCCCTGTCCCCCATGCAA
CAGGCCAGTGGCTCCCTCAGGGCACAACAGGCTGAGAAGGGATGACGAGGACATCAGAGGTTAG
ACTTACCTCCCTCCCTTCCACCAGCTGCCAGTCAGGGCAGTGGGATCTCGATGGAGCCTCCCCCCCC
CCCCACCCATGCCCTCTCCCTCTTCCCTCTCTTGTGTAGCGGTTGAATGTTGGTTCCA
TGCCTGCCAGCCCCACCTCAGTCTCAGGACATTCTCCAGCCTGGAGGGAAAGGGACAA
AGACCCAGGAGGCCAAAGGGCTGCAGTCACCCCTGTGCTACCCATAGTGTGATGCCACTGGTATA
TCGCTCCCTCCATGCCAAGGACAGGACTTGGACCGCTTCAGCCTGGCTGGAGCAGCCCTAAGGTAGA
GGCCTCATGGCCCAGGAGACCCACCTCTGGCAGAGCCACATTACCTACCTGTGATGGTCTGGGAG
CAAGGAAGAAGCTCAGAGGGTGGGGAGAAGCATGAAGCAGTGCAGAGCAGACTGGGTGAGAGGGAGAAGAC
CTTGGTTCCCTAGCAGCCCTGCTAATGTGCTGTGGCCTCTGTAAGTCCCTGCCCTCTGGCCTGGC
CTTCCTCATTCTGAGCTGAGGCCCTCGCTTGGCATTGCTCTCCAGATTGGGTGTGAGCTCTGTG
ATTCCAGGTGGATATGTGGGAAAGCTCTGGTACCCCTGGCTCGCAGGGTAGATCCAGGACTGGCA
GTGGATGGGATGCCAGTCAGGTTAGGGTAGCAGAGACTCAGAGTCCAGGGCAAGGTTCAAGGCAG
ACTAACCTCATGCATGGATTGAAAAACAGCTCCCTTGATCAACCCAGCCTGGCACCCCTGCTGTC
TGAGAGTGTCTCAAAGGGCTGATGGCTCCCTGGTCCCTTGAGTCATCACCAGCTCCCCAAGAGAGTGT
AGAATCTTAAGAGCTGAGAGGCCGGCACGGTGGCTACGCCTGTAATCCAGCAGTGGTAAACCC
CAGGCAGATCACTTGAGGTAGGAGTTCGAAGTCAGCCTGCCAACGTTGAAACCCATCTTCA
AATAACAAACTTAGCTGGTAGGTGGCAGTGCCTGTAGTCCCAGCTACTGGGAGGCCAGGAAGAA
TCTCTGAACTGAGGAGGTGGAGGTTGAGATCAGGCCATTGCACTCCAGCCTGGCAACAG
AGCAAGACTCCATCTCAAAAAATAATAATCTTAAAGATGAGAAAAGCCACCCATCTGGCACCA
CTGCATCTGCTTGAGAAATGGGAAGAGTTCAGGGAGGACACGTGACCTGCACAGGATCACAGAGCAT
GGGGCAGGCCAGGACTAGAGCTCAGGGCATCTGACTCCCTCTCAGTGTCTTCCCCCTCCATGTTGCC
GCCCTGAAGACCTTGAGTCAGTCTACACCTAAGCAGGTAGACATCCCGCAGGTCAAGTGT
ATGACACCTGAACATCTCCTTATGCAACACCCAAACATCTGGCATCCCCACCCAGGAAGTGC
GGAGGTTATGATCCCTGGCGCTTCGGCAGAATGGAGAGCTGAGGTGTCCCTCCCTGCTAGTC
AGGTGCTGAGCAGCTGCATGCTCCCTGGCTCAAGTGGCACTGTACCTTTGCTGCC
ATCTCCACTCCCTGAGGCCACTAGCCTGAGACATGATGCAAGAGCTGAGGCC
GGAAGCTACTCCAAGTGCATTGCCCTCCGCCAGATCCTGCTTCCATTGAGAACATA
GACCTACCCACTGGAAGAAAAGGCCAAAGCAGACATCAACCCAGTGCAGGCC
TAGGTTGCCACGGTAACAGCCAAGACATCAACCCAAGTGTGGGCACTGTCTCAT
TTGCTGGGTGACGGCAGAATTCAAGACTTCAGTGCAGCCAAGCTTGT
TTTGAAGAAGGTAGCTTGTGGAGGACTGGGAGAAGGATGGGTCTAGGAAG
CATGGTCACTTGAGCCCACACACGCTCAACCCAAGTC
TCTGACGTCCAAGCCTGTTCTGTGCTGACCCACTCAGCCTCAAAGGG
CCCAGCACCACGTGGTGGGCCAGTCAACACGCTGAGCTGCTCTGG
AAGAGGAGTCAGACATGGCCAGTCACATCCTGAGCTGCTCTGG
TGTCTGCCATCTGGCACTGCACTGAGTGTGGCAGGGC
GTTAGGACGTTCTGGCTCCGTTAGACAGGTGAGGAA
TGTCAATCAGCTAGCAAGTGTGAGGCCAGATT
CAAACCAAAGGGGTTACGTCAGGGCTGAGTT

CACTCACCTGTAGAGTGCATCTGGCACCATTGCTCCAGACGTGTTCCAGCCCTTCCCAGCCCACA
GGGCTTGAAGTGAAGAACAGAGGCAGGGGTGGCCAGCCCCAGGGCAGGGTCCCCTGGTGAAGCCGT
GCCAGGGGGCTCAGCTGCTTCAGGAAATGTGTCCTCCACCATGGGCAGAGCTTCAGCCCTTTAGC
TCAGCTAGAGTTCACAGGAGGCCAAAAAAGAAAAGGAAGCTGAGCATCTCCGAGTCCTGGCAGGGAA
GGGAGGGAAATTGCTGCTCTCCAACCTTGCTGGGGCCAAGCCCTGCACCAGTTGCTCCAGCTGTTA
TCTGCCAGATCTCCATCTTGTGGCATGTGGTGCACCAACATCCAAGGGACCAATCCCTTGCC
ACCACTTGCATCACCTGGGACCACAGATTGGACAGGAAGGGCTCTGAGAAGAGGCCAAAGCCCTCATTT
TACAGATGAGGAAGCTGAAGCCGGGGAGGGAGCGACCCTCAAGGCCACCCAGCTGGACACGGGAGACTT
GAGCCCAGCCTCTGACTGCATTGCAGCCTCTAGGACGCAGCAGCCCTCTCCAGCACTGAGTCCCCC
TCCTTGTGTCCAGCACCTTGGCTGAGTAAACTTGGAAAGGGCTCCCTCCCAGAGAAGGGACTAC
TCTCTTACCCCTTATTCCAGCTGCCACCCAGACCCCCACCTCCACCCCTGACCCCGACCCCTG
GGTGGGAAGGGCTCACATGGCCCAGGCTGAGTGTGAGCATGTCAAGTTGTCTGACACTGTGACA
TTAGTGCACCCCTACTGACAACCCCTCCCCAGCCTGCCCCCTCTCCCTGTGTTGTACATAATTG
ACATGAGCTGCAACATGTGTGCGTGTGCGTGTGTGTATGTGTGATCTGTGTCATG
GTTTGTACCTTTGTTGTAAACTGAATGTTCAAATAACATGCTGTTACTCTGA

LKFLAVLLAAGMLAFLGAVICIIASVPLAASPARALPGGADNASVASGAAASPGPQRSLSALHGAGGSAGP
PALPGAPAASAHPPLPPGPLFSRFLCTPLAACPSGAQQGDAAGAAPGEREELLLLQSTAEQLRQTA
LQQAERIRADQDTIRELTGKLGRCESGLPRLQGAGPRTDMADGPWDSPALILELEDAV
RALRDRIDRLEQELPA
RVNLSAAPAPVSAPVTGLHSKMDQLEGQQLLAQVLAKEVALSHSSRRQRQEVEKE
LDVLQGRVAELEHGS
SSAYSPPDAFKISIPIRNMYARVRKALPELYAFTACMWLRSRSSGTGQGTPFSY
SVPQANEIVLLEAG
HEPMELLINDKVAQLPLSLKDNGWHHICIAWTRDGLWSAYQDGELQGSGENL
AAWHPIKPHGILILGQE
DTLGGRFDATQAFVGDIQFNLWDHALTPAQVLGIANCTAPLLGNVLP
WEDKLVEAFGGATKA
AKA

+++++

#3

NM_021182

No conservation found.

HMHB1 histocompatibility (minor) HB-1

Synonyms: HB-1; HLA-HB1; HMHB1

CCACATCCCAGGAGGCCAGGGCTTGCCCCCATCTCAGAACGCCAGGGCAGGCCCTGAGCCTCTGACCT
CACATCCTCTGCCACACCACAGTGGAGAAACCAGAAGTGAGGAGCAGCCAGAACATGCAGAGAAGAAAAAG
AGGTTCTCTGCATGTTGGAGTCGAATTGGTTGAAGTTGAAGATGATGTATCTGAGGCACAGCTTT
CCCTGACTTATAGGCTTGACACTGCTGTTGAGGTTGACTCGAAGCCAGAGTTGGTGGATGAGCA
GGGACAAATTGCTGAGCATGAAGAAGAGTAAATTAGCAAGTGGACATATGCCCTTGCTCTG
CACAGTGAATGAAAAGTCAACCTTGAAAAA

+++++

#4

NM_001169118

The first annotated AUG (not conserved) actually precedes the supposed UUG initiation codon. In any case nothing upstream of the second in-frame AUG is conserved beyond humans.

STIM2 stromal interaction molecule 2

Synonyms: FLJ39527; KIAA1482; STIM2

GGCGGGAGCGTGGTACTACGACCAGCGCGGCCGGAGGGGGCGGGGGATGCGCCGCCGGCGCGCG
GGAGCTGGGTTGGTGTGCGCGCAGAGCAGCGGATCCGGTCTGCCGCAGCAGCAGCGGGGTGT
CGTGCACCGCCTGAAGACGCCGTACCTTCTACCCCCCACCTTTTTTTTTAAATAACCGGA
CCAATGAACGCAGCCGGATCAGAGCTCCGGAGGCCGCGGTGCCGATGGGACCAGGCTGGGCCCGG
GAGCCCGTGTCTGAGGCCGGGGCGCCGGAGGAGTCGCCGGCGGGTGGTGGCGCTCGCGAGCCG
GCGAGCTGCAGGCCGGCGCCGCGCTGCGCTTCACCCGGCTTCTCCGCCGCTTCATCCGCCCG
ACTCCTGGCCAGCGTGGGCTGGCTGCTGCCGGCGCGTGGCTGCCTGGCTGCCTGGCTGCTG
GGTAGCCGGAGCGCGACGGATCGAGCTTGTGCCCGGCACCTCCGCCGGCGGGCGACTGGCTG
CCGCAACTGCCGCTCCTCTCCGCCGGCGCGATAGCCGGCTCATGACAGATCCCTGCATG
TCACTGAGTCCACCATGCTTACAGAAGAACAGATTAGTCTGGAAGCTCTCAAACAATACATAAAC
AATGGATGATGACAAAGATGGTGAATTGAAGTAGAGGAAGTGATGAATTCAAGAGAACATGAA
ATAAAGATGCTACTAATAAACACAGCCATCTGCACAGAGAACATATAACGATTGAGGATTATGG
AAACGATGGAAAACATCAGAAGTTCATAATTGGACCCCTGAAGACACTCTCAGTGGTGTAGAGTTGT
TGAACCTACCCAAATATGAGAAGAATTAGAGACAACAATGTCAAAGGAACGACACTCCCAGGGATAGCAG
TGCACGAACCTTCATTTATGATCTCCAGTTGAAATCAGTGACCGGAGTCACAGACAAAACCTCAGCTC
AAGGCATTGGATGTGGTTGGACCTCTAACACGCCACCTCATAACTGGATGAAAGATTATCCT
CACAGTTCTATAGTAATTGGTGTGGAGGCTGCTGGTTGTTCTTACCGAGAACATAGAACATCAAAGAAC
ATGTTGCAAAATGATGAAAGATTAGAGAGCTTACAAACTGCAGAGCAAAGTCTAATGGACTTACAAGAG
AGGCTTGGAAAAGGCACAGGAAGAACAGAAATGTTGCTGTAGAAAGCAAAATTAGAGCGAAAATGAT
GGATGAAATCAATTATGCAAAGGAGGGCTGTCGGCTGAGAGAGCTAACGGAGGGAGCTGAATGTGAAT
TGAGTAGACGTCAGTATGAGAACAGGAATTGGAACAGGTTCGATGGCTCTGAAAAGGCCAAAAGAA
TTTGAACCTGAGAACAGTTGGTCTGTTCCAGATGCACTTCAGAAATGGCTCAGTTAACACATGAAAGT
AGTCAACTACAAATTAAAAGACAAAAGCTGAAATGAGCTAGCTATTGCTAAAGATGAGGAGAAC
AAATTAAAAAGAAGAGAACAGTCTTGGGACTCTGCACGGTGCACACAGCTCCCTAGATGAGGT
GACCACAAAATTCTGGAAGCAAAGAACGCTCTGAGTTGACAACCTGTTACAGAACGACTTTTCG
CTGGCAACAAATTGAGAACAGTCTGTTCCAGATAGCCATAACTCAGGACTCCCCAGCCTGACCTCT
CCCTTATTCTGATCACAGCTGGTGTGATGCCAGAGTCTCCATTCCACCCCTATCCAATTGCTGGAGGA
GTTGATGACTTAGATGAAAGAACACACCCCCAATAGTGTACAATTCCCGGACCAGGCTAAACCTCCTGG
ATCATTAGCCAGAACAGCAGCAGCCTGTGCCGTTACGCCGAGCATTGTCCTCGCTCAGCCTCAGC
GAGCTCAGCTGCTCCACACGCCACCCGTACACCCCTGGCACCCCTACCAACACACCCA
CACTCCTGCCTTCCCTGATCCAGATACCTCTAGTGTCAAGTTGCCCTGCGCTTATCGAAATGAAGA
GGAGGAAGAGGCCATTACTCTGCTGAAAAGCAATGGGAAGTGCAGACACAGCTTCAGAACATGTACT
CCTTAAATTCTTCATTGAAAGAACAGTCTCCCTTAAGCCTCGAGATATAACCAAACATTATCTCCG
CGAAAGATATCAAGAGATGAGGTGTCCTAGAGGATTCCCTGGGAGTCCTGTAACGTGGATGT
GTCCTGGGTTCTCCGACTGTGTAGGCTGACAGAAACTAACAGTATGATCTCAGTCTGCAAGAAC
TGTACAATGGCATTGGAGAAATCTGTAGCATGAACCGAGCTTCCAGTGGCATCCGGTGCTAAACCT
CGCCACACATCATGTTCTCAGCTGGCAACGACAGTAAACCAAGTTGAGAACCCCCAAGTGTGCCAGAA
AAGCAGCATCCACATGACCTTGTCTATGGAGAGAACAGCAAAAGCCATCAAACAGCCTT
TTAAGAAGAAATCTAAGTGAACAGCTGGTACTGTGATGGAATCATGTTCAAGTGGCATCTGAAACTATTATC
CCCCACCCCTCCACTCCCCACCTTTTTGGTTAATTAGGAAATGTAACCTCATTGGGCTTCCAGGC
CGGATGCCATAGTGGAACATCCAGAACAGGGCAACTGTCTACTGTCTGCTTATTTAAAGTGA
AATTCAAGCCAGTTATTACTGAAAATCATTGAAATGAGAACAGTTACAGTCTGCTTATTT
TTCTGCTTGTCTCAGTGTATATGCAACATTGTTGAAAGGCCAGATGGACTTACAAGCTTTAATG
GACTCGTAAGCCAGCATGGGCTGCAAAAATTCTGTGTTACAGAGCATCTCTTATCTTCCACAGAGC
TATTACATCCTGGACTATATAACTTAAAGAAGTAAACAGTAATTGCACTACTGTTTCCAGACTGGAAA
AAAAAAAAATCTGCAAGTGAACAGTGTATAGAGTTATAAAAGACTATGGATAGGGACTGTTCACT
TTTAGATCAAATGGGTTTAAGTAGAACCTAGGGTTCTAATTGACTTGATTTCTGAAATGAAAACCC
GCGCTTTATTATGGGAAGCTTCTGAACTGCATTACTATTGTAAGTTCAAGTCCCGCTGTAAGATC
ATGTTGTTGTTCTCCCAAGGGCTTCACTGTGATTACTGCATTGAGCTGTATGATAAAACACAT
AATTAAAGAGAGAACGGCTCTGATTCTTATGCAAGTGGAAAGAGTTGAAACTTGATTGAAGGACTTAAA
CATTACAACCTTAAGCGAGGTGGGGGATATGGGGATTCAAGGCAATTGTTACACACTTGAATAACTG
CAAAGGATTACGGTTGTGAAAATGTTACTGTGAAAGAGATAATAATTGAAGAACATTGTTG
ATTGTGCTGATTGTTGTTGATAACACAAAAACATATGTTCTGAGAGCTGTGTAAGCTGCTTGTG
CTTAGITGCAATATAAGAAATAGTGTGATGTTGGACGTAAGTTGTCACAAATTCTATTATATTGTTA
TATTGTTATGTTGAAATGTTCTAATATCAAGATTAACAAATATAAAATTATGGTGCATT
TAGAAAAAAAAAAAAAA

++++++

#5

NM_005354

From the GenBank description:

In addition, three non-AUG codons also support translation, an ACG codon (in-frame with JunD) and a CUG are positioned in the 5'-untranslated region, and a CUG codon (also in-frame with JunD) is located downstream of the short open reading frame.

In reality the non-AUG sites were observed in the rat sequence. They are not conserved in human.

JUND jun D proto-oncogene

Synonyms: AP-1; JUND

```
AGGAGCCGCCAGTGGAGGGCCGGCGCTGCGCCGCCGGGGCGAGGGCGAGCGGACGG  
GGGGCGGGCCCCCGGGAGGCCACTCCCCCCCAGGCCAGCTGGCAGCGGGAGGAGGATGG  
AAACACCTTCTACGGCGATGAGCGCTGAGCGGCTGGCGGCCAGTGGCAGCGGGAGCCTTC  
GCGTCCCCGGCCCTTGTTCCCCGGGCGCCACGGCCGGCCAGCATGATGAAGAAGGACGC  
GCTGACGCTGAGCCTGAGTGAAGCAGGTGGCGCAGCCTCAAGCCTGCGGCCGCCCTACCCCC  
TGCGCCGACGGGCCAGCGCGACCCCCAGGCCACGGCTGCTGCCCTCCCCAACCTGGGCTGCTG  
AAGCTGCCCTCCCCAGCTGAGCGCTCATCATCCAGTCCAACGGCTGGTCACCACGCCAGAG  
CTCACAGTTCTTCTACCCCAAGGTGGCGCCAGCGAGGAGCAGGAGTTGCCAGGGCTCGTAAGGCC  
TGGAGGATTACACAAGCAGAACAGCTCGGCCGGCGCCGCTGCCGCCGCCGCCGCCGG  
GGGCCCTGGGACGGCCACGGGCTCCGCCACCCCCGGCGAGCTGGCCCGGGCGGCCGAAGC  
GCCTGCTACCGAACCTGAGCAGCTACGCCGGCGCCGGGGCGGGGGCGCCGACGGTCGCCT  
TCGCTGCCGAACCTGTGCCCTCCGCCGCCACCCCCAGGCGCTGGCCGCTGGCCCGCCTGGCTGCG  
CTCAAGGACGAGCCACAGACGGTGCCGAGCTCGCGAGAGAGCCGCCGTTGCGCCATCGA  
CATGGACACGCGAGGAGCGCATCAAGCGGAGCGCAAGCGCTGCCAACCGCATGCCGCCCTCAAGTGC  
GCAAGCGCAAGCTGGAGCGCATCTCGCCCTGGAAGAGAAAGTGAAGACCCCTCAAGAGTCAGAACACGG  
CTGGCGTCCACGGCGAGCCTGCTGCCGAGCAGGTGGCGAGCTCAAGCAGAAAGCTCAGCCACGTCAA  
CAGCGGCTGCCAGCTGCTGCCCAAGCACCAGGTGCCCGCTACTGAGTGGCGCCGCGATGCGCG  
CCACCCCTCCCAAGGGCGGGCTCGCGGGGGTGTGCGTGGCGCCCCGGACTTGGAGAGGGTGC  
GGGGACCCCCCTCCCCGAGTGTGCCAGGAACCTCAGAGAGGGCGCGGCCCCGGGATTCCCCCCCC  
AGGGTGCCAGGACTCGACAAGCTGGACCCCCCTGCTCCCGGGGGGGCGAGCGCATGCCGCCCTCG  
CGCTGCCCTTCCCCCGCGCGGCCGCCGTGTTGACAAACCCGCGCTCGGCTGCCCTTGACA  
CCCGCGCGCGGAAGGGGCTCCGAGGGGCGCAGCTCAAACCCCTGCCCTTCTTACTTTACTTTTT  
TTTTTTCTTGGAAAGAGAGAAAGAACAGAGTGTGATTCTGCCCTATTATGTTCTACTCGGG  
CGTTGGTTGTGTGTGTGTTCTGTTAAAGAAATGGGAAGAAGAAAAAAAATTCT  
CCGCCCTTCTCGATCTCGCTCCCCCTCGTTCTTCGACCGGCCCCCTCCCTTTGTTCTG  
TTGTTTGTGTTGCTACGAGTCCACATTCTGTTGTAATCTGGTTGCCCGTTCTGTTCTGAGT  
AAAGTCTCGTTACGCCAGCTCGGCAAAAAAAAAAAAAAA
```

++++++

#6

NM_001098579

Region conserved from zebrafish to human but it is very difficult to analyze because there are at least four human transcript variants with one of them having 5' AUG codon.

MRVII1 murine retrovirus integration site 1 homolog

Synonyms: IRAG; JAW1L; MRV1

GGCGCTGGCCCAGGGTCTTCAGGCTCACATGCTGCTGGAGGCTCCAGGC GGTTGGGAGATCTGCAG
GGTCCCTTGTCTCCGTCCGCCAGGACAGTGTCCAGTGGCAGAGTTGCATGGGACTGCCATCTGCTG
TGTGGATAACCCATCGGAGGGCCCAGGGT GAGCGACTGCCAAGAGGCCAGGCCAGGGACTCCAGAGGAGCTCTGG
CCAGCTGGTCTGGCTTCCCTGCCATTATCCCCGAGCCAGGCCAGGGAGTGGGCTCCAGAGGAGCTCTGG
CTCTAAGTCCCAGCCCTCTCAGAGGAGGCGTGCGGCCAGGTTCTGCAGCCAGAGGCTCTCGCTCTGA
GCTCCGGAGCCAGATGTAACATTGACCTTAAATGGTAAAGCTCCCAGAGTGAAGAGAGGCTGGCCAGAG
GAGGAAAGGAGAATAACTCAGTTAGACAGGGCTCGCACTGTCACCAGGCTGGAGTGCAGTGGTGCAA
CCCCACCTCCGGGTCAAGCAATTCTCATGCCTCAGCCTCCAAGTAGCTGGGATTACAGGGCTCTCCTA
CCGCAGGAGTATCTGCAGTCCAACCTCCACGATTGTCTGACTGGGATGCCACTTCACCAAGAAGGAGAA
ACCGACAAAACCTGGCCAACAGAGTTACAGTCCCCACAAGAGGTTCTCACCGACACTTGAAGGTGTC
CACTGCCTCCCTGACATCTGTGGACCCCGGGCACATCATTGACCTGGTGAATGCCAGCTGCCAGACA
TCAGCATCTCAGAGGAGGACAAGAAGAAAACCTGGCGCTGTTGAGAAGGCCAAGTGGTGAGTGAGCGA
TTCCTGACCCGGCGTGGGAGGAAGTCCAGGAGCAGCCCCGGAGACTCCCCATCAGTGTTCGGCGAACCT
CAGCCCCAGCGCTCTCCTACATCCTCTCGGAGCAACTCAGTCCCCACCCGCCAGGGTGGATG
TGTGCAGTGGCCGCCATCCCTCTGCCTGGAGCACACCAGAAGGGGATGAGGCCAGGTCTTCA
CCTCACCTGGCGAGCCTAACGTCCCCAAAGGGTAGCTGACAGGAAGCAGAATGACCAGAGGAAAGTGT
TCAGGGCAGGCTGGCTCTCGTCTCCAGTTGAGAAGTCAAAGAGATTGCAATAGAACAAAAGGAAA
ACTTCGATCCCCTCAGTACCCGAGACCAACACCAAGGCTAGCTCTGTACAAACAGCAGTGGGAA
ATGCCCCCTGAACAGCCCTAGCCTGGCCCCGTGGAGAGCGAGCTGGGGAGCAGCTTGTAAAAGGGCTG
GGAGGGCAGCCCTTGCGAGAAGTCAAACCCAGGATGCGGAGGAGTGGTCCCCAGCCTCCAGGGGA
GAGGCCAGCTGGAGAGCCATGGGCCCAGGCTGGCTCAAAGCTGAGCTCCACCCACTGTGTCCGG
CCCCCGCTGTCGGAGGGCTCTCTGGACAGTGGCCCTGAAGAACCTGGCCCCGGCTGCAGAAAGTGT
TGCCAAGCTGCCACTGGCAGAGGAAGAAAAGCGTTTGCAAGGCAAGGCCGGCGCAAGCTGCCAGGCC
CTGGTCTCAAAGACTTCAGATACTAGTGGAGCAGCCGTGCGGATGCAGAAACTGACCAAGCTCCAGAGGAG
CACATCCTGATGAGAAATCAGAACTTACTAGTGGGCTCAAGCTTCAAGCCTTAGTGAAGCAGCTGAGCAGGA
AAAAGGGCTTCTCTGAACTCTCCCCAGCTATTGAGGAAGAAGAGTCAAAGAGTGGCTTAGATGTATGC
CTAATATTCTGATGTGCTGCGCAAACCTGCGGGTCCACAGGAGTCTCCCTGGAAAGTGCCCTCCACTC
ACTGAAAAGGAAGTTGAGAACGTTGCTGCAACTGCTCTGGCTTTAGAAATGACAGCTACACTCTGGA
ATCTAGAATTAACCAGGCTGAAAGGGAACGCAACCTGACAGAGGAAACACTGAGAAAGAACACTGGAAA
TCAAAGCTCCATTACGTCCTCAGCTTCACTCTGGCACCACGTGAGCACCGGGAAACCTACAGAACGTT
CTGGAGGACATCGCTGTCCTGCACCGCTGGCTGCCGCTCTCCAGCCAGCTGAGGTGGTAGGCGCCG
CCGCCAGGAAAAGCGCATGTCGAAAGCAACGGAAAGTGTAGTGCAGTATGTGGAGAATCTAAAGAGGACGT
ATGAGAAGGACCATGCGGAGCTCATGGAGTTAAAAGCTTGCACATTCAAGCCGAGCTGTGGC
CCCTCTGAAGATGGGCTCCCTCGCACGGCACGGTCCATGTCCTCACGCTGGAAAGAATATGCCCTGCC
GAGGGTCAGCGTTGCTGTGGTCTTAAGTTAATGCCCTGAATCTGCCCTGGCAAACCTCCAGCTCATCAT
CCATTCCCTCCCTTACAGCCTGCGGAATCACCAATGGGAAAGGAGCCTACCTGTCACTTCAGCACTG
CCTGCACTTTGGAAATGGAAAGACAAATGGGACCCAGATTGTGAAGCCTCTGCTCGCTGACCC
GAGCTGCCCTGGAGGAGCTTAGTCAGGAGACCAAGGCCAGGATGGAGGAAGAAGCCTACAGCAAGGGATTCC
AAGAAGGTCTAAAGAAGACCAAAGAACCTCAAGAACCTGAAGGAGGAGGAGGAAGAACAGAACAGTGGAGAGT
CCTGAGGAACCTGAAGAGGTAGAAGAAACTGAGGAAGAGGAAAAGGGCCAAGAACAGCAGCAAACCTGAAGA
ATTGGTCCATTCTACAAGTCATGTATCCCAAACCTGTGTCAGCACTGCCAGTGTGATCTGGATGATGGCTG
CACTGATGCTGGCTTGACTGTTGCTGGGCTCTACAATTCTATAACTCTTGTCAGAGCAGGCTGAT
GGGCCCCCTGGAAAGATCCACTTGCTCGCAGCCAGAGGACTCTGGTGGAGCTCAGGACTCCAGCATGA
GCAGCCTACAGAGCAGTAGGAAACCTCACACCTAGCCAGTGCCCTGCTGTGAGACACTCAGACTACCACCC
TTTCCCCAAGTATAACGTCAGGCCAAGTGTGGACACACTGCCGCCATCCCACAGGTATGAGGAAGGG
TTCTTTAACACTCGGCACTCTGTGGAGCTATTCAACACAGTGAACCTGATGTTCTGGAGGATCAACA
AAACTGCCCTGGAAAGCATCCAGTGGATGAAGAACGTCACCTTCACCAAGGAACACTTATTGGAAAGGGAGG
TCTCCTGCCCTAGCTCAGGTGGCTGGGAGAACTAAAACACCTTCACGGTGGTTGGGGTAAGGAGCGG
GGCACGGGGAGGAGGAGGTAGGGGAGTAAAAAAACTTACTCTCTTTCTCTGTAAATTGGTTATC
AGGAAGAATTGCTTAATGACTAACACCCCTAACGATCAGACCTGGAATTGGAGTTGCAAAGTGACTATCT
TCCCATTCCCATCTCATTTCAATAACTCAGCCTCCATTCTTCCCTTGGAAATGAGAGTTCTTTTA
CAGAAGTAGGAAAGGCTCTCAGAAAAAAAGTATAGGCTGAATTAGCTCAGTGTGCTGAAATGG
GAAGATATGAATTATTATATACGCATCTGTCCACACATACACATACTGTTGTGACACACACACAACAT
GCCTGTGCACAGAGCCAACAACCCCTCAAAAGTGTGCTCTGGGTGTGACCTCTGGATAAAATAAGATGCAT
GCCAAGCCAACCCACAGATTTCACCAAGTGTGGGCAGTCACCAGGCACCTGTCATGAGCTGTCCACAT

GGATTGAAGATTTAAAAACACAGAAAACATGGCTTCAATGGCAGACTTACTAGTCTCCATTCAAATGCCAACCTGAGCTGTACAGCACAACTATTCCCTATTCTCTTTGAAACAGTTAACCCACCTCACAGGTGAATGAGGAGAGAAGATGTGCTTCTGCTTCAGTCTTACTCTGTGTGACCACATGCAAGAGTAAACTTGCACCTCAGTGCTTCAGTCAGTCAAATGGGGTTCCAACCCCAGTATAATTAGGGGTGTTTCAGAGCATCCCCAGTTTACAGCCACTGCCTTCATCAGTTGTAGAGGTTGATTCCATGTGGGTTGTTGTCATTGTTGCATTTTTGATTTTGTTGATATTGATATTGTTGCTTCAATTGCTAAACACTCATATACGACTTACTATGAGCCAAGCACTGTTCTCAGTATTACATAGGTATGAATTCAATTGCTGAAGAAAAAGAAAAAAATACGAAGTGGATATTACCCCTCCATTTCAAATAAGGAAACTGAAGCACAAGAACAGTAATTGACAAGGACACCCGGATGTAATCATGGGGCTGGAGCTCAACCCCAGGGTAGGCTGGCTCCAGAGCTGTGCTCTCCTGACTCTCTGATGGCTCCTAGCTGGAGCCTCACATTCACTGCTTCAGTGGACCTCACTCACTAGTTTTTCCAGCTGTTCTCTCTCAGTCAGTCTTCAGGTCACTCTCGACCGAGTGCAAAATATTACCCCTCCATACCAGCTTGATGACCTTCCTCCATACCTCACCAGACACAACATAATAGGTACACACTCCTCTGTGCTTCTGGCACGTTAAACATTATTATTATTGACCTTACCTATACTGATACCATGGCTATTATGTATCCATCTCCCTAGCATTTTCCCTCAAAGACAAGAACATGTCTTACCCATCTTGGTAAGTGCCTAGCATGGGGCTGACGCTGGAGGGTCAATTAAATGTTGCTAAAAGAACAGCAACATTAAAGGTGGAGAGCAGCCTGGGACAGCTGACATGCTGATGCTTCAGTCTACTCTGTGAGAGAGCAGGAATGAGACCAGACTAGCAACACCATTGCCAAGCTCAAAGACTGGGCTCAATGCAGTCACTCCTTCAGAGAGACCCCCACCCAAAGCATGCCCACTTTAAATAGCATGTTATTGAAGGGGGCATCTTACAGTAGCTAGAAAATGACTGAGGCCAAGCAGGGTTGATCAAGGATGTGCCATTAAAGTAAAGAGTTACAGAGCAGGGCAGAGGACTCTGGGGCAGAAGTGGATGATTGCCCGGCCTCTCCAGGGGTCTGGATACAACACTGAAGGAGCTTAGCTACATGAGGCCCTAGAGCCTAAAGACAGGATGCAAATAGAGTCTAGAGAGTGGCGTGAAGCAGAACTCCAGGTGGGAATGTTCAATCTGCCTCCCTTAAAGCAGGGCAGGCTCAGTGGCCCCATTGTCACTTGGTCACAAGTTCTACCTTGTGGAAGTCAAAGGCCAGGCTGCAAAGACTTCCAGGCACAGGCCCTCTCCACCCATTGTAAGCAAGGGAGGTCGAGGAAAGGACAATCTCCAAAGGGAGCATAAGAATAGCCATATATAACAGGGGCTGAAAATGCTCATGAGTCATCTTCTGGAAAGCAGCAGCCTGAAGCAGTGGAGCTGCTGCCAAGCGGTAGTGAACCTGGAGAGAAACAGGCCCTGGGATTTCACTGAGGATGTGGATCTGAAGAGTCCCTGAAGTCTGACATCTCTGCTTAGCCCTAGGAGTCTGGTCCCTGCCTTCAGTTGCAGAGTGATGTTGTGTCATTGTTGATGTCACCTCCTAAAGACCTTCACCTGGCTGCCACAAAGCCATATGTGTTGCTCCCATATACAGCCTGACAGAGTAAATGGAGAGGAAGTGGTGGATTGTGATACAGGCTATCAGTCCTCATGTTGTTAAGCCTCACACAGGTGTCAGTCATTGAGTTGAAAATAAAGCACATTCCAACCTCAAAAAAAGAAC

LITGDATSPGETDKNLANRVHSPHKRLSHRHLKVSTASLTSVDPAGHIIDLVNDQLPDISISEEDKKKNILLEEAALKVUSERFLTRRGRKSSSPGDPSAVSPNLSPSASPSSRSNSLTVPPTPPGLDVCSGPPSPLPGAPPKQGDEADVSSPHPGEPNVPKGLADRQNDQRKVSQGRLAPRPPPVEKSKEIAIEQKENFDPLQYPETTPKGLAPVTNSSGKMALNSPQPGPVESELGKQLLKTGWEGLPRSPPTQDAAGVGPPASQGRGPAGEPMGPEAGSKAELPPTVSRPPLLRLGSWDSGPEEPGPRLQKVLAKLPLAEEEKRFAGKAGGKLAKAPGLKDFQIQVQPVRMQKLTKLREEHILMRNQNVLVGLKLPDLSEAAEQEKGGLPSELSPAIEEEEKSGLDVMPNISDVLLRKLVRHSLPGSAPPTEKEVENVFVQLSLAFRNDSYTLESRINQAERERNLTEENTEKELENFKASITSSASLWHCEHRETYQKLLEDIAVLHRLAARLSSRAEVVGAVRQEKRMSKATEVMMQYVENLKRTYEKDHAELMEFKKLANQNSSRSCGPSEDGVPRTARSMSLTLGKNMPRRVSVAVVPKFNLNPQQTSSSIPLPALSESPNGKGSLPVTSPALLENKGTDPDCEASAPALTSCLELSQETKARMEEEAYSKGFGQEGLKKTKEQLDQLEEEEEQKSESPEPEEEVEETEEEKGRSSKLEELVHFLQVMYPKLCQHWQVIWMMAAVMLVLTVVLGLYNSYNSCAEQADGPLRSTCSAAQRDSWSSGLQHEQPTEQ

+++++

#7

NM_001085386

No conservation found.

NF-E4 transcription factor NF-E4

Synonyms: NF-E4

GACAATTCCCTGTTACGGAAGACTATAAAACCTGCCCTACTCATTGGTGCTGACGCCATTTAGGCC
TTAGCCTGCCTGCACCCAGCGCTCATTAACAGCAGGTTGCTCCACA**CTG**CCTCGTGTCTGTTGGC
ACACTCTCAAGAGTTGAACGGATACAAGAATCTTCATCTGGTGC~~CC~~AAACCCGGGAGGGCTCCGGTCT
TCGCCCCCGTGGACCTACCCCTCCGCCAGAACAGCAGGCCACAGCAGCCGGACAAAGGAAGCTCCTCAG
CCTCCAGTTGCTTCTGTGC**ATG**CACATCAGTCAGTCACTGATCTCACCTACTGGGCCCTGCAGGCC**ATGGG**
CCACAGCTCCACACAGAACGCCCTAGCAATCCACCTCCACCTGGTGCCTGCTTCAGTGC~~GG~~CAAT**GAAG**
GCCACTGGCCCACACAATGCCAACACCAGGTAAACCCACGAGGCCATGCCCTCTGCGGAGGACCCAC
TGGAAAGTTGGACTGTGAGCGCCCCCTGCAAGGACCACCCCATCCCTTGAGCCAATCAAACCCCTCTA
CTCGGATCTCGTCA~~G~~CC~~T~~GGC~~G~~CTGAAGACTGATAGTGC~~CTT~~GGAACAGACACCCAGCAACTACC~~ATCG~~
CTTCATCTGAGCCAAGGG**TA**AC~~CT~~GTGGCAGGCCAGT~~CC~~AAAGGCCACTTG~~G~~CC~~A~~AG
AGTGTGAGTCAGCAAGATAGCAGAACAGGAGAGGCCAGAACAGACACCTACTCTGACTGGGAGACA
CGTACCCCTGAAGATTGAGAAAGAGGCCATCCAGGTACACATAGCAGTTACATCAGACTGGGACATTCC
TGT~~TT~~ACAGGAGACTATAAAACCC~~T~~GCCCTACTCATTGGTGCTGACGCCATTAGGCC~~T~~CAGCCTG
CCTGCACCCAGGTACTCATTAAAACAGCGTGTGCTCCAAAAAA

LPRVVCWHTLKS~~LN~~GYKNLSSGAETREGLRSSSPV~~D~~LPLRPRKQATAAGQRKLLS~~QL~~LLCACTSV~~D~~LTY
WGPAGHGATAPHR~~S~~LLAIHLHV~~P~~ASSAAMKAT~~G~~HNAQTQVNPRGHAPS~~A~~DPTGS~~W~~T~~V~~SGPCKDHPH~~P~~F
LSQSNPPTRISSALPLKD~~S~~ALEQTPQQLPSLHSQG

+++++

#8

NM_199072

No conservation found.

MDFIC MyoD family inhibitor domain containing

Synonyms: HIC; MDFIC

CCCAGGCCGGCTCGCCCTCTGACCCAGACAGCGCAGGGCGCAGGGATCGCGCGGCCAGGCCGGGTCG
CGCCGCTCCCAGCATCGGGCCGCTAGCAAGAGTTGAGGCCCTCCGATCCGGATGTGATGAAAAAGAG
CAACAGAGGGAGAAGT~~TT~~CAGGATTG**TAG**GAGT~~GG~~AAAGAGGGAAAGAGAGGCCAGAGAGGGGAAGGCC
CCCTCGCAGGGAGGCCGGCTGGAGTGAGCTGGCTGGAAAGAGGGGGCGA**GTG**CGGGAGTCAGAGCCGCC
~~ACCG~~CTGCCGAG~~TT~~GGCCACTCGGGCTCTGGCTGAGCCGGAGGGAGGCCAGGCCAGGGG
GGCCGCCGCCGTCGT~~C~~AGGCCACCGGGG~~C~~**AA**~~A~~**AT**GC~~GG~~CC~~G~~CTGCCGGAGGCTCGCTAAC~~TT~~CCGGGC
GGAAGAGGAGGAGGAGGAGGAAGGGCTGGAGCGACTACGGGG**GG****AT**GC~~GG~~AGAAGCAGTCAGTTCC
CTGCACCCAGCAC~~CT~~CACGCC~~T~~CCGTGCC~~C~~TGCCGGCGAGCTAGGCCAGGCCGG
CGC~~GG~~CTCGGCCAGGCC~~AT~~**GT**CCGGCGGCCAGCTGGCTCCACAGCCAGGGAAATGTGATAAAGACAATACTGAG
AAAGATATAACTCAAGCTACCAATAGCCACTTCACACATGGAGAGATGCAAGACCAGTCATTGGGAAA
TCCTTCGGATGGTGA~~ACT~~ATTAGAAC~~CC~~AC~~CT~~CAGCG~~CT~~GC~~CT~~CAGCTTCAGACTTCAGCCAGGTG~~C~~
CAAGTGGT~~G~~AGGAAATAGGCAAGATAAAGAACGCCACACAGGTCTGAGCAATGGAAATGGAATT~~C~~ACCAC
GGGG~~CC~~AAACACGGATCCG~~C~~AGATAATCG~~AA~~ACT~~TT~~CAGCAC~~CT~~GT~~TT~~CTC~~AA~~AAATGCATAGAAAAAAT
TCAGTCAGCTTG~~T~~GTAAACAGCGATATCAGTAAGAAGAGCAAAGTAAATG~~CT~~GT~~TT~~CC~~AA~~AGA
CAGGCTTCACCTGAAGATTGTTG~~T~~CCACTGTATCCTGGCTTG~~T~~CTG~~G~~GAATT~~C~~ACC~~TT~~
TGCAACATTG~~C~~CTGGACAAGCGT~~C~~ATG~~GG~~C~~A~~T~~T~~GCAC~~CT~~CAGAAC~~CT~~G~~T~~GT~~TT~~GT~~G~~GAAT~~C~~AT
CGAGATGGGGGATATTGTA~~ACT~~GGCC~~TT~~GTGATATGGACTG~~GG~~C~~A~~T~~T~~GGAT~~G~~C~~T~~GT~~TT~~GT~~G~~GAAT~~C~~AT
CAGACTGCTGGAAATCTGTATGGAATGCTGTGGAATTG~~TT~~CC~~TT~~C~~A~~AAATATT~~T~~AT~~TT~~GT~~TT~~GT~~G~~GA
TGT~~TT~~AAA~~A~~CTGGAGAGTGT~~TT~~AAA~~A~~TT~~C~~CT~~TT~~GGGGGGAAAGAAAAGC~~A~~ATTG~~T~~GAAGATT~~C~~T~~C~~ATG~~AA~~
ACAACATGGAATT~~T~~GC~~A~~T~~T~~GT~~A~~ACT~~T~~GT~~A~~ACTCATTATTGT~~A~~AGTAATCTCTG~~A~~AAGC~~C~~TTT~~T~~TACT~~T~~TAACCAAA~~T~~
CTACATGGTTAATATGTGAAATT~~T~~TA~~A~~CTACT~~T~~TA~~A~~CTAG~~TT~~TATAA~~A~~TTCTTAATATG~~T~~TAACAA~~T~~
CTTAGGGACATT~~T~~GACACCCCCCTCC~~AA~~ATTG~~T~~AA~~A~~GT~~C~~CT~~T~~CT~~TT~~TACCG~~A~~ATT~~T~~CT~~G~~T~~TT~~G

TTTTAACCGTTCTCAGGAGCACTTGCTCAAATATATTATTTTCAGTGTGTATTAACGAGGCAGTT
ATTTTGATATGTATCTATTGATTGAAAGGAAGCAGCTGGCCAGGCACGGTGGCTTACACCTGTAAC
CCTGGCATTGGGAGGCCAAGGTGGCAGATTGCTGAGCTCAGGAGTTGAGACCAGCCAGGGCAACAT
GGTGAACCCCCATCTACTAAAATACAAAAAGTTAGCTGGCTTGGCGGTGCGCCTGAGTCCCAGCT
ACTCAGGAGGCTGAGGAGAATTGCTGAACCCGAGAGGGAGAATTGAGTGCAGTGAGCCAGATTGTGCC
ACTGAACCTCCAACCTGCACTCCAGCCTGGCAACAGAGCAGACTCCATCTCTAAATAAAATAAAATAA
ATAAATAAAATAAAATAAAACAAACCAGTCTTATTTAAAAGAAACTTAGGAAACAAACCCACA
TAATAGTTGGAACCCAGTGTGATCTCTCCCTACCTTCTCCACTGTTCAACAGACTCTGAATGCCA
CTGTGGACTCTTCTCAGACTGTGGGACAGATACTTCACTCTGTCCACAGGAACATGAGATT
TAGCAGACTAAGGAGATCTGTAAAGAATGAACCATACCACAAGGCATACTGAAGTGAGGATTATAAGAGAA
ATAAACTCAAATGCTGTGAATATGCAGAGAATTGCTACAGAATATTCAAGTAAAGGTTCAGGGAGAAT
GTGGCATTGAGGACTCTCTAGAATGAGTGATTCACCTGCTATTAAATGAATTATTAGATTGGACA
AAGATTAGGTGGACACCCCTAAACTGTGTGCTTAAACAGTTAAAGAACAGTGCCTTCAGCATACTT
TTTATTAGTTGAGAATACAGCTTTGAAAAAGCTATAAGTTAACTAAAATATGCATT
CTTACACATAATTAAATGTTATCATACTTTTGATGAAAACATAATGCCTTAGTAAAATAGCTCTATT
AATAAAGAAGATTGAGTACTCTGACACATTCAATTAAATTAGGAAATTAAATATTAAAATCCAGTGT
TCTGAGTTATTGAAAGGCTTCTTTATTTGAGAGCTTAGGTCTTTGGATGAGAACATTAGTTG
TTAGTTGTTCTTAAGCAGTGTCTTAAACAGATAATGGAAACATTCTTCAATGCA
AAGAAATCTAGATATCCCCTACTGTGACCAAATTCTGTATTACGATTATGTTAAATTAAACTAATATG
GCAGGTTATAATGATCCTTAAGTGTAAAGAACAGTCAATTACAAGAGTAATTGTATAGTTGAGACC
TATAGTGTGGCTTAGATGAAAGGGAGGTAATTTCATACCATGCTCTCCTACTCAGTTGATCTC
TCTAAAATTGTAGTTGGTTGATTTAATATAATTCTTAGTAGAAATTGAAAGTATGCTTGGATTAA
TAATTATTTAATTTCTGGCTGAATATCAAATTGATAGTAACAACAGAACGATAATTAGGAAAGGCT
TTCGCAACCTAGCCTTTAAGAGGGTTAACCTGAAGCATGAGAATATACACCTGTGGTTTCT
TTGAGATGAAACGTAGTTCTAGTTATCATTACTAAAGGGCTAAAGAAAACCTAGCAAACATT
TGAATCTTCTTTATTGCTATTACACATACACATACAAACCTTAAATTGGATCTGAAT
ATAATTCTGGTAAACAGCTGCTTCATTTCCTCTAAAGAACTTAATTGTTACATAAAATATA
AGGAAATCTTACTATTACAGTAACCACAATCTAAATATTACATATACCCAAATTAACCTTATGCT
CATATATTAGGATGTGAGAATATCATCTGTTATGGACACATGAAACCTCTAATGACCTGGAATTGTTAG
AATATTGACTTTTATATGCAAAGTTTCAACCAAGTGGTTGTCTAATATTAAACATGACTGGCAC
AATTGTGATGAAATATTAGCACATTGCAATAATTGTTCTCCATAACAGAGAACGTTAATTGGATACAG
AATTATTTTGTATTTATGTTCTAGTACTTTCTCTGTACTCCAGACAGTTATTCCATAAAAGCA
TTTGTATAATTAAAAGGAAACAGAAAAAGGAAAGTAGGCAAATGTGAAATAGTTCAATATATCTTAT
GATTCTTAATGTAAATGTTGTTGAAGTATATGGCTATCATGACTAAGTGTCTAGATTATGTTACA
GGCGGTGTCTTTAAATGTGGAAAGGGTTAAATATTAAACTGGACCTGTATTATCCTGAATACA
CTATTGAAATTTAAATGACTCTTTATTGCTTACCGTATGTTATATCTAATTGACATATT
GACTAATGTTGAAGAACATCAACCATAAGTTAAATCTGAAGGTTATCTTATCATGTTCATCCCTGTC
TGAAGATTCTCTAGTCTTATGTAATACACATGACTCATGTCGTAAATGAAACTATGAAAGATATGAT
CAGTTATGATCATTGACATGTGATTCAAAACACAGTGTCTTTAAATCTATAATATGTC
AAAGTTTTTTTACATCGTTTAGTAAGTTAATTCTATTACTTGAGGCTATATTCCACTTA
GAAAAACTAAGGTAATTACAATATGCTGAGATTAAAACCAAGGAAATGATCAAACATATATGA
AATTGAGTCTTAGATTAAATGAATTCACTCGAAAATTAATGATCAGAACATTTCATCTAA

VRGVRAATAAAVAAATAASGLSRREAGGRAGAAA VVRPPGRKCGRCRRLANFPGRKRRRRRKGLGATTGG
CGEAVSSLHPAPHSPPSSVRPAGRRARRQRRGAGSAERPM MSGAGEALAPGPVGPQRVAEAGGGQLGSTAQK
CDKDNTEDITQATNSHFTHEGMQDQS I WGNPSDGE LIRTPQRLPQLQTSQVPSGEEIGKIKNGHTGLS
NGNGI HHGAKHGSADNRKLSAPVSQKMRKI QSSLVNSDISKKSKVNAVFSQKTGSSPEDCCVHCILACL
FCEFLTLNCNIVLGQASC GIC TSEACCCCCGDEM GDDCNCPDCMDCGIMDACCESSDCLIECM ECGICFPS

+++++

#9

NM_005583

Not well conserved beyond humans.

LYL1 lymphoblastic leukemia derived sequence 1

Synonyms: bHLHa18; LYL1

CAGCCTGGCCCTTATCTGCACTGGGCCAGCATCCTCCGGCCCTGCGCCGCCAGGGGTGAGAGGGAGGAAA
CCGGGCCGCCGGGGCGGGGAGAAGGCGGCCGGGCCAGGCCACTTTCCCTGGGGGACCTACG
CGGAGACCTCGGCTATCCTGCCCTCCAGGGCCCACGAGGAGGCCAACGCCGGGCCCTGGAGCAT
TGAGGCCGGACCCTCGCGAGACAGCAGAGCCTGGCTGACGCTGGAAACCACACCCTGGCCCAGACTGCCA
GCCCTGACGGGACAGAGCCAGGGACTCACCAAGGCTGAAGAACAGTGTGGGTGAGTACCCCCACGTCG
GGGTCCATGTGCCGCCCTCAGGCACAGGCAGAGGTGGGCCACCATGACTGAGAAGGCAGAGATGGTGTG
TGCCCCCAGCCCAGGCCCTGCCCTCCCCACCCCTAACGCTGCCCTGCCCTGGGCCAGGTGGAGGAGGTGG
GCCACCGAGGAGGCTCTGCCCTCCCCCAGGCTGCCACCTGGTGTACAGTGTACGCCCTGGGCCACAGCAGG
CCCCCAGGGGTAGGCATGCCACCACAGAGCTGGCACTCTGCCCTGCCCTGCTGCAACTCTCCACCC
GGGAACGTCCCCGCCACTTGGCCCTGCACTACCAACCTCACCCCTTCTCAACAGTGTACATTGGC
CAGCAGGACCTTTAGCATCTCCCTAGCAGCCGTTGAAGCGGAGACCAAGCCACTGTGAGCTGGACCTG
GCTGAGGGCACCAGCCCCAGAAGGTGGCCGGCGTGTACCAACAGCCGGAGCGCTGGCGCAGCA
GAACGTTAACGGGCCCTCGCGAGCTGAGGAAGCTGCTGCCAGCACCAGGGACCGGAAGCTGAGCA
AGAACGAGGTGCTCGCCCTAGCCATGAAGTACATCGGCTTCTGGTGCCTGCTGCGCAGCAAGCCGCA
GCTCTGCCCGCAGGCCACCCCTCCGGCCTCGCAAACGCCGGTGACCAGGGTCCCAGACGACGGCGC
CCGCCGGGATCCGACGCAAGGCCGGAGGCCAGCGCTCGCAGCCCGGCCGGCAGCCCGACG
GCAGCCCCGGTGGAGGCCGGCCATCAAGATGGAGCAACCGCTTGAGGCCAGGGTGGTGA
CACGCGCAGCACCTCTGAGCCGGAGGGCACCAGGGACTCGGCCAGGCCGTCAAGGAAGGGCAGTGG
CGTGCCTGCATGTTGGAGCGAACTCCCCGAAGAAGGACCAAGTGAAGACGTCAAGGGCAAGGTCTGG
GGTCCGGAAGGGTGTACATGCCACCCCAAGGGACCCCGAGACCTTAAAAAAATACCCACAACCTCTG
GAAGTGGCCTGCCGGTCCCTCCAGGGCGAGGTGGCAAAGCAACATGGCAGAGCAGTCATAGGAA
AAAAAAAAAAAAAA

LAQTASPDGTEPGHSPGCKNSAGVSTPTSGSMCPQQAEVGPTMTEKAEMVCAPSPAPAPPKPASP
QVEEVGHRRGGSSPPRLPPGPVVISLGHSHRPPGVAMPTTELGTLRPLLQLSTLGTA
PPTLALHYHPFLN
SVYIGPAGPFSIFPSSRLKRPSPHCELDLAEGHQPKVARRVFTNSRERWRQQNVNGAF
AELRKLLPTHPP
DRKLSKNEVRLAMKYIGFLVRLLRDQAAALAAGPTPPGPRKRPVHRVPDDGARRGS
GRRAEAARSQPAP
PADPDGSPGGAARPIKMEQTALSPEVR

+++++

#10

NM_004936

Apparently conserved in primates only.

CDKN2B cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)

Synonyms: P15; MTS2; TP15; CDK4I; INK4B; p15INK4b; CDKN2B

GGCTCCCACCTCTGCCAGAGCGAGGCCAGTGAGGACTCCGCAGCGTCCGCACCCCTGCCAGAG
CGGCTTGAGCTCGGCTGCCGCTAGGCCTTTCCAGAACGCAATCCAGGCCGCCCTGGTTC
TTGAGGCCAGGAAAAGCCGGAGCTAACGACCGCCGCTCGGCCACTGCACGGGCCCAAGCCGAGAA
GGACGACGGGAGGTAATGAAGCTGAGCCCAGGTCTCTAGGAAGGAGAGTGCCTGGAGACGCGTGG
AAAGAAGGGAAAGAGTGTGTTAACGGCCAACGGTGGATTATCGGCCGCTGCCGTCTGGGGCT
GCCGAATCGCGAGGAGAACAGGGCATGCCAGTGGGGCGGCCAGCGATGAGGGTCTGCCAGGCCCG
GCGCGGGACTAGTGGAGAAGGTGCACAGCTCTGGAAAGCCGGCGGATCCCAACGGAGTCACCGTT
CGGGAGGCCGCGCATCCAGGTCTGATGATGGCAGCGCCCGTGGGGAGCTGCTGCTGCCACGGCG
CGGAGCCAACCTGCCAGACCCCTGCCACTCTCACCGACCGGTGCATGATGCTGCCGGAGGGCTTCTG
GACACGCTGGTGGTGTGCACCGGGCGGGCTGGACGTGCGCGATGCCCTGGGGCTGCTGCCCGT
GGACTTGGCGAGGAGCGGGGCCACCGCGACGTTGCAGGGTACCTGCGCACAGCCACGGGGACTGACGCC
AGGTTCCCCAGCCGCCACAAGCACTTATTCTACCAATTCCACCCACCCACCTAATTGATG
AAGGCTGCCAACGGGAGCGCGGAAAGCCTGTAAGCCTGCAAGCCTGCTGAGACTCACAGGAAGGAGGA

GCCGACCGGAATAACCTCCATACATTTTCTTGTCTTATCTGCCCTCGACACTCACCATGAAGCG
AAACACAGAGAAGCGGATTTCCAGGGATATTAGGAGTGTGACATTCAGGGTCGTTGCTTCAGG
GTTTCTGAGGGAAAGTCATATGAAATCCTGACTGGACCTGGCTACGAATCTCCGATGGATGAAT
CTCCCACCTCCAGCGCTGAGTGGAGAAGGCAGTGATTAGCACTGGGTGACGGCAGTCGATGCGTCACTC
CAATGTCTGCTGAGGAGTTATGGTGAACCCACAACCTAGGCCCTAGCAGAAAGGAAACCTGAAGACT
GAGGACAAAGTGGAGGAGGGCGAGGTTGGCTTCAGTAAGTCCCCGGCGCCTTAGTTGAGCGCATGG
CAAGTCACATGCGAACGACACTCTTGAAGCCTGGAGACCCCTGCCAACCTCACCAGATAGCAGAG
GGGTAAGAGAGGATGTGCAAGCAGACAGATGCTAAAATCCCTGGATCACGACGCTGCAGAGCACCTTG
CACAGGATGCTGCCCTTGTCTTACTACACTGAGGAGAGATTCCCGGGGTTCCGCAGGCAGACTACACA
GGATGAGGTGGTGGAGTGGAGTGAAGAGAATTGTAACGGTTAAGTGTAACTGTAACGTTCTTCACACACACA
CACACACACACACACATGCTAGGATGCGGAATCCCCTATGACTTGTACTTTTGATTGTGATAT
TTTGTACTTTAGTTGTCAGCAACTGTCTTATTTAATGGGAGATTAAAGTAACATAACTAGTGGCTC
TCAGTTAAATGTGAGGAAGAACTACAGCTCTTAAATGTAGCAATGGCACTGTTGCAAACACTGCAAAC
GCCTAGATTGCTTCTTAACTATTATTCTTGTAAATTCTGATTGTTAGAGTGTGACTTATAGAGTG
TCTCAGGGTGCAGAGGTCAAGAAATATTCCAATGCTTTAGAAGATAGATGCACTTATGCACTA
AATTATCTGGATAGTTCCAAAAGATTGCTGAAAAAGTAGATTGAGTATAAAACTTGAAAATATATGA
TGGCTCGTGGGATGTCCTACTATCACTGAACAAACTAAAGGTGCACTGTTGGGATTTAATTCCAGGGT
TGCTTGATCATTATATCATTGGAACAACGTGATACTTCACTACTTAAATAAGAATTAAACAGAGATTGAACT
CCAAGAGGTGGTAATTGGTTAAAATACATGTTCATGGGTTACCAACTCCTGAGAAATGTTAAA
GGTTCACAGGGTCCCTCTCAATGTTGTAATAATTGCTCATAGCAATACAGCAATTCAAAAAA
CTGCTTACTTATGCCATAGAAAATTAAACACAAAGTGTATACTGTTATGCTTCTAAATGCTCATTCTA
CCAGATACACATTAAAAGAGAAAAAGGAACAGAAACAGTCATTGAGAGTGGAGACTTAAAGAGGA
GTACATTGAGTTGAATACACAAATCTTACTTCTCATTAACCTTAAATTCCCAAATGAAACATATTACTG
GGGAAAGTTAGTTGAGAATCAGAGCATATGTTATTGGGAAAGGATATGTTATTGACACATAATCTGTAC
CAGGTATGCATTAAAATATTGTTAATTAAATTTAAACCTGAGAGATAGGTATTGTTCCAGATGA
GGACAATGAGGCAAAGAAATATCAAGTAACCTGCCAAAGGTTACAAGATATTCCATGGATGCACAAA
GAAGTGCATCTAGTCCACAGCTGATTATGGTTGTCCTGCTTCTCCATTGCAACCAGCTGCTCCTCCA
AAATCATGAATGATAACATGAAGATAACTTTTAAAAAGCAGAAATACACAATGATCTCCCTGT
AAGCTCTAAGGTGGCTTTCTTAACCTCTAGTAAATATAACGGTTGTTGAAACACTATTAA
AATGTCAACAATATGGAGAATAACCCCCCAACACACCTATAAAACCCAAATTGGAAACAAAGATAA
TGGAACCTCCATTTCAAACTGAAGCACAGGGACAGAAAATATATTCTAGTTATCACTTAAGCACTCAAT
CATTAGAGGCTACAAGAATAATTGTTAAAGTACAGTATTACAATTATTAGAAAACATTCTATATAA
AAGAAGTCAGTTGATACTTTAAAATCTCCATTGGTTATAAAATCCCTAATTGACCTCTATATCTTA
AATTCCAAGATGTTAAATTGCTAGTTGCAATTACTGGGTCACTGAAACATTATCCCTGAAATAGATAT
GAAACATGTTACTTCATTCTGGTTAAATAACTTGTGGAATCTTCTTAATGACAACCTGATATTAGGG
AAACTAAAGAAAATGTTATTGAGTCCCACAGTACTATATTACACTGTTTTTTGTTGTTAGT
TTTTTTATTAAAGCAAACCTCAAACATTATTGAGTCAATTACCACTGGGTGATTGAAATAGTAAC
TTATCAATGCCATGTTAAATTCAATTCTGAAGGCCCTGGCAGACAGGTTAGCTGTTCATCAG
CAGCCTAATATAACTGTTAAATTGTTAAGGATTCACTTGAAGGATACTGAAAACATATACTTACT
ATTTCATGAGTCCTGCTCTAGCTCCATTGTGGAATACAGAAAATTAAATACCTGTTAAGTTCGTATC
TAAACCTAAGACATTACCAAGGTTGTACAAATTCTACTACCTGACATTATTCCAAGAAGATCTGGAAAG
TTAAATAAAATTATAAAATTAAACAAAAAA

TVDYPGRCASGGGMREENKGMPSGGSDEGLASAAARGLVEKVRQLLEAGADPNGVNRFGRAIQVMMMG
SARVAELLLLHGAEPNCADPATLTRPVHDAREGFLDTLVVLHRAGARLDVRDAWGLPVDLAERGHRDV
AGYLRTATGD

+++++

#11 NM_001085386

Apparently conserved in primates only.
NFE4 transcription factor NF-E4

Synonyms: NF-E4

GACAATTCCCTGTTACGGAAGACTATAAAACCTGCCCTACTCATTGGTGTGACGCCATTTAGGCC
TTAGCCTGCCTGCACCCAGGCCTCATTAACAGCAGGTTGCTCCACACTGCCTCGTGTCTGTTGGC
ACACTCTCAAGAGTTGAACGGATAACAAGAATCTTCATCTGGTGCCTGAAACCCGGGAGGGCTCCGGTCT
TCGTCCCCCGTGGACCTACCCCTCCGCCAGAACAGCAGGCCACAGCAGCCGGACAAAGGAAGCTCCTCAG
CCTCCAGTTGCTCTCTGTGCATGACATCAGTCACTGATCTCACCTACTGGGCCCTGCAGGCCATGGG
CCACAGCTCCACACAGAACGCTCCTAGAATCCACCTCCACCTGGTGCCTGCTCAAGTGCAGGCCATGAAG
GCCACTGGCCCACACAATGCCAACCCAGGTAAACCCACGAGGCCATGCCCTCTGCAGGCCAC
TGGAAAGTTGACTGTGAGCGGCCCTGCAAGGACACCCCATCCCTTGAGCCAATCAAACCCCTCCTA
CTCGGATCTCGTCAGCCTTGGCCTGAAAGACTGATAGTGCCTTGGAACAGACACCCAGCAACTACCATCG
CTTCATCTGAGCCAAGGGTAACCTGTATGGTGGCAGGCCAGTCCAAAGGCAGGCCACTTGTGCCAGC
AGTGTGAGTCAGCAAGATAGCAGAACAGAGAGGCCAGAACAGACACCTACTGACTGGGAGACA
CGTACCCCTGAAGATTGAGAAAGAGGCCATCCAGTACCATAGCAGTTACATCAGACTGGGACATTCC
TGTTTACAGGAGACTATAAAACCCCTGCCCTACTCATTTGGTGTGACGCCATTAGGCCCTAGCCTG
CCTGCACCCAGGTACTCATTAAAACAGCGTGTGCTCCAAAAAA

LPRVVCWHTLKSLSNGYKNLSSGAETREGLRSSSPVLDPLRPRKQATAAGQRKLLSLQLLLCACTSVTDLY
WGPAGHGATAPHRSLLAIHLHLPASSAAMKATGPHNAQTQVNPRGHAPSADPTGSWTVSGPCKDHPHPF
LSQSNPPTRISSALPLKTDSEQTPQQLPSLHLSQG

+++++

#12

NM_004364

Partial conservation of the non-AUG initiation in mammals but 7 of 17 mammalian orthologs examined have either broken frames or are missing the non-AUG codon including: *Monodelphis domestica*, *Macropus eugenii*, *Oryctolagus cuniculus*, *Loxodonta Africana*, *Sus scrofa*, *Dasyurus novemcinctus*, *Pteropus vampyrus*. As a result it did not pass our criteria for conserved non-AUG initiated candidates. Nevertheless, the non-AUG initiation is actually present in frogs and fish.

Cebpa CCAAT/enhancer binding protein (C/EBP), alpha

Synonyms: CEBP; C/EBP-alpha; CEBPA

CGGAGCTCGCGGGCGCGGGCGAGCAGGGTCTCCGGTGGCGGGCGACGCCCGCGCAGGCTGGAGGC
CGCGAGGCTCGCATGCCGGAGAACTCTAACCTCCCCCATGGAGTCGCCACTTCTACGAGGCCAGCC
GCGGCCCGATGAGCAGCACCTGCAGAGCCCCCGCACGCCAGCAGGCCCTCGGCTTTCCCC
GGGGCGGGCCCCCGCGCAGCTCCGCCACCTGCCGCCGGAGCCGCTGGCGCATCTGCAGCAC
GAGACGTCATCGACATCAGGCCCTACATCGACCCGGCCCTCAACGACGAGTTCTGGCCGACCTGTT
CCAGCACAGCCGGCAGCAGGAGAAAGCCAAGGCCAGGGCGCTGGGCCCCACGGCGGGCGCGCGACT
TTGACTACCCGGCGCGCCCGGGCCCCGGCGCGTCATGCCGGGGAGCCACGGGCCCCCGCC
GGCTACGGCTGCGCGCCGCGCTACCTGGACGGCAGGCTGGAGCCCTGTACGAGCGCTGGGCGCC
GGCGCTGCGGCCGCTGGTATCAAGCAGGAGCCCCCGAGGAGGATGAAGCCAAGCAGCTGGCGCTGGCG
GCCTCTCCCTTACCAAGCCGCCGCCGCCGCCCTCGCACCCGACCCGACCCGCCCGCGCAC
CTGGCGCCCCCGCACCTGCAGTTCCAGATCGCGCACTGCGCCAGACCACCATGCACCTGCAGCCGGTCA
CCCCACGCCGCCACGCCGTGCCAGCCGACCCCGCCGCCCTCGGTGCCGCCCTGCCGG
GCCCTGGCAGCGCGCTCAAGGGCTGGCGCCGCCACCCGACCTCCCGCGAGTGGCGCAGCGCG
GGCAAGCCAAGAAGTCGGTGGACAAGAACAGCAACGAGTACCGGGTGGCGCGAGCGAACAAACATCGC
GGTGCAGCAAGAGCCGCGACAAGCCAAGCAGCGAACGTGGAGACGAGCAGAAGGTGCTGGAGCTGACCA
GTGACAATGACCGCCTGCGCAAGCGGTGGAACAGCTGAGCGCGAACCTGGACACGCTGCGGGCATCTTC
CGCCAGCTGCCAGAGAGCTCCTGGTCAAGGCCATGGCAACTGCGCTGAGGCGCGGCTGTGGGACCG
CCCTGGGCCAGCCTCCGGCGGGACCCAGGGAGTGGTTGGGTGCGGGATCTCGAGGCTTGGCCGAGCG
GTGCGAGGCCAGGACTAGGAGATTCCGGTGCCTGAAAGCCTGGCCTGCTCCCGGTGCCCCCTCC

TCTGCCCGGACTTGGTGCCTAAGATGAGGGGCCAGGCGGTGGCTCTCCCTGCAGGAGGGAGAAT
TCTTGGGCTGAGCTGGAGCCCGCACTCTAGTATTAGATAACCTTGTGCCTGGAAATGCAAAC
ACCGCTCCAATGCCTACTGAGTAGGGGAGCAAATCGTGCCTGTCATTATTTGGAGGTTCCCTGCCTC
CTTCCCAGGCTACAGCAGACCCCCATGAGAGAAGGAGGGAGCAGGCCGTGGCAGGAGGAGGGCTCAGG
GAGCTGAGATCCCACAAGCCCAGCCCCAGCCGCTCCTCACGCCTGTCCCTAGAAAGGGTGGAAAC
ATAGGGACTTGGGCTTGGAACCTAAGTTGTTCCCTAGTTCTACATGAAGGTGGAGGGTCTCTAGTTCC
ACGCCTCTCCCACCTCCCTCGCACACACCCACCCAGCCTGCTATAGGCTGGCTCCCTGGGGCGG
AACTCACTGCAGGGGTCAACCAGGTGACCAGTGGAGCCCCCACCCAGTCACACCAGAAAGCTAGGT
CGTGGGTCACTGAGGATGTATAACCCCTGGTGGAGAGGGAGACCTAGAGATCTGGCTGGGGCGGC
ATGGGGGTGAAGGGCCACTGGGACCCCTCAGCCTGTTGACTGTATGCCTCAGCATTGCTAGGAACA
CGAACGACGATCAGTCCATCCCAGAGGGACCGGAGTTATGACAAGCTTCAAATATTTGCTTATCAGC
CGATATCAACACTGTATCTGGCTCTGTGCCAGCAGTGCCCTGTGCAATGTGAATGTGCCGTCTCG
CTAAACCACCATTTATTGTTTGTGTTGCTCGGATACTGCCAAATGAGACTCTC
CGTCGGCAGCTGGGGAGGGTCTGAGACTCCCTTCCCTGGGATTACTTTGATCCTGGGG
ACCAATGAGGTGAGGGGGTTCTCCTTGCCCTCAGCTTCCCCAGCCCCCTCCGGCTGGCTGCCACAA
GGCTTGCCCCAGAGGCCCTGGCTCGGAAGGGAGGTGGCCTCCGCCAACGCATCACTGGGG
TGGGAGCAGGAAAGGACGGCTGGTCTCTTGGGAGAACGTAGAGTCTCACTAGATTTTAT
GTATTATATCTATAATAAACATATCAAAGTCAA

VRGRGRAGSPGGRRRPAQAGGRRGSPCRENSNSPME SADFYEAEPRPMSHLQSPPHAPSSAAGFPRG
AGPAQPAPPAAPEPLGGICEHETSIDISAYIDPAAFNDEFIADLFQHSRQQEKAKAAVGPTGGGGDFD
YPGAPAGPGGAVMPGAHGPAGGYGCAAAGYLDGRLEPLYERVGAPALRPLVIKQEEREDEAKQLALAGL
FPYQPPPPPPSHPHPHPPPAHLAAPHLQFQIAHCQTTMHLQPGHPTPPPTVPSPHPAPALGAAGLPGP
GSALKGLGAAHPDLRASGGAGKAKSVDKNSNEYVRERNNIAVRKSRDKAQRNVETQQKVLELTSD
NDRLRKVEQLSRELDTLRGIFRQLPESSLVKAMGNCA