

**UCC Library and UCC researchers have made this item openly available.
Please [let us know](#) how this has helped you. Thanks!**

Title	Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences
Author(s)	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 http://dx.doi.org/10.1093/nar/gkr007 Access to the full text of the published version may require a subscription.
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
Item downloaded from	http://hdl.handle.net/10468/5024

Downloaded on 2021-01-24T01:01:35Z

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

CGACGTCAAGTCGAGGCCGCCCGCGGGCCTGGTTATCGCCGGTTCAGCGCAGCCCGGAGTCGCCAGG
CCTGAACTCCTACCCAGCCTAGACTCAAGTCTGGGTTTCAGCTGCCGCCAGCCCTATTGCTGCTGTTGCTG
TTCTCTGTCCTTGGCCCAGGGGCTGGAGGCCTTTTCTGACTGATTACTCCACCTGCTCACCCGCAAGCT
GAGTCCTTTCCGCTCCTTTGCCAGCACCAGACTCCTTCCACTTCCATGTTTCTGAGGACACATTCTGGCTG
TTTGAACCTCATCATCTTCAAGGAGCAAGGGGGAACTTTGGGGACCACTGCCAGACCAAAGTGTGACT
GTGTATTTCCGGTCCGGGGCACCCCTGTCAATCCCCTGCATACACACTTCCCAGGGGACACAGCTGT
GCCTGGGGTTTTCTCACTGACCCTCAGCTGGACACTGCCAACCACCTCAGGCATCTTTAACGTCAGCA
GCCCTTACCTGGGGACTGGTTCTTGGCTGCCACCTTCCCAGGCCACGCCACATCTCTGTCAAGGGT
CTCCAGGATGAGTGTGAGTACCTCCTTCCAGCCGAGCTGATTGTCCGGCGTTTGTGACGTCGCTGTGCT
GGTTCTGGCCGGCCCTCAGAGCAAACCCTCTCCACACAATCGCTCAGCCCTGTACAAGGTCTTTGTGC
CCAGCTTCACTTACAGGGTTTTCAGCACAGCTGGTGTGTGGGGGGCCGTGGGGTATCTGCCTGCCCTG
TCACTGCGTCTGCGTCCCAAAGCCCCACCCCTGCACAACCTCAAGCTCTGTGGCCTGTGGAGGTGCCTCAGG
ATGCCAGCTGGAGCTGGCACTGCCCCCTGGGGGCACTGGGTCTACGTGCGTGTGGAAACATCATCCCGG

GCCCTGGTAGGACCATCCGCTTCCAGCTGTGTGTGCGGTTGCAAGAGTGCCACAGCCCGGCTGCTCCGA
GCCCTGGTCCCTGGAGCTGCCATGAACATGCCCCAGTCCCTGGGCAACCAGCCACTGCCCCAGAACCGCC
ATCCCTTGGAAACCCCTGCGGAGGGGCTGGGACCACGTCCCCACCCGAGCACTGCTGGCCAGTGCGCCCGA
CTCTGCGCAACGAGCTGGACACCTTCTCTGTCCACTTCTACATCTTCTTTGGCCCAAGTGTGGCCCTTCCC
CCTGAGCGCCCAGCCGTGTTGCGCATGAGGCTGTTGCCAGTGTGGACAGTGGAGGCGTCTCAGCCTGGA
GCTCCAGCTCAATGCGAGCTCCGTGCGCCAGGAAAACGTGACGGTGTGGATGCTTGACTCACGAGGTGC
CCTTGAGCCTGGGGGATGCAGCAGTACCTGTTCCAAAGAGTCCCTGGCCGGCTTCTCCTCTCTGTCACT
GCCACCACCAGGGTTGCCAGGCTGCGAATCCCATTTCCCGCAGACGGGGACCTGGTTTCTGGCCCTCCGCTC
CCTGTGCGGGGTGGGGCCTCGGTTCTGTGCGGTGCCGCAACGCGACGGCCGAGGTGCGGATGCGCACCTTCC
TGTCCCCATGCGTGGACGACTGCGGGCCCTACGGCCAGTGAAGCTGTGCGCACACACAATTATCTGTAC
GCAGCTGCGAGTGCACAGGCCGGGTGGAGAGCTGGGCTGACCCGACAGTGCAGATGCGCTCACCTATGG
ATTTCCAGTGTCTCCACACTCCTGTCTGAGGCTGAGCAACCTCATGTTTCTGCCACCTGTGGTCTGGCCA
TTCGGAGTCGATATGTGCTGGAAGCTGCAGTCTACACCTTACCATGTTCTTCTCCACGTTCTATCATGCC
TGTGACCAGCCAGGCATCGTGGTTTTCTGCATCATGGACTACGATGTGCTGCAGTTCTGTGATTTCTGGG
CTCCTTAATGTCCGTGTGGGTCACTGTATTGCCATGGCTCGTTTTACAGCCCGTGGTCAAGCAGGTGCTGT
ATTTGCTGGGAGCTATGCTGCTGTCCATGGCTCTGCAGCTTGACCGACATGGACTCTGGAACCTGCTTGGGA
CCCAGTCTCTTCGCCCTGGGGATCTTGGCCACAGCCTGGACAGTACGCAGCGTCCGCCGCCGGCACTGCTA
CCCACCCACGTGGCGCCGCTGGCTTTTTCTACTTGTGCCCTGGCAGCCTTATTGCAGGCAGTGCCGTCTGC
TTTATGCTTTTGTGGAGACCCGGGACAATACTTCTACATTACAGCATTGGCATATGCTCATTGCGGGC
AGTGTGGGCTTCTGCTGCCCCCTCGTGCCAAGACTGACCACGGGGTCCCATCTGGAGCCCCGGGCCGGGG
CTGTGGTTACCAGCTATGCATCAACGAGCAGGAGGAGCTGGGCCTCGTGGGCCAGGAGGGGCCACTGTCA
GCAGCATCTGTGCCAGCTGAGAGGGCTTTGGGCCTGGCCCTGAGGGGATATGAATGCTTCTAGAGTTCT
TTCTGGGGGTGTGGAGCCCTCTTAGAAGGAGACAGGCTGTATTTCTTGAGGACATGGAGTCTTTCTCAAGG
ACACAAAACCTTCCAGGGACCTGGAGCCCTTCCCAGGACATGGAGAATCTCTGAGGGCCTGGAGTCCCC
CTGCATCATGGAGTCTTCTTAAGGACTGGAGCCTATGCAGGCACAGAGTCCCTCAGGACCAAGGAGTCCC
TCCTGCAGGTGTGGAGCCTTCTTGGGATGCAGAGCCTTCCCAAGACATGGATTCTTCCCAGGGAGACAA
AGCCCTGTGAGGACACAGCATCTTCCAGAGGAGGTGGAGTCTATCTTGGGAAAACCAAATTTCCAGATT
TTCCCAGAGGCTCAGCAACTCTGGCCTCAGGCTTCTTCCCAGAGGCAGCGTCTGGGCTGTGCTGTGCTGT
GGAGGAGGATTGAGGATGGATGGAGCTGGGACTGGGCTGTCTGGGTGGCTGTGATCTCGTTTGTGATA
AGGTGGAGTCTGTGTCTCCAGTATTGATTGGTTGAGAATGGTTCTGTGATGCCTTTTTTTCCCCCTGGG
GTCAGGGGTGTGGGAAAGTGGGAAAGAGGGCCCTCAGGGAATCAGCAGGGCTGATGGGAGCTACTGCCG
GAGGCTTTTTGTACACCCTGTACACAGTCTGATCCCGCTCATCTGGGCCCTGCATTCAATTTCTAAACAGTT
TCTAATGCCTATTCCCCAATTCCTATTGAGCCCGATTTGCAGTATCTGAGGGGTGTGTGTGTGTGTGTG
TGTTTATGTATGTATACGTATGCTGAGATGATTTAAATCAGTGAGACAGACTTTCCTTATGATGCCACAC
CCAACACAGAAGGAAGCCGAGGTCCCAGGAAATTGGAATAGCAGGTACACGTCTCAGGTGTGCAAGAAATA
TCACAAGAATGTAACCTTCTGTTTTGAGCCCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LLLLFSVLGPGAGGLFLTDYSTCSPRKLSPFRSFASTELFHFHVPEDTFLAVWNLIIKFKEQGGTFGDHCPD
QSVTVYFRSGAPPVINPLHTHFPGDTAVPGVFSLTLSWTLNRTSGIFNVSSPLPGDWFLAAHLPQAHGHI
SVKGLQDECQYLLQPQLIVRRLLDVAVLVPGRPSEQTLSPHNRSALYKVFVPSFTYRVSAQLVCVGGRGVS
ACPLSLRLRPKAPPLHNSSSVACGGASGCQLELALPPWGHVYVVRVETSSRGPGRTRIRFQLCVRLQECQPQ
GLLRALVPGAAMNMPQSLGNQPLPPEPPSLGTPAEGPGTTPPEHCWPVRPTLRNELDTFSVHFYIFFGPS
VALPPERPAVFAMRLLPVLDSGGVLSLELQLNASSVRQENVTVFGCLTHEVPLSLGDAAVTCSKESLAGFL
LSVSATTRVARLRIPFPQTGTWFLALRSLCGVGRFVRCRNATAEVRMRTFLSPCVDDCGPYGQCKLLRTH
NYLYAACECKAGWRGWGCTDSADALTYGFQLLSTLLLCLSNLMFLPPVVLAIRSRYVLEAAVYFTMFFST
FYHACDQPGIVVFCIMDYDVLQFCDFLGLMSVWVTVIAMARLQPVVKQVLYLLGAMLLSMALQLDRHGLW
NLLGPSLFFALGILATAWTVRSVRRRHCPPTWRRWLFYLCPGSLIAGSAVLLYAFVETRDNYFYIHSIWHM
LIAGSVGFLLPPRAKTDHGVPSGARARGCGYQLCINEQEELGLVGPGGATVSSICAS

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

AGAATCGAAACTGAGAGCTCCTGGGCAGGCTCGGCAGGGCAGGCAGCTCCAGGAGGGCTTCGAACCGTGGC
CAACAGTTCCAGTGGACTGCGTGGACCCGTTGAGCTCAGGAGCCTCAGACGCCTCCCTGGAGAGCCAAGCTG
GTGTTCCGAGTTGGCGCCTCCAGGGTCCACCCTGCTGCCAACAGCCCCGCGGCCACCAGAGGGCCCTCCCT
GGCCCCGCTGTGTGCCCTGGTGGACCTGTGTCTGGGCTGCTCCCGCTGCACCCAGCGGCTCAATGAAAGCA
CCTACGTCTCCGTAGGGTGGAGCATGACTGCTCCCGCGAGATCCTGCTGGCCCCGCTTTAAGCAGGCCACC
AAGAGCAAGGTCTGGCGCTGGTGGGCTGCCGGCCACCTTCCCAAGGCCCTGTGCTACCAAGTCTGCCA
CTACTACAGCCCTGGGCTCGGCTGCCGGCGCCACCGAAACCGGTGCACCTTTGCCCGCAGTCGCGAGGAGG
CCCTGGTCTGGACCTTCGAGCGTCAGCACAACTCCAGCGCCTATGGCTGAAGGCGGAGGTGCAGGGCAGC
GGGGCCCAGGGAGGGGCGAGCCGGGCGGCCAGCCATCCTTACGGAGTTTGGCGGCCGCTTCGAGCTGCT
TTGCTCCCTCTGCTTCAGGCGCTGTCCCCATGTCATCTGTGCGTGGACCCCGAGGGGCGAGTGCCTGAGC
ACGGAGCCTGCCCCCTCCCTCCTGGCCACGTGAGCGCCGAGGGCCGCCGCAAGCAACAGTTTGTGGTGGTG
AGGCCGCGGGCCCCGGGCCAGCCTCCTGCCTACTGCAGGTTTGTGGGGCGTGGGCAGCCGTGCTGGCG
TGGGGAGTCCCCTGCCAGTTTGACACAGCGCCGTGGAGATGGCTGTGTGGGAGGCCGAGCAGCTGGGTG
GCCTCCAGCGGGGGGACCTGCTCACACCCCTGCCCTGATGGCGACGGGCGCACGGCCCCCTTGGCCAG
CCCCCTGGGGCCCAGCTGTACTGCCCGGCTGCTTGGTCACTTCCACTCTCAGGAGGCCTTCGAGAACCA
CTGCGCATCCTCGGAGCACGCACAGATGGTGGCCTTCGACCAGGCCCTGCCCTGGGAGCACCGTTCCCCAC
CCCCGAGCTTCCAAGTTCGAGCTCTGCCAAAGCTGACCTCTGTGAGTATGGGACGCCTGCACCAAG
GCACACTCAGCACAGGAGCTGCAGGAGTGGTCCGGCGCACGCAGGCTGTGGAGCTGCGGGGGCAGCCGGC
CTGGCAGGACGGGCTGGTGCCTACCAGGAGCGGCTGCTGGCCGAGTACCAGCGCAGCAGCAGTGAAGTCC
TTGTGCTGGCAGAGACCCTTGATGGAGTGCCTGTCACCTGCAACCAGCCCCTGATGTACCAGGCCAGGAG
AGGAAGACCCAGTACAGCTGGACGTTTGCCTCCTACTCTGAGGAGCCCCTGCTACACGTGGCCCTGCTGAA
GCAGGAGCCAGGAGCCGACTTCTCTCTGGTGGCTCCCGGCCCTCCCGCCAGGCCGGCTCTACGCACGGGGTG
AGCGCTTCCGTGTGCCAGCTCCACTGCCGACTTCCAGGTGGGAGTGCCTGTGCAGGCTGCCTCGTTCGGC
ACCTTTGAGCAATGGGTGGTCTTCGACTTTGGCCGCCGGCCGGTGTGCTGCTACAAAAGCTGGGGCTGCAGCT
GGGCCAGGGGCGTCGCCAGGACCCTGCAGGAATCTGGCGCTCGGCCACCCTGAGGAGATGGAGCGCTGGC
ACACTGGCAACCGCCACGTGGTGCCTGGCGTGGAGCGGACGGCCGAGCAGACGGCCCTGATGGCCAAGTAC
AAGGGCCCTGCCCTGGCCCTGGAGTTCAACCGCAGCAGCGTGGCCTCGGGCCCCATCTACCAACCAACTA
TCGGCAGAGGATGCACCAGTTTCTATGAGGAGGAGGCGGCTCAGCAGCAGCTGGTGGCCAAGCTGACCC
TGCGGGGCCAGGTGTTCTGAAGACGGCATTGCAGACGCCAGCGCTGAACATGCTCTTCGCGCCTCCGGGA
GCACTGTACGCAGAGGTCCCCGTCCCCTCCTCCCTGATGCCAGACACAGACCAGGGCTTCTGCTGGGCCG
GGCGGTGAGCACAGCCCTGGTGGCCCCCTGTACCTGCACCCGACAATACGGTGTTCGAGGTGCGGCTGGAGA
GGCGGGCCAGCTCAGAGCAGGCGCTGTGGCTGCTGCTTCCGGCCCCGCTGCTGCTTGGCCCTGGGGCTGCAG
CCTGAGGCCCGCCTGGTCTGGAGGTGCAGTTCAGATTGACCCGATGACCTTCCGCCTCTGGCACCAGGC
AGTGGACACACTGCCTGAGGAGCAGCTGGTGGTGGCCGACTTGGCCACCTGCGCCCTGCCAGACCTTGGT
CTGTCCCACCTTGGCGGTGGCAACCGCAGCAGGAGCTGGCCGTGGCGCTCATCGCGGCTGGGGCCCT
GGGATGGGAGGCGTGTCCCCCGCTACTCATCTATGGCCCCCTTGGCACCGGAAGACCTACACGCTGGC
CATGGCCTCCCTGGAGGTATCCGGAGGCCTGAAACCAAGGTGCTCATCTGCACACACACCAACAGTGGC
CCGACATCTACATCCGGGAGTATTTCCACAGCCACGTGAGCGGCGGCCACCCCGAGGCCACTCCTCTCCGT
GTGATGTACACGGACCGGCCGCTGAGCCAGACGGACCCAGTACGCTGCAGTACTGTTGCCTGACCGACGA
CCGCCAGGCTTTCGCCCCGCCACACGGGCAGAGCTGGCGCGGCACCGCGTGGTGGTACCACCACCTCCC
AGGCCCGTGAGCTCAGGGTGGCGGTGGCTTCTTCTCCACATTCTCATCGATGAGGCGGCCAGATGCTG
GAGTGGAGGCCCTCACCCGCTGGCCTATGCCTCGCACGGCACCCGCTCGTGTGGCGGGCGACCAT
GCAGGTACACCCCGGCTGTTCAGTGTGGCCAGGGCCGGGCGGCCGAGCACACGCTGCTGCACCGCCTCT
TCCTGTGCTACCAGCAGGAGACTCACGAGGTGGCGCGGCAGAGCCGCCTGGTCTTCCACGAGAATAACGC
TGCACGGACGCCATTGTGAGCTTCTATCTCGCGCACTTCTACGTGGCCAAGGGCAACCCCATCCACGCCAG
GGCAAGGTTCCGCCCCACCCCGGCACTACCCGCTCATGTTCTGCCACGTGGCGGGCAGCCAGACCGGG
ACATGTCCATGGCGTCTGGCTGAATCTGGCTGAGATTGCGCAGGTGCTCGAGAAGGTGCAGGAGGCCTAC
AACACCTGGCCCAGCTGCTGGGGCGGCCGCGAGCAGAGGTGCATCTGTGTGCTTTCCACGGTGGCCAGGT
CAGTGCCTGAGGAGGAGCTGAGGAGGCGGGACCTAGGCCAGGTGTCTGTGCGCAGTTTGTAGATCCTGC
CAGGGCGGCAGTTCGGGTGCTGGTGTGCTCAGCACTGTGCACACCTGCCAGAGCCTGCTCAGCCCTGGGGCA

CTGGCCCCTGAGTTCTTACCGACGCCCCGCGTGCTCAACACCGTCCTGACCCGCGCCCAGTCCCAGCTGGT
GGTAGTGGGGGACGCCGTGGCCCTCTGCTCCTTCGGGGCCTGCGGCAAGCTCTGGGAGAGCTTCATCCGTG
AGTGCGTGGAGCGGCACAGTGTCTGCCCCGAGGGCCTGTCCATGGAGCAGGTGAGCAGGGTGTGGCGCAG
AGACGGCGCTGGCCTCCCCGAGGCACACAGGCTGGGGCAGCGGGGAAGTGGGAGGCTGCCCCAGAGCCAGT
AGGGGACCTGGCCGAGGAGCAGGCGGCTGTGGTGACGGCCATGGTGAAGGCAGAGCCGGGAGATGAGGCTC
TGAGCCCAGCATCCCCTGACATCACGGCAACCACAGCGCAGACGGAGGCTGCGGCAGCACCAGCAGGAGAC
GCAGTGAAGGAAGACGTGGTGGCCGGGGCCTGTGCGGCAGGAGCGGCTGCTGCAGCGGGCGTGGAGTCCAC
GGAGGCTGAGGATGCAGAGGCTGACTTCTGGCCGTGGGACGGGGAGCTCAACGCTGACGACGCCATCCTAC
GGGAGCTTCTGGACGAGAGCCAGAAGGTGATGGTGACCGTCGGGGAGGACGGGCTGCTGGACACTGTCGCC
AGGCCCCAGTCCCTGCAGCAGGCCCCGGCTGTACGAGAACCTGCCCCGGCTGCGCTACGGAAGCTGCTGCA
CGCGGAGCCCCGAGCGGTACCGCCACTGCTCTTTCGTGCCAGAGACCTTCGAGCGGGCGTCCAGCCATCCCCG
TGGACGATGCCTCCTCGGGCCCCATCCAGGTGAGGGCGCCTGGACTGTGGGATGGCCTTCGCCGGGGAT
GAGGTGCTGGTGCAGCTCCTTTCGGGAGACAAGGCGCCCGAGGGGCGGCTTCGGGGCCGCGTGTGGGCGT
GCTGAAGAGGAAGAGGCACGAGCTGGCGTTTTGTGTGCCGCATGGACACGTGGGACCCGCGCATCATGGTCC
CCATCAATGGCTCCGTGACCAAGATCTTTCGTGGCCGAGCTGAAGGACCCATCGCAGGTCCCCATCTACAGC
CTCCGGAAGGGCCGGCTGCAGCGTGTGGGGCTTGAGAGGCTCACCGCCGAGGCCCCGGCACAGCCGGCTCTT
CTGGGTCCAAATTGTCTGTGGCGGCAAGGCTTCTACTACCCGCTGGGCATCGTCCGGGAGGTGCTGCCTG
AGGCCAGCACCTGGGAGCAGGGCCTCCGCATCCTCGGCCTGGAGTACAGCTTGAGGGTGCCTCCGTCGGAC
CAGGCCACCATCACCAAGGTGCTGCAGAAATACCACACGGAGCTTGGCCGGGTTGCCGGCCGCGAGAGGA
CTGCCGCGCCTTCTTGACCTTCACTGTGGACCCCCAGGGCGCCTGCAACCTCGATGATGCCCTCAGTGTCC
GAGACCTGGGTCCCAGGTGCGAGGTGGCTGTGCACATCACTGATGTGGCCAGCTTCGTGCCAGGGACGGG
GTGCTGGACGTGGAGGCGCAAGGCAGGGCGCTGCGTTCTATGCCCCGGCAGGGAGCCAGTGCCCATGCT
GCCGGCCAGCCTCTGCCAGGACGTCTCAGCCTCCTGCCTGGCCGGGACCGCCTGGCCATCTCCCTGTTCC
TCACCATGGAGAAGGCCAGTGGCCAGCTGAAGAGCCTGCGCTTTGCACCCTCCGTGGTCCAGTCTGACCGC
CAGCTGTCTACGAGGAGGCGGAGGAGGTGATCAGGCAGCACCCGGGTGCCGGCCGTGAGCTGCCGGCCCCG
CCTGGACTCCGTGGACGCCTGCGTGTGGCCGCGTGTACTTCTCTCGGCTGCTGCGCCGGCACCCGCTGC
GGTCCGACTGCTTCTATGAGCAGCCGGACGAGGACGGCACCCCTGGGCTTCCGCGCGGCCACATCATGGTG
AAGGAGTACATGATTCAAGTTAATAGGCTCGTGGCTGAGTTCCTGGTGGGACGCGATGCACGCGGACGGT
CACGCCCTGCGGTGGCAGCCAGCACCCCGCAGCAGCTCAAGGCCCTGTGTGAGAAGCATGGGAGCC
GGGTCCCCCTGTCACTGCACCTCGGCCACCACCTGCACGGCGGGGGCAGTCCCCCGACAGCGGGCTG
CACCTCCTGGCCTCCCTCTGGAAGCAGGTCCAGTTTGTGCTGCCCGCACCCAGGACTACGAACAGATGGTGGA
CTTGGTCAACACGGACGACATGCACCCATTCTGGCTCCTGCAGGCCGCGACCTCCGCAAGGCCCTGGAGC
GCTCGGCGTTTCGGCCGCTGCGCCCCGGGGCCACCAGCAGCAGGGCGGCCACTACTCGTGCAGGTGGACTGG
TACACGTGGGCCACCTCGCCCATCCGCAGGTACTGGACGTGGTGTTCAGCAGCGGCAGATCCTGCTGGCGCT
GGGCCATGGGGGCTCTGCCTACTCTGCCAGGGACATCGATGGGCTCTGCCAGGCCTTCAGCCTCCAGCACG
CACTTGGCCAGAGCTATCAGCGGCGGGCGCGCAGCCTGCACCTGGCCGTGCAGCTCAAGGCCAGCCTCTG
GACAAGCTGGGCTTCGTGGTGGACGTGGAGGCGGGCTCCCGCTGCTTCCGGCTGCTCTTCCCCAGCAACCG
GGAGACGCTGCCTGACCCCTGCCCCGTCCCTACGGCTCCCTGCAGCTGGCCGAGCACCCCCACGCCCTGG
CAGGCCGGCCGGGCTGCGGCTCCTGTGGCGGCGCCGTGTCTACTCAGCGCAGGGATCCAGCCCCGCCCTG
CCACTGCCTGGCACTGTGCCGACCCACACACCCTGGCCGTGGAGACGGCCCTGTGGAAGCAGCTGCTGGA
GCTGGTGGAGCTGCAGCGCTGGCCGGAGGCGGCTGCTCTCATCCAGGAGAAGGGCGAGGCGTCCCAGCGGC
GGGAGCTGGTGCAGGTGCAGCGGAGCCACTGTGGCCATTTCTGGAGGTGGCCCGGGAGCTGGGCAGTGGG
GACACCCTGCAGGTGCAGCTCGGCACCAGCCTGCAGCACGGCTTCTGGTACCGAGCCCTCAGCTCTGGAC
GGTGGCACCCGGCTTCAGCCTCTGCCTGGAGCACGTGGAGCGGCCCCGGAGACTGCTTCTCAGGCCGTGTGT
ACCGGGCCCCGAGGGACCGGTACCGCGACGTGGATGAGTACGCCTGCGTGTGGGAACCATTCTGCGCCCTG
GAGTCCGGCCACCGGCGCGGTTGCCGAGAATGACTCCGTCAACTTCAGCACCTGAGTGTCTCCTGGGAGGC
GTACGGACCGCCGAGGGCAGCTGCAGGGCGCCTTCGCGCTGGAGGCCGCTTCTCGAGGAGAAGCTGTG
CCGACATCAACTTCAGCTGCTGCTACCTCTGCATCCGGCTCGAGGGGCTGCCGGCTCCCACGGCCAGCCCA
CGCCCTGGGCCCAGCAGCCTCGGCCCTGGCCTGAATGTTGACCCCGGCACGTATACTGGGTGGCCACGG
GCAGACGCAGGACTGGGACCAGGAGCGCCGGGCAGACCGGCAGGAGGCTCCCAGACGGGTGCACCTCTTCG
TCCACCACATGGGCATGGAGAAGGTTCCGGAAGAGGTGCTGAGGCCGGGCACCCTGTTACCCGTTGAGCTG
CTGCCCAAGCAGCTTCTGATCTCCGCAAGGAGGAAGCCGTGCGTGGACTAGAGGAGGCGTCCCCGCTGGT
CACCAGCATCGCACTGGGCCGGCCTGTCCCGCAGCCCCTCTGCAGAGTGATCCCCAGCAGGTTCTGGAGC
GGCAGACCTACAACATCCCCGGAGGCGGCCACAAGCTGAACCCAGCCAGAAGCTGGCGGTGAGGGAGGCT
CTGGAGAAGCCTTTCACGGTCATTACGGGCCACCAGGTACAGGGAAGACGATCGTGGGCTCCACATCGT
ATTCTGGTTTCATAAATCAAACCAGGAGCAGGTGCAGCCCGGAGGCCCCCCCCGTTGGGAGAAGCGGCTGG
GGGTCCTGCATCTTGTACTGCGGCCCTCCAACAAGTCGGTGGATGTCCTGGCAGGACTGCTCCTGAGA

AGGATGGAGCTGAAGCCCCTCCGTGTGTACAGTGAGCAGGCTGAGGCCAGCGAGTTCCCAGTGCCGCGTGT
GGGCAGCAGGAAGCTGCTCAGGAAGAGCCCCGGGAGGGGAGGCCGAACCAGAGCCTCAGGAGCATCACCC
TGCACCACCGGATCCGGCAGGCCCCCAACCTTACTCGTCGAAATCAAGGCCTTTGACACCCGGCTGCAG
AGAGGGGAGCTCTTCTCCAGGGAGGACCTGGTCTGGTACAAGAAGGTCTTGTGGGAGGCTCGGAAGTTTGA
GCTGGACCGGCATGAGGTCATCCTCTGCACCTGCTCCTGTGCAGCCTCTGCCAGCCTCAAAATCCTGGACG
TGAGGCAGATCCTTGTGACGAGGCAGGCATGGCCACGGAACCTGAAACCCTCATCCCCCTGGTGCAGTTC
CCACAGGCCGAGAAGGTGGTTCTTCTCGGAGACCACAAGCAGCTGCGGCCTGTGGTCAAGAATGAGCGGCT
GCAAAACCTGGGTCTGGACCGGTCTCTGTTTCGAGCGGTACCACGAGGACGCACATATGCTGGACACTCAGT
ACCGCATGCATGAGGGCATCTGTGCCTTCCCCTCTGTGGCGTCTACAAGAGCAAGCTGAAGACGTGGCAG
GGCCTGAGGAGGCCCGCCAGTGTCTGGGCCACGCTGGCAAGGAGAGCTGCCCTGTATCTTTGGCCACGT
GCAGGGCCACGAGCGGAGCCTGTGGTGTCCACGGACGAAGGGAATGAGAACTCCAAGGCCAACCTGGAGG
AGGTGGCTGAGGTGGTCCGTATCACCAGCAGCTGACCTGGGGAGGACCGTAGAGCCCCAGGACATCGCC
GTCCTCACGCCCTACAACGCGCAGGCCTCTGAGATCAGCAAGGCCCTTCGGCGAGAGGGCATCGCCGGGGT
GGCCGTGTCTCCATCACCAGAGCCAGGGGAGCGAGTGGCGCTATGTGCTGGTGTGAGCACCGTCCGCACCT
GTGCCAAGAGCGACCTGGACCAGCGGCCACCAAGAGCTGGCTCAAGAAGTTTTCTGGGCTTCGTTGTGGAC
CCCAACCAAGTGAATGTGGCTGTACGCGGGGCCAGGAGGGGCTCTGCCTGATCGGAGACCACCTCCTTCT
GCGCTGTGCCCTCTGGCGTAGCCTCCTGGACTTCTGCGAGGCTCAGCAGACCCTCGTGCCTGCCGGCC
AGGTGCGCGTCTGCAGGAGGCCAACTATGCCTTCTGAGAGCCCTCTCCACCTGCAAGGTGCCAGGACTG
GGAGGGAAAGTCCAGGGCCCCCACCCTCCCCACCACCTCCGCCATCAGCCTGCCTGGGCCAGCCTGCC
TCCTGGCCTCAGGATCTCCCTGGGGCAGGAAGTGGGTGGACTTGGAGGAGGAACGGCAAGGGCCGGTCCCC
ACACTCTTGCTCCTCCTGGGCATCTGGGGACAGGAACACCTCTCAGGGTGGCCAAGAGGGGCTTCTTGGG
ACCAGTGAAGGAAGTTCTGCTGTCCCTGGGAAGTCTCAGAACCAGGAAGGAAACCCTGTGGAGGCTGCC
TTTGACCTTTGGACGCCACCTGTAATCACTCTGGATGGTGTCTGTGGAGTTCAGGTTTTATTTTCTATTTAA
ATGAACTAAAAGCCAGCATCCTGCCGGGCGTGGTGGCTCACGCCATAATCCAGCACTTTGGGAAGCCGAG
GCGGGCAGATCACCTGAGGTGAGGATTCGAGACCATCCTGGCCAACATGGTGAACCCCGTCTCCACTAA
AAATACAAAATTAGCTGGGCGTGGTGGCGGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGA
ATCGCTTGACTCCAGGAGGCGCAGGCTGCAGTGAGCTCTGATGGTACCACTGCACTCCAGCCTGGGTGACA
GAGCGAGACCCTGTCTTAAAAAAAAAAAAACAAAACAAAACCAGCATCTGTTACTGGAACAGAGACCTCAGCC
CAAGCTCAGGACAAGGAGCCTCCCTGGGAGAGGGGGCTTTCTCCACCATCACCCTGTGTCTTCCAG
GGGCTGTGGAGGGCAGCCCAAAGCCAGGGGCCACCCCGCCACCCACCCACCCACCCCTCTGAAAT
GTGAAAAGCCTGCACTCTTCTGTGGGCTGCGCAGTGGGCTCGGGGGTGGGGGGTGCCGGGGCGATGCTT
CCACCCTGGACTGAGTGTTTCGGGGCAGCCTGCTCGGCTGTATGTTTCTCGCACGCTCACCCGCTTCAGTGC
GGCTCTGTCCCTGCTGCCGCTCCAGTCTGAGGGAGGGGCTTGGCCACCTTCTACATTCCAATTTTTTAT
ATCTTTGAATTATGTGATTAGATATTAATTTAATGATAAAACCCTCTGAAAGCTCTTCTCA

VAPPGSTLLPNSPAATRGP SLARL CALVDLCLGCSRCTQRLNESTYVLRVVEHDCSREILLARFKQATKSK
VWRVVGCRPTFPRPLCYQVCHYSPGLGRRHRNRCTFARSREEALVWTFERQHNLRQLWLKAEVQSGAQ
GGAGRAADAILTEFGGRFELLSLFCFRPCPCICRVDPQGQCPEHGACPSLLAHVSAEGRKQQFVVVRPR
PRAGQPPAYCRFVGRGQPCWRGESRCQFAHSAVEMAVWEAEQLGGLQRGDLLTPPAPDGDGRTAPLQPPG
AQLYCPACLVTCHSQEAFENHCASSEHAQMVAFDQALPWEHRSPPPGLSKFELCPKPDLCYEGDACTKAHS
AQELQEWVRRTQAVELRQAAWQDGLVPYQERLLAEYQRSSEVLVLAETLDGVRVTCNQPLMYQAQERKT
QYSWTFVAHSEEPLLHVALLKQEPGADFSLVAPGLPPGRLYARGERFRVPSSTADFQVGVVRVQAASFQTFE
QWVVFDFGRRPVLLQKLGLQLGQRRPGPCRNALALGHPEEMERWHTGNRHVVPVGVERTAEQTALMAKYKGP
ALALEFNRSSVASGPI SPTNYRQRMHQFLYEEEAQQQLVAKLTLRGQVFLKTALQTPALNMLFAPPALY
AEVVPVSSLMPDTDQGFLLGRAVSTALVAPVPAPDNTVFEVRLERRASSEQALWLLLPARCCLALGLQPEA
RLVLEVQFQIDPMTFRLWHQAVDTLPEEQLVVLDLPTCALPRPWSVPLRRGNRKQELAVALIAGWPGDG
RRVPLLIIYGFPGTGKTYTLAMASLEVI RRPETKVLICTHTNSAADIYIREYFHSVSGGHPEATPLRVMI
TDRPLSQTDPVTLQYCLTDDRQAFRPPTRAELARHRVVTTTSQARELRVPVGGFFSHILIDEAAQMLECE
ALTPLAYASHGTRLVLAGDHMQVT PRLFSVARARAAEHTLLHRLFLCYQQETHEVARQSRLVFHENYRCTD
AIVSFI SRHFYVAKGNP IHARGKVP PHPRHYPLMFCHVAGSPDRDMSMASWLNLAIEIAQVVEKQVEAYNTW
PSCWGGREQRICV VSHGAQVSALRQELRRRDLGQVSVGSFEILPGRQFRVVVLSTVHTCQSLSPGALAP
EFFTDARVLNTVLTRAQSQLVVVGVDAVALCSFGACGLWESFIRECVERHSVCPEGLSMEQVEQGVARRR
WPPRGTQAGAAGNWEAAPEPVGD LAEEQA AVVTAMVKAEPGDEALSPASRDITATTAQTEAAAAPAGDAVK
EDVVPGACAAGAAAAGVESTEAEADAEADFWPDGELNADDA I LRELLDESQKVMVTVGEDGLLDTVARPE
SLQQARLYENLPPAALRKL LHAEPERYRHCSFVPETFERASAIPLDDASSGPIQVRGRLCDGMFAFAGDEVL
VQLLSGDKAPEGRLLRGRV L GVLKRKRHELA FVCRM DTWDPRIMVPI NGSVTKIFVAELKDPSQVPIYSLRK
GRLQRVGLERLTAEARHSRLFVWQIVLWRQGFYIPLGIVREVLPEASTWEQGLRILGLEYSLRVPPSDQAT

ITKVLQKYHTELGRVAGREDCRFLTFTVDPQACNLDDALSVRDLGPRCEVAVHITDVASFVPRDGVLD
VEARRQGAIFYAPGREPVMPPLPASLQCQDVLSLLPGRDRLAISLFLTMEKASGQLKSLRFAPSVVQSDRQLS
YEEAEVIRQHPGAGRELPARLDSVDACVVAACYFSRLLRRHRLRSDCFYEQPDEDGTLGFRAAHIMVKEY
MIQFNRLVAEFLVGSSECTRTVTPLRWQPAPRSQQLKALCEKHGDRVPLSLHLGHHLHGSGGSPDTRLHLL
ASLWKQVQFAARTQDYEQMVLDLVTDDMHPFLAPAGRDLRKALERSAFGRFCARGHQQQGGHYSLQVDWYTW
ATSPIRRYLDVVLQRQILLALGHGGSAYSARDIDGLCQAFSLQHALAQSYQRRARSLHLAVQLKAQPLDKL
GFVVDVEAGSRCFRLFPNRETLDPDPCVPYQSLQLAEHPHALAGRPLRLLWRRRVYSAQGSPPPLPLP
GTVPDPHTLAVETALWKQLLELVELQRWPEAAALI QEKGEASQRREL VQVQRSHCGHFLEVARELGS GDTL
QVQLGTSLQHGFLVPSQQLWTVAPGFSLCLEHVERPGDCFSGRVYRAPRDYRDVDEYACVWEPFCALESA
TGAVAENDSVTLQHLVSVWEASRTPQGQLQGAFLRLEAAFLAENCADINFSCCYLCIRLEGLPAPTASPRPG
PSSLGPGLNVDPGTYTVAHGQTQDWDQERRADRQEAPRRVHLFVHHMGMEKVPEEVLRPGLTFTVELLPK
QLPDLRKEEAVRGTLEEASPLVTSIALGRPVQPPLCRVIPSFLERQTYNIPGGRHKLNPSQNVAVREALEK
PFTVIQGGPGTGKTIIVGLHIVFWFHKSNQEQQVPGGPPRGEKRLGGPCILYCGPSNKSVDVLAGLLLRME
LKPLRVYSEQAEESEFPVPRVGRKLLRKS PREGRPNQSLRSITLHHRIRQAPNPYSSEIKAFDTRLQRGE
LFSREDLVVYKKVLWEARKFELDRHEVILCTCSAASASLKILDVRQILVDEAGMATEPETLIPLVQFPQA
EKVVLLGDHKQLRPVVKNERLQNLGLDRSLFERYHEDAHMLDTQYRMHEGICAFPSVAFYKSKLKTWQGLR
RPPSVLGHAGKESCPVIFGHVQGHESLLVSTDEGNENSKANLEEVAEVVRITKQLTLGRTVEPQDIAVLT
PYNAQASEISKALRREGIAGVAVSSITKSQGEWRYVLVSTVTRTCAKSDLDQRPTKSWLKKFLGFVVDPNQ
VNVAVTRAQEGCLIGDHLRLCCPLWRSLLDFCEAQOTLVPAGQVRVCRRTMPS

AUG initiation = 294.6 kDa
GUG initiation = 322.3 kDa

+++++

#3

NM_00101858

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

TC TAGCGAGGTGACAGCGTAGAACAGGTGCGCGTCCCCGGCGTTGGCGTCTTCGTCCTGTTGCTGGTCTC
CGTCCGGTGC CGCGCCGTCTAGGTCTCCGCCCCTCCCCAGCCGCTCCTGCGCCCTTGCCGGCCCCCGCC
CGCAGCC CTGGCGCTCCCTGCGGGCCCCGCCAGGCCGCTGCGCCCTGTGCCAGCGCGCCCCGGGAAC
CGGTGCGCGCCGACTGCGGCCACCGCTTCTGTGCGGCGTGCCTGGTGCCTTCTGGGCCGAGGAGGACGGG
CCCTTCCCGTGC CCGGAGTGC CCGACGACTGCTGGCAGCGCGCCGTGGAGCCCGCAGGCCCCCGCTCAG
CCGCCGCTTCTGGCGCTCGAGGAGGCGGCCGCGCGCCCGCGCGACGGCCCGCCAGCGAGGCCGCGC
TGCAGCTGCTGTGCCGCGCCGACGCCGCCGCTCTGCGCCGCTGCCGT ATGGCTGCGGGCCCCGAGCCG
CCCCGAGTGGGAACCGCGCTGGAGGAAGGCGCTGCGCGGCAAGGAGAACAAGGGGTCTGTGGAAATCATGAG
AAAGGACTTGAATGACGCCCGGGACCTGCATGGCCAGGCAGAGTCAGCAGCTGCAGTGTGGAAGGACACG
TGATGGACCGTAGGAAGAAGGCACTGACCGACTACAAGAAGCTGCGGGCCTTCTTTGTGGAGGAGGAGGAG
CATTTCTGCAGGAGGCTGAGAAGGAGGAGGGGCTCCCTGAGGACGAGCTGGCTGACCCCACTGAGCGGTT
CAGGTCACTGCTGCAGGCGGTCTCGGAGCTGGAGAAGAAGCATCGCAACCTGGGCCTCAGCATGCTGCTGC
AGTCA TGGCGCCAACCCGTGGCAGTCCCAGAGCTGGAGGCAGGAGGATGGATCCTCATCTCCATGGGAAGT
GTCAGCGTGTGGCTGCCAGGGAAGCGTGGCAGGCGCCTGGCCTTGGGTCCATCTACATAGTTGCGTGTTC
ACAATGTCCATTTATCCTTACCCCGAGGCGTGT TTTGGGGGCTGCAAACACCTCCCGGTAGAGGCTGGA
CCTGAGGACCCTTCCCACCTGTGCCCCGCTCCCTTCTGAAGTCTTAGCCACAGCCCATCCTCCATGAGTCCC
GGCAGCTCTGGGT CATGCCCTTCCCTGGTCAACCATCTGCCCCACCTCGTCATCCAGGGACCCAGACCC
TGCACCTTCCATGTGGGCCACAGATCCTTGGCAGGTACCTGAGGTGCACCATTTGAGTGTGGGATTTGGG
TTAGCATCCAGAAAGAAGATGCGCATGACGCTGTGTAAGGCTGGAAGTCAAGTCTTTCAGGGAGAGAAAG
GAAGACTGGATTGCACCTTGTATGCCTCCTGAGGAGCGGCCCCCTCTTGGAGTGGGCGTGGGCCCGGCC
AGCCTTATCCAAGTCGCTCTGTCCACCTCCCCCTTCTGGCCCCACCCACTCCTGTGCCTCCCAGGAGC

CCTCCCTGTGCTCCACCTGCCTCCGCAGAAGGAAGCCTCTTTCTCTGTTTCCCTGGGTGAGGGGGCTGGCA
GGTGGCTAACCCCATTTAGCATCTCCAGGCCCTGCCATCGTGTCTCATCTTGCTGTTATCTCTAGCTCTTT
CCCTCCTCCCATTTCTTTAGTAGTTGAATTTTGCAAAGCTTGTAGCAGTAGCTCAGTTGCCTGCAGCATC
CTTGTGTGTAGATAAATTAGTCGACAGAACTCAGCACTGGGGACAGGATTGCAAAGTCGGGGACATAGAT
GCAGACAGTTGTTGAGATTTGGGGATAGCCGGGCTTGTGAGCGGTGCCATTTCCAGATGAAGCCTTTCAG
CCCTTCTGAGTCCCCGGCCCTTGGTGCGATGTCTGTGAGTTTGACCTGCCAGCGTGTGGGCTGGCTCAAT
GCTGAATAAAGTGGGTTTGTGTGAGCTCGTTTGTCTCCGTGTGTCCACCTGGCCTCTTCCCCCTGC
CCTGGCCACCCTCCAGTGTCAAAGGAACTTCTCGTGACACGTGCTAAAGCATGGTGAGGAGGACTTTGA
TTGGGACCATTGAGATGGGTGTGGGACCCTTTCCTTGGGGCCTGGGGGGAGATGGGGCTCCACCCGACGT
AGCAGGGCAGGGGTTGGAGGAGCGAGGAGCAGTATAGGGTCCATGGGTGGGAATGACTGTGAGGAGACATC
AGGGCTGAGGGGCTCTGGCTAAACCCACCTCACAGATCCTTGTGCAGGCAGGCAGGCGATCAGACAT
TGGCTGCAAACGGTCAGAGAGGAACCCAGTCAGGTACCATTGAGGGTGGTCAGATATTATGGTTAACAAA
TTAGGGTTCTTGCTAAAAGTGGATTTTATAAGAAAGGGCAAAGAGGGCCCTAGGAGAAGATTCCAGAGCCT
GGCCAGAGTTTGGCCAAGTAGAGAATCTTTGTGACGACGCAACAACATCCCGACCCTGAGACCTCCAGTT
TGTCTTTCTCACTGTCTCCGCCTGCTGCAGTCTGCTGTGATCCCTGAGCATCCCTGCCCTGCCCTGCACA
CCTGTGATGCTTGGCCGACAGGTCCTGATGGCAGAGTCTCCACAACATCAGTGTCTCCACATCACCAGG
TCCGACAGTGGCTTACCATCCTCACCTAACCTAGCTGACCAGCAACATCCACCCCTGTCAATCACAACCT
CTTTCTATTTAAGAAAATTATATATTTATGGGGCACAGTG
CGTCCTGTTGCTGGTCTCCGTCCGGTCGCCGGCCGTCTAGGTCTCCGGCCCTCCCCAGCC

LALPAGPAEAACALCQRAPREPVRADCGHRFCRACVVRFWAEDGPFPCPECADDCWQRAVEPGRPPLSRR
LLALEEAAAAPARDGPASEAALQLLCRADAGPLCAACRMAAGPEPPEWEPRWRKALRGKENKGSVEIMRKD
LNDARDLHGQAESAAAVVWKGHVMDRRKKALTDYKKLRAFFVEEEHFLQEAKEEGLPEDELADPTERFRS
LLQAVSELEKKHRNLGLSMLLQ

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

CTCGGGCGCGCTGGCCCTGGGGACGCCGAGGGCGGCTGCGACGCGCCGAGAGGGCCGCGGCTCTCCCACCT
GTCACCCTGGCCCTTCTCTGCTTGGATGGTGTCTTCTCTCCTCAGCCGAGAATGACTTCGTCCACCGGAT
CCAGGAGGAAGTGGACCGCTTTCTGCTGCAGAAGCAGCTGTCAAAGGTTCTTCTTTTCCCCCACTCTCCA
GTCGCCTCCGGTACCTGATCCATAGAACAGCAGAGAATTTGATCTCTTGAGCAGCTTCTCCGTTGGGGAG
GGCTGGAAGAGGAGGACGGTCATCTGTACCAGGACATCAGGGTACCCAGTTCCGATGGCCTCTCTGGCCC
CTGCCGCGCTCCTGCCTCCTGCCCCAGCAGGTACCAGGTCTCGGCCATCTCCAACCAAGGAGCAGCTG
CGGTTCCCCGAGGTGCCCGGCTGGCCGGTGGTATCGTGACGCAAGCCTGACCAGCCTTTGTATGTGCC
CGGGTGTGCGCAGGCAGGAAGAATGGGGCTGACCTCTACCTCGGTGCTCAAGAGAGAGGCCCCAGCTGG
CAGGGACCCAGAAGAGCCTGGAGATGTTGGTGTGGAGACCCCAACTCTGATCAGGGACTCCCTGTGCTGA
TGACTCAGGGAACAGAGGACCTAAAGGGCCCAGGACAAAGGTGTGAGAATGAGCCACTGCTGGACCCTGTT
GGCCCTGAGCCTCTGGGGCCTGAGAGTCAGTCAGGGAAGGGAGACATGGTGGAGATGGCCACACGGTTTGG
GTCCACCCTGCAGCTAGACCTGGAAAAGGGGAAGGAGAGTCTGTTGGAGAAGAGGCTGGTGGCAGAGGAGG
AAGAGGACGAAGAGGAGGTGGAAGAGGATGGCCCCAGCAGCTGCTCGGAGGACGATTACAGTGAGCTGCTG
CAGGAGATCACAGACAACCTGACGAAGAAGGAGATTCAGATAGAGAAGATCCATTTGGACACATCCTCCTT
CGTGGAGGAGCTGCCTGGAGAGAAGGACCTTCCCCACGTGGTAGAGATCTATGACTTTGAACCAGCAGCTCA
AGACGGAGGACCTGCTGGCAACGTTTTCTGAGTTCCAAGAGAAGGGGTTCAGGATTCAGTGGGTGGATGAT
ACTCACGCACTCGGCATCTTTCCCTGCCTGGCCTCAGCTGCGGAAGCCCTGACCCGGGAGTTCTCGGTGCT

CAAGATCCGGCCCCCTCACACAGGGAACCAAGCAGTCAAAGCTCAAAGCCTTGCAGAGGCCAAAACTCCTGC
GTCTGGTGAAGGAGAGGCCACAGACAAATGCGACTGTGGCCCGGGCTGGTGGCCCGGGCCCTGGGACTC
CAACACAAAAAGAAAGAGCGGCCTGCTGTCCGGGGTCCGCTGCCGCCCTGAGGCCTGGAGACCCAACTGGC
CTGGATCTGCGTCCCAGCTAGCTGGCGCCCCAACACCATAAGCCTTACAGACGCCAGAGCAGCCCCGC
ACCACCCTCGAGCTTACCATGGGGTGTGGTGGGCTTTAGTTTTAGTCCAGAAATGGAGAAAAATAAAAA
CTCACGTTGTTCTAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LALLCLDGVFLSSAENDFVHRIQEELDRFLLQQLSKVLLFPPLSSRLRYLIHRTAENFDLLSSFSVGEW
KRRTVICHQDIRVPSSDGLSGPCRAPASCPSRYHGPRPISNOGAAVPRGARAGRWRGRKPDQPLYVPRV
LRRQEEWGLTSTSVLKREAPAGRDPEEPGDVAGDPNSDQGLPVLMTQGTEDLKGPGQRCENEPLLDVPGP
EPLGPESQSGKDMVEMATRFGSTLQLDLEKKGESLLEKRLVAEEEEDEEEVEEDGPPSSCEDDYSELLQE
ITDNLTKKEIQIEKIHLDTSFVEELPGEKDLAHVVEIYDFEPALKTEDLLATFSEFQEKGFRIQWVDDTH
ALGIFPCLASAAEALTRFVSVLKIRPLTQGTQSKLQALQRPKLLRLVKERPQTNATVARRLVARALGLQH
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

TCACGCGTCCGCTGCTCGTGTGAGTGAAGAAAATCCACCGGCATCGCCTGAGCCCCGCTACCGAGAAGGGCG
CCGCTTCCTCCGGGGAGGGGATAAAGATCCCCCGCCCGGCCATGAGGATATTGCCGTGAAAGGCACA
GCGACTGCAGCAGGAACCGGACCCGGCACCAGGAGCGGCGGGCGGCAGCAGCGGTACCGCCTCCTCA
CCCGGGCGGGCAGCAGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGCAGCGGTCCCCCTCCTCACC
CGAACATCAGGGCCCTCCAGACTCAGGCGCCCCAACAAATTCCTAGAGGACCTGTGCAACAACCTCTT
GAGGATCGAATCTTCACTCCCGCTGTCTCAGCAGTCTACAGCAGGTAACACAAGTGGCAAGACAGCCGG
AACCCCTACCCCATCCCCTTATTTCAGCACATGAAATAAACAAGGGGCATCCAAATCTTGCGGCAACG
CCCCCGGGGACATGACATCGTCCCCTGGACTCTCTCAAACCCCTTATCCCTCTGGACAGAATGCAGGT
CCAACACGCTGTATACCCCTCAAACCCCTCAGACAATGAATTCACAACCTCAAACCCGTTCTCCGTT
TTTTCCAGAGGCCTCA AATACAGCCTCCTAGAGCTACCATCCCGAACAGCAGTCCCTTCCATTTCGT
CCTGGTGCACAGACACCCACTGCAGTGTACCAGGCTAATCAGCACATCATGATGGTTAACCATCTG
CCCATGCCGTACCCAGTGCACAGGGG CCTCAGTACTGTATACCACAGTACCGTCATAGTGGCCCT
CCTTATGTTGGGCCCCCAACAATATCCAGT TCAACCACCGGGGCCAGGTCTTTTTATCCTGG
ACCAGGACCTGGGGACTTCCCCAATGCTTATGGAACGC CTTTTTACCCAAGTCAAGCCGGT
GTATCAGTCAGCACCTATCATAGTGCCTACGCAGCAACAGCCGCCTCCA GCCAAGAGAGAG
AAAAAACTATAAGAATTCGGGATCCAAACCAGGGAGGTAAAGACATAACAGAGGAGAT TAT
GTCTGGAGGTGGCAGCAGAAATCCTACTCCACCATAGGAAGACCCACGTCCACACCTACTCCT
CCTC AGCAGCTGCCAGCCAGGTCCCCGAGCACAGCCCTGTGGTTTATGGGACTGTGGAGAGCG
CTCATCTTGCT GCCAGCACCCCTGTCACTGCAGCTAGCGACCAGAAGCAAGAGGAGAAGCC
AAAAACAGATCCAGTGTAAA GTCTCCTTCCCCAGTCCCTTAGGCTAGTCCCTCAGTGGAGAG
AAGAAAGAAACAAGAAGGCCAGACATCTGAAA CTACTGCAATAGTATCCATAGCAGAGCT
TCCCTCTGCCTCCATCACCTACCACTGTTTCTTCTGTTGCTCGA AGTACAATTGCAGCCCC
CACCTCTTCTGCTCTTAGTAGCCAACCAATATTCACCACTGCTATAGATGACAG ATGTGAA
CTCTCATCCCCAAGAGAAGACACAATTCCTATACCCAGCCTCACATCTTGACAGAAACATCAG
ACCCTTTACCAACAAATGAAAATGATGATGATATATGCAAGAAACCCTGTAGTGTAGCACCT
AATGATATT CCACTGGTTTTCTAGTACTAACCTAATTAATGAAATAAATGGAGTTAGCG
AAAAATATCAGCCACGGAGAG CATTGTGAAATAGTAAAACAGGAAGTATTGCCATTGACT
CTTGAATTGGAGATTCTCGAAAATCCCCAG AAGAAATGAAACTGGAGTGTATCCCAGCT
CCCATCACCCCTCCACAGTTCCTTCTTCTCCAACCTCCT CCAACTCCTCCAGCTTCTCCT
CCTCACACTCCAGTCATTGTTCTGCTGCTGCCACTACTGTTAGTTCTCC GAGTGTGCCAT
CACAGTCCAGAGAGTCCTAGAGGAGGACGAGAGCATAAGAACTTGCCTTAGTGAAGATG
CAAAAGAGATTCAGAACAAAATAGAGGTAGAAGCAGATGGGCAACAGAAGAGATTTTGGATT
CTCAAAAC

TTAAATTCAAGAAGGAGCCCTGTCCCAGCTCAAATAGCTATAACTGTACCAAAGACATGGAAGAAACCAA
AGATCGGACCCGAACCACTGAAGAGATGTTAGAGGCAGAATTGGAGCTTAAAGCTGAAGAGGAGCTTTCCA
TTGACAAAGTACTTGAATCTGAACAAGATAAAATGAGCCAGGGGTTTCATCCTGAAAGAGACCCCTCTGAC
CTAAAAAAGTGAAGCTGTGGAAGAAAATGGAGAAGAAGCTGAGCCAGTACGTAATGGTGCTGAGAGTGT
TTCTGAGGGTGAAGGAATAGATGCTAATTCAGGCTCCACAGATAGTTCTGGTGATGGGGTTACATTTCCAT
TTAAACCAGAATCCTGGAAGCCTACTGATACTGAAGGTAAGAAGCAGTATGACAGGGAGTTTCTGCTGGAC
TTCCAGTTCATGCCTGCCTGTATACAAAAACCAGAGGGCCTGCCTCCTATCAGTGATGTGGTTCTTGACAA
GATCAACCAACCCAAATTGCCAATGCGAACTCTGGATCCTCGAATTTTGCCTCGAGGACCAGACTTTACAC
CAGCCTTTGCTGATTTTGGAAAGGCAGACACCTGGTGGAAAGAGGCGTACCTTTGTTGAATGTTGGGTACGA
AGATCTCAACCTGGCCAAAGAAGAGAACCAGAAAGATCATCACAGTTTCTGTAAAAGAAGATGTACACCT
GAAAAAGGCAGAAAATGCCTGGAAGCCAAGCCAAAAACGAGACAGCCAAGCCGATGATCCCGAAAAACATTA
AAACCCAGGAGCTTTTGTAGAAAAGTTTGAAGTATCTTAAATAAATTGACACCACAGATGTTCAATCAACTG
ATGAAGCAAGTGTGAGGACTTACTGTTGACACAGAGGAGCGGCTGAAAGGAGTTATTGACCTGGTCTTTGA
GAAGGCTATTGATGAACCCAGTTTCTCTGTGGCTTACGCAAAACATGTGTGATGTCTAGTAACGCTGAAAG
TACCCATGGCAGACAAGCCTGGTAACACAGTGAATTTCCGGAAGCTGCTACTGAACCGTTGCCAGAAGGAG
TTTGAAAAAGATAAAGCAGATGATGATGTCTTTGAGAAGAAGCAGAAAAGAACTTGAGGCTGCCAGTGCTCC
AGAGGAGAGGACAAGGCTTCATGATGAACTGGAAGAAGCCAAGGACAAAGCCCGGCGGAGATCCATTGGCA
ACATCAAGTTTATTGGAGAACTCTTTAAACTCAAATGCTGACTGAAGCCATCATGCATGACTGTGTGGTG
AAGCTGCTAAAGAACCATGATGAAGAATCCCTGGAGTGCCTGTGTGCGCTGCTCACCACCATTGGCAAAGA
CTTGGACTTTGAAAAAGCAAAGCCACGTATGGACCAGTACTTTAATCAGATGGAGAAAATTGTGAAAGAAA
GAAAAACCTCATCTAGGATTTCGGTTTCTGCTTCAAGATGTTATAGACCTAAGGCTGTGCAATTGGGTATCT
CGAAGAGCAGATCAAGGGCCTAAAACATCGAACAGATTACAAAGAGGCTAAAATAGAAGAACAAGAAGA
GCAAAGGAAGGTCCAGCAACTCATGACCAAAGAGAAGAGAAGACCAGGTGTCCAGAGAGTGGACGAAGGTG
GGTGGAACTGTACAAGGGGCAAGAACAGTCCGGTACTGGACCCCTCAAATTCCTAAAAATCACTAAG
CCTACAATTGATGAAAAAATTCAGCTGGTACCTAAAGCACAGCTAGGCAGCTGGGGAAAAGGCAGCAGTGG
TGGAGCAAAGGCAAGTGAGACTGATGCCTTACGGTCAAGTGCTTCCAGTTTAAACAGATTCTCTGCCCTGC
AACCTCCAGCACCCCTCAGGGTCCACGCCATCCACGCCTGTAGAGTTTGATTCCCGAAGGACCTTAACTAGT
CGTGGAAAGTATGGGCAGGGAGAAGAATGACAAGCCCTTCCATCTGCAACAGCTCGGCCAAATACTTTTCA
GAGGGTGGCAGCAAGACCTGCTAGACAATCAGTCTCAAGAAGAGCAGCGGAGAGATGCTGGAGA
CCGTGAAGCAGCTCACAGGAGGTGTGGATGTGGAGAGGAACAGCACTGAGGCTGAGCGAAATAAAAACAGG
GAGTCAGCAAAACCAGAAAATTTAGCAATGTGAGCTCATGACAAGGCTGCATTATCAGAAGAGGAACTGGA
GAGGAAGTCGAAATCTATCATTGATGAATTTCTACACATTAATGATTTTAAAGGAAGCCATGCAGTGTGTGG
AAGAGCTGAATGCCAGGGCCTACTACATGTTTTTGTGAGAGTGGGAGTGGAGTCCACCCTGGAAAGGAGC
CAGATCACCAGGGATCACATGGGCCAATTACTCTATCAGCTGGTACAGTCAGAAAAACTCAGCAAAACAGGA
CTTTTTCAAAGTTTTTTCAGAAAATTTGGAATTGGCAGATGACATGGCCATTGATATTCCCCATATTTGGT
TGTACCTTGCTGAACTGGTGACCCCATGTTAAAAGAAGGTGGAATCTCCATGAGAGAACTTACCATAGAA
TTTAGCAAACCTTTACTTCTGTTGGAAGAGCTGGGGTCTTGCTATCTGAAATATTGCACCTACTATGCAA
ACAAATGAGCCATAAGAAAGTGGGAGCCTTATGGAGGGAGGCTGACCTCAGCTGGAAGGACTTTTTACCAG
AAGGAGAAGATGTACATAATTTTCTTTTGGAGCAGAAGTTGGACTTCATAGAGTCTGACAGTCCCTGTTCC
TCTGAAGCACTTTCAAAGAAAAGAACTGTCTGCCGAAGAGCTGTATAAGCGACTCGAGAAAATCATTATTGA
GGACAAAGCGAATGATGAACAGATCTTTGACTGGGTAGAGGCTAATCTAGACGAAATCCAGATGAGTTCAC
CTACATTCCTTAGAGCTTTAATGACTGCTGTTTTGTAAAGCAGCTATTATAGCCGACTCTTCTACCTTCAGA
GTGGACACTGCTGTTATCAAGCAGAGAGTGCCGATCTTACTCAAGTACCTAGACTCAGATACAGAGAAGGA
ACTGCAAGCACTTTATGCACTACAAGCATCGATAGTAAAATTTGATCAACCTGCCAATTTGCTGCGGATGT
TTTTTGATTGTCTATATGACGAGGAGGTGATCTCCGAGGATGCCTTCTACAAATGGGAGAGCAGCAAGGAC
CCTGCAGAGCAGAATGGGAAGGGCGTGGCTCTGAAATCTGTACGGCATTCTTACGTGGCTGCGGGAAGC
AGAAGAGGAGTCTGAGGATAACTTAAACTTCAAATACACAAAATGAAACAAAAGAAAACAATTTAAGTATTT
TTTTAAAAAGTTTTACGTCTTTCGCCAATCACAGTGCAGCAAGGCCAATTTCTCGCAGAAAACCCACAGTGTG
CACGAGTGGGAGAGGGGAAAAGAGAAAAAAGGTGATCATGGAGGAAAAAGGTAAGTGGATAAAAAGTAACTT
CAAACCTTAGGGCGGGAGCACTAAAACCAAATACATGTATTATTTATAGAAAATATTTTCTGTTTTAATC
TTTTCTTTTTTAAACAAGGACTCATACTTAAAAAATGTTTAGCAAAAAAAAAAAAAAGTTGAGAACTTTTTAA
TTTTTTTTAAGGACTGCAAATGCCAGTGAATTTTTTAAATTTGCAGTTTCTGTAAACAACCTGTATAATAG
AAAAGCAGAGAAATAAATTTCCCTCCCCTTCAAGATGCACCTCATGTTTGTTTTTAAGGTATAGCATTTAGT
CCAGATTTGAGAAAAGTTTGGGGTGAACAAGGTAAGAAAGATTTTTTTTTTTTTTGGCATCAAATCTTTCTGC
CTGCCTCTCAGCTTGCTTCAAAAATTTAAAAAATCACAATAGTAATCAAACATACATAACATTGAAACA
GAAGGAAATGCTGTGGACCACAGAACTCCAAGAATTTGTTTTAAAAAAGTGTCTACCCTGAGAAAAGT
ACTCTTAATACTCTTGAATCTTTAGAGCAACTTTAAGGCTTGTAAATACATAGAACAATATTTAAAAAA

ACAAAAAGAAATTGACTCAGTACTATTTCTTTTCACTTTGAAAATATAAAGAACAAAATAAAGACAAACAT
TGCAAGTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

IHRHRLSPATEKGAASSGEGDKDPPPPAHEDIAVKGTATAAGTGPGTGAAAAAAAAAVPPHPAAAAAAAA
AAAAAAAAVPPHPNIRALQTQAPQQIPRGPVQQPLEDRIFTPAVSAVYSTVTQVARQPGTPTSPYSAHE
INKGHPNLAATPPGHASSPGLSQTPYPSGQNAGPTTLVYPQTPQTMNSQPQTRSPFFQRPQIQPPRATIPN
SSPSIRPGAQTPTAVYQANQHIMMVNHLMPYPVPQGPQYCIQYRHSGBPYPVGGPPQQYYPVQPPGPGPFYP
GPGPGDFPNAYGTPFYPSQPVYQSAPIIVPTQQQPPPAKREKKTIRIRDPNQGGKDITEEIMSGGSRNPT
PPIGRPTSTPTPPQQLPSQVPEHSPVVYGTVESAHLAASTPVTAASDQKQEEKPKPDPVLKSPSPVLRVL
SGEKKEQEGQTSETTAIVSIAELPLPPSPTTVSSVARSTIAAPTSSALSSQPIFTTAIDDRCELSSPREDT
IPIPSLTCTETSDPLPTNENDDDICKKPCSVAPNDIPLVSSTNLININGVSEKLSATESIVEIVKQEV
PLTLELEILENPEEMKLECIAPITPSTVPSFPPPTPPASPHTPVIVPAAATTVSSPSAAITVQRVL
EEDESIRTCLSEDAKEIQNKIEVEADGQTEEILDSQNLNSRRSPVPAQIAITVPKTKPKDRTRTTEML
EALELEKAAEELSIDKVLSESEQDKMSQGFHPERDPSDLKVKVAEENGEAEAPVRNGAESVSEGEIDANS
GSTSSGDGVTFFPKPESWKPTDTEGKKQYDREFLLDFQFMPACIQKPEGLPPI SDVVLDKINQPKLPMT
LDPRILPRGPDFTPAFADGRQTPGGRGVPLLNVGSRSSQPGQRREPRKIITVSVKEDVHLKKAENAWKPS
QKRDSQADDPENIKTQELFRKVR SILNKLTQMFNQLMKQVSGLTVDEERLKGVIDLVFEKAIDEPSFSV
AYANMCRCLVTLKVPMDKPGNTVNFRKLLLNRCQKEFEKDKADDDVFEKKQKELEAASAPEERTRLHDEL
EEAKDKARRRSIGNIKFIGELFKLMLTEAIMHDCVVKLLKNHDEESLECLCRLTTIGKDLDFEKAKPRM
DQYFNQMEKIVKERKTSSRIRFMLQDVIDLRLCNWVSRRADQGPKTIEQIHKEAKIEEQEEQRKVQQLMTK
EKRRPGVQRVDEGGWNTVQGAKNSRVLDPKFLKITKPTIDEKIQLVPKAQLGWSGKSSGGAKASETDL
RSSASSLNRFSALQPPAPSGSTPSTPVEFDSRRTLTSRSGSMGREKNDKPLPSATARPNTFMRGGSSKDLLD
NQSQEEQRREMLETVKQLTGGVDVERNSTEAERNKTRSAKPEISAMSAHDKAALSEELEKRSKSIIDF
LHINDFKEAMQCVEELNAQGLLHVFRVGVVESTLERSQITRDHMGQLLYQLVQSEKLSKQDFFKGFSETLE
LADDMAIDIPHIWLYLAELVTPMLKEGGISMRELTIEFSKPLLPVGRAGVLLSEILHLLCKQMSHKKVGAL
WREADLSWKDFLPEGEDVHNFLLQKLDIESDPCSSEALSKKELSAEELYKRLEKLIIEDKANDEQIFD
WVEANLDEIQMSSPTFLRALMTAVCKAAI IADSSTFRVDTAVIKQRPILLKYLDSDTEKELQALYALQAS
IVKLDQPANLLRMFFDCLYDEEVI SEDAFYKWESSKDPAEQNGKGVALKSVTAFFTWLREAEEESEDN

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

GTGTGTGAGAGTCCAGCCTTGGGCCAGAGTGCCGGGTCTAGGCCTGGAAGCGCGAGGCGGGCATTTCGGAA
GAGAGTGGCGCGGCGTGGGGCGCCGCGAGGGCACTACGAGCCCAGGAAGTCCGCGCCGCGCCGATTTTCC
GCGGCTGTATAAGGACTAGCGCTTCGGTAGCCGGGAGCTGGAGGGAACCTGATTGAGATCAATTCTTGG
AGAACACAAAGAAGTGTGATGATTGAACAGCTGCCCTTTGACCCTGCTAACAACAACGAGCCCCTGCAG
TTTGGTAGTGCCAGTGGCCCTCTGGTCACAGAAGGCCTCATTGAGAAATGGAGGGGAATCAAGCAAGAAAAG
AAAGAGAACAATACTCCTTCAGCCGATAACAGTAGAACTCTGAACGTGGATTCCACTGCAATGACACTAC
CTATGTCTGATCCAACCTGCATGGGCCACAGCAATGAATAATCTTGAATGGCACCGCTGGGAATTGCCGGA
CAACCAATTTTACCTGACTTTGATCCTGCTCTTGAATGATGACTGGAATCCACCAATAACTCCAATGAT
GCCTGGTTTGGGAATAGTACCTCCACCAATCCTCCAGATATGCCAGTAGTAAAAGAGATCATACTACTGTA
AAAGCTGCACGCTCTTCCCTCCAAATCCAAATCTCCACCTCCTGCAACCCGAGAAAAGACCACCAGGATGC
AAAACAGATATTTGTGGTGGTCTGCCTGAAAATGGGACAGAGCAAATCATTGTGGAAGTTTTGAGCAGTG
TGGAGAGATCATTGCCATTCGCAAGAGCAAGAAGAACTCTGCCACATTCGCTTTGCTGAGGAGTACATGG
TGGACAAAGCCCTGTATCTGTCTGGTTACCGCATTTCGCTGGGCTCTAGTACTGACAAGAAGGACACAGGC

AGACTCCACGTTGATTTTCGCACAGGCTCGAGATGACCTGTATGAGTGGGAGTGTAAACAGCGTATGCTAGC
CAGAGAGGAGCGCCATCGTAGAAGAATGGAAGAAGAAAGATTGCGTCCACCATCTCCACCCCCAGTGGTCC
ACTATTAGATCATGAATGCAGCATTGTTGCTGAAAAATTTAAAAGATGATTCCAAATTTCTCAGAAGCTGTA
CAGACCTTGCTTACCTGGATAGAGCGAGGAGAGGTCAACCGTCGTAGCGCCAATAACTTCTACTCCATGAT
CCAGTCGGCCAACAGCCATGTCCGCCGCTGGTGAACGAGAAAGCTGCCATGAGAAAAGATATGGAAGAAG
CAAAGGAGAAGTTCAAGCAGGCCCTTTCTGGAATTCTCATTCAATTTGAGCAGATAGTGGCTGTGTACCAT
TCCGCCTCCAAGCAGAAGGCATGGGACCACTTCAAAAAGCCCAGCGGAAGAACATCAGCGTGTGGTGCAA
ACAAGCTGAGGAAATTCGCAACATTCATAATGATGAATTAATGGGAATCAGGCGAGAAGAAGAAATGGAAA
TGTCTGATGATGAAATAGAAGAAATGACAGAAACAAAAGAACTGAGGAATCAGCCTTAGTATCACAGGCA
GAAGCTCTGAAGGAAGAAAATGACAGCCTCCGTTGGCAGCTCGATGCCTACCGAATGAAGTAGAATGCT
CAAGCAAGAACAAGGCAAAGTCCACAGAGAAGATGACCTTAACAAAAGAACAGCAGCTGAAACTCCTGCAAC
AAGCCCTCAAGGATGACAAGTTACAGGTGGAAAAATGTTGGAAAATCTTAAAGAAAAGGAAAGCTGTGC
TTCTAGGCTGTGTGCCTCAAACCAGGATAGCGAATACCCTCTTGAGAAGACCATGAACAGCAGTCCTATCA
AATCTGAACGTGAAGCACTGCTAGTGGGATTATCTCCACATTCCTTCATGTTCAACCATTTGGAGCAAGC
ATTGAATACATCTGTTTCTACTTGCACCGTCTTGATAATAAGATCTGCACCAGCGATGTGGAGTGTCTCAT
GGGTAGACTCCAGCATACTTCAAGCAGGAAATGACTGGAGTTGGAGCCAGCCTGGAAAAGAGATGGAAAT
TCTGTGGCTTCGAGGGCTTGAAGCTGACCTAAATCTCTTTGCCTAACAACTGGGATCCTGAAGATAAATA
TGTGTTGGACAAGCATAGAAAGTGATTTATATTTTTAATGGTTTTCAAGTGAAGTTCCTTTGAATTTGTC
AGTTCATTCTGGAAAATCTTTTGGATTAATAAGGATCTTAGGACAGCACCTCGAACTACAGGCCCTAA
AGAGAAATTGCCTCAAACCACAAGTGTGTAACCTCCTCCCCTTTCTGTCAATTGGTTGTCTTTAAATATT
GCAAAAGTCTGATGCTAAACAGTATTTGGAGTGTTCAGTGTCTGTACTACTGTTGTAGACCTTGGTAT
TTTTTTAAACACTGTTAACTGAAATGTTTTGATGATTTGTATGTGATTTGTGTTTTCTAAACTTCTCTTTAC
ATTAATGTTGTTACTGGTGAAGGCATGAGAGCAGCACTAAGTCTCTGTGTAACCTGCCATTGTCTTTCCA
ATCCCCAGTAGACCAGTAAATAAATAACACATCAGTGTCTTCTAGAAGGTGCCTGACCAGGTTACCTTTT
AAACGACAAAGCATGGTTTTGTGGCTTTTTGCAAAATTACTATGAACCAAAAGTTGACAAATGTTCCAAAGT
TATTTTCTCTAACATATCACATTAAGATCTGTTTTCAGAATTGTAAAAAGTACATCTAGATGTGTTTACAG
AAAGCAAGTATCCAGTATGACTGGCATGTGTTTATGCTATTGTAATCACTTGTAAATAGTCTGCTTTTAA
AGGAGGCATGTTTCAAGTTTCTGTGAATTAATAATGCTCATTGTGTGGGCACACAGCACAAACACACACA
CGCAGCGCACAGTGGCAGAGGGATTTATTAATATTCTTTCCCCTCTGGCCTTCTTACAGTCTGTTGG
TCCCTTTGCTTCTGTTGTGAGTGTGTTGAATTGCAACCAGTACTGCTGTAAATACTATGTTTACTTTCAT
GCTGAATGTTTGCAAAGACTTGATATAAGTATTAATAGTAATGAATCAATGAATAAATAATGAGCTAGGGT
TTGTGAGGCTTTCTACAAATAGGTGAGCTCCACCTGGAGTGCGAATTGCCAGAGACACCTTGGTAGTGCC
ATCGGCAAAATCGCAATGGCAGCATGTGAGTGGACCATTGAGAACTTCTGCTTGGTGGAAAGTAAACAGAG
AGGATGGAGGTTTTGGGGCGAATGTCCTGAGGCAGAGATGGTCTTTATTGTGTGTGGTGGTGTGTTGTTAT
TTATAATAATGCAAGCATACCCTCCCTTGGAGTCTCAATTGAAGATAAAAAGAAATGTACTGAGCAAGCAAAGC
CAATGGAGAGTATTTACAAAAATACTTTGTAATGAGATGCCAGTAGTGTCAAAGTTGTATTTTTAAAA
GATAAATATTCCTTTTTTATACCTCAGTTTTGTGTCCTGTTTTAATGACTTACGCTCTAAGTAATCCATTA
GTAGTTATCTCAGTTCCTTCTTTGGGTTACTAGAATGTTGGAAAAGATGCCAAGTCTGTCTTGACAACCT
GGAAACAGGGTTCCACAGCAGCCATTCTGTGCTGAAAAGTGGCTTCCCCTGAAGCACCTGCTGTGGCAC
CAGCAGGAAGCTCAGGTTAATTTTACACTAGCTTGCTCACTGATGCATCTCTCATCAATGCTACGGAAGGC
TTTGATTATCAGTCTCGGGCTCTTGGAAATACCTAATTTTTAATAATATCTATGAAATCAAGGGAACTTTC
CATTTACAGTTATTTCTTGTTTAAATAAACTAAATTAATTTTTAGGGGAGAGCAGTAGGAAAAAGAGCTAA
TGCATGCGGGGTTTTAATACCTAGGTGATGGGTTGAGGTGCAGCAAAACCACCATGGCACACGTTACCTAT
GTAACAAACCTGCACATCCTGCACATGTACCCCGGAACCTTACTTAAAA

LNQLPFDPANNEPLQFGSASGPLVTEGLIENGGESSKKRKRNTNTPSADNSRTLNV DSTAMTLPMSDPTAW
ATAMNNLGMAPLGIAGQPILPDFDPALGMMTGIPPIIPMPPLGIVPPPIPPDMPVVKEI IHCKSCTLFPP
NPNLPPPATRERPPGCKTVFVGGLPENGTEQIIVEVFEQCGEIIAIRKSKKNFCHIRFAEEYMVDKALYLS
GYRIRLGSSTDKKDTRGLHVDFAQARDDLYEWECKQRMLAREERHRRRMEERLRPPSPPPVHVHSDHECS
IVAELKDDSKFSEAVQTLTWTIERGEVNRSSANNFYSMIQSANSVRRLVNEKAAHEKDMEEAKEFKQA
LSGILIQFEQIVAVYHSASKQKAWDHFTKAQRKNI SVWCKQAEERIRNIHNDEL MGIRREEEMEMSDEIEE
MTETKETEE SALVSQAEALKEENDSLRWQLDAYRNEVELLQEQGKVHREDDPNKEQQLKLLQALQGMQQ
HLLKVQEEYKKKEAELEKLDKDLQVEKMLENLKEKESASCASRLCASNQDSEYPLEKTMNSSPIKSEREALL
VGIISTFLHVHPFGASIEYICSYLHRLDNKICTSDVECLMGRLQHTFKQEMTGVGASLEKRWKFCGFEGLL
LT

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

CAGCCACCCAGCCCCGAAGCAGGGCCCAGCGAGACAGGGCCCAGCCCCAAGCGTGA⁺CCAGCTGCAGAGAA
GCTGAGACGGTGATTCCGACTGACTGGACTCTGGCCTCCAGGTGTCTGCACCTCACA⁺CTGGACCCACCTGG
AGACTCCACGGGCTCCCGGCAGCTCCGGCTGCTTGGGTGGATCCGGGTGCCCGCAGGGGCTGTGGGAGGGC
CCCGGGCCTGCACTGCTCCCCGATGGCCTGCTCTTCTCACGGCCGGGGCTGCACCCTGTGTCCACGTG
CTAGATCTGGAGGGACGCCCATCTGCCTCCTGCCCTGCCGCACTCCGGGGAGCGGGGCCTTCGTTCCAGA
GGACGTGGCTGTGACAGCGTCAGGGCTTGTGGTGGTCAGCGATCCCATCCATGGGGCTGTCCATGCACTCC
AGCACACAGCCCGGGACCCCGGGGGCCACTGGGTGACAGTGGGCACCTTCTGTCTCCCCGAGGGCTGGCT
GTGGATGCCCTCAACCGCCTCCTGGTGACGGACTACTTGCCTGGGGCTGTGCACAGCTTCTCGTTGGGTCC
TGCTTGGGAGCCCCTGGCCCCAGCCTCCATGCTGGGTCTGGAGGGCCCCTGCTGGGTGGGCCCAGGGCCTG
ATGGGGGCTTGTGTGAGTGAGGAGTTTGGGGATGTGAGGCTGTTTGGCAGTGCCCCGCAACCCCTGGGC
TCCCTGGGGGGCTGGACGGGGCACACTTTCGGCTGCCAGCGGGCATCTGCTCCAACCTCAGAGGGCAATGT
TATTGTGGCAGACGAGCAGAGGGCCAGGTGACCTGTTTCCCCGGGCTGGGCCACCCATCTGCCTGGTGT
CAGAGGGGCTTGGGCAGCCCTTGGGAGTGGCCTGTGCACCCAGGGCCAGCTCCTGGTGGCTGATGCCAAG
GACAACTCCATCAAGGTGTACCAGGGCCTCAAGGAGCTGGCC⁺TGACCTGAGGCTGGGTTGGAGCAGCCCTC
CTGTGCCTGAGGCCAGCTCCCAGGCCCTTGGATCACCGCGGGAGGAACCCTCAGGATGGGTGGAGCCTCCA
GGCTATGGGCATTGCCTGCCTGATGCCAGCACCACTGGGCTGGGGCCCTGGGCTTGGCTCGAGTTCTCCTG
CTGGTGAGGCTCCGGATCTCAGGAGCAGCCCTGAGTCTGCTTCCCAGGCTGCCCTGCCAGGCCTGCAGCC
TCCCCAGCCAGGGCTGCTCTCTGCTGTCCCCATTAGTGCCTGGCCCTGCATTATGCCCCCACACCC
CCTCAGGCCCTGTGCCTGGACTTTGGGGCTGGCAGCTGAAGCCTTGGATCCTGGGCCAGCTGCCGGCACA
CAGCTAGGCAGACTCTCCACCAGGTGCCCTGCCAGGCCTCCTAATCGGGGGCAGACAGGCAGGGAGGG
TGTGGCTGGGCTGGGCTGGGCGGGGCGGCCTGGGGCAGGGGTGTGGCCCCTAAATGTCCCCAACCTCAGAG
GGACCTAGAGTCTGAGCCTCCAGTAGCTTCTCTGGGCCTGGCAGAGGTAAGGGGGAGGCAACCCTGGAGT
GTCTGGAGGCCATGGCTGGCTGAACCCTGGATGCCTTTTCTTCCGCGTCCCATGAATGAAAGCTGTCTG
GGCCTTCACTTGCAGACAGGGACAAACAGCTCCATGCTGTTTGTCTCCAGTGCAGCCGTGCTGGGAGG
GTCTGGGGGAGCTTCTTACAAGGAGAGACTCCTGCTGCTTTGGAAAAGTGAAGAAAAATAGGGGTCTAACCC
CTCTCCTCCCATTTTACAAGTGGGGAAATGAGGCGTGAAAGGAGAGGCGTCTGGGTACTCCGTGGGTCTG
GGGTCCAGGGAAGGGCCTGTATGGGGGAGGGAGCTGGGAGGGGACGGTGTCTGGCTCTACCCCTGTGGGGT
GGGGAGGTGGGGCTCCCCTGTATCACAGGACATCCCCCTGAGAGGTCCCTCATATGTCTGGGTCTGTGG
GTGGGGGACTAACTGCGCAATGTAGTTAGGTGCTCAATAAACGGAGTTGCCGCTGAAAAAAAAAAAAAAAAA
AAAAAA

LDPPGDSTGSRQLRLLGWIRVPAGAVGGPRGLHCS PDGLLFLTAGAAPCVHVL DLEGRPICLLPCRTPGSG
AFVPEDVAVTASGLVVSDPIHGAVHALQHTARDPGGHWVTVGTF LSPRGLAVDALNRLLVTDYLP GAVHS
FSLGPAWEPLAPASMLGLEGPCWVGPGPDGGLAVSEEF GDVRLFGSARQPLGSLGGWTGHTFGCPAGICSN
SEGNVIVADEQRRQVTLFPRAGPPICLVSEGLGQPLGVACAPQGQLLVADAKDNSIKVYQGLKELA

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

GCGGCCCGCCGGCGCCGAGCCGCGTCCCCCTGCGCCGCCCGGGCCTGCCGGCCGAGGAGCCACC^{ATA}ATAC
ACCCCGGGTTCGCCGAGCGCCTGGCCGCCCGCGCCCGCGCCGCTCCGCCTGTGGCTGGGCTGCGTCTGC
TTCGCGCTGGTGCAGGCGGACAGTCCCTCAGCCCCAGTGAACGTCACCGTCAGGCACCTCAAGGCCAACTC
TGCAGTGGTGAAGCTGGGATGTTCTGGAGGATGAGGTTGTATCGGATTTGCCATCTCCCAGCAGAAGAAGG
ATGTGCGGATGCTGCGCTTCATCCAGGAGGTGAACACCACCACCCGCTCATGTGCCCTCTGGGACCTGGAG
GAGGATACGGAGTACATAGTCCACGTGCAGGCCATCTCCATTGAGGGCCAGAGCCCAGCCAGCGAGCCTGT
GCTCTTCAAGACCCCGCGTGGAGGCTGAGAAGATGGCCTCCAAGAACAAGATGAGGTAACCATGAAAGAGA
TGGGGAGGAACCAACAGCTGCGGACAGGCGAGGTGCTGATCATCGTCTGGTCTGTTCATGTGGGCAGGT
GTCATTGCCCTCTTCTGCCGCCAGTATGACATCATCAAGGACAATGAACCCAATAACAACAAGGAAAAAAC
CAAGAGTGCATCAGAAACCAGCACACCAGAGCACCAGGGCGGGGGCTTCTCCGCAGCAAGATATGA^{AA}AAAC
CTTTTTAGTGTCTTGCCTCAGCAGCTAAGAAGACAGACAGTAGAGAATGTGAGAGGATCTCATGGTTCTGA
TGATGATTATCCAACAAACATCTGGCCCTCTCTACATCTCTTCCCTCCATCTCCTTGTACCCCTCTGGCTTA
CTGTCTCTCTCTGGCGCACTTCTCTGAAGCCTCTTATTAACCTCCATCTCCAGAAGCACCTCAACAATGT
CAGTGGCTGAGGCTGCACTCAGAGGGATGACTGCTGGGGGTAGACTGGGTGCCAGGGGCCATGGGCCCAGG
ACCCAGTCTTGGCCATTGAGTTGAGTGAAGGCTGGGTTTGAAGGCAAAAAGACAAGACATCCA
GGCAGGCTTCTCTTTTCTTCCACAAGGGACAAGAGCTTGGCTTATTTAGGCTACAGCCCTGCTGCTGCT
CCCTTCTTCTCTGTCTGTTTCCAGCCTTGCAAGAAGCTATTACAATTAGGCCTGCTTCTCTCTC
ATTTTTTCTCTCCAGTTTCCAACAAGCCCTCAGTGAACATCATTGAAGCGTACTGCCTGTCTGCAGGG
AGAAGGATTCATTTTTCTTCTCAGCTGGTCCCCAGGCCACGGGCACAGGGAGAGGGACAACCTGCAGCAG
TGGGGAGGAGGCACAGCTAGCTGCACAGTTCTCTTCTCTTCTGCTTAGTCAAGGAGGCTGCACT
ACAAACCCAAATTCTGCAAAAAAATAAAAAATAAGCCACAAAACATAAAGGCCTGGCCCCATTCTGGAAAAG
GCAAAGCTGCATGAGACACAGCCTTCTGCCTCTCGCCTCTCTGACTGGCTTCTCTTTGAGAAAATGC
ACAAAGCCCTGGGAGATGACAAGCACAAGGACTGACTCAAGCTGTGTCTTTTCCAGACCAAGGAACATCAGAG
AAGCTGTGGGGCTGCCTGCCAGGCAGGATCATGGCTGCCATCAAGCCTTTTCTGGATCCAGCCATCAAGGA
CATGTTTGTGGTGTGATGCACACTTTTGAAGCGTGAAGATGTTACCTGGTTTGTCTCTTTTGGAAAACA
AAAATCAGAAGGCTGCATTCTAGAGGGCAGAGAAATCCCCCGAAGACTGAGCTGGTTGCCTGCATCCTCT
ATCTTCTTTGACCCTTATGACTGAAAGATCATCAGTTTGAAGGTAAGTGGTCCAATTTATTTAGGAAGTAT
CTCTTGGAGTTTCAAAAATGCTAGCTTGGACAACCTGAAAAGTCAATCACAGCTGGCATTCTGGGGGCTAC
CAAAACACCCCTTCTGGAGTAGAAGCTGCTGGAAGGCAGGCCTGAGCCATTACCACGGACAGGAAGAGCA
GCTCTGGCTATCACCCTGGCCTCTGGGGTCTTCATATCTTGGCATCTCATCCAGGGTTCATGAAAGTTA
CCCAGGGTCTCATGTCTTCTTCTAGAGCCTGAGTGGTGTGAGGTGACAGGTCTCTCTCTCCACTGCCCT
TTCTGGTTTAAAAAATGGTGTCTGATGAGGGAAGGTAGACTCTTCCCTAGGACTGACGAGTTACGGCTGC
CAGATGCCTGCATGGGAAGAGGTGGACATCTGCATCTTCCATTGGTGGTCAAGGATGGGTGTGGGAGAACC
ACACCTAGTGCAAGCCTGGTACTCAGTAAATATTTGTTGAAATGAATGATAAGAGCATTGGTCCCCAAGCC
AGAGAGCCAGAAGCCATCACCCAATGACCCCTTCTTCCGGTCTACAAGAGCTCTCAAGGCTGGGTCT
GCCACCCTCTGCTTTGCCAAGTGTGACAGCACTGGGGAGGAGAGACAGGATAAAGGGCAGATGTGAGCA
ATACTAAGGGCTTCTCATGGGAGGGCATGAGGCTCCACTCATTGTCTTGTGACTTCCATCCCTGCTGAAT
GGGGCTGCAAGGCCAAGGCTCCTTAGGGGAGAGGTCTTACCTCTGATCCACTTAGAGCAATAACCACTTT
TTAAATGTAAAATAAAAAAGACAAATGAAAAGGCCAAAAA

IHPGSPSAWPPRARAALRLWLGCVCFALVQADSPSAPVNVTVRHLKANSVVSWDVLEDEVVIGFAISQK
KDVRMLRFIQEVNTTTRSCALWDLEEDTEYIVHVQAIISIQQSPASEPVLFKTPREAEMKASKNKDEVTMK
EMGRNQQLRTGEVLIIVVFLFMWAGVIALFCRQYDI IKDNEPNNNKEKTKSASETSTPEHQGGLLRSKI

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

GCGAGGCAGGAGCCGGCGGCTGGGCTCCGCAGCGCAGCCAGCGCAGCGCGGGCGCCCCGGGCCCCGAT
 GCCCGCAGCCCCCGCGACCGTCTCTGAGCGCGGGCGCCTAGCCCCGCGCCCCCTGCCCGCCGCGACCATTG
 CCCCAGCGCGCGGCCGGGGCGGCCCGGCGCTCCCCAGGCTCCGCGCGGGCCGAAAGACGCTGCTAGCGGCC
 GCCCGGGGTGTGGTGAATGCTGCTGGTGTGGTGGTGTGCATCCCCGTGCTGGTGAGCTCGGGCGGGCCCGGA
 AGGCCACTATGAGATGCTGGGCACCTGCCGCATGGTGTGCGACCCCTACCCCGCGCGGGGCCCGGGCGCCG
 GCGCGCGGACCGACGGCGGGCAGCCCTGAGCGAGCAGAGCGGGCGCGCCCCCGCCTTCCACGCTGGTGCAG
 GGCCCCCAGGGGAAGCCGGGCCGACCCGGCAAGCCCGGCCCTCCGGGGCCTCCCGGGGACCCAGGTCTCTCC
 CGGCCCTGTGGGGCCCGGGGGGAGAAGGGTGAAGCAGGCAAGCCGGGCCCTCCGGGGCTGCCGGGCGCGG
 GGGCAGCGGCCATCAGCACTGCCACTACACCAGGTGCCGCGCGTGGCCCTTCTACGCCGGCCTCAAG
 AACCCCCACGAGGGTTACGAGGTACTCAAGTTTACGACAGTGGTCAACAACCTAGGCAACAACCTACGACGC
 GGCCAGCGGCAAGTTTACGTGCAACATTTCCGGCACCTACTTTTTTACCTACCATGTCTCATGCGCGGCG
 GCGACGGCACAGTATGTGGGCAGACCTCTGCAAGAATGGCCAGGTGCGGGCCAGTGCTATTGCCAGGAC
 GCGGACCAGAACTACGACTACGCCAGCAACAGCGTGATCCTGCACCTGGACGCCGGCGACGAGGTCTTCAT
 CAAGCTGGATGGAGGCAAAGCACACGGCGGCAACAGCAACAATAACAGCACGTTCTCTGGCTTCATCATCT
 ACTCCGACTGAGCTCCCCACGTCTCCCTCCACCCACGTCCCTCACCCGCGGGGTCCCCTCCGGGCGGGGC
 AGACGATGACTCGCCCCCTCGCCACCCGCTCGCTGCCCGGCCCTCCCCGGCTATGACGCCCCCGGCCGTG
 CTCAACACCGCCTGGGCCACAGCTAGGCCCTCCACCGGCTCGCTGCAGAGCCGGGCCAGCGCGCCCTGT
 CCCCCTGCCAGGGAACCGGGGTTGACCGCCCCCGCCAGCCCGCGCTATATATTTGTACAATAGGACTGTT
 TACTGCCACCTCCGCTGCCAGCCACCCAGCCTGGGGAGAGGTGCGGGCGGGGGTTTGTCTTCTGCG
 CTCTGAGATGAGCTGCCCTCGGCTCCCTCCGGGGTGGCGCGCCCGGGGAGGGGGGAGTTGGGGGCTGGAT
 AGCTTCCAGCACCTCAGAGCCCCCGCCGGCTGTGCCCGTCTGACCAAAGTTATAATAAAAAACATTTT
 CACCCCGCAAAAAAAAAAAAAAAAAA

IAPTARPGGPALPRLRAGRKTLAAAAGVVMLLVLVVLIPVLVSSGGPEGHYEMLGTCRMVCDPYPARGPG
 AGARTDGGDALSEQSGAPPPSTLVQGPQKPKGRGTGKPGPPGPPGDPGPPGPPVGGPEKGEKPGKPPGLPG
 AGGSGAISTATYTTVPRVAFYAGLKNPHEGYEVLKFDVVTLNGLNNYDAASGKFTCNIPGTYYFTYHVLMR
 GGDGTSMWADLCKNGQVRASAI AQDADQNYDYASNSVILHLDAGDEVFIKLDGGKAHGGNSNKYSTFSGFI
 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

AGTGTTATGATGCAGTTCACAACACACAGCCACATTACCCACAGACCGAGGTACAGAACGAGAGACAAC
 CTCTGCCCCCCAGCAGCTGGCCAGCTTTGCAGCCCCAGTCTTGAGCCCCCACTACCTCCCCCACCCA
 CCCCCATCCCCCTTCCCAATTGAAGGAGCGGAAAGAGAAGAGAGAAGAGTGAAGAGTGAAGAGAGATTGAGAGATTG
 AGAGAGAGAGAGAGATAGACGGAGATCTCTGGAGCAGACCTCAAGGTGACTTCTATTTCTATCTGTTT
 TCGTCTGGGGGGGCCCTGGCCGGCAGCCCCCAACACTTCTCCTGCCCTGAAACACGGCTCTAGCCAACC
 TGCTCCGCTGCTTACCTGCGACCGTCTCTGCGGGGGCTGCACGGCGCCAGCCCCCTCCAGCCCACCAGGGC
 ATTGTCTCCAGCCCGTCAAGCCAGCTGTGACCCCGGTCCGGGCCCTGCCTGCCTCCCCACCAAGACTTT

CCGCAGCTATCTGCCCCGCTGTCACCGCACTTACAGCTGTGTCCACTGCCGTGCACACCTGGCCAAACACG
ATGAGCTTATTTCCAAGTCCTTCCAAGGGAGCCATGGCCGAGCCTACCTGTTTAACTCCGTGGTCAACGTG
GGTTGCGGGCCAGCTGAACAGCGCCTCTTGCTCACGGGGCTCCACTCGGTAGCTGACATTTTCTGTGAGAG
CTGCAAAACCACACTGGGCTGGAAATATGAGCAAGCTTTTGAGACGAGCCAGAAGTACAAGGAAGGGAAAT
ACATCATTGAAATGTACACATGGTGAAGGACAACGGCTGGGACTGAGGGGCTCAGGCAGGGTGTGCCCTT
CCTCCGCATGCCCCCTCCCTCCCCACGGCCCTGCCAAGCAGTCTATAACCAGCATGAGTACTGCCCCACCCCT
GGGGGAAACCTGGCTCCAACCAACCCCTCCCCCTGCCTCCACCATATCCACTACCAGGCACCCCTTTAGAACA
GGGGTCTGGGGGTACCCAGGGGTGTTAAGGCTCAGGAGTGGGCAGCAGTCAGGGAGAGACAGAACTGGGG
GAAAGGGATGGTTGTGGGTCTTTCTGTTCCCAAGATCCTGAACATGGAAGCGATGGCAGGGCATAGACTCA
GGCAGAGGGATTGTGGGAGGAATCCGTTTTTGTCTCCACCTCTTTTTGAGTGAACAGAGGACAAACCTTGG
GTCACAGGGCAAGTAGATCATGGACCACAGAACAGCAGATGAGAAAAGACTTGGGTTGGAGTGAATTTCTG
GTCTCAGACACCAGGAGACCAGAGTCTCTGAGGATGAAGTTTTCTACCCCTATTTGTAGGGAAAAGGACTT
GAGTGCAGGGAAAACCTCAAATCCCAGGCCCTGGGAAATAGTAAAATAATCAAAGGGTTTTCCATTTCACTC
CACTTGTTAGTTTTATCTTGGCACTGAAGAGGCACCTTTCGAGTATCTAACTTTTTGCCATTGGGTGGGGTGGG
GACAGCTGCTCGCGGAACAGCCCCTAGTCGGCTGCTTCCAGAGTAAGCAGTCTTTATGGGCTTTCTCTGAG
GCCCAGTCACTGCTCCTGGGACCCAGTCCCCTGGAGGGGAGGTGGAAAATCAGTGCTACGGGGCCAGTCTT
TCCCGTGGCTGCCACCAGCGAATGAAACTTTTTGTATGATACATAAAGTGCTTGAGTCTATTTTTTAATAAAA
AGGGAAAAAGCAACTTGAAAAA

TALANLLRCFTCDRLCGGCTAPAPPAHQIVLQPVMPSCDPGPGPACLPTKTFRSYLPRCHRITYSCVHCRA
HLAKHDELIKSKFQSGHGRAYLFNSVNVVVGCGPAEQRLLLTGLHVSADIFCESCKTTLGWKYEQAFETSQK
YKEGKYIIEMSHMVKDNGWD

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

GCTGATGACATGAGGGCTCCGTCTCCCAGTGATGGCAGCGCGCTGCTTCGCCGCCTCCGCCGCTCAGC
CCCGGACTCCTTACGTCAGGGTAGCGGGGTCCCCCTCCGCGCGGGAGCCAGCGAACAGCGAGAGAGCACA
GCAGAGCGCGCCGCGGAGCCGGGGCGCCCTCACTGCGCTAGGAGCCCCACTAAACCCAGCGGAGCGGAGCC
TGGCGGGAGGCAGCGCCGCGGAGCCAGCGCCGACGCCAAGCAGACTCCCCGGCCAGCGCAAGCACTCC
CTGGCCGGCGCCGACCCCTCGGGGCGCCGATTCTGTTGTGTGCCCCACGTCATGCGCGCGGGCGTCCGC
GGTCTCCCAGCAGCCCCAGCGCGGTGGGCCAGGTCAGAGGTCGCGGCCTCTCGCTGAAGTAGTTG
GGTACCCGGGCTGGGGTGCACAGTCCGGGGCGCGCCAGGACCCGCGGAGCCGTTCCCGAGCGCGGG
GAGCGGGCCGCGCCGCGCCCCACCATTACCTCCCCGGGCGCAAGGAGGAGCTGGTGGCGGTGCGCTC
CCGGCTGTGGCAGCGGCGGCGGTGCCTGCTGCGGCGCCGTCGGCGTACTCTTGCCATGGCGCTCGGGC
TGCTCATCGCCGTGCCGCTGCTGCTGCAGGCGGCGCCCCGAGGCGCCGCGCACTATGAGATGATGGGCACC
TGCCGCATGATCTGCGACCCCTTACACTGCCGCGCCCGGCGGGGAGCCCCGGGTGCAAAGGCGCAGCCACC
CGGACCCAGCACCGCCGCCCTGGAAGTCATGCAGGACCTCAGCGCCAACCCTCCTCCTCTTTCATCCAGG
GACCCAAGGGCGACCCGGGGCGACCGGGCAAGCCAGGGCCGCGGGGGCCCCCTGGAGAGCCGGGCCGCT
GGACCCAGGGGCCCTCCGGGAGAGAAGGGCGACTCGGGGCGGCCCGGGTGGCAGGGCTGCAACTGACGGC
GGGCACGGCCAGCGGCGTGGGGTGGTGGGCGGCGGGGCCGGGTAGGTGGCGATTCCGAGGGTGAAGTGA
CCAGTGCCTGAGCGCCACCTTACGCGCCCCAAGATCGCCTTCTATGTGGGTCTCAAGAGCCCCACGAA
GGCTATGAGGTGCTGAAGTTCGATGACGTGGTACCAACCTCGGCAATCACTATGACCCACCACGGGCAA

GTTTCAGCTGCCAGGTACGCGGCATCTACTTCTTACCTACCACATCCTCATGCGCGGGCGGCGACGGCACCA
GCATGTGGGCGGACCTCTGCAAGAACGGGCAGGTCCGGGCCAGCGCCATTGCACAGGACGCCGACCAGAAC
TACGACTACGCCAGTAACAGCGTGGTGCTGCACTTGGATTAGGGGACGAAGTGTATGTGAAGCTGGATGG
CGGGAAGGCTCACGGAGGCAATAATAACAAGTACAGCACGTTTCTCGGGCTTTCTTCTGTACCCGGATTAGG
GGCGCGGGGGGTGCGAGGCGGGGTGGCTGCAGGCCGCCCGGTCTCCGCCCGGGCGCGGCTCCTTGGCAAAG
GCCACTCTCGATTATAACTTCTGACATCTCCTTTGGAAAAGACAAATCCCTGCGTCTCCTGCCCC
GCTCCTGGCCTCAGTGCCTGCGACCCACCAGCTCAGGGCTGTGCTCCTGGTCTCCATCCCCATCCCGG
CAAGGGAGGAAGGGACGCCGAGCCCTTGAGGCGGCGGCACAGACTTTGCAAACCTGATTAGACTGGACAG
GCAGGGCCGGGAGCCTGCCCTCCTCAGACAGCCTCCTCCCAGTGCCTAGAAGCGGAGGGCTCCGGGCCCTG
GCCAGGGAGGTAGGCCAGAGGGAGCGCGGGCTTCTGGGGCGTCTTCTTTGTGACCCGAAATACTTGTGC
AGATTTCCCTGTCCATCAGCCAAAACCCACCCACAGCAGAATTCCAGCAAACAGAAAATTCACCTCTCCA
CACCGCATTCCCTCCTGACTCAGACTCACCGCATGCATTAATTATGTTTTTACTATG

ITSPGGKEELVAVASRLWQRRRRACLAAVGVLLAMALGLLIAVPLLLQAAPRGAHYEMMGTCRMI CD PYT
AAPGGEPGAKAQPPGPSTAALVEMQDL SANPPPF IQGPKGDPGRPGKPGPRGPPGEPGPPGPRGPPGPK
GDSGRPLPGLQLTAGTASGVGVVGGGAGVGGDSEGEVTSALSATFSGPKIAFYVGLKSPHEGYEVLKFD
VVTNLGNHYDPTTGKFCQVRGIYFFTYHILMRGGDTSMWADLCKNGQVRASAI AQDADQNYDYASNSV
LHLD SGDEVYVKLDGGKAHGGNNNKYSTFSGFLLYPD

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; TEPI; MMAC1; PTEN1; 10q23del; MGC11227;

PTEN

CCTCCCCTCGCCCGGCGCGGTCCCGTCCGCTCTCGCTCGCCTCCCGCCTCCCCTCGGTCTTCCGAGGCGC
CCGGGCTCCCGGCGCGGCGGGAGGGGGCGGGCAGGCCGGCGGGCGGTGATGTGGCGGGACTCTTTATGC
GCTGCGGCAGGATACGCGCTCGGCGCTGGGACGCGACTGCGCTCAGTTCTCTCCTCTCGGAAGCTGCAGCC
ATGATGGAAGTTTGTAGAGTTGAGCCGCTGTGAGGCGAGGCCGGGCTCAGGCGAGGGAGATGAGAGACGGCG
GCGGCCGCGGCCGGAGCCCTCTCAGCGCCTGTGACGAGCCGCGGGGCGAGCGCCCTCGGGGAGCCGGCC
GGCCTGCGGCGGCGGCGGCGGCTTTCTCGCTCTCTTCTTCTTCTTCTAACCCTGCAGCCTCTTCTCCT
CGGTTCTCCTGAAAGGGAAGGTGGAAGCCGTGGGCTCGGGCGGGAGCCGGCTGAGGCGCGGCGGCGGCGG
CGGCACCTCCCGCTCCTGAGCGGGGGGAGAAGCGGCGGCGGCGGCGGCGGCGGCGGCTGCAGCTCCAGG
GAGGGGTCTGAGTCGCTGTACCATTTCCAGGGCTGGGAACGCCGAGAGTTGGTCTCTCCCCTTCTAC
TGCCTCCAACACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
GCCGCGCACCCCCCGTGGCCCGGCTCCGGAGGCCGCGGCGGAGGCAGCCGTTCCGGAGGATTATTCTGTC
TTCTCCCCATTCCGCTGCCGCGCTGCCAGGCCTCTGGCTGCTGAGGAGAAGCAGGCCAGTCGCTGCAAC
CATCCAGCAGCCGCGCAGCAGCCATTACCCGGCTGCGGTCCAGAGCCAAGCGGCGGCGAGAGCGAGGGGCA
TCAGCTACCGCCAAGTCCAGAGCCATTTCCATCCTGCAGAAGAAGCCCCGCCACCAGCAGCTTCTGCCATC
TCTCTCCTCTTTTCTTTCAGCCACAGGCTCCAGACATGACAGCCATCATCAAAGAGATCGTTAGCAGAA
ACAAAAGGAGATATCAAGAGGATGGATTGCAGTTAGACTTGACCTATATTTATCCAAACATTATTGCTATG
GGATTTCTGCAGAAAGACTTGAAGGCGTATACAGGAACAATATTGATGATGTAGTAAGGTTTTTGGATT
AAAGCATAAAAACATTACAAGATATACAATCTTTGTGCTGAAAGACATTATGACACCCGCAAATTTAATT
GCAGAGTTGCACAATATCCTTTTGAAGACCATAACCCACCACAGCTAGAATTATCAAACCTTTTGTGAA
GATCTTGACCAATGGCTAAGTGAAGATGACAATCATGTTGCAGCAATTCAGTGTAAAGCTGGAAAGGGACG
AACTGGTGTAAATGATATGTGCATATTTATTACATCGGGGCAAATTTTTAAAGGCACAAGAGGCCCTAGATT
TCTATGGGGAAGTAAGGACCAGAGACAAAAAGGGAGTAATACTATCCAGTCAGAGGCGCTATGTGTATTAT
TATAGCTACCTGTTAAAGAATCATCTGGATTATAGACCAGTGGCACTGTTGTTTACAAGATGATGTTTGA

AACTATTCCAATGTTTCAGTGGCGGAACTTGCAATCCTCAGTTTTGTGGTCTGCCAGCTAAAGGTGAAGATAT
ATTCCTCCAATT CAGGACCCACACGACGGGAAGACAAGTTTCATGTACTTTGAGTTCCCTCAGCCGTTACCT
GTGTGTGGTGTATATCAAAGTAGAGTTCTTCCACAAACAGAACAAGATGCTAAAAAAGGACAAAATGTTTCA
CTTTTGGGTAAATACATTCTTCATACCAGGACCAGAGGAAACCTCAGAAAAAGTAGAAAATGGAAGTCTAT
GTGATCAAGAAATCGATAGCATTTCAGTATAGAGCGTGCAGATAATGACAAGGAATATCTAGTACTTACT
TTAACAAAAATGATCTTGACAAAGCAAATAAAGACAAAGCCAACCGATACTTTTTCTCCAAATTTAAGGT
GAAGCTGTACTTCACAAAAACAGTAGAGGAGCCGTCAAATCCAGAGGCTAGCAGTTCAACTTCTGTAAACAC
CAGATGTTAGTGACAATGAACCTGATCATTATAGATATTTCTGACACCACTGACTCTGATCCAGAGAATGAA
CCTTTTGATGAAGATCAGCATAACAAAATTACAAAAGTCTGAATTTTTTTTTATCAAGAGGGATAAAAACAC
CATGAAAAATAAACTTGAATAAACTGAAAACTGGACCTTTTTTTTTTTAATGGCAATAGGCATTTGTGTGAGA
TTACCAGTTATAGGAACAATTCTCTTTTCCCTGACCAATCTTGTTTTACCCTATACATCCACAGGGTTTTGA
CACTTGTGTCCAGTTGAAAAAAGGTTGTGTAGCTGTGTATATACCTTTTTGTGTCAAAGGACAT
TTAAAATTC AATTAGGATTAATAAAGATGGCACTTTCCCGTTTTATTCCAGTTTTATAAAAAAGTGGAGACA
GACTGATGTGTATACGTAGGAATTTTTTCTTTTGTGTTCTGTCCCAACTGAAGTGGCTAAAGAGCTTTG
TGATATACTGGTTCACATCCTACCCCTTTGCACTTGTGGCAACAGATAAGTTTTGCAGTTGGCTAAGAGAGG
TTTCCGAAGGGTTTTGTACATTCTAATGCATGTATTCCGGT TAGGGGAATGGAGGGAATGCTCAGAAAGG
AAATAATTTTATGCTGGACTCTGGACCATATACCATCTCCAGCTATTTACACACACCTTTCTTTAGCATGC
TACAGTTATTAATCTGGACATTTCGAGGAATTGGCCGCTGTCACTGCTTGTGTTTGCGCATTTTTTTTTAA
AGCATATTGGTGTCTAGAAAAGGCAGCTAAAGGAAGTGAATCTGTATTGGGTACAGGAATGAACCTTCTGC
AACATCTTAAGATCCACAAATGAAGGGATATAAAAAATAATGT CATAGGTAAGAAACACAGCAACAATGACT
TAACCATATAAATGTGGAGGCTATCAACAAAGAATGGGCTTGAACATTATAAAAAATTGACAATGATTTAT
TAAATATGTTTTCTCAATTGTAACGACTTCTCCATCTCCTGTGTAATCAAGGCCAGTGTCAAATTCAGAT
GCTGTTAGTACCTACATCAGTCAACAACCTTACACTTATTTTACTAGTTTTCAATCATAATACCTGCTGTGG
ATGCTTCATGTGCTGCCTGCAAGCTTCTTTTTTCTCATTAAATATAAAATATTTTTGTAATGCTGCACAGAA
ATTTTCAATTTGAGATTCTACAGTAAGCGTTTTTTTTCTTTGAAAGATTTATGATGCACCTTATTCAATAGCT
GTCAGCCGTTCCACCCTTTGACCTTACACATTCTATTACAATGAATTTTGCAGTTTTGCACATTTTTTAA
ATGTCATTAACTGTTAGGGAATTTTACTTGAATACTGAATACATATAATGTTTATATTA AAAAGGACATTT
GTGTTAAAAAGGAAATTAGAGTTGCAGTAACTTTCAATGCTGCACACAAAAAAGACATTTGATTTTTT
AGTAGAAATGTCCTACATGTGCTTTATTGATTTGCTATTGAAAGAATAGGGTTTTTTTTTTTTTTTTTT
TTTTTTTTTAAATGTGCAGTGTGAATCATTCTT CATAGTCTCCCCGAGTTGGGACTAGGGCTTCAA
TTTCACTTCTTAAAAAATCATCATATATTTGATATGCCAGACTGCATACGATTTTAAAGCGGAGTACAA
CTACTATTGTAAAGCTAATGTGAAGATATTATTA AAAAGGTTTTTTTTTCCAGAAATTTGGTGTCTTCAA
TTATACCTTACCTTGACATTTGAATATCCAGCCATTTTGTCTTAAATGGTATAAAAATTCATTTTCAAT
AACTTATTGGTGTGAAATTTGTTCACTAGCTGTGGTCTGACCTAGTTAATTTACAAATACAGATTGAATAG
GACCTACTAGAGCAGCATTTATAGAGTTTGATGGCAAATAGATTAGGCAGAACTTCATCTAAAATATTCTT
AGTAAATAATGTTGACACGTTTTCCATACCTTGTGAGTTTCATTCAACAATTTTTAAATTTTTAACAAAGC
TCTTAGGATTTACACATTTATATTTAAACATTGATATATAGAGTATTGATTGATTGCTCATAAGTTAAAT
GGTAAAGTTAGAGACA ACTATTCTAACACCTCACCATTGAAATTTATATGCCACCTTGTCTTTCATAAAAG
CTGAAATTTGTTACCTAAAATGAAAATCAACTTCATGTTTTGAAGATAGTTATAAATATTGTTCTTTGTTA
CAATTTCCGGCACCAGCATATTA AACGTAACCTTTATTGTTCCAATATGTAACATGGAGGGCCAGGT CATAA
ATAATGACATTATAATGGGCTTTTGCAGTGTATTATTTTTCTTTTGGAAATGTGAAGGTCTGAATGAGGGT
TTTGATTTTTGAATGTTTTCAATGTTTTTGAAGCCTTGTCTTACATTTTATGGTGTAGTCAATGGAAATGGA
AAAATGGCATTATATATATATATATAAATATATATTATACATACTCTCCTTACTTTATTTTCAGTTACC
ATCCCCATAGAATTTGACAAGAATTGCTATGACTGAAAGTTTTTCGAGTCTAATTA AAACTTTATTTATG
GCAGTATTCATAATTAGCCTGAAATGCATTCTGTAGGTAATCTCTGAGTTTCTGGAATATTTTCTTAGACT
TTTTGGATGTGCAGCAGCTTACATGTCTGAAGTTACTTGAAGGCATCACTTTTAAAGAAAGCTTACAGTTGG
GCCCCTGTACCATCCCAAGTCTTTGTAGCTCCTCTTGAACATGTTTGCACACTTTTAAAAGGGTAGTTGA
ATAAATAGCATCACCATTCTTTGCTGTGGCACAGGTTATAAACTTAAAGTGGAGTTTACC GG CAGCATCAA
TGTTTTCAGCTTTAAAAAATAAAGTAGGGTACAAGTTAATGTTTAGTTCTAGAAAATTTTGTGCAATATGT
TCATAACGATGGCTGTGGTTGCCACAAAGTGCCTCGTTTTACCTTTAAATACTGTTAATGTGTGCATGCATGC
AGATGGAAGGGGTGGA ACTGTGCACTAAAGTGGGGGCTTTAACTGTAGTATTTGGCAGAGTTGCCTTCTAC
CTGCCAGTTCAAAGTTCAACCTGTTTTCATATAGAATATATATACTAAAAAATTT CAGTCTGTTAAACAG
CCTTACTCTGATT CAGCCTCTT CAGATACTCTTGTGCTGTGCAGCAGTGGCTCTGTGTGTAATGCTATGC
ACTGAGGATACACAAAAATACCAATATGATGTGTACAGGATAATGCCTCATCCAATCAGATGTCCATTTG
TTATTGTGTTTTGTTAAACAACCCTTTATCTCTTAGTGTATAAACTCCACTTAAA ACTGATTAAGTCTCAT
TCTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LERGGEEAAAAAAAAAAPGRGSESPVTISRAGNAGELVSPLLLPTTRRRRRRIQGPGPVLNLPASAAAAPP
VARAPEAAGGSRSEDIYSSSPHSAAAAARPLAAEEKQAQSLQPSSRRSSHYPAAVQSQAAAERGASATAK
SRAISILQKKPRHQQLLPSLSSFFFSHRLPDMTATAIKEIVSRNKRRYQEDGFDLDTLYIYPNI IAMGFPAE
RLEGVYRNNIDDVVRFLLDSKHKNHYKIYNLCAERHYDTAKFNCRVAQYPFEDHNPPQLELIKPFCELDLQW
LSEDDNHVAAIHCKAGKGRGTGVMICAYLLHRGKFLKAQEALDFYGEVTRDKKGVTI PSQRRYVYYSYLL
KNHLDYRPVALLFHKMMFETIPMFSGGTCNPQFVVCQLKVKIYSSNSGPTRRREDKFMFYFEPQPLPVCEDI
KVEFFHKQNKMLKKDKMFHFVWNTFFIPGPEETSEKVENGLCDQEIDSICSIERADNDKEYLVLTTLTKND
LDKANKDKANRYFSPNFKVKLYFTKTVEEPSNPEASSSTSVTPDVSDNEPDHYRYSDTTSDPENEPFDED
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

TTCCGCGGAGCGCGACCCGGGGACTCCCAGGCCTGTGGGCGGGCCCTGCCAGGACTGGGCGGTGCCATAA
CCCCTAGTTTAAAAACTCGCGGGTACCGGACCCAAGATCGGGGACCCGGCGGGCTCCGCGGGGAAACA
GCGAGGCTGGCGCAGCGCCGAGGCCCGCGCCCTGGGGCCCGCAATCCACGCCAGGAATCCCCGAGTGAG
CAGGGGTGAGCGCAGCCACTGCCAACGCAAACCGTGAAGAAGCTTCTGGAAGAGCAGAGGCCCGCCAGC
AGCAGCAGCCCGACGCTGGCGGGGTGCAGGGACAATTTCTCCCTCCCCAGAGCAGCCCCTGACCCCATCT
GTGAATGAGGCTGTGACTGGCCACCCTCCCTTCCAGCACACTCGGAGACTGTGGGTTCTGGACCTAGCAG
CCTGGGCTTTCCAGACTGGGACCCCAACACGCTATGCTGCCTACACTGACAGCCCCTACTCTTGCCCTGCTT
CTGCTGCCGAAAATTTCTGCTCCTGACTTCTACCCACCCTCGGACCCAGGGCAGCCGTGCCCATTTCCC
CAGGGCATGGAGGCTGGACCCTGGAGAGTTTCTGCACCCCTTCCAGGACCCCAACAGTTCCCCGCTGTGGT
CCCTGGACCATCGCTGGAGGTGGCCCGAGCTCACATGCTGGCTTTGGGGCCACAGCAGCTGCTGGCCAGG
ATGAGGAGGGGGACACGCTCCTTACCTGTTTGCAGGCTCGGGGGCTGCGCTGGGCGGCATATGCTGCGGCT
GAGGTGCTCCAGGTGTACCGGCGTCTTGACATTCGTGAGCATAAGGGCAAGACCCCTCTCCTGGTGGCGGC
TGCTGCCAACCCAGCCCTGATTGTGGAGGATCTGTTGAACCTGGGAGCAGAGCCCAATGCCGCTGACCATC
AGGGACGTTCCGGTCTTGACAGTGGCCGCTACCTACGGGCTCCCAGGAGTTCTCTTGGCTGTGCTTAACTCT
GGGGTCCAGGTTGACCTGGAAGCCAGAGACTTCGAGGGCCTCACCCCGCTCCACACGGCCATCCTGGCCCT
TAACGTTGCTATGCGCCCTTCCGACCTCTGTCCCGGGTGTGAGCACACAGGCCCGAGACAGGCTGGATT
GTGTCCACATGTTGCTGCAAATGGGTGCTAATCACACCAGCCAGGAGATCAAGAGCAACAAGACAGTTCTG
CACTTGGCCGTGCAGGCTGCCAACCCCACTCTGGTTTACGCTGCTGCTGGAGCTGCCCCGGGGAGACCTGCG
GACCTTTGTCAACATGAAGGCCACGGGAACACAGCCCTCCACATGGCGGGTGCCTTCCCCCTGGGCCGG
CCCAGGAGGCCATCGTGCAGCACCTGTTGGCAGCTGGGGCGGACCCCACTGCGCAACCTGGAGAATGAG
CAGCCCGTTACCTGCTGCGGCCCGGGCCGGCCCTGAGGGGCTCCGGCAGCTGTTGAAGAGGAGCCGTGT
GGCGCCCGCAGGCTGTCTCTTAGGACTCAAACCCAGACCCTGGACTGATTTTCCAGTCCCCACCGTCTCT
GCGGGACAGCCAGCGTATGCTAATGTTGCAAACCCATGATAATGTATGTGAATATCCTGCCATTGGGGTT
TTACATTAATAAACCAGAAATGGCTGCAGAGGGGTGAACAGGCCCAATATTTGGGGTGTGTGATAACCCCT
CTTCTACCCACAAGGAGCCCTCTTGATGATTTCTGTGAAATCGAGGCCCTTGATTGTTTCTGTGAAACAC
CCTGCACCCCTAGTCCTTTCCCCACTGAGATCTTTCCGGTCTCTCCCCCTAACTCAGT

LAQRRGRGPGGPQSTPRNPRVSRGERSHCPTQTVKLLLEEQRRRQQQPDAGGVQGFLLPPEQPLTPSVN
EAVTGHPPFPAHSETVSGPSSLGFDPDWPNTAAAYTDSPIYSCPASAAENFLPPDFYPPSDPGQPCFPQ
MEAGPWRVSAAPPSPGFPAVVPGPSLEVARAHMLALGPQQLLAQDEEGDTLLHLFAARGLRWAAYAAAEV
LQVYRRLDIREHKGTPLLVAANQPLIVEDLLNLGAEPNAADHQGRSVLHVAATYGLPGVLLAVLNSGV
QVDLEARDFEGLTPLHTAILALNVAMRPSDLCPVRLSTQARDRLDCVHMLLQMGANHTSQEIKSNKTVLHL

AVQAANPTLVQLLLELPRGDLRFTVNMKAHGNTALHMAAALPPGPAQEAIVRHLLAAGADPTLRNLENEQP
VHLLRPGPGPEGLRQLLKRSRVAPPGLSS

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

AGCATTGTGGGAGCTCCTCGGTCCGGTGCCGGTCCGGTGGCTGCCTATTGCGGCCTGCGGTGATCAACGAGGC
CCGGGGAGCGCGTCCCCAGTCTGCGCGCCGGTCTGCGGCAGCTGGCCCAAGACCCGAGCCGAAAGGAAG
TGTTGGAGCCTGAGGTGCCTCCGCGCCGCTAGGAGGACGCTGTGCCTGGCCTGGGACCTCCGCTCCCGCC
ACCGCCCTGGAGCCGCTGAGGGACGTCCACGTGGGCCTGTCCCGCCGAGCCGCGGCCCTGTCCGCTGGC
GCTGCTCTCGGGCCACTACCTCTACTACCACTACGGCTGCGACGGCCTGGACGACCGCGGCTGGGGCTGCG
GCTACCGCACTCTGCAGACGCTGTGCTCGTGGCCAGAGGGCCAGCCCGCGGGCGTACCTGGACTGGCCGCC
GTACAGGCGGCCCTGGAGGACATGGGCGACAAGCCCCCGGCTTCCGGGGCTCCCGGGACTGGATCGGCTG
CGTGGAGGCCAGCCTCTGCCTCGCTCACTTCGGAGGGCCCCAGGGACGCCTCTGCCACGTACCCCGGGGAG
TGGGGCTGCACGGGGAGCTGGAGAGGCTTTACTCGCACTTCGCAGGGGGTGGGGGCCAGTCATGGTTGGG
GGGACGCAGATGCCAGGTCCAAGGCCTTGCTGGGAGTCTGCGTAGGGTCAGGCACGGAAGCCTATGTCCT
GGTATTGGACCCTCACTACTGGGGCACTCCAAAAAGCCCCAGTGAACACTACAGGCTGCTGGGTGGGTGGGCT
GGCAAGAGGTGAGTGCAGCCTTTGACCCCAACTCCTTCTACAACCTGTGCTTGACCAGCCTTAGCTCCCAA
CAGCAGCAGCGCACCTTGGACTGAGGACGAAGTTACAGAAGTACAGATTCTCGGGTCCAGACACGCACCTA
TGTACCTCCCACTGGTGTCCCTGCAAAGCCTGGCGCTTTTGGACATCAATAATAAAAGTGGCAGGGCTGAGC
AAAAAAAAAAAAAAAAAAAAA

LEPLRDVHVGLSPPSRGPVRLALLSGHYLYYHYGCDGLDDRGWGCGYRTLQTLCSWPEGQPAGVPLAAVQ
AALEDMDGDKPPGFRGRSDWIGVEASLCLAHFGGPQGRLECHVPRGVGLHGELERLYSHFAGGGPVMVGGD
ADARSKALLGVCVGSGETEAYVLVLDPHYWGTPKSPSELQAAGWVWQEVSAAFDPNSFYNLCLTSLSSQQQ
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

GTGACTCCTTGATCTTCTTCTGCGTGGAGAGCCTTCGCGGGTGAGGCTTAACGCGCAGGAGGTCTCACGA
GAGTGGGAAGCAACTCTCGCGAATTTTAAAATTTATCTTTTTGCCTAGCGACTGACAACAGGCTGGTTGCTT
GGCGTGGAATCCTAAAGTGGCCTGGCTTTGAGACTGGAGTGAGACCCAGCCCTAGGCTGGGGTTCTTTCC
ATTATAGAGGAGACGGATTGAGAAGGGCTACAGACCAAGGTTGTTGAAAACCAAACATATGATGAGCGTCT
AGAGATTAACGACTCCGAAGAGGTTGCAAGTATTTATACTCCAACCCCAAGACACCAAGGACTTCCTCGTT
CTGCCCATCTTCTAACAAGGCTATGGCTGATAACAGCAGTGATGAGTATGAAGAGGAAAATAGCAAGGTC
CTAAGGGAGGGCATGCCACAGGCTCCAGGCCACAGAGGCCAAAGACATGGACCCTGTTCCACCTGCCCTGC
AAGTCTCAAGTGCCACCAGACCCCTTCTCATGTACTTGAGAGGGTGGGATGGTACAGAGAGAGCCAGAAGC
ACAGAAAGGAGAAGAAGGAGACCTCACAGTTGACACCTCAACGGGGCTTTAGTAAAATGAGGATGACGAT
GATGATGATGATGATTGATCTGAAACTGATTCTGATTCTGATGATGATGATGAAGAGCATGGAGCCCCTCT
GGAAGGGGCTATGACCCTGCAGACTATGAGCATTGCGCAGTTTCTGCTGAAATTAAGGAACTCTTCCAGT
ACATCAGTAGGTACACACCTCAGTTGATTGACCTGGACCACAACTGAAGCCTTTCATTCCCTGATTTTATC
CCAGTGTGGGGATATTGATGCATTCTTAAAGTCCCACGTCCTGATGGAAAGCCTGACAACCTTGGCCT
ATTGGTATTGGATGAACCTTCTACAAGCAGTCAGACCTACGGTGCTCTCACTCTGGTTAACAGAGAATT
CTAAGCAGCAACATCACACAACATATGAAAGTAAAAGCCTAGAAGATGCAGAAAAGAATCCCAAAGCC
ATTGACACGTGGATTGAGAGCATCTCTGAATTACACCGTTCTAAGCCCCCTGCGACTGTGCACTACACCAG
GCCCATGCCCGACATTGACACGCTGATGCAGGAATGGTCCCCGGAGTTTGAAGAGCTTTTGGGCAAGGTAA
GCCTGCCCACGGCAGAGATTGATTGCAGCCTGGCAGAGTACATTGACATGATCTGTGCCATTCTAGACATC
CCTGTCTACAAGAGTCGGATCCAGTCCCTCCATCTGCTCTTTTCCCTCTACTCAGAATTCAGAAGCTACA
GCATTTTAAAGCTCTCGCTGAAGGCAAGAAAGCATTCACTCCTTCATCCAATTCCACCTCCCAAGCTGGAG
ACATGGAGACATTAACCTTCAGCTGAGACTTCCCAAGCTGCTGTTTCAAGGCTGAGCTGGCCCCTCTGC
CCCAGCTGAGATGGACAGATCGTTGTCAGCTACTTGATGTCCTTGCCCATGCCACAGCTTGGCTCAGGGC
AGTGCATGTCCTGCTGCCCTCTCTGCCAGAGGGCACAGAACATGTTTGTGTTAATGAACCTGCCTGCCTCAG
ATTGCTGTCCCCGGGGAGTTAATGCATCTACACCCTGTGGGGATTTGAGTTATAAGAATTGGAATTTCTG
AGATCCCATGGAGTTAGATTGGGAGGAAAGCTTAAAAGATGTCCTTTTTTGTGAGAGGGATGGAATTTGTT
TCTTTTCAATTCGTAAGTTAGTGAGTAAAGATTTTATAAATCAAAAAAAAAAAAAAAAAAAAAA

IEETDSEGLQTKVVENQTYDERLEINDSEEVASIYTPTPRHQGLPRSAHLPNKAMADNSSDEYEEENSKVL
REGMPQAPGHRGKMDMPVPPAPASLKKCHQTPSHVLERVGVYRESQKHKRKEKKETSQLTQRFSENEDEDD
DDDDSSSETSDSDDDDEEHGAPLEGAYDPADYEHLPVSAEIKELFQYISRYTPQLIDLHKLKPFIPDFIP
AVGDIDAFLLKVPKPDGKPDNLGLLVLEPSTKQSDPTVLSLWLTENSKQHNIHQHMKVKSLEDAEKNPKAI
DTWIESISELHRSKPPATVHYTRPMPDIDTLMQEWSPFEFELLGKVSLEPTAEIDCSLAEYIDMICAILDIP
VYKSRIQSLHLLFLSLYSEFKNSQHFKALAEKKAFTPSSNSTSQAGDMETLTFS

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

GCCGCGGCGGTGGCGGAGACTGTGGCTTTAAGAGCGTGCCGGGAGCCGAGCCCCAGCCGGGCCGCGCTTC
GCCGCTGCGCACCCAGCGGAGCCAAGCCCCACGCTGGCCGGACAGGGCCGCTGTGCGCGGGCTGCTGAG
AACCAGCCCTAGACCTCTGCGTGGGGTTCTTCTGCCGAAGACATCACCAGTGTGTGGAGCCTGCCACACC
CACCCGCTGCCAAACCACGGCCTTTACCTGTGTCTTCCGGTGTTTTCCCGTGCGACCCATCTGTGGGAGTG
CCTCGTGGGCTGCCCCAGAGTTACCCACACTCAGCAGCACCAATGGTGAAGATGACAAGATCGAAGACT
TTCCAGGCATATCTGCCCTCCTGCCACCGGACCTACAGCTGCATTCACTGCAGAGCTCACTTGGCCAATCA
TGATGAACTAATTTCCAAGTCATTCCAAGGAAGTCAAGGACGAGCATACTCTTTAACTCAGTAGTTAATG
TGGGCTGTGGGCTGCAGAAGAGCGAGTGTGCTAACAGGACTGCATGCAGTCGCAGACATTTACTGTGAA
AACTGCAAAACCCTCTGGGCTGGAAATACGAACATGCTTTTGAAGCAGCCAGAAATATAAAGAAGGCAA

ATACATCATTGAACTAGCACACATGATCAAGGACAATGGCTGGGACTGATTGGACAGCATCTACCCAACCC
AGTGTCCACGTGAACGCCATTCAACCGAACATTCTTCCCAAGCGTGAGAGAGTACTGACACTTGGTTCCA
TCCATTTAGGGGCCTTGCCATCCGGGGCATCCTCCCACCCTGACGCCATCTTTCTGGTGACCGGCCTCTAA
ATCGCTGTCTCTGTCTCTTTGCTTTGTATCTGTTTGTGAGTTGATCCTGGCTTCTCTCTCTGTTCTAGT
TTTGGCTGAAAACAAAACAACAAAAGGAACAGATCCTTGACCGCATGGCGGCAGCCCACCTTGGTAAGGGC
CCCAGGGCCCATGCGAGAGCTGCCTGATGGCCTCTTGTGAGGAGAGCAGTGGCACGGGGGCGTGAGGAAGA
GGGAAAGGGGAAACTCTAAGGGTCTTGGCGCGGGGAAGGGGTGGAAGGGTGGAGGTAGGAACAAAATTGCG
CCGCTCCTGGAGACCTGATAACTTAGGCTTGAATAAATTGACTTGTCTAAAAGGACAAAGAGAAAAAAAAA
ATACCTCATGACTGCATTCTCTGACTAGAAGCTTCTGTTCTGACACCAAATGTGCCAGGTTAGCAAAT
GAGCACAAAGATGTGGCCCTGATTCTAGTTGGTGGGGCAAGGGCCTGGTTCTCCTGGGCTGAGTGGGGGAGT
GTCCTGGCAGCAGCAGTGGCAGTGGCCAGTGGCCAGGTGGGTGCGATGACTCTGATGCCTCACTCAGTCTC
TGGCAATCATCATCTTTGCTCTAGCCACCCTAGATAAGGTGTGAAGGGACTGCTGTTTGAATGGGCTT
ACCATCAAATATCCCAAAGGCTTTGACCAGCAACCAAGTAAAATCAGTAATTGAGGAGAGCAGGGCACAA
AGGGGCTGCAGTTTGGGAGCTCCTGAAGAAATGGCTCAGATATTGAGTCAGAGAAATAAAAAGTAGGATCA
GTTAGCAATTCTAACTGCCCTTCTTCTGACCCCTCATAAGAGGAGTGTGGTGAGGGAGGGGACTGGGTAG
GGGTATCCCAGGAGGAGGGGTTTACATTGGAACCAGTTCAGGTTCCGGTGCATCTTCTCTTCCGTTTTTA
CAGTGGCTTCCGTGGGATCGTCAATTTCTTGTCTTAGAGTTTTCCGGTGTTTTTTCTCCAGTCTTGTACTG
TAGACTGTAGAAAGCACGGGCCCCAGGCTCTGAGCTTAGTAATAACCTGGCTGGTAGATTCTCATGCCCC
TAATTGTCCCCTTAGGCTGAATGTCTTGCATGGAGAGAAATCTCCTGTCAGTGTGGTCCAGCAGCAGGG
AGGAGTTCTGCCCAAATTCCGATATCACCCCTTCCCCATCCAAGCATCCTTCGATTAGGGAAGTGGAGAG
CACATCCCTGTAAGGCCATAAGAGAAAGAGGAGTTTTGTTACATTTAATCAACACTGTGAAGTCTGTTCTA
CAGCAATTGAGCCATTACACAGTATATGACTGAAACTCATTTAACTGGGTTAATTTCAATTTCTTAGACTGA
ATATATTATTGTTAAGATACGTGTGCGTGTAGGTAATTTCTCAGCATCTCCTCCAAGTAGGCCGACCTTCT
CGGAAAATTCACCCTAAAAGTCTCACAAAAGAATGAGTTTATGGGGAGATTCTGTAAAGTGAAGTGAAGTGA
ATGAAAGCAGCCAACAGCCAGGAGCTTTTCCAGAAATAGCGTCTGCAGCAGAACCAGTTTTCCATTGAGAGCG
CGTCTTGGTGGAAATGCTTTTTTGTGTGTCTCCACGCGCTGATGGTGGAAATGGGAGCCCCAAGACGTGTG
GGCTTAGAAATCAACTTTTTGTTCCCAAGGCTTCTTGTCCAGATCTTCCAGTGTCTTCCATAGCCCTGGGA
GATCAAGTTGTTCTCCCCACTTTACTGCAAGGTAGACTGAAGTTTCCAGAAATACTGAATTTCTGCTCCC
AGAAGAATAGTTTTCTCTGGCTCACAGGCCAAAGTTCTCAATGAAATCGTTTTTTAACTTTACATTTCTTAA
GCTGGCTTCCCGGCAGAGAAGCCATGGATTTCCCTCTCTCCCTTCCCCCTCCTCAAGGAAATAGTCTTCC
TTTTATGGATTTTCAATGGACTCTTCTCAGCGATTGTCTGGCTGTTTATTGATAGTCTTCCATAAGA
AAATGGGGTTAAACATGGGGTAGGTATTTTGTCTTTCAAACACAAATGGAATGTGGTGACATAAACTAGA
CATGGGGTGCCCTCAAGTTTTCCAAGGGACCAATGTGCCACTGTTCTTCTTGGGGATGAGGCCTTTGACT
GTTGGATGGATCAGAGCAGGCTCCAGTCAGACCCTGGTCTGAATGTTTTTTTTTTTCCGGTGACTATCCAGT
GAGCCTTCCAGTGGGTGCAAGGCGCCATACTTGTGTGAGAGAGCTGAGTAGAGTGTGGTTTTTTCCATAAC
TACAGGGGGAAAAAAGTCAATTAGGCTTTCCCTTTGTGTGAGTAAACAAAAGTGTCTTCTTACAACGTTT
GCTCTGTTTATGGGTTGTCTATCTAACATTGAGCAGCATTGGAGAGGCCACAGCTGAGCTATGGAGATGCT
AAATTAACTCATGGCCTCAGTCAGTTTATTCTTTAATTTCTCACCAAATTAATTGACTTAGAGCATAACCA
AAGACCTCATTATTACCCAGGTGGGTTGGGGTAATTGGAGTTTGTGGTGAAGTTTGGGGGCGGGGTG
TTGGGAGTAGAGACAGGGTAAGGGGACGTGAGAAAGGAAAAGGCATGAAGTTCTATACCTCAGCCAGCAGC
TGCCTTCCGTTTGGAACTGAAGTCCAGCCAGCAGACTCTCTAGCTCCATCTCCCCTGTGCCACCCTAGGTCA
TATGACCTTGGCCACCTTGGAGTAGACCCAGACCCTCGGGACCCGGGACATTAGTCTCAGGCTGCTGATG
GATTGATTTGACATGAACCAAACACAGCCAAACTCGATACCCACAAGCTGTCAGCTGAACCTGACTGAGTG
TTCTTCTGAGTTCACGAGGATAGGCTAGAGTGCATTTTTACTGGTGGATCAGTGTGTGCGAAAGAGATGA
CCCTTTATAAAGAGATTTTTCAAGTGGATATATATAAAAAGAAACAGTTGCTTGTAAAATATACTTTTTGAAA
TAATATTTAATTTTTTAAATAATATATTTGGTGTGTTTTCTCAGATCCCCTGAGAGCACTTTTTATTTTCT
CTTTTAAATTTCTATGGTTTTCTTTGCATTTCTTGAAGTATATTTTTAAGGGAAACAGTGATCACCATAACAT
GTTTTAGTTTTTTTTTTTTTAAAGTCTCTATCACTTAACTGAGTCAAGGCTTTGAAGCAATGCCTCT
CTGCATTTTTTCCCAGTGGAAACAGACTCTGCAGTACATTAATCAGGTTGAGAATTGAAATATTTTCTTGC
ATCAGTATTGGCTAGAAAAGAAAATAAATAAAACCAAGTTAATTTAGTAGTAACAACCTTACAGTGATTCTT
CCTGTTGGAAGAATTTCCAACAAATCAGAATCACGTTTTTGTGTTGTGCGTGTGCGCGCACACGTGTGTA
AAGCACTTTCGATTGTGCCTCCTGTTTTCTCGAGTGGGGACACTTTAACTACAGTTTACACCTCGGGCGCA
TAAAGTTTTTCTTCTTCTTCTGTTTGTGTTTCTGTTTCTGAGTGGACCAACAGCAGAACCCACGAGGAT
TTGTTTTGAGTATGGAGCTGTTGCGGGTTTGTCTCTTTTTTCTGCTTTGCGTGCTCAGTTTTTACAGACTG
TAAAGGAGATGTGTTGTTTGTGAAGATGGAGCAGAGTCAAATCTGTGCTTCTAACTGAGATGAGAGTGTAT
TAATCACGTATCGAGGGCTCCAGCTGTTTTAGAAGCCACATCATGTTAAACATTAAGTGGTTTTGATTAA
AAGAACATTAATATTATAATACACATATCTTAGTGGTAAACAGCTTTTTTTTTTTTTTAAAGTTCAGATTGCCTC

AGGTTT TAGAAAGAGGCTGAGAAATCAAATCTTGAACACAATCAACTTACATATTTTAAAGGAATCTGCCTC
AAATGAGAAAATATGCTAGTTATCTAGATAGAGGAAAGAGATATTTACTTTTTTAAAAATTTAAAATAGTTA
TGAAATCTGGCAGAAAAGGTAAAGCCTAGAAGAACTATGAAAGCTATTCTCATGTTACCAAATTCTATCT
GCGCATATGTTTTTGTATAACATTTCCGGTGACAGTGGGAGTCGGTTCCCTTTCCCAACCTGCAGAGACTAT
CTTCCAATACAGAATCTGTCTATTTATGCTTGTGTTTACAACTGTATTTGTTGGGTTTTGGGTTTTTGT
TCTTTGGTGGCATTTTTTCAGGTCACTTTGCTTCTATAACAAAGGTAATTGTTTTCAAATAATTTGTCTTCA
CCTTTTCCTGTATTTGTACATAGTGATTGATTAGAGAAAAGTGCATTGTTTCTGTATATTTCCAATC
TGTGTTGGTGTCTATTTGAGAAAATAAAAGTTTTCAAATATTAATCTTAAAAA

TAFTCVFRFCPCDPSCGSASWAAPFTPHSAAPMVKMTRSKTFQAYLPSCHRITYSCIHCRAHLANHDELIS
KSFQGSQGRAYLFNSVNVVNGCGPAEERVLLTGLHAVADIYCENCKTTLGWKYEHAFFESSQKYKEGYIIEL
AHMIKDNWD

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

GGATCCC GATCTGATTCCTGATCCTTGATCCTTGATCCTTGGTCCC GCCATCGGAGCCTGAGCGCCCCCT
ATCCCCCCTGGCCCCCAGCCCCCGGGCCCTTGAGGGGGAAGAGGCAGCGGTCTGGGACGGAGCAGGGGGT
GACCAGACTCAAGAACCCCCCCTCAACATCCCCATCGCGCGCTGCCTGTCCAGGAGCGCCGAGTTCCG
GAGCGACCCGGAGCGCTGCGGATACAAAGGCGACGGGCCGAGCGGGCGCCCGCGGAGCCACCCGGCAGT
TCGCGAGCGGGAGCGTGCCTGATTTTTCTCTGAGACAAGCCACCCGTCCAGCAAAATAGAGTCCCTCA
GGGTGACAGTTGACTTCTGAAGGTGCCTCTTGGCCTAAAGAAGCCGGTGCTGAAGGAGGTGGCTGTGGGG
CCCCCAAGAGGCCCCAGCCTGCGGCCCTGGAGCGCTACAAGGCGCGCGTTTCAGACGCCATGGACACCGA
GTCCCAGTACTCGGGCTATTCTACAAGTCCGGGCCACTCCCAGCTCCCGCAAGCACAGGGACCGCCGGG
ACCGACACCGCTCTAAGAGTGCAGATGGGGGCCGAGGGGACAAGTCCGGTGACAATCCAGGCTCCCGGGGAG
CCCCTGCTGGACAATGAGTCCACACGAGGGGATGAGCGGGATGACAACCTGGGGGAAACGACGACAGTAGT
AACGGGCACCTCAGAGCACAGCATCTCCATGATGACCTCACACGCATCGCCAAGGACATGGAGGACAGTG
TCCCTCTGGACTGCTCCCGTACCTGGGTGTGGCAGCGGGGCCACCCTGGCACTGCTGTCTTTCTCAGC
CCTCTGGCCTTCTGCTGCTGCCCCACTGCTGTGGCGGGAGGAGCTGGAGCCTTGCGGGACGGCCTGCGA
GGCCTCTTCATCTCTGTGCCTTCAAGCTGCTCATCCTGCTACTGGGCAGCTGGGCTCTGTTCTTCCGCC
GGCCCAAGGCCTCGCTGCCCGCGTCTTTGTGCTGCGTGCCCTGCTTATGGTGCTGGTTTTCTGCTCGTG
GTCTCCTACTGGCTCTTCTATGGTGTGCGCATCCTGGATGCTCGGGAGCGCAGCTACCAGGGCGTGGTGCA
GTTCCGCCGTGTGCTGGTGGACGCCCTTCTTTTCGTGCACTACCTGGCCGTGGTCTGCTGGAGCTGCGCC
AGCTCCAGCCTCAGTTCAGCTCAAGGTGCTGCGCTCCACCGACGGCGCCAGCCGCTTCTACAACGTTGGC
CATCTCAGCATCCAGCGCGTGGCAGTGTGGATCCTGGAGAAGTATTACCATGACTTCCCTGTCTACAACC
TGCCCTCCTCAACCTGCCCAAGTCCGTCTGGCCAAGAAAGTGTCTGGCTTCAAGGTGTATTCCCTCGGAG
AGGAAAACAGCACCAACTCCACTGGCCAGTCTCGGGCTGTGATTGCAGCGGCAGCTCGGAGGGCCGGAC
AACAGTCAACATGAGTACTACTATGAGGAGGTGACTGAGCGAAGGGTGCGAAGAGGAGGAGGAGGAGGAGG
TGAGTGGCGGTGGAGGAGGCCCTTCACTCACATTAAGCGGCTGCAGGAAGAGGAGCAGAAAAACCCAGGG
AGGTGATGGACCCCGGGAGGCAGCCCAAGCCATCTTTGCATCCATGGCCCGTCCATGCAGAAGTACCTT
CGGACCACCAAGCAGCAGCCCTACCACACCATGGAGAGCATCCTGCAGCACCTTGAATTCTGCATCAGCA

TGACATGACGCCCAAGGCCTTCTTGGAGCGATACTTGGCGGCTGGACCTACCATCCAGTACCACAAGGAAC
GCTGGCTGGCCAAACAGTGGACATTGGTGGAGCGAGGAGCCGGTGACCAACGGCCTCAAGGATGGCATCGTT
TTCCTCTTAAAACGCCAGGACTTCAGCCTGGTGGTGCACCAAGAAGGTCCCATTCTTCAAACCTCTCCGA
GGAATTTGTGGATCCCAAGTACACAAGTTTGTGATGAGGCTGCAGTCTGAGACCTCAGTGTGA

CTGTGCA
ACAGCAGGGGGAGTGGGAAACTCTGGGGGGTCTGAGGGGGTGGGAGGGGGCTTGGTTCTCAGGCCAGCC
ACATTCCTGCCACCCTTCTTCTTCTTGGCTCTTTTTTTTTTACTTGAATTAACGCACCCCCACCTTCTCTCC
TCGCTTCTTCTTATTTTACCCCATGTGAACCTGGAGAGACCATCCTGCTGTCAACAGTACCTGGGAAGGA
CTCCCACCTCACCAACAACCTTTTGTATTACTCTAGGCCCTGCAGGAATCAGTGCCTCTCTCCCTCTTCTTT
CCCTAGTCTTTTCCCAGATTACAGTCTCTCTGAAAGGGCACAGGGCCCTGCTGATTGTACTTTCCCCTCC
TGAGCCCCGACTCACAAATCCAAGTTCTTAAAACATTTCTCTTTCAGTGGCCCAACAGGGTTTCTCTGGGGC
ACATGGACACTGACTCCAGAGAGCCACAGTGCCAAACTCTCCAGGGCAGCAACTGGCCCTCTGTCCCTCA
CCCCAGCCACAACAACCTGGGTTCTAGGGCAGGGATACTCCTGCCACACAGCCCCGAGTTAGAAATCTCC
TTGCTAGGAGCATTGCTTCCACATATATTTAGAGCAAAGAAGGATCCCATCCTTTTCCCAGAAATCTCCA
CCTAATGTTTTTGGTTTTGTATGGTCACGTGACCATAGGCAACCACGTGGAAACCCTCTGTGACCACTTTTT
CAGGGACTTAGGGGAAGGTACCTTTCTTCCAATGTGTCTTCTTAGGCAGCCCCCTGAGGAGGAGGGCTGAAT
AGATCCCTGAGGTTTTGGAGAGACCCCCATCACTGACTCCTGCTCCCTAACCTACCCTCACTTTTCGTCCC
CGCTCTTCCCAGTGAAGGATGGTATGTAGACTCCTGTACAGACATAGTGGCTTGCAGACCCTGACCCAGCC
CCTGTGGTCTTAGACAAATGTTTTTATTTTTGTACCAGCCACCCCTGTCTGCCGCCTTCTCTCGACTCC
AGAGACCTGTTGCCTCATCTCTTTTGGGGAAGAGCCGGCAGCTCCTCCTCATCCCCTGCCTTAAGTCCAGT
TCTTTGCCTCAGGGGTCTCGTTTCTTGGCCTTCCAGGGTCCCCACCCCTTTTCTCCCTGCCTGATTCTCT
GAGCTCTGGGCTCCGTCTGTATTGGGTTGAGGGGAAGGATTACTGCCTTTTGTAGGTACTTCACCCCTCA
CCCCATTTTAGCTTCCATAGTCTTTCACCAAATCCAAATCTTGATAATTTAGATCTCATTTTGGAGCAA
ATTTGCTGGCCCTCTAATAAATATTTTTCAATATAAATCTGAGCCTTTGACTCAGACATTTTTGCCAAGGAG
AGTAGAATTAGGAAGTACCCATATACATCCAGCCAGGATCCACATGGAGGACCTTTCTGATGGCTGCAATG
ACTAGGCCATTCTCTGAGTAACTCACAGTGTCTTTTGTAGGCCCTTCTTTTCCCTGAAAGACTGGTTGG
TACTTACCTTGAGAGCACATCCTGGGATAAGATCCCCAGTGTCTCCCCTGGGAGGCTCCCCCTCTGTGTA
GCACCAGCCCTGGGAATGATGGAGCCTAGTGTATCGGGTTTTCTCCTGCTGTCTTTTCTGAAAAGTTCACT
TGTTTTACCACCCGATGCTAGAGAGGAGCTCATTGGCCAATGCTTACCTTGTCCCCAAAGGGGTGGGTTGT
GGAGCTCACTTAGGCAGGGCCTCTGGCTGGGGCCAGGGTTATGAGATAGGCCCTGTATGAAATATGCTCTGT
TCTGGGGGTCTGTCTTTTTCTTCTCTTCAA AAAACTTTGTGTAGAGAGTCCCTTCTGAGTACATAAATA
CCTCACTATCCTGAAAACAGGGCCTGGATGGTACTGGGGTATTGCCTTTGTGGACAGGATGGAGTGTG
GTGTGGTCTGAGGAGCAGGTTGGGGTGGGGGAGAGGGAAAGGATTTGGGATCTTAGTTGCTGCCCTAGGTT
AGGGGCTGGGGAGTGTATTTTTAAGATCCTGCCATGTTTTTAATCACTGTGATTTTTTTTTTATTCCCCT
TTCCTAAAAAAATTTTTTCTCCTCAACTCTCTAAGCACTAAGGGCTGTGCCTGAGAATGGTAGCATTTTG
GTCTTTTGTCTTCAAGACTGTGGTATCTTTGTCTTTTTTATTATTATTATTATTATTATTATTATTATTAC
TATTGTTTTTTAAAATGTCAGGATGAATTGTGAGACATATGGCCATGTGTTTGTCTCTGCTTCTCCCCTG
TGGGAAGTTGTCTCCATGCTGTGAACCTGCTGTGGGGTGTGCAGCTGACTCAGTCCCTCTGAGCAGTTTCCC
CACTGTGTCTGTCCCATCATGCGCTGGATCTGCTCATTCTCCTGCTGTGGGGTATGCCACCTCTTACCC
CCTTGACACCATAGGGCTGCTGTGGCTGGGCCTCACCAGCACTGTCTTTTGTGTGACTCATGGCATCCTCG
TTCATCCCCACCGTGCCTAGCAGGCCTTCTTTTACCACCTCGGAACGCTTGCCTTTCTCCCTCCACAA
CAGGACGCTGTGCCTCAGTCTTACCTACCTCGCCACTCTGCCACTGTCCCCATTGGTCTTTTCTCCTAA
ACTGGTCTTTGTGCTCTCTTTGTTTTTTCTTATTTCCCTCTTGTCTCTCATTTTTTTCTTCCCATTCCCCTC
CCATTTACAGCCCTTAACTTTTCTCTTTCCCATCTCCACTCAGTATTCCAATGGCAAACCCTGATGATGTAA
CACCTGCGATGAGACATCGGACTCTCCGAACTTTCTCATCTGACACGTCTTTTTCCCAGGGTTTCGTA
CTCCTCCATTGGTCCCAGGCTAACTCCCCTGTTCTCTGTGGTGTCTGTGAGTCCGTCTGTCTTCTCTTT
CTCTGCCCTTCCCACAGGGCAGTATCTGCTGATGGATTCACTCCTGGTGTGTGATTGTTGTGATTGTTCT
TCCGTGCGCAAAAGGAAGAGGGCTTTTTGAGTCCCTTCCAAGTGAAGATTGTAATGTAGAATTTTCCACTG
TTGGATCTAGATTTTTTTTCTTTTTTTTTGGGGGGTGGGGTTACAGAGCTGAGACCTTGTGCATGCATGT
AGAAAATTGAAAATGAAAATTTTTTTTTAATATATAAAAAGCTTGTCTTACAGTTTGCAGTGGATCTAAA
CATTACGGCAATTTTAGGATTTTTTTCTTAAACATAGGAACTAAAACCTGTACAAAATTTTTTTTTATATAAAA
TAAAGACATTTGACTTTTGTGGGAAAAAAAAAAAAAAAAA

IESLRVTVDFLKVPLGLKPKVLEKAVGPPKRPQPAALERYKARRSDAMDTESQYSYSGYSKSGHSRSSRKH
RDRDRHRHSKSRDGGRGDKSVTIQAPGEPLLDNESTRGDERDDNWGETTTVVVTGTSEHSISHDDLTRIAKD
MEDSVPLDCSRHLGVAAGATLALLSFLTPLAFLLLPPLLWREELEPCGTACEGLFISVAFKLLILLGWSA
LFFRRPKASLPRVFLRALLMVLVFLLVVSYWLFYGVIRILDARERSYQGVVQFAVSLVDALLFVHYLAVVL
LELRQLQPQFTLKVVRSTDGASRFYVNGHLSIQRVAVWILEKYYHDFPVYNPALLNLPKSVLAKKVSQFKV
YSLGEENSTNNSTGQSRVIAAAAARRRDNHNEYYYEEAEHERRVRKRRARLVVAVEEAFTHIKRLQEEEEQ

KNPREVMDPREAAQAI FASMARAMQKYLRTTKQOPYHTMESILQHLEFCITHDMTPKAFLERYLAAGPTIQ
YHKERWLAKQWTLVSEEPVTNGLKDGIVFLLKRQDFSLVSTKKVPPFKLSEEFVDPKSHKFMVRLQSETS
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

ATACCTACACCGGCTTTTTGTACGACTGTTGGCCCTGGAGAACGATCCTTTGGTGGCGAGGGCGGGGAGGA
CGAAAGCGCCCACTGTGGATTGGACAGTGTCAAAAAGAGGGGCGGTCCCTACTGAAGGGGCGGTTGGGCGA
CGAAGGGAAGAGTCTTTTCAGCGCTGAGGACTGGCGCTGAGGAGGCGGCGGTGGCTCCCGGGGCGTTTGG
CGGGCTCACCCGAGCCCGCGGGCCAACGCGGATCCAGGCCCGACTGGCGGGACCGCCCCGATTCCCCGCG
GGCCTTCTAGCCGCCATGGAGGACGGCGTCTATGAACCCCCAGACCTGACTCCGGAGGAGCGGATGGAGC
TGGAGAACATCCGGCGGGCGGAAGCAGGAGCTGCTGGTGGAGATTAGCGCCTGCGGGAGGAGCTCAGTGAA
GCCATGAGCGAGGTGGAGGGGCTGGAGGCCAATGAGGGCAGTAAGACCTTGCAACGGAACCGGAAGATGGC
AATGGGCAGGAAGAAGTTCAACATGGACCCCAAGAAGGGGATCCAGTTCTTGGTGGAGAATGAACTGCTGC
AGAACACACCCGAGGAGATCGCCCGCTTCTGTACAAGGGCGAGGGGCTGAACAAGACAGCCATCGGGGAC
TACCTGGGGGAGAGGAAGAAGACTGAACCTGGCAGTGTCTCCAGTCTTTGTGGATCTGCATGATTCACCGA
CTCAATCTGGTGCAGGCCCTCAGGCAGTTTCTATGGAGCTTTTCGCTACCCGAGAGGCCAGAAAATTG
ACCGGATGATGGAGGCCCTTCGCCAGCGATACTGCCTGTGCAACCCTGGGGTTTTCCAGTCCACAGACAGC
TGCTATGTGCTGTCTTCGCCGTATCATGCTCAACACCAGTCTCCACAATCCCAATGTCCGGGACAAGCC
GGGCTGGAGCGCTTTGTGGCCATGAACCGGGGCATCAACGAGGGCGGGGACCTGCCTGAGGAGCTGCTCA
GGAACCTGTACGACAGCATCCGAAATGAGCCCTTCAAGATTCTGAGGATGACGGGAATGACCTGACCCAC
ACCTTCTTCAACCCGACCGGGAGGGCTGGCTCCTGAAGCTGGGAGGGGGCCGGGTGAAGACGTGGAAGCG
GCGCTGGTTTTATCCTCACAGACAACTGCCTCTACTACTTTGAGTACACCACGGACAAGGAGCCCCGAGGAA
TCATCCCCCTGGAGAATCTGAGCATCCGAGAGGTGGACGACCCCGGAAACCGAACTGCTTTGAACTTTAC
ATCCCCAACAACAAGGGGAGCTCATCAAAGCCTGCAAACTGAGGCGGACGGCCGAGTGGTGGAGGGAAA
CCACATGGTGTACCGGATCTCGGCCCCACGCAGGAGGAGAAGGACGAGTGGATCAAGTCCATCCAGGCGG
CTGTGAGTGTGGACCCCTTCTATGAGATGCTGGCAGCGAGAAAAGAAGCGGATTTAGTCAAGAAGAAGCAG
GAGCAGCCCTGACCCCTGCCCCCAACTCCATTATTTATTACGGAGCTGCCCCGCCTGGGTGGCCGGACCC
CTGGGCCTTGGGGCTGTGGATCCTGGTTCCCTGTTTGGAAAATTACCACCTCTAGCTCCTCACTGTTCTT
TGTAATTAACACGCTGTTGGTAATCTTATTAATTATTTAACCACCTTGGCCTGCTGACCCCTCATTCTTG
GGTTGACAGAGTCGAGGTGCTCCGTGGAGCCAGCCTGTTTCCCTGGACAGGGGCCTGGACCCGCCTGTCT
CTGGGTGCTGCCTGGGCTGTCCCGGTGGGTCTGTTCTGGTTTTACCCCGAGCCAGCAGGAGTGGAGTAAA
GGGAGAAGGTTAATATGGTGGGAGTCTGGAGTTGGAATTGGCCGGGGACAGAGTTTAGAATGCAGGGATTC
AGGTC AAGTCTAGCATTCTTAGAAGTTACTTGACAGCCATCAGCCAGTTACGCTACAGCCATCAGC
TAGTTCCCATCGTTGCTTTTCATGGGGCTTGGAGTCTTTGAGGGGAGGAGATGAAGCTGGTGGGTTGTGA
GCTGGGCCAGGGCTTTGAGGACAACCTGGAGCTGGAAGAACATGCGACCACCTCAGGGAGGGT CAGGGAAG
GATGAGTGGGGAGGTGGCCATGTCCTGCAAGGGCCTTGCTGATGGGATGTCCTGAAGGGCTGGGCAGCCT
TAGATCGGGCTAAGAGGGCAGGACTGTGGGCCAGTGCCAGAGCCAGGCTTGTCTGTTCTCAAAGGATCAGC
CTCCTTTGGAGGACATTTTGTGTCAAGGATAGGGCTGAGGACCTGCGTTCTGAACGTCTTTCTGGGATTC
TTGACAGCCAAAGAAATTTCGAGGATCGCCTCTATGAAGGTGAAGTCTTGGCATGTGGGTACCCCTGAATC
CTGGGCTAAGGTCTGTGACCTTTTGGCCCTCCAGTTGGGGTATGTGTGGTGTCCCCAAAGAACAAGGGCTCG
TGGATCCCGAGCAGGGTGCATGGCGGATTGGGAGGTCCCATGTCACTCTCCCATGCCCGCCTTTGAAGCT
GAGGCGCTCTTTAGTTAACAACACTACAAGTCCAGCAGGGACCGGGACGCGGGTGGGAGAGGCGCCTGTGG
CCCCGAGGCGTGCCGCGAGTTGTAGTCCCTCCTGCCCGCTGTGTTGCTTTTGGAGCTCTCCGATGGGATGC
GGCGCTTCGGAATTTGGGCTTTGATCCCTGTCCCGCCTTGGCCACAGGCACCTGCCGGCCTGAAGGCC

CCGCGGTGGGGGTACCCTGCGCCCCTCCGCGGGAAGGTGGACTACAGTTATCGGCAGGCTGTGCGGGCCCA
AAGCCACGGTGACCCAGACCCGAGGTTTTTCCGGGCGTGCAGTTTTCCCGAGACTCCGTGGCGGCGTTTTGT
CTTCTTTTTCTTAGTCAGATCCCCTACTTTTTGTGGAGGGTAGAGGAGGCTTTGACCGCCGCGGCCCGGGG
GCTGGTGGGAAATGGAGTTCCAAATGAGAAAATAGAATTCCCCTTCTCTTTCCACAGGTGCCAGGGAA
ACGGAGTCATCGACCCAAGAAGGTCGTGGGAGATGAGGTCCCAGGGTAAACAGCGGGTCCCGCCACTATGT
CACCTTTTCTGCGCCCTCCCCGGATGAACTGCATGCAGGGCGGCCGGCTCCGTGGCAGGCAGAGGCAGGA
AGAGGCGCGGAGCCCGCTGGTGGCGAAGAGGGGAGTGCGGCGGGAGGGGAGGGCAGAGGCAGGCGGCCGGT
CGCGTGGGGCCTGGGCCGCCCCAGGAGGGCCTCTGGCTGGATTCTTAGCAGATGGAAGCCGTGCAAGGGCA
GGAGGCAGGGGCCTGACGTGTTTTGGATTGAGGTTGCAGGAGGGGCCCTGGCTGCTTCAGGGAGAATAATT
TGGAGGCGAGCGGGGAGGTGAGAGGAAGCTGTGCGGTGCAGGATGGCGTGGCTGTGGAGGTGGAGAACGTT
GGGCCACGCTAGGAGTCTTGAAGAGGAGGCCAAAAGGACTTGCTGAGAGCCCGCCGAGCTGGTGAATGACG
GTGCCATGACTGAGGTGGTTCGAAAGGGAGTATTTAAGGGGAAATCAGGCATTCCGTTTTGACCAAATTAAG
CTGGAGATGCCTGTGGAACATTCCAGCCAGGCTGCATCACGTCACTGACCTTCAGCATTGCCACAGTC
CATCTTCCCCCAGGACCTGAGGATTTTTGCGTCCGGCTCCCTCCTCGCCCAGGACCCCCAAGCTCCCAGCAC
GCTTCTGATTTTTTTTTGTAGGTTTTTTTTTTTTGTTTTTGTGTTTTGTTTTGTTTTGTTTTGAGAGGGAGTC
TCACTTTGTGCCCTAGACTGGAGTGCAATGGCGCCATCTCGGCTTACTGCAACCTCCACCTCCCAGGTGCA
AGCGATTCCCCTGCCTCAGCTTCCCAGTAGCTGGGATTACAGATGTGAGCCACCGTACCCAGCTAATTTTT
TGTATTTTTTAGTAGAGACGGGGTTTTACCATATTGGCCAAGCTGGTCTCGAACTCCTGACCTCAGGCGATC
CATCTGTCTCAGCCTCCCAAAGTGTGGGATTACAGGTGTGAGCCACCGCGCCAGGCCTCAGCCCCTTCT
TTTGGGTGGCAATGGTTTTGGATATCGTTTTGCTCTACTAAAATTTCATGTTGAGATTGAGCCCCAGTGTGG
CAGGTGTTGGGATGTGGGGCCTCATAAGGAGCTGTGTGGTCTTGGAGGTGGAGTCCTCATGGATAGATTAA
TGCCTGCCTTAAGGGGTGAGTGAGTGTCTCACCTTGGGAATTTGGGAATTTGGTTTTCTCTCCCTTGCTTC
CTTCAGCACCATGTGATCTCTGTGCACACAGCTGCATCCTTCTGCCCTCCCCAGAAGCAGCAGCAGCCTC
AGGGATGAGGGAGGTCCTCAGTGAGTGCACCTGCCAGTCTTGAACCTTCCAGCCATCAGAATCTGAGCCA
AATCAACCTCTTCTTTATAAAGACCCAGCCTCAGGTCTTCTGTGAGAGCAACACAAAATGGACTCAGCAC
GGATCAAATTGTGTCTCCCCACCCCCACAAAAAATTTATATTTAAATCCTAACCCCCAGCA

LRTGAEAAVAPGAFERAHPSPRANADPGPTGGTAPDSPRAFLAAMEDGVYEPDLTPEERMELNIRRK
QELLVEIQRLREELSEAMSEVEGLEANEKSKTLQRNRKMAMGRKKNMMDPKKGIQFLVENELLQNTPEEIA
RFLYKGEGLNKTAIGDYLGEREELNLAVLHAFVDLHEFTDLNLVQALRQFLWSFRLPGEAQKIDRMMEFAFA
QRYCLCNPGVFQSTDTCYVLSFAVIMLNTSLHNPVNRDKPGLERFVAMNRRGINEGGDLPEELLRNLYDSIR
NEPFKIPEDDGNLTHTFNPDREGWLLKLGGRVKTWKRRWFILTDNCLYFYFYTTDKEPRGIIPLNLS
IREVDDPRKPNCFELYIPNNKQLIKACKTEADGRVVEGNHMYRISAPTQEEKDEWIKSIQAAVSVDPFY
EMLAARKKRISVKKKQEQP*

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

TCAGAGTAATGCCAACTCTCTCTGAGTGGGATGAGCAGAGCAGATGCTGCAATGAGATGCCAAAGCGGCT
CCCCTTCTCTGTGCCTTGGGTGCCTATAAATTGCTCCGGCGCGGTTTTGTGAGCCTCCTCTTCTCCTGGC
AGGTGGTACCCAGGCAGAATTCTGCCTTCAGTCTCTCTCTCGCTCCGCTCCCGCCGTGAGGCGCTCGCCG
CTGCTCGCTCGCTCCTCCGCCCCAGCTCTGAGCCTCGCCGTGCCGACCGTGCCCGCCGCCGCCGCTGG
GCGCACCCGGGGACGCCCGGGCCCACGCGGGGCTTTGGGGTGCAGCTTATTGAGTTGTGGTGGTGGCA
AAGGAGGAAGAGAGAAAGGAGGCAATAAAAAAAAAAAGGCAGCGGACGGGCGAACTGAGCGAGCGAAAGAAG
AGGAGGAGGAGGCAGAAAAGGCAACTTCAGACGGAAAGTTGGTGCGAACAGGCGCAGTCTGCAAAAACGG
AAAAGTCGATCGCAGGCGGCGGCGGCATAAAGTGTGAGCTGCCCGGGCGAGCTCAGGAGGCGGCGGCTGG
CTCTGCCCTCCCGGTGGCCGCGCCGGCGCCGGCTGCAGCCACAGGTGCGAAGGAGCTCGCGGGGGCGAGG

GCGCCCCGCGCACCCCTCCCCGGCCCCACCCCGGGCTCGGGACTTCGGCTCAAGTCACTGGGCGCCCCG
GCTCCCTCCCCAGCCGACGCTCCGCGGGGGAGCAGGAGTCCGGCAGCAGCGGCCGGCGCACGCAGCCCCA
GGGGAGTTGGGGTTCGACGGGGTGTGTTTTCGGCTCTGAAGAGGTCCCCGCCAACCTTCAAATTCTGTC
CAAAGCAGACAAGAGGATCGCCCCGCGCTGAGCCGGCTGGTGGGCAGCAGGAGGCGCCTGATCGCCGCCG
GGGCGCTGGGGGTGGTGAATGGTGCTGCTGCTGGTGATCCTCATCCCGGTGCTGGTGAGCTCGGCCGGCAGC
TCGGCGCACTACGAGATGCTGGGCACCTGCCGCATGGTCTGCGACCCCTACGGGGGCACCAAGGCGCCCAG
CACCGCTGCCACGCCCCGACCGCGGCCTCATGCAGTCCCTGCCACCTTCATCCAGGGCCCCAAAGGCGAGG
CCGGCAGGCCCGGAAGGCGGGTCCGCGCGGGCCCCCGGAGAGCCCGGGCCACCCGGCCCCATGGGGCCC
CCGGGCGAGAAGGGCGAGCCGGGCCGCCAAGGCCTGCCGGGCCCGCCCGGGGCGCCCGGCCTGAACGCGGC
CGGGGCCATCAGCGCCGCCACCTACAGCACGGTGCCCAAGATCGCCTTCTACGCCGGCCTCAAGCGGCGAG
ATGAAGGTCACGAGGTGCTCAAGTTCGACGACGTGGTCACCAACCTCGGAAACCACTACGACCCCAACC
GGCAAGTCCACTGCTCCATCCCGGCATCTACTTCTTACCTACCACGTCTGATGCGCGGAGGGGACGG
CACCAGCATGTGGGCTGATCTCTGCAAAAAACAACCAGGTGCGTGCTAGTGCAATTGCCCAAGATGCTGATC
AGAATTACGACTATGCCAGTAACAGTGTGGTTCTTCAATTTGGAGCCGGGAGATGAAGTCTATATCAAATTA
GATGGCGGGAAAGCCCATGGAGGAAACAACAATAACAGCACGTTTTCTGGATTTATTATTTATGCTGA
CTGATAATGCAGAACTAAGCTTATTATTCTGAGTTTGAACACTGGATTTCGTATGGCTAACGTCAGTGAAT
CAAGGATCCCAGGGGATGCCAATGGCAGGGCACCTCAGTTGTGTATATGTGGGGAAATCAAATGCTACCTG
ACTCACATCTGTATCACTCAGAAACATTATGTAATAAATATCAAAGCAAGATAAGCAGATGTGTGATCCAC
TACCGCCAAAGCAAATACTCCTTATCGTTAGTGTCCATGTGAATGAAGTCTATATAGATCACAAATTTTT
ATAGACAAATCTAAGACATTGAATTATTTCTTCTATATATATGATACTTTGGTGTACTGTGATCTTGCTGC
TTTTATCCATATGTGAGCTTTGGTTCTTGTGAGTTTACCTGCTTATTATGATACTTGGAGTCCATTTCATAG
TGTGGGAAGAATGATTTTTGCCCTGCAGGAGAAGGTCTAATTGAAATAATGCTGCTTGTCCCCAAAGAAA
TTGTTTGCCTTGTACTCTTGTAAACCTTAGAGCTAGACCTGGGAATGATTCAACTTCAAGCCTTAACCTGG
AATTTTCTGGATTTGAGGGAATCCCAAGCCTATGATCTTTTTTACATTTTTCTTTTTCTTATATGAATTTT
CTTTTCTCTTTTTCTGGTGATAGTCTTTAAAAATAGAGATTGACAGTGTATCATAGGATTACCCTCTAAG
TGTGAAAATGTGTAATTATGTACGGTATTTAGAAAAAGCCCTTATGTGTCCATTTTGTCAACTGAATACAT
TTAGATTC

ILSKSRQEDRPALSRLVGSRRRLIAAGALGVVMVLLLVILIPVLVSSAGTSAHYEMLGTCRMVCDPYGGTK
APSTAATPDRGLMQSLPTFIQGPKEAGRPGKAGPRGPPGEPGPPGMPGPPGEGKEGEPGRQGLPGPPGAPGL
NAAGAI SAATYSTVPKIAFYAGLKRQHEGYEVLKFDDVVTNLGNHYDPTTGKFTCSIPIGIYFFTYHVLMRG
GDGTSMWADLCKNNQVRASAI AQDADQNYDYASNSVVLHLEPGDEVYIKLDGGKAHGNNNKYSTFSGFI I
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

ATACACCCGCGCCCCGGGAGGGGAGGGGAGGAAGGTTAGGGAGGCGGAGAGGGACCCGCCCCGAGGAGAGA
GGCGCGGGCCAGGGCTCTAGCAGGGACTGGGGCCGCGGCAGGGGTAGCAAGGTGAGTCGGTGCTTGCCAAG
AGGCAGAGCGCAAAACCTACTAGGAGATCGCGCCCGGTGAGCAGCACCCGCGAGCTCAGAGCCCGGGACGTC
CGGAGCGCGGGGAGCAGTCCCCTCTCCATCAGGGAGTGGTCTATCTGGGCAGTCTGGGACCCAGGCACCGC
GCCATCCCTGAGAGAGCAGCAGTCTGGAGAGCAGGCATCTCAGATCCCTAAGAAACCAGCCGTCCGAGAAG
CCGCGGATCTCAGGTGCCCAGGATCGTTAGGACTGAACGGGAGGGTACTAGAGGACCACTGGCTCTGGACC
GTCGGGAGCTGCCCTGACGTAACCCACGAGGGGCCTCCCCTTGACGGACGGCTTGGGGAGCGGCACCGCC
GCGGCTGGAGCCCGCAGAGGCAGGGTAAGGGGAGCGGGGGCAGCCGTGCGGGGAGTGCAGACCCAGGCC
AAGGCGGGTACC GCCTCCTGGCCCCGCGGAGAGCCCCGGCCCCGCGAGCCATTGCGCCCAAGAGTGAGGAA
GATTTGCTGGCCCTGGCAGCGTCGCGGCTGAGCCGCCGAAGAGGGTGGCGGGCGCGCCGTGCGAGTGGC

CATGGTGTGCTGCTGCTGCTGGTGGCCATCCCGCTGCTGGTGCACAGCTCCCGCGGGCCAGCGCACTACGAGA
TGCTGGGTGCTGCCCATGGTGTGCGACCCGCATGGGCCCCGTGGCCCTGGTCCCGACGGCGCGCCTGCT
TCCGTGCCCCCTTCCCAGGCGCAAGGGAGAGGTGGGCCGGCGGGAAAGCAGGCCTGCGGGGGCC
CCCTGGACCACCAGGTCCAAGAGGGCCCCAGGAGAACCCGGCAGGCCAGGCCCCCCGGGCCCTCCCGGTC
CAGGTCCGGGCGGGTGGCGCCCGCTGCCGGCTACGTGCCTCGCATTGCTTTCTACGCGGGCCTGCGGCGG
CCCCACGAGGGTTACGAGGTGCTGCGCTTCGACGACGTGGTACCAACGTGGGCAACGCCTACGAGGCAGC
CAGCGCAAGTTTACTTGCCCCATGCCAGGCGTCTACTTCTTCGCTTACCACGTGCTCATGCGCGGGCGG
ACGGCACCAGCATGTGGGCCGACCTCATGAAGAACGGACAGGTCCGGGCCAGCGCCATTGCTCAGGACGCG
GACCAGAACTACGACTACGCCAGCAACAGCGTCATTCTGCACCTGGACGTGGGCGACGAGGTCTTCATCAA
GCTGGACGGCGGGAAAGTGCACGGCGGCAACACCAACAAGTACAGCACCTTCTCCGGCTTCATCATCTACC
CCGACTGAGCCCGCCCCCGCTGCCCGCTCGCCCCCTTCTCTCCCGTCCCTCACCCACCTCCTGCCCGC
CCCACCCGAGGCGCCACCCACCTTTGAGAGCCTGGCGGTGGGGTGGACCCCTCCGTTCCCGGAGGCGGC
CTAAATGGGCGAACTCTTGGTGTCAAGGGTATAAGTGGCCGGGAAGAGGAGGAGACCCGGCCAGAGGAGC
AGAGCGACTTCCGGAGGGATCACCCGCACCCAAGTGCAGCGCTGGACCCCATAGGGGCAGAGGTCGTGGCTT
TCTCTTTTGTACAGAGATGGGGAGCAGTTTTAATAGCGGGACTCAGAGGCCAGAAAGCCGGAGGGAAGCC
CCCGCAGCTTGCAGGGAAATAACAGAAACAGGAGGAGCCATTTAGGCAAGAGAAGACATTAACAGGG
TAGTGCAGGTTCTCCGTCAAACTTTCTCTCGCCACCCTCTCGTCCCCTCGTCTCCACTTTTCAGGCTCAGG
CTCCAGCCTTGGCAGCCTTCTGTGAACCTGGAGGAACCAGTGAATTCTTTCTGGCATTTAACAGCATTTC
TGTACAGTCCCCATTCCCCCTATCCGGACTAGGCCCTGGGGCTACAGCTGCTGCTGCCTCTTCTAATAAA
GTGAGG

IAPKSEEDLLALAASRLSRRKRVAAGVAMVLLLLVAIPLLHSSRGPAPHYEMLGRCRMVCDPHGPRGP
GPDGAPASVPPFPAGAKGEVRRGKAGLRGPPGPPGPRGPPGEPGRPPGPPGPGVAPAAGYVPRIA
FYAGLRRPHEGYEVLRFDDVVTNVGNAYEAASGKFTCPMPGVYFFAYHVLMRGGDGTSMWADLMKNGQVRA
SAIAQDADQNYDYASNSVILHLDVGDVEVFIKLDGGKVHGGNTNKYSTFSGFIIYPD

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

ATTAGAGGCTCCAGCCCCGCCGACTTGCAGACGTGAGATCGGGCACACCTGAGCGGCGGGCGGGGCGGTTCGT
GGCCACATCCGGGGCGACGTGCCTGAGTACCCCGTCCCGCCAGCGTCTGCCAGTCCAGCCAGTCCGCCCA
GTCTCTCGCGTCCGAGACTCGCCTCCAGCCTCCACCTCCGCCCCGGGCGCGGAGCCTCGCGGGGGCGGG
GGCGGGGCGCAAGGGGCGGGGCTGTCTTTAAAGGGCCCCGGGCCGCTGCCCTTAGGCCACTTCTGGGG
GCGGAGAGGACCTCAGCGGCTGCGGCGACACCCAGGGAAGGCGGCGCGGCCGGGTCCCGAAACTCCTGGCT
GTTTCCATCAGAGCCCTCGGACACTCCAGCCCGGGCTGAGCACGCATCGTTCGCTCCCGGGCGGATAACAAG
GGGCTCCGCCATCCGCTCCCGTCACTTCGGCCTCCATCTCTGGGACCCGCGCCGGCAGCCAGGCCAGGC
CTCTGAGTGGCCCCAGAGCCCTGGCTGGACTCGTCCACGGCGGCAGCGATCTGCCCGGGTCTCGGAGGCC
ATCCCTTCAGAGTCGGCCCTGTGCTCGCCACCCTCACCTGCTGGTTGGATTCCGAAACCCACTGTCTGAA
GACCACAGAGGGGTGTGCTGACACCCCAAATCGGATACGTCCAGACCTCAAGCTCCCTTCCCCTCTCTG
GCTGCCCTCTGCTCTTTTCTCTCTCAACCTTTTGGGGATTTCTGTGCTCCTGACACCACCTCCCCA
TCCACCACAAAGTAGCCGGGGTGGAGCCCCAAACCTTACTGGGTGTGCTCCACCTGTGCCTCCAACCCAGC
GAATCTGACAGCTTCGACCCAATTCTGCACACACCCAGGAAGTTCTGCCTTTTCTTTTCTTTTCGGTGTCTC
CTGTACTTCCAAAATTCTCCTCCTCCTGTGCCCTCTTCGCCCCCTCCTTTGGGGCCCCGTGACCCTG
AATCTGGGGGGCACACTATATTCCACCCTTTGGAGACCCTGACCCGCTTCCAGACTCTATGCTGGGGGC
CATGTTTAGGGCCGGCACCCCATGCCCCCAACCTCAATTCCAAGGAGGCGGCCACTACTTCATCGACC
GGGATGGCAAGGCCTTCCGGCACATCCTCAATTTCTGAGGCTGGGCCGCTGGACCTGCCCGTGGGTAC
GGAGAGACAGCGCTGCTCAGGGCAGAGGCTGACTTCTACCAGATCCGGCCCCCTCCTGGACGCGCTGCGGGA

ACTGGAGGCCTCTCAGGGGACCCCTGCACCCACAGCTGCCCTGCTCCACGCAGATGTAGATGTCAGCCCCC
 GCCTGGTGCACCTTCTCTGCTCGCCGGGGACCCCATCACTATGAGCTGAGCTCCGTCCAGGTGGACACCTTC
 CGAGCCAACCTTTTCTGCACCGACTCTGAGTGTCTAGGTGCTTTGCGGGCCCCGATTTGGTGTGGCCAGTGG
 GGATAGGGCAGAGGGGAGGCCACATTTTCATCTGGAGTGGGCCCCCGCCCCGTGGAACCCCCGAGGTGG
 AGTATGGGAGACTGGGGCTGCAGCCGCTGTGGACTGGGGGGCCAGGAGAGCGGCGGGAGGTGGTGGGCACC
 CCAAGCTTCCTGGAGGAGGTGCTGCGGGTGGCTCTCGAGCACGGCTTCCGACTAGACTCTGTCTTCCCCGA
 CCCCCAAGACCTGCTCAACTCCAGGTCTCTGCGCTTTGTCCGGCAC TGA GGATGCTGTTCTCAGTTTACT
 GTGGGGAGGAGAGAGAATGGGGTACTAGCACCCTGAAGCCTCTTTCCAGCTCTGCTTCCAGGACTATGAG
 AGTCGGGACTCTCCTGCACCTGACTGGAGCTCAGATGTGGGCAGGAATCCCAAACCTGAGCCCACCAAGG
 ACTCACAAGTGGTCCAGAAGGTCTCAACCTGTGCTGACCTGGGAGGGGTAGGGAAGGTTCTCTCAGCTTG
 TTCTTGCCTTAGCCTCAGCCTCCAGTCTCCTTCTGATTTGGAGCTCAGTGTTTAAGGGCTTGGAAAAGG
 GGGAACATCTCTTACCCAGACTAGACCTAGCAAAAACCTGGAAGGATATTGAGGTCTGGGAAAAGGGA
 GACTTTGCATTTTCCCAATGCGGTCTCTTGACCATGGCTTCTACTCCTGAAGCTGGGTGGCCTGGCCTG
 GCCTGACCAATGAGAGGCCAGAACACTCTGGAACATCGGAAGAGGAGTTCTTTGCTATGTTCCAAGCCATC
 TACTGAGGGAGGAGAAAGGCCACAACCACCCTAGGTTGATGATGGGAGCTAGGACAGTCCCCATGGCA
 ATGGGGCTGGAGCATCCCTCATCTGGAAGAATCCCATACTGATGGCAGGGCTGGCCAGGGGGAAGAGGGTA
 GTATCTGTGGGTCTGGCCTTCTTTCATGTGTGCGTGCATATCAGCCCGTGTGGCTGACTGATGTATAGGT
 CCCTGGCATCCTGGTTCATATCTGTGTTGCTGACTACAGTGTCTGTGATGTCCGCATGTCCAGGCCTGTTT
 GGGGTTGCTTAGCGACTCTTCTGGCACAGGGTGTGTCTGTGGTATACTGTGAGGTGGTTGACAATTAGTA
 GTTTAAATCACAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
 GTAGCCAGGAGGGGCTGTGGGGTTTTGAGTCACTGGGATCTTCCCTGGTGGGAGGTAAGAGAAGTCACTGG
 GCTTAGCTGGGCCTCTGAGGCCTGTATGGAACCTTTGGTTGCTGAGGCAACCATGGACCTGTTGCTAGGAG
 ATAGCTGGGGAAGGCCAAGGCCCGCCAGGGCAGAGAGAGGAGACGAAGAGTTTGGGACAGTGGGGGAGGA
 GATGGGAAGGGATGGGATTTCTGGGTCCCAGAGCGGGTGGGATACTCACGCACAGCTTCTTCACTGGTGGG
 GGGTGGGGCACACATTATTTCTCACTGGTCAATGATTTACAAGAAGAAAAATAAAACTGCTTTTGGAACCA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

ISPPVPSSPPSFGGVPVLNVGGTLYSTTLETLTRFPDSMLGAMFRAGTPMPPNLSQGGGHYFIDRDGKA
 FRHILNFRLLRGLDLPRGYGETALLRAEADFYQIRPLLDALRELEASQGTAPTAALLHADVDVSPRLVHF
 SARRGPHHYELSSVQVDTFRANLFDSECLGARFRFVAVSGDRAEGSPHFHLEWAPRPVELPEVEYGRLL
 GLQPLWTGGPGERREVVTSPFLEEVLVALEHGFRLDVFPDPEDLLNSRSLRFVRH

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

CCCGGCCCTCTCCCCGCCCGCCCTCCCCTCCCCCGCCTCGCCCCACCGGCTTCCCACCACGGCCTCTCT
 CGGCGAGGAAACTCTGCGCCTCCGCTTCCCTCCTCCGACTCGGACACCGGCGGAGCCTCCCCGCCCGC
 GGAAGAAACCCCGAGCCTCGGCGGCGGAGGGAGTAGGAGAGCCCGGGCTTCGGCAGGCAGAGCAGGCCTC
 TCCCCTCCGTCCTCGTCGTCGTCCTCCTCCTCCTCCTCCTCCTCCTCCGTGGTGGTAGTGGTGGGACTTCC
 CCCGGCTGCTGCCCTCCC GCCGCGCTGCTGTCCCCACCGAAGTAGCGGCCACAGCCTGGTCTAGCGGCA
 GCATCATG CAGGCAATGGGGCAGGAGGAGGAGGAGGTGGCGGCGGCGGCGGTGGAGGAGGAGGGGGC
 GGCGGGGGCCAGGGACAGACCCCGGAACTCGCCTGCTTGTGCGCCAGAACGGGGAGTCGTCCCCCTCGTC
 GTCGTGTCGTCGCGGGGGACCTGGCCACGCCAATGGGCTCCTGCCTTCCGCCCCCTCCGCCCCAGCAACA
 ATAGCAACAGCCTGAATGTCAATAACGGGGTCCC GGCGGGGCGGCCGCCGATCCTCAGCCACCGTCGCA
 GCTGCCTCCGCCACCACCGCCGCTCCTCTTCTTGGCCACCCAGAAGTGGGCAGCAGCCTCAAGAAGAA
 GAAGCGGCTCTCCAGT CAGATGAGGATGTCATTAGGCTAATAGGACAGCACTTGAATGGCTTAGGGCTCA

ACCAGACTGTTGATCTCCTCATGCAAGAGTCAGGATGTCGTTTTAGAACATCCTTCTGCTACCAAATTCCGA
AATCATGTCATGGAAGGAGACTGGGATAAGGCAGAAAATGACCTGAATGAACTAAAGCCTTTAGTGCATTC
TCCTCATGCTATTGTGGTAAGAGGCGCACTTGAAATCTCTCAAACGTTGTTGGGAATAATTGTGAGGATGA
AGTTTTTGCTGCTGCAGCAGAAGTACCTAGAATACCTGGAGGATGGCAAGGTCCTGGAGGCACTTCAAGTT
CTACGCTGTGAATTGACGCCGCTGAAATACAATACAGAGCGCATTGATGTTCTTAGTGGGTATCTGATGTG
TAGCCATGCAGAAGACCTACGTGCAAAAAGCAGAATGGGAAGGCAAAGGGACAGCTTCCCGATCTAAACTAT
TGGATAAACTTCAGACCTATTTACCACCATCAGTGATGCTTCCCCACGGCGTTTACAGACTCTCCTGCGG
CAGGCGGTGGAACTACAAAGGGATCGGTGCCTATATCACAATACCAAACCTTGATAATAATCTAGATTCTGT
GTCTCTGCTTATAGACCATGTTTGTAGTAGGAGGCAGTTCCTCATGTTATACGCAGCAGATACTTACGGAGC
ATTGTAATGAAGTGTGGTTCTGTAAATTCTCTAATGATGGCACTAAACTAGCAACAGGATCAAAAAGATACA
ACAGTTATCATATGGCAAGTTGATCCGGATACACACTGCTAAAACCTGCTTAAAACATTAGAAGGACATGC
TTATGGCGTTTCTTATATTGCATGGAGTCCAGATGACAACCTATCTTGTGCTTGTGGCCAGATGACTGCT
CTGAGCTTTGGCTTTGGAATGTACAAAACAGGAGAATAAGGACAAAAATGAGCCAGTCTCATGAAGACAGT
TTGACAAGTGTGGCTTGAATCCAGATGGGAAGCGCTTTGTGACTGGAGGTGAGCGTGGGCAGTTCATCA
GTGTGACTTAGATGGTAATCTCCTTGACTCCTGGGAAGGGGTAAGAGTGCAATGCCTTTGGTGCTTGAGTG
ATGGAAAGACTGTTCTGGCATCAGATACACACCAGCGAATTCGGGGCTATAACTTCGAGGACCTTACAGAT
AGGAACATAGTACAAGAAGATCATCCTATTATGTCTTTTACTATTTCAAAAAATGGCCGATTAGCTTTGTT
AAATGTAGCAACTCAGGGAGTTCATTTATGGGACTTGCAAGACAGAGTTTTAGTAAGAAAGTATCAAGGTG
TTACACAAGGGTTTTATACAATTCATTCATGTTTTGGAGGCCATAATGAAGACTTCATCGCTAGTGGCAGT
GAAGATCACAAGGTTTACATCTGGCACAACGTAAGTGAAGTGCCTAATTGCGGAGCTGACAGGGCACACAG
TACAGTAACTGTGTGAGCTGGAACCCACAGATTCATCCATGATGGCCAGCGCCTCAGATGATGGCACTG
TTAGAATATGGGGACCAGCACCTTTTATAGACCACCAGAATATTGAAGAGGAATGCAGTAGCATGGATAGT
TGA

TGGTGAATTTGGAGCAGACGACTTCTGTTTTAACTTAAAATTAGTCGATTTTTAATGGCTTGGGATTTG
GTGCAAAACAAACATGATTGATAGCTGGACAGACATGCTCGTCATGAAAAAGAACCATTTCTGAAGCCCGA
TTGGGGCCAAACATTTACACCTTGCTTCATAGTAACCAGTTGAGATGAAGCACGTCGTTAGAACGTTGTTG
GACACCATGTTGAATTTCCCCCATCGGTTGTGAAGAAGTGTGCTACATTCAGGCTTACCCATTGAACTC
AGTATATATATTTTTTCTTCTGTCTTTTGTCTGGCAGGATACCATTCTTGTGCTCTTCTGTGTAATG
AAGTTTAAATGCTTTGTTGGAAAACCTTTATTTAACAGTTTTAGAAGGCTTGATAGAAAGAGTGCATTAGTCT
GAAGAGTATACATTTGGATAGGAAAGAATTTCTTCTTTTCTTCCAAATCTTTCCGCCTTATTTAGCTT
GAGATCTTTGCAGCTTTGGTTTATGGATTTAGCTTTGCCCGTTGCGCAGTATATACTGATCCAGATGATAA
ACCAGTGAACATGTCAAAGCACTCTCAATATTACATTTGACAAAAAGTTTTGTAAGTTTTTACATAGCTT
GTTGCCCGTAAAAGGGTTAACAGCACAATTTTTTAAAAATAAATTAAGAAGTATTTATAGGATTAAGTG
ACTTCATTTGTATACATTTGGAATCTAAACCAGCTTAAAAACAGTTTTCTCAATGACTTAGATACACAGTA
TAACTGATGCTCTTCTGGAATACCACATGAGACATGGTCAGAAAACAGTGCTTGAAGGACATTACACAAGA
AATTCAGAGTAATGCTTTGAAGATTTCCCCCTTTTGTTTTTATTCTGAAGGAACATCAGTACCCGATCTT
GAAGAAATTCAGATTTCAAAAAGAATTTTAAATACACCAACATGAGACATCAGTAGTCAGTTGGTTTTT
TAAAGCTTGTTCAGATTTGTTCTCAACTTAGGAAGTAATTTGGTGTGATCTAGCAAAAAGAGTAGGAATCA
GCGATACAACCACTTTGGAAGTTTATAGTATAATTGAAATTTATAGAAGAATTCAGCAGGTTACAGACATA
CTTAAACTGGGATTAACCTCATAGTCATTTTTCTTAAATTGCCCTTAAATTTTTGACATATAGGGATACA
TAAATTTAAAGAATATTTTTTCTCAGTTTTTTCAGATATTGCCATACTGAACCTCATTCTAAACTGGTGCT
GTGGATAGTCTTTCCCTCTCCCTCCTGTTTTAGTTTTAAGGAAAGGTTTTCTTTCATGGAAAAGCACTTTTA
ATTTTGAGAATTTACATTTAAAATAAGCCAAAAGCCTGATGTTTACATTTGTGCTTTTTAACTCCGTAACA
GCTGAATACAACATAACATAGTCTCCCTTGAAATTCCTCCCTTTTTTTCAGTAGAGGATAAATAGGGTGACT
CCAGCTGTTGGAATGAGAATGGGGATACCCCCAAAATAAGATTCTGGTGAAGAAAAAGTCTTGTACTAAGT
TCCCCTTTAGGTGTAGCCAAGTCTATAATATAATTATGATCATGAATAGACTCCAGAATCCACAAAGAGAA
GAGGGCTTGGATGTCAAAAATTTCCATTGTGGGTAGGTGGTCGGGGATGGGAGAAAGGGAAAAGTCAG
GGAAGGGGAGAAGAAAAGAAAACACATTTGGCCTTAAAATAGCATGCATTCTCCAGAAGCCCATAGTAAA
GAAAATGACAGATCGTGGTCAGCAGACTCTTATCAAATTTGGTATGTAAGAATTTCTTCAAGCTCCATTTT
GCAGAGACCACCCTTAGAGAATCAAAAATTTCTGTTTTGATACTTCTAAGTTAAAATATATGTTACAGT
TTATCTGGTACTTCATTTTTCTTAACTAAAATTACTTTTTACTTTTAAAGCTTGAATAAAAATCTTCAATGGT
AACTGTATGTAATTAAGTGGCAACTTCTCTTTACCTCAGTATAGTTAAAAACCATTTCATTTTTAAGTGAT
TCATGTATACCCTGAAGCCGTTAATACACCCAGTCTCTATGGTAACATCACTAAATGTGGGCATAGAAAC
TCCACAGCTGTTTACAAATTTGGTTACAGTGTGCTACGGTGCATTTGTAACACTCAGATTTGTGCCCATTC
AGGGGCAATTTGGCACTCAGATTTCTACTCTACTTCACCTAACCCCTAGATCTGAGTTTTTTCAGAGTGCT
TCTGAAGTACAGTTTTAAAACACTTTTTTAAAAGTGGAGTAAAAGTGAGGCACATTTTACAAGAACATAAC
TCCTATTTAAACGGAGTAACAACATGCAAAGGTTTCTATAGCAGCTAGGTGAGTTTTGTTTTCCGGGTCTG
TCTTAACTGGCAGCTTCTGTACATACTGGTACTTATTTGCTGTAAACGTCGTTTTCATACATTTGCCATG

CAGTGA...
 CAGGAT...
 AGTGAG...
 TTGCAC...
 CTTAAAA...
 CCAAATA...
 AATTGG...
 TAACAG...
 CAGAGA...
 TTCCCT...
 CTTAGT...
 CATATT...
 TGTTGG...
 TTGAAAG...
 AAGACAT...
 AATTTCT...
 CATATCT...
 AAACAG...
 AAAAATA...
 GTATAAA...
 AAACAG...
 AGGTTTG...
 TTATCAG...
 CCTTGAA...
 TAGAGAT...
 GTGGAAC...
 AAGATTG...
 TAGTTTAA...
 TTACTAA...
 AATTTT...
 CCATGTC...
 CCTAAATA

TASLGEETLASASSSSSDTGGASPPPRKKPRASAAEGVGE...
 VGLPPAAAPPAAAAPHRS...
 SPSSSSSAGDLAHANGL...
 LKKKKR...
 LVHSPH...
 YLMCSHA...
 LDSVSL...
 EGHAYG...
 QFYQCD...
 LALLNV...
 GHTRTVN...

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

AGTCCCTATTTGAACTGCTCTCGCGGCAGTTTCAGACCTCGTGCTCGTCCCCTTCGCCTGTCTGTGTGTGG
TATCCGTAGGTCCGGGGCACTTTTTTTTTTTGGTGTGTGTGTGTGTGTGCGTGCAGCGCTGTGTGTGTGGG
TGTGTGTGTGTGTGGGGTTTGCCGGGGCGCGGAGAGAGAGCCGGGCCGAGGGGAAGGAAGGAAGCCGTAG
AGCTCCCCGGTCTCGGGCTGCGTCTCCCCTCCAGCGCTCTCCTGCCTGTGCGCTCCGACGCGGCAGCGGCG
GTGGCGCTGAGCGCTCCTCGGTGGGGACGCGGTGCTCAAGCTGGGAGCAGCGAGAACCCTTTGCGGAGGCT
TCCCGTGCCCGCTGGACTTAGGATGCCAGGGCGTTTGATTTGCACCTACTGAGCCCGGAGTTTCTCCCCGG
CCCTGGGAAGGGCTTGCAGCGGCGGCGGCGGCGGCGGCGAGCAGTAGCAGACGGAACAGCAGGCTCTCC
TCTCCGAGGGGGGGCGTGAAGCGCCCCGCGCCGGGCTCCCCGCCCTCACCTCCGGCGCTCCTGCGGCCCA
AGCCACCACCTCGTAGCCCGCCTCCTTCTTCTCGCCGCGAGCCCCAGGCCCGGACGCCTCTAGGATCG
GACCCGCGCGCCCCGGATGCTTGGGGCTGCCCGGCGCGCCCCCTACAGGCGCGACCCGCTGAGTGACCC
GGAGGGGGCCCGAGCTCGCCCCAGGAGGGGCCCGGCTCGGGACCCCGCGCCGGACCTGCTGGGGAGCC
GACGATGTCCGCGCAGTCCCAGGCATTTGACGCGCTCGACGACCCAGGGAAGGCGGCGGCGCTGCCCG
CAGAGCCCTAGGCGCACAAAGCCCCGCGCCCGCCGCGCCCGGCGCCCGCGCGCCCGACGACTTTGCCGCTGC
TCCGCGGCTCTTTGTCTCCACTTGGGGCGGGCGCCGACTCTGGGATTTGCTGCGAGAACGAGCTGGGG
GGCCGGGGCGAGCTCTCGTTTTCCCGGCCGCCCGGCTCGGGCTCGGCTCCCCCTCCCCGCACCTCC
CCGGCCCCGGCTCTCGGCGCTTCCACGCTCTCGGAATCACGACCCCTCCCTGCCATGTATCCGCAGGGCAG
ACATCCGGCTCCCCATCAACCCGGGCAGCCGGGATTTAAATTACGGTGGCTGAGTCTTGTGACAGGATCA
AAGACGAATTCAGTTTCTGCAAGCTCAGTATCACAGCCTCAAAGTGGAGTACGACAAGCTGGCAAACGAG
AAGACGGAGATGCAGCGCCATTATGTGATGTACTATGAGATGTCTATGGCTTGAACATTGAAATGCACAA
GCAGACAGAGATTGCGAAGAGACTGAACACAATTTTAGCACAGATCATGCCTTTCTGTCAAGAGCACC
AGCAGCAGGTGGCGCAGGCAGTGGAGCGGCCAAGCAGGTACCATGACGGAGCTGAACGCCATCATCGGG
CAGCAGCAGCTCCAGGCGCAGCACCTCTCCATGCCACACACGGCCCCCGGTCCAGTTGCCACCCCAACC
GTCAGGTCTCCAGCCTCCAGGAATCCCCCAGTGACAGGGAGCAGCTCCGGGCTGCTGGCACTGGGCGCCC
TGGGCAGCCAGGCCATCTGACGGTGAAGGATGAGAAGAACCACCATGAACTCGATCACAGAGAGAGAGAA
TCCAGTGCGAATAACTCTGTGTACCCTCGGAAAGCCTCCGGGCCAGTGAGAAGCACCAGGGGCTCTGCGGA
CTACAGCATGGAAGCCAAGAAGCGGAAGGCGGAGGAGAAGGACAGCTTGAGCCGATACGACAGTGATGGAG
ACAAGAGTGATGATCTGGTGGTGGATGTTTCCAATGAGGACCCCGCAACGCCCGGGTCCAGCCCGGCACAC
TCCCCTCCTGAAAATGGGCTGGACAAGGCCCGTAGCCTGAAAAAAGATGCCCCCAACAGCCCTGCCTCGGT
GGCCTCTTCCAGTAGCACACCTTCTCCAAGACCAAAGACCTTGGTCATAACGACAAATCCTCCACCCCTG
GGCTCAAGTCCAACACACCAACCCCAAGGAACGACGCCCAACTCCAGGCACCAGCACGACCCCAAGGGCTC
AGGTGATGCCGGGTAAACCTCCGGGCATGGACCCGATAGGTATAATGGCCTCGGCTCTGCGCACGCCCAT
CTCCATCACCAGTCTTATGCGGCGCCCTTCCGCTATGATGAGCCACCATGAGATGAACGGCTCCTCCACCA
GTCTCTGGCGCCTACGCGGCTCCACAACATCCCACCCAGATGAGCGCCGCGCCGCTGCTGCAGCCGCT
GCCTATGGCCGATCGCCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
AGGCCTCCCCTCAAGCCTGGCCTCCATTCTGGAGGAAAACCAGCGTACTCATTCCATGTGAGTGCTGATG
GGCAGATGCAGCCCGTGCCTTCCCCACGACGCCCTGGCAGGCCCGGCATCCCAGGCACGCCCGGCAG
ATCAACACACTCAGCCACGGGGAGGTGGTGTGTGCCGTGACCATCAGCAACCCACGAGGCACGTCTACAC
AGGTGGCAAGGGCTGCGTGAAGATCTGGGACATCAGCCAGCCAGGCAGCAAGAGCCCCATCTCCAGCTGG
ACTGCCTGAACAGGGACAATTACATCCGCTCCTGCAAGCTGCTCCCTGATGGGCGCACGCTCATCGTGGGC
GGCGAGGCCAGCACGCTCACCATCTGGGACCTGGCCTCGCCACGCCCCGCATCAAGGCCGAGCTGACGTC
CTCGGCTCCCGCTGTTATGCCCTGGCCATTAGCCCTGACGCCAAAGTCTGCTTCTCCTGCTGCAGCGATG
GGAACATTGCTGTCTGGGACCTGCACAACCAGACCCTGGTCCAGGCAGTTCAGGGCCACACAGATGGGGCC
AGCTGCATAGACATCTCCATGATGGCACAAACTGTGGACAGGGGGCCTGGACAACACGGTGCCTCCTG
GGACCTGCGGGAGGGCCGACAGCTACAGCAGCATGACTTCACTTCCAGATCTTCTCGCTGGGCTACTGCC
CCACTGGGGAGTGGCTGGCTGTGGGCATGGAGAGCAGCAACGTGGAGGTGCTGCACCACACCAAGCCTGAC
AAGTACCAGCTGCACCTGCACGAGAGCTGCGTGCTCTCCCTCAAGTTCGCCTACTGCGGCAAGTGGTTTCGT
GAGCACTGGGAAAGATAACCTTCTCAACGCCTGGAGGACGCCTTATGGAGCCAGCATATTCCAGTCTAAAG
AATCCTCGTCTGTCTTGAGTTGTGACATTTAGCGGATGACAAATACATTGTAACAGGCTCTGGTGAACAAG
AAGGCCACAGTTTATGAGGTCATCTACTAAACAAGAATCCAGCAGGCTGTCAAATCTGGGAGAAACCCG
ACTCGGCTCTGACAGGAGACCCCCAGGCGAGGGCCCCGAGGATGGCGGAGGATGGGCGGAGGATGGGCGCCG
AGCGTTCCAGGGCTGCGCTCCGGCCGGCTGAGAGGGCACGTGCCCGCTCACAGTCTGGACTCCTGGCCCTGG
ATTGATGTGTCTCACAGACTCGGAAGGGTTCTGCTCCTCCTCCTCCCCCTGAACAATGCTGGCAGTTGCTA
CAAATAGATTTTATTGGAGGCTTATGGCTCCGGTTCCCCCACAGACCCGCTCATGAGTCTCTGTTTTGTTCTT
CCCTTTTCTTTTGGCCTGTCCCTCACCTTGGGTGGGGGTGCTGGAGTGGACCACAATGTTGTGCTGGGGG
ATGGGGGGGTCTCTCTTTGCCGATTGTGACGTGCACAAGATTTGTGAAAATGTAAATAACAGACTCCTAT

TGCGGACTGATCAGTGGGAGAGGAGGCCCTTCCCACCGGAAACTCTGAGTGTGTATTTTCGCTGCTGTAT
TTGTAATCCACTCGTGGTGGTGGCTTTTTTTTTTTTTTTTTTTTTTAAATAAACAGATGCTCTCACCTGGGAA
GAGGAGACAGGGAGGGGAACCAATTGAAGAAAGAGGAGAAAAGTCTTAGAGTGTGGAAAAGGCAACCAGGT
TGGCCGTAAGGTGCCTGCTGGAATGCGTGTGCCTCCACACGGGTCTGGGCATCCGGACTGATAACCAGCCG
GCCAGACTGAGGGATGGAAGGCACTGAGATGGGGGCCCGTCCAGGCGGACACCCGCAGAAATGGAGCTTTC
TGTGGTCTCTTGCCTGCTGCTTGGCTGCTTGGCCCTCTCTGTGTCTCTCTTTCTTGGTCTCTCCCTCTCTCCT
CCTCAGCCTGGTCTTTCTCTTTGGTGCACACTTAGTTATTGTTGTGAGCAATGGAAGTTCAAAGGAACTCC
CTCTCCAGCTCTTCTGAATCTTGGGACACAGCCTAAAAAGGACAAAAAGTTAGAAGACAGCATAGCAACTC
AGCTCAGGGAGCTACCAGAGAAAAATAGCAACTGATGTGGGTGCTTTTTTTTTTTTTTTAATTTGAATAAAA
AAGAATTAGAAGTGTGCTTTTATAAAATGCCTTCTCCCCCTTCCCGCCTACAGTCTCTTCTCTCCCC
TTAGAGGGGGGAAAAGTGTATAAACCTACAGGGTTGTGAGTCTGAAAAGAGGATCCCCCTCACCCCCACCCT
GGCAGAGCAGTGGGGTTGGGGGGTGGGAGAGGGGGACACAGATCCTGGCACACTGTGGATATTTCTTGC
AGATTGCAGTCTCTTGTGGCCAAACAGGTTAGGTAGACTATCGCCTCTGGCAGGTGCCACCTTTTGGTAC
CAACATGTTCTGAGGTGTTAGGATTTGGGTTGGGTTTTTTTTTGTGTTGTTTTTTTTTCTTTTGGTCTTTTT
TTTTTTTTCTCTTTTAAAGAAAAGCTAAAGGCCGCTGTGAGTCTGGTGGCAGGCTCTCCATGGATGTAGC
ATATCGAAGATAATTTTTTATACTGCATTTTTATGGATTATTTTGTAAATGTGTGATTCCGTCTGCTGAGGAG
GTGGGAGGGGCTCCAGGGAAAGCCACCCACCTTCAAGTGGGTTGCTCCCCAGCTGAGCGCACCGGGCATGG
GATGTGGAGGCTGGCGACACACCCTGTGCCTCTCCAAGGCTGGGCGCGTGGGGCGTCCAGAGTCTCTCTGG
GTCTCAGATGTCCATCTGCCACCTCTTGTAAAGGCTCTAGCCAGAAGGGAGGGTGAGGGTAGAAGAAAGTT
ATCCCGAAGAAAAAAGAATGAAAAGTCATTGTACTGAACTGTTTTTATATTTTTTAAAAGTTACTATTTA
AAGTTAAAAAATAAAAAAAAAAAAAAAAAA

LGFRCENELGGPGASSRFPGRPPLGLGSPLPRTSPARALGASTLSESRPLPAMYPOGRHPAPHQPGQPGFK
FTVAESCRIKDEFQFLQAQYHSLKVEYDKLANEKTEMQRHYVMYYEMSYGLNIEMHKQTEIAKRLNTILA
QIMPFLSQEHQQQVAQVERAKQVTMTELNAIIGQQQLQAQHLSHATHGPPVQLPPHPSGLQPPGIPPVVG
SSSGLLALGALGSQAHLTVKDEKNHHELDHRERESSANNSVSPSESLRASEKHRGSADYSMEAKKRKAEEK
DSLRYDSGDGKSDLLVVDVSNEDPATPRVSPAHSPPENGLDKARSLKDDAPTSPASVASSSSTPSSKTKD
LGHNDKSSTPGLKSNTPTRNDAPTPTGTTTTPGLRSMGKPPGMDPIGIMASALRTPISITSSYAAPFAMM
SHHEMNGSLTSPGAYAGLHNIIPPQMSAAAAAAYGRSPMVSFGAVGFDPHPPMRATGLPSSLASIPGGK
PAYSFHVSADGQMOPVFPFDALAGPGIPRHARQINTLSHGEVVCVAVTISNPTRHVVYTGKGCVKIWDISQ
PGSKSPIQLDCLNRDNYIRSKLLPDGRTLIVGGEASTLTIWDLASPTPRIKAELTSSAPACYALAI SPD
AKVCFSCCSDGNIAVWDLHNQTLVRQFQGHDTGASCIDISHDGTGLWTGGLDNTVRSWDLREGRQLQOHDF
TSQIFSLGYCPTGEWLAVMESNVEVLHHTKPKDKYQLHLHESCVLSLKFAYCGKWFVSTGKDNLLNAWRT
PYGASIFQSKESSVLSCDISADDKYIVTGS GDKKATVYEVY

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; hKCa4; KCa3.1; hKCa1; KCNN4

GTCCTTCGGTGTCTGGGTGTGGTGTGAGTAGAGGTGTGTGTGCACAAAGTACAGACCATTGTGTGTGACAAAGC
CCATCGTGTGTCTGTGTGTGTCTTTATCCACGTGGATGGACGTCTCTTTCTTGGCTCTGCCCAAGACACAC
CCTAGCCCCCTCCTTATTCTCAAAGGGGGAGCTGGGGAGCCTCCCCCTACCCTGGGGCCTCCCCCTGCCCT
CCCCGCCCTGCCTGGCCGTCACTACTCCCCAGAGGGCAGGGGCTCTGCTGTGCCTCAGAGCAAAAGTCCC
AGAGCCAGCAGAGCAGGCTGACGACCTGCAAGCCACA GTGGCTGCCCTGTGCGTGTGCGAGGTGGGGGAC
CCTGGGCAGGAAGCTGGCTGAGCCCCAAGACCCCGGGGCCATGGGCGGGGATCTGGTGTCTGGCCTGGGG
GCCTTGAGACGCCGAAAGCGCTTGTCTGGAGCAGGAGAAGTCTCTGGCCGGCTGGGCACTGGTGTGTCAGG

AACTGGCATTGGACTCATGGTGTCTGCATGCAGAGATGCTGTGGTTTCGGGGGGTGTCTCGTGGGCGCTCTACC
TGTTTCCTGGTTAAATGCACGATCAGCATTTCCACCTTCTTACTCCTCTGCCTCATCGTGGCCTTTTCATGCC
AAAGAGGTCCAGCTGTTTCATGACCGACAACGGGCTGCGGGACTGGCGCGTGGCGCTGACCGGGCGGCAGGC
GGCGCAGATCGTGTGGAGCTGGTGGTGTGTGGGCTGCACCCGGCGCCCGTGCGGGGCCCGCCGTGCGTGC
AGGATTTAGGGGCGCCGCTGACCTCCCCGAGCCCTGGCCGGGATTCTGGGCCAAGGGGAAGCGCTGCTG
TCCCTGGCCATGCTGCTGCGTCTCTACCTGGTGCCTCCGCGCCGTGCTCCTGCGCAGCGGCGTCTGCTCAA
CGCTTCTACCGCAGCATCGGCGCTCTCAATCAAGTCCGCTTCCGCCACTGGTTTCGTGGCCAAGCTTTACA
TGAACACGCACCCTGGCCGCTGCTGCTCGGCCTCACGCTTGGCCTCTGGCTGACCACCGCCTGGGTGCTG
TCCGTGGCCGAGAGGCAGGCTGTTAATGCCACTGGGCACCTTTTCAGACACACTTTGGCTGATCCCCATCAC
ATTCCTGACCATCGGCTATGGTGACGTGGTGCCTGGGACCATGTGGGGCAAGATCGTCTGCCTGTGCACTG
GAGTCATGGGTGTCTGTGCACAGCCCTGCTGGTGGCCGTGGTGGCCCGGAAGCTGGAGTTTAAACAAGGCA
GAGAAGCAGTGCACAACCTTCATGATGGATATCCAGTATACCAAAGAGATGAAGGAGTCCGCTGCCCGAGT
GCTACAAGAAGCCTGGATGTTCTACAAACATACTCGCAGGAAGGAGTCTCATGCTGCCCGCAGGCATCAGC
GCAAGCTGCTGGCCGCCATCAACGCGTTCCGCCAGGTGCGGCTGAAACACCGGAAGCTCCGGGAACAAGTG
AACTCCATGGTGGACATCTCCAAGATGCACATGATCCTGTATGACCTGCAGCAGAATCTGAGCAGCTCACA
CCGGGCCCTGGAGAAACAGATTGACACGCTGGCGGGGAAGCTGGATGCCCTGACTGAGCTGCTTAGCACTG
CCCTGGGGCCGAGGCAGCTTCCAGAACCCAGCCAGCAGTCCAAGTAGCTGGACCCACGAGGAGGAACCAGG
CTACTTTCCCCAGTACTGAGGTGGTGGACATCGTCTCTGCCACTCCTGACCCAGCCCTGAACAAAGCACCT
CAAGTGCAAGGACCAAGGGGGCCCTGGCTTGGAGTGGGTTGGCTTGGCTGATGGCTGCTGGAGGGGACGCT
GGCTAAAGTGGGTAGGCCTTGGCCACCTGAGGCCCCAGGTGGGAACATGGTCACCCCCACTCTGCATACC
CTCATCAAAAACACTCTCACTATGCTGCTATGGACGACCTCCAGCTCTCAGTTACAAGTGCAGGCGACTGG
AGGCAGGACTCCTGGGTCCCTGGGAAAGAGGGTACTAGGGGGCCCGATCCAGGATTCTGGGAGGCTTCAGT
TACCGCTGGCCGAGCTGAAGAAGTGGGTATGAGGCTGGGGCGGGGCTGGAGGTGGCGCCCCCTGGTGGGAC
AACAAAGAGGACACCATTTTTCCAGAGCTGCAGAGAGCACCTGGTGGGGAGGAAGAAGTGTAACTCACCAG
CCTCTGCTCTTATCTTTGTAATAAATGTTAAAGCCAGAA

VAALCVLRGGGPWAGSWLSPKTPGAMGGDLVVLGLGALRRRKRLLEQEKSLAGWALVLAGTGIGLMVLHAEM
LWFGGCSWALYFLVKCTISISIFLLLLCLIVAFHAKVQLFMTDNGLRDWRVALTGRQAAQIVLELVVCGL
HPAPVRGPPCVQDLGAPLTSQPWPFGFLGQGEALLSLAMLLRLYLVPRAVLLRSGVLLNASYRSIGALNQV
RFRHWFVAKLYMNTHPGRLLLGLTLGLWLTTAWVLSVAERQAVNATGHLSDTLWLIPIITFLTIGYGDVVP
TMWGKIVCLCTGVMGVCCTALLVAVVARKLEFNKAKEKHVHNFMMDIQYTKEMKESAAARVLEAWMFYKHTR
RKESHAARRHQKLLAAINAFRQVRLKHKRLREQVNSMVDISKMHMILYDLQNLSSSHRALEKQIDTLAG
KLDALTELLSTALGPRQLPEPSQSK

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

GAGGGAGGAGGAGGAGTGGGGACCGGGCGGGGGTGGAGGAAGAGGCCTCGCGCAGAGGAGGGAGCAATTG
AATTTCAAACACAAACAACCTGCACGAGCGCGCACCCACCGCGCCGGAGCCTTGCCCCGATCCGCGCCCGCC
CCGTCCGTGCGGCGCGCGGGCGGAGACGCCGTGGCCGCGCCGGAGCTCGGGCCGGGGCCACCATCGAGGC
GGGGCCGCGCGAGGGCCGGAGCGGAGCGCGCCGCCACCGCCGCACGCGCAAACCTTGGGCTCGCGCTTCC
CGGCCCGCGCGAGCCCGGGCGCCCGGAGCCCGCATGTCGCGATCCAACCGGCAGAAAGGAGTACAAA
TGCGGGGACCTGGTGTTCGCCAAGATGAAGGGCTACCCACACTGGCCGGCCCGGATTGACGAGATGCATA
GGCTGCCGTGAAATCAACAGCCAACAAATACCAAGTCTTTTTTTTTCGGGACCCACGAGACCGCATTCTGG
GCCCCAAAGACCTCTTCCCTTACGAGGAATCCAAGGAGAAGTTTGGCAAGCCCAACAAGAGGAAAGGGTTC

AGCGAGGGGCTGTGGGAGATCGAGAACAACCCTACTGTCAAGGCTTCCGGCTATCAGTCCTCCCAGAAAA
GAGCTGTGTGGAAGAGCCTGAACCAGAGCCCAGCTGCAGAGGGTGACGGTGATAAGAAGGGGAATGCAG
AGGGCAGCAGCGCAGGGAAGGAAGCTGGTCATTGATGAGCCAGCCAAGGAGAAGAACGAGAAAGGAGCG
TTGAAGAGGAGAGCAGGGGACTTGTCTGGAGGACTCTCCTAAACGTCCAAGGAGGCAGAAAACCTGAAGG
AGAGGAGAAGGAGGCAGCCACCTTGGAGGTTGAGAGGCCCTTCTATGGAGGTGGAAAAGAATAGCACCC
CCTCTGAGCCCGGCTCTGGCCGGGGCCTCCCCAAGAGGAAGAAGAAGAGGAGGATGAAGAGGAAGAGGCT
ACCAAGGAAGATGCTGAGGCCCCAGGCATCAGAGATCATGAGAGCCTGTAGCCACCAATGTTTTCAAGAGGA
GCCCCACCCTGTTTCTGTCTGTCTGGGTGCTACTGGGAAACTGGCCATGGCCTGCAAACCTGGGAACC
CCTTTCCACCCCAACCTGCTCTCCTCTTCTACTCACTTTTCCCACTCCAAGCCCAGCCCATGGAGATTGA
CCTGGATGGGGCAGGCCACCTGGCTCTCACCTCTAGGTCCCATACTCCTATGATCTGAGTCAGAGCCATG
TCTTCTCCCTGGAATGAGTTGAGGCCACTGTGTCTTCTTCCGCTTGGAGCTATTTTCCAGCTTCTGGG
GCCTGGGACAACCTGCCACCTCCTGACACCTTCTCCCCTCTCCTAGGCATTCTGGACCTCTGGGTTG
GGATCAGGGGTAGGAATGGAAAGGATGGAGCATCAACAGCAGGGTGGGCTTGTGGGGCCTGGGAGGGCAA
TCCTCAAATGCGGGGTGGGGGCAGCACAGGAGGGCGGCCTCCTTCTGAGCTCCTGTCCCCTGCTACACCTA
TTATCCCAGCTGCCTAGATTAGGGAAAGTGGGACAGCTTGTAGGGGAGGGGCTCCTTTCCATAAATCCTT
GATGATTGACAACACCATTTTTTCTTTTGGCGACCCCAAGAGTTTTGGGAGTTGTAGTTAATCATCAAGA
GAATTTGGGGCTTCCAAGTTGTTCCGGCCAAGGACCTGAGACCTGAAGGGTTGACTTTACCCATTTGGGTG
GGAGTGTGAGCATCTGTCCCCCTTTAGATCTCTGAAGCCACAAATAGGATGCTTGGGAAGACTCCTAGCT
GTCCTTTTTCTCTCCACACAGTGTCAAGGCCAGCTTATAGTCATATATATCACCCAGACATAAAGGAAA
AGACACATTTTTTAGGAAATGTTTTAATAAAAGAAAATTACAAAAAAATTTTTAAAGACCCCTAACCT
TTGTGTGCTCTCCATTCTGCTCCTTCCCATCGTTGCCCCATTTCTGAGGTGCACTGGGAGGCTCCCCTT
CTATTTGGGGCTTGTGACTTTCTTTTTGTAGCTGGGGCTTTGATGTTCTTCCAGTGTCAATTTCTCATCC
ACATACCCTGACCTGGCCCCCTCAGTGTGTACCAGATCTGATTTGTAACCCACTGAGAGGACAGAGAGA
AATAAGTGCCCTCTCCACCCCTTCTCCTACTGGTCTCTCTATGCCTCTCTACAGTCTCGTCTCTTTTACCC
TGGCCCTCTCCCTTGGGCTCTGATGAAAAATTGCTGACTGTAGCTTTGGAAGTTTAGCTCTGAGAACCCT
AGATGATTTTCAAGTTCTAGGAAAATAAAACCCGTTGATTACTATAAAAAAAAAA

REEEEWPGGGWRKRPRAEEGAIEFQTQTTARARTHRAGALPRSAPAPSVRRAGGDAVAPELGPATIEA
GAARGPERSGAATAARANLGSRFPARRGARGARSPAMSRSNRQKEYKCGDLVFAKMKGYPHWPARIDEMPE
AAVKSTANKYQVFFFGTHETAFLGPKDLFPYEESEKEKFKGNPKRKGFSGLWEIENNPTVKASGYQSSQKK
SCVEEPEPEPEAAEGDGDKKGNAEGSSDEEGKLVIDEPAKEKNEKGALKRRAGDLLEDSPKRPKEAENPEG
EEKEAATLEVERPLPMEVEKNSTPSEPGSGRPPQEEEEEEDEEEEAATKEDAEAPGIRDHESL

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: FKS3; MGC64992; YPEL1

CAGGCGGCCCGCGGGGCGGCGGGGAGGATGCCTCGCGCCCTGCCTAGGCGTTAACGGCCTCAGCGCGTC
CCGGGCCCGCCGGGAACGCCTGAGAGCCGAGCCCGCGCTGACCGGGGCCCGGGCCGGATGGGCGCTGCGG
GCCGGGGCGCGGACCGCGGAGCGGCCGTTACGTTTTTCTCTCGTCCCAGCTGTGTGGACAGTGCCACACGC
CCTCCTGGACAACGACCGCTCTTGCCTGATTTTCCAGTGTGCTTTTGTGAGAGAGGCCAGCAGAGCCAC
GGTTCTTCCAGAACCAGCCCTGAGCTGAGTGAGGAGTGCCAGGAGAGATGGTGAAAATGACAAAGTCCAA
AACTTTCCAAGCGTATCTGCCGAACGTGTCACCGAACGTACAGCTGTATCCACTGCAGAGCACACCTGGCCA
ATCATGACGAGCTCATCTCCAAGTCTTTCCAGGGAGCCAGGGACGCGCCTACCTCTTCAATTCCTGGTG
AACGTGGGCTGCGGCCCTGCAGAGGAGAGGCTCCTTCTCACCGGGCTGCATGCGGTTGCCGACATCTACTG
CGAGAACTGCAAGACCACGCTCGGGTGGAAATACGAGCATGCCTTTGAGAGCAGTCAGAAATATAAGGAAG

GAAAATTCATCATTGAGCTTGCTCATATGATCAAAGACAATGGCTGGGAGTAA TGTGCGAACTTTCCCTTC
TCCTTTGAATGCTGTTTTGTGAAAGAACTGTGAATGTAATGGAAACGTAGGAGCATCTGGTGACAGCCTT
TCTTGGCCTCTGACCTCAAAGGCTAGCTGCGCATAGCTCTTGACACTCTCGCCATCTCTGTGGGTAAGGT
GTCCCTCGGATCTGTCTCTTCGTGTACACAGTTGTTTCTGAAAATTTTCAATGAGCTTTTTCTAACTTCT
CAAGTTCTAGAGAAAAGAAATTAACCAACTGATGACTTACCTGCCTAGTTAATATCTTCTTTTACCTTTGTC
TTCAATATAGTTGGGCTCTGCTTTTTTAAAGTTTCAGTTGAAAACCAAACCTGGGGCCGGGTGCGGTGGCTCA
CGCCTGTAATCCCAGCACTTTGGGAGGCCAAGATGGGTGGATCACCTGAGGTCAGGAGTTCTAGATCAGCC
TGGCCAACATGGTGAACCCCATCTCTACTAAAAATACGAAAATTAGCCGGGCATGGTGGCGAGTGCCTGT
AATCTTAGCTACTCAGGAGGCTAAGGCAGGAGAATCACTTGAACCTGGGACACGGAGGTTGCAGTGAGCTA
AGATCATGCCATCGCACTCCAGCCTGGGGGACAAAAGTGAGACATCGTCTCAAAAAAAAAAAAAAAAAAGTG
GGTATGGTGGCGCATGCCTTAAATCCCAGCTACTCGGGAGGCTGAGGCACGAGAATCACTTGAACCCAGGA
GGCGGAGGTTGCAGTGAGCCAAGATCGCGCCACTGCACCTCCAGCCTGGCAATAGGGCGAGACTCCGTCTCA
ATTTAAAAACAAAAGAGAACCAGACTGAGTCTCTGAAGACCACAGGGACAGGGTCTCTTTAGATAGCAAGTC
TCACCATTCCCTTTTTTGTAGAGAAAAGGTATTGTAGCCACCCTCCACCCGCTGTTTTTCTTAAATTTGCA
GAACTTCAAATTTGGCTATTCTCTTGTCAAATGAACCTTAAAGTACAGTGTTATTTAAGAATCTTCCAGAG
GCAGTCAACAGACTTATACTAAGGGCATTTTTTGGTTTTTGTAGCTTGTTCAAAAACAGAGGCCAGCACAGA
TGACATTTTAGATACACTCTAAATTTGAGAATGGTGTCTAGTGGAACATGTTTATTTAAGCCAGTAGATTCC
TTATCTAGAAAAGCAGGTGAGCTAGCCCTTAGAGAAGGCTGTCCCGGGGCCGAGAGGTGCCCTTACTGAG
GTGACAGCCTCACAGGGTCTGGTACCAGGGGTTGTGCCCTCAGCAGTGACAGCAGCTTAGGTGTCAGGCAG
TTGCTGAGTGGCTGGTCCATGTCTATAGAGTAACACACTGGACCGAGGAAAAGTCAGATTTTCATTTTCTAC
CCTGGATGTACTTTGAAGAAAAGAATTATTTTTGCATATGAAAGAGGCCAGAACCACAGGAAAACCTTCA
AACTTGACATTTGCCAGAATGTTTAAATTTGTTTCAAAAAGGTTAAAGCAACAAGTTTAGCCTTTGTGC
ATGAAGACGCCTGGCCTGCTAGACGCGTTGCCCGTCCCTGCGTGGTGTCTGCCATGTCACTTGAACCTGAT
AGAGGGGCTGTGCAATCTCCTAAGGCCTGTGTTTTCTGCCATATATTTTATTATAAATTACAATCCACTCA
TCCACCTGCCCTCCACCAGGAGTGGGCACCCATAAGGGTTTAGGCCACTTTGCAGAGGATGGAGGTCAA
AACCCTCCAGATAAGTTTGGTTTTTCAACATTTAGTAACTTGTCTCAGGGCAGAGGGCAGGCAGGGGGAC
CGAGGGGCAGCAGATAGGAGAGCACTGAGCCCGGATAGTTCTCAGCCTGGCAAGTGGCTCTGAAGCTGCCT
TCAGACAAGGCTAGTCTAGGGGCAAGAGTGCAGCTGGCTGACAATAAGAACGTGGCCACCTGCCAGCTT
CACACTCCCCGACTTCCAGCCCTTCTTAAACCCAGACTGCGGTCCAGGCAGGCAGGCAGGCAGGCAGGC
CTCAGACTCACTGCCACACAGCATGCCTTTGGGTGCCATCTCTTTGCCCAAGCCTGGAAGCCTTTGGCAG
GTGGGAAATGCCGCTGCCCTGGTGGGCATGGCACTGAGATGCATCCACTCAGCAGGAGTGACAGAGGCAGA
AGTTCTTTTAAAGCACATCTTCCACTTAGGAAAGGAAGGAAATCTTTGTACTGTCTTGAAGCCTCCACAT
CCGGCTATGGCCCTGCAAGCTGCTTTATCCCTGCGCTAGTCTCCCCCGAGGGTTTTAGGCTGGCCCAGCACA
TCCTGTCTCTGAGCTCGCGTGCAGCCACCCAGAGCGCAGGGGTCACTGCACGCTGCAGGGCTCTTGCTG
CCATGGTCTCAAGCCTGAAGAGGCTCCGCCACAAGCTGGCCCATGAAGTTAGCAATGCCTGTGGCTTCCAG
TCAATTGTCTTGAGACTGTGAAGAGGCTGAAAGACACCTTCCCGGGTGAAGAAGGAGTTCACTGAAAAC
TATCTTAAACTGACCCTTCCCTTTGAGTGAGTCTTCACTTCTCTCCATGTGGGAACCCAGCCTCCGATGC
CCCGGGGACTAGGGGAAACAGTTGGAGGTTCTGTCGCTCCCGAGCCTGCCACGGGTGCGAGGACAGCCAAG
TCCTGAGTGACTCAAGATGCTTCACTTACATGGAAGAACTTCTAAAACCTTACCAGAGTGGTTTTTGTATA
TACTAAAGTTCTATTTAGAGCTTTTTCTGTTTTGGGCAAGTTCGCTGCTCCTTCTATTTGGGCACCTTTGGTT
TTTGTACTGTCTTTTGTGACGGCATTGATTGAACATTTTTTACTAGTAGTCTTATGACTTTTTGTATTTTTT
TTTTTTTTTTGTAATTTATACCAACAACACTTTTTATCACTTTTTTTTTTTGTTGGGCTTCTGCAAAATACAAGC
TCATTTTTTAAACCAATGAACAGACCATGAGCTGGCTTCCAGGGGAAGTGCTATTACAGGACCATATCCAC
CACCTCTTAAATTCCTAAACAATATCATCTAGGACTTCTATTTAAGTTATTTAAAATAAATCTTCCCTTGA
GAGCCTTGGGAGGTGATGTGAGGTTATAAATGGCACAGTGCATTTGCTGTAGGAATGTGGTTTGGCATTG
TTTTATAACACACAGTATTTTTTATACCTTAATGCTTATTTCTTGATGGCATCTGTGAGATATTAGAATTGAA
AATAAGAATCTTCCCAAAATCCTTTAATTTACCTGATGCCCTCATCAGGTCGTTAAAATTTCAAATGGTTTT
TAATAGCTAAAAACTACAATTAAGCTCTAAAACAAACAACTACAGAAATGTAACCTTCAATTTGCCAA
AGGTCTTGGTGGCCTGTCCCCTGCCCTGGGAGCAGATGGCCCTGAAGCCCTTCCCTCACTGTGCAGGCCA
CCGGGTGAGGCTGGACGGTCAACCATGGTGGCTTCACTGCAAGGAGCAGGACTGCCGAGCTCAAGCACGGG
GCCTTCACTTCCCCTGTCTCTGGCCACACCGCCAGCCCTTGGTCCTTATCTGTGTGAGGTTTACAAATA
AAGCTTCTGATGTCAAATGTTTAAAAAAAAAAAAAAAAAAAAA

TALACIFQCCFCQRGQOSHSSRTSPELSEECPGEMVKMTKSKTFQAYLPNCHRYSIHCRAHLANHDEL
ISKSFQGSQGRAYLFNSVVNVGCGPAEERVLLTGLHAVADIYCENCKTTLGWKYEHAFFESSQYKEGKFI I
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

TCAGCTCAGCCAGGGAGCTCAGCGGAGCTGCGCGCCTCCGCCTCCAGCTCCCCTGCCGCAGCGCGCCGCAG
CCGGGCGTCCCCGCGCGGGCGCCGGAGAGGAAGGAAGGCTGGCAGCCTCGTACGTGTCCGCTGCAGTCGC
GAAACAGTTCGGTGGTGAGGAGACCTTTCCAATATAAGAGGAATAAGAAGTCACCTCCCCAGCTGTCA
TCATCTCCAGCAGATTGACCAAGAATTTTTGAGCACTACAGGAAAGACAGTCCATCAAACCCGAGATGA
TGATCAGCCACGTGATTTTTCAAGAAGAGGAATAGGGTGAATGAATCTCATCAGAAAAGCAGCAATATGA
ATGCTGGCCCATCTTGGAAATAAAGTGCAACATTCAAAGAATTCTTCAGGAAAAAGGCAGAGTAAATCCCAA
GTACCCACGCTTCTTCCAGCCGAGAAGCAGCCTCACAGCTGTCAACCAGCCTACTGAAGAAAACTTAA
AGAAAGCATTTCGCCGAAGCAAGACGCAAAAGGAATCCACTCGGTTCCAGGTGTGAGGGGGCCTCAGGGA
ATAAACTGTTTCTTGATTTTCAGTCAATGAAAATTATTAAGAGAATGCTGATGAGGACAGTGCAAGTGAT
CTCTCTGATTCGAAAGAATTCCCATTCCTCCTTCTCCCCTCACACCTCCAGATCTCAATCTTCGAGCTGA
AGAAATTGATCCAGTTTACTTTGATCTTACCCTGGTCAGGGCCATACAAAACCTGAATACTATTATCCTA
ATTTCTTCCATCCCCTTTTCAGCTCCTGGGACCTACGAGATATGGCCCTGCTTCTGAACGCAGAGAACAAA
ACGGAAGCCGTGCCCGAGTGGGAGGACTTCTTGGGAAGTATATCGATAGACTTATTTCAGCTTGAGTGGCT
GCAAGTCCAGACTGTACAGTGTGAAAAAGCAAAGGGGGGCAAAGCAAGGCCCCCTGCCCCTGGGACCT
CAGGGGCACTGAAAAGCCCTGGGAGAAGTAAGCTAATTGCTAGTGCTCTGTCCAAGCCACTACCTCACCAG
GAAGGGGCTTCAAAGTCAGGCCCTTCCGAAAGAAAGCTTTTACCATGAAGAAATCCACCCATCACATTA
TGCATTTGAGACTTCCCCTAGACCCATTGATGTGCTTGGTGGTACCAGGTTTTGTTCTCAGAGGCAAACCC
TTGAAATGAGGACAGAAGAAAAGAAAAGAAATCAAGTAAGAGTACGAAGCTGCAGCGCTGGGATCTGTCC
GGCAGTGAAGCAGCTCTAAGGTGGAACCAGCGGTACATTCGAGTTCCCAAACAGGCAGCTGTGATTCT
GGACTCAGCAGATTCTGTAAAGCCCTCCAAAACACAAGCACATGCACATCCTAGGAAAAAGGAAAGGCAG
AGAGCTGTGGTTCATGCCACTGTATCGAGTGAGAAAAAACTGAAAACAAACGGAGTAAAGCAAACACATAT
AAACTAAAAATAATCTAAAATGCTGAATTTGCCAAGACCTGCAGGTACCTCAATGTTAGAGCGCTTTCAA
AAGTCAAATACTGTGAATTTTAAGGAATTTTACAATACTGACATTTAAGTAGTTGACTGGCATTTTTGT
CCACCTTTATTTCTACCCTGAGTGGGGTTATTTTCAAAGGAAGTGTCTTTCAATAAGCCTTTCTTTGTAT
TGTCAGTCTTAGGCAAATGAGAGCCCTTTAGATAAAAATTATGTAATAATATGTCCATATAAAGGAATAAA
ATGGCACCTCTCCAGGAAAGTGTGAGTAAACCTCAGCTACAGTAGCCGGTCTGTGTAGAGCAGCTAGTG
GTGTTACCTCCCCATTTTACATGCACGTAAGTATATGAAATAGTGCAGACTGTTTCAAATGGTGTGGAAT
CCTAAATGTTTAAATAAGGTCTTCTTGCCCACTCCCTCGCTTACTTTTTTATAAACTCCTCAAGCAAAA
TTTCTGTTCATTTTACCCTTAGGAGAAGCTTTAGTTCTTCTCAAGTCAGGGAGTAGTGAGTTTGTATTTT
GAGTAGTCATTTCTACTAAGCTGGTTGCTTTCTAGAGAGACAGTGAATCTAGTACTTTAATACATTTTC
TCTGACATGGTTTTTTTTTTCTTTTTTGGGGGCATTTTAACTTAGAGGTGGTGGTAAAACCTACTTTTG
AGTTCTCCGAAGTGAAGTTAAAATAACTTGCAGAATTTTCAAAGTCAATGGGCTTAGCATGATTACTGCT
GTTTGGTGGGGCTGAGAATGAAATATTTGACATTCTGGAATTGCTGGCATGTAAAGCTTCTCCAGAGAGGC
ACCCAGGGAAATCACTCTTTACAATTTGTAAGGAAGGGCCTGTAAGGATCAAACACATGGACCTAC
ATTCAGTGAATAGTTACAAAGTTACTGATTTGGGTTCCACACCCTGTGGTCTTAGTCAAAAATAATGAT
CTGTTTCAGTTTGAAGAGCAGGATTTTATTATTTGCTTGGGGTGAGGGGCGGGAGAGTGGAATATGAGT
AAGGTTGCTGAATGAATTCTAAACTCGCTTATCTGGTCTTCAGGCTTCCCAACTCTCTCCAAGCCTTCTTA
TTTCACTGCAGTTAAATAACATCTTCTGTTCTATAGTTGTGCTGTGAGTTTTCTGTTTCAATTTGCGCA
GTGTATTTTAAATACGGCCCATGTCAATTATAGTTGATTTTATCCCTTTAAACAATTACTGTATTTGTTTTG
ACGTAGAGGTTTCAATTTTTTACCTTGGGGCAAATGAAAACTTGGCATTTTTTCAATTTGGGAACATATA

ATAGCTTGTAAGCTTTTCAGACAGCAGTAAATGTCTGAAAAAATATCAAAAAACAGCATAAAGACAAGATTA
TGTAGCTCTAATTATACGTATATAATTATAAAAAACAATGTGCAAGGGTTATATTTTAAAGGTCTTTTAAAA
TCTGATTTTGATCATACCAATGACATAATATTTTTTATGGTAGCCTTTTACTTTCAAGACTTAATTTTCA
GACTTGTAACAAGTTCCTTCTTACATTCTTTCCCTCTCACACCATCCTACTGGAGAAAGCATACTTTTATGC
TAAGATCTTACTTTAAGCTTTTATGTGAACAAAAGATGTACATATAGTAAGTATTACTTCCGTAGTCCCTC
AAATTTACTATAACTTTTGTACTTAGTATATGTTTTATATTTGGAAAACAGCACTACGCTTAGTTTTCTGTG
TAGTTCCTGAGTGATGTCTGTGTGTTCCCTTGCCCTGCCCCTTTTTTGTGAGCACAGATTAGTCTGTTATCCAT
GGCTGGCACTTCACTTATGATCCTTTCTCTGCTAGATTTTTATGCAGCTCTCTATGAAGTTTCATGGCCCA
TAGATATTCAAAAGCAAGATATTCTATACATATGTGTATATGTATATATACTCCTTATGTTAATACTAAAG
TGTTTTATGCTGAGTTGCTGCCTTTCCCGTCAATGTATCCATGTGCATGCTCTTAGAGACCTTGAATGGTTG
AGGGTAAAGTGATTTATTAGTAATTCTACTTGCCTTGTGTATGTCTGAGCTGAAAACAATCGTGATTAAGA
AATTTAGAGGTGGCTGGCGGTGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGA
TCACCTGAGGTCGGGAGTTCAAGACCAGCCTGACCAACATGGAGAAAACCTGTCTCCACTAAAAATACAAA
ATTAGCCGGGTGTGGTGGTGCATGCCTGTAATCCCAGCTACTCGGAAGTTTGAGACGAGAATCTCTTGAAC
CCGGGAGGCGGAGGTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CCGTTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

IRGIKKSPPQLSSSSSRLSKNILSTTGKTVHQTRDDDDQPRDFFKRNRVNESHQKSSNMNAGPSWNKVVQHS
KNSSGKRQSKSQVPHASSQPRSSLTAVTQPTTEKLKESISPEARRKRNP LGSRCQGASGNKLF LDFQSMKI
IKENADEDSASDLSDSERIPIPPSPLTPPDNLNRAEIEDVPYFDLHPGQGHKPEYYPNPLSPFSSWDL
RDMALLLNENKTEAVPRVGGLLGKYIDRLIQLEWLQVQTVQCEKAKGGKARPPTAPGTSALKSPGRSKL
IASALSKPLPHQEGASKSGPSRKKAHFHEEIHPSHYAFETS SPRPIDVLGGTRFCSQRQTLERTEEKKKS
SKSTKLQRWDLSGSGSSSVKETS GHIRVPKQAAVILDSADSKASKTQAHAPRKKGKAESCGHATVVSSEK
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

ACAGCTCATTGTTGGCAGCTGCCGGGCGGTCTGCGGAGCTGTGAGGGCAACGGAGGGGAAATAAAAGGGA
ACGGCTCCGAATCTGCCCCAGCGGCCGCTGCGAGACCTCGGGCGCCGACATCGCGACAGCGAAGCGCTTTGC
ACGCCAGGAAGTCCCCTCTATGTGCTGCTGAGCCGGTCTGGGACGCGACGAGCCCGCCTCGGTCTTCGG
AGCAGAAATCGAAAAACGGAAGGACTGGAAATGGCAGACCATATGATGGCCATGAACCACGGGCGCTTCC
CCGACGGCACCAATGGGCTGCACCATCACCTGCCACCGCATGGGCATGGGGCAGTTCCTCCGAGCCCCCAT
CACCACCAGCAGCAGCAGCCCCAGCACGCCTTCAACGCCCTAATGGGCGAGCACATACACTACGGGCGCGG
CAACATGAATGCCACGAGCGGCATCAGGCATGCGATGGGGCCGGGACTGTGAACGGAGGGCACCCCCGA
GCGCGCTGGCCCCCGCGCCAGGTTTAAACAACCTCCAGTTCATGGGTCCCCCGGTGGCCAGCCAGGGAGGC
TCCCTGCCGGCCAGCATGCAGCTGCAGAAGCTCAACAACCAGTATTTCAACCATCACCCCTACCCCCACAA
CCACTACATGCCGATTTGCACCCTGCTGCAGGCCACCATGAACGGGACAAACCAGCACTTCCGAGATT
GCAACCCCAAGCACAGCGGGCGGACAGCACCCGCCGGCTCGGGCGGCAGCAGCACCCCTGGCGGCTCT
GGCAGCAGCTCGGGCGGGCGGGCGGCAGCAGCAACAGCGGGCGGGCAGCGGCAGCGGCAACATGCCCCG
CTCCGTGGCCACGTCCCCGCTGCAATGCTGCCGCCAATGTGATAGACACTGATTTTCATCGACGAGGAAG
TTCTTATGCTCTTGGTGTAGAAATGGGTTTGGACCGCATCAAGGAGCTGCCCGAACTCTGGCTGGGGCAA
AACGAGTTTGATTTTATGACGGACTTCGTGTGCAAACAGCAGCCCAGCAGAGTGAGCTGTGACTCGATCG
AAACCCCGGCGAAAGAAATCAAACCCCAACTTCTTCGGCGTGAATTAAGAAACATTCCTTAGACACA
GTATCTCACTTTTCAGATCTTGAAAGGTTTGGAGAACTTGAAACAAAGTAAACTATAAATTTGTACAAATT
GGTTTTAAAAAAAAATTGCTGCCACTTTTTTTCTGTTTTTTGTTTTGTTAGCCTTGACATTCACC

ACCTCCCTTATGTAGTTGAAATATCTAGCTAACTTGGTCTTTTTTCGTTGTTTGTGTTTTACTCCTTTCCCTC
ACTTTCTCCAGTGTCAACTGTTAGATATTAATCTTGGCAAAGTCTTAATCTTGTGGATTTTTGTAGATGG
TTTCAAATGACTGAACTGCATTACAGATTTACGAGTGAAAGGAAAAATTGCATTAGTTGGTTGCATGAACTT
CGAAGGGCAGATATTACTGCACAACTGCCATCTCGCTTCATTTTTTTAACTATGCATTTGAGTACAGACT
AATTTTTAAAATATGCTAACTGGAAGATTAACAGATGTGGGCCAACTGTTCTGGATCAGGAAAGTCAT
ACTGTTCACTTTCAAGTTGGCTGTCCCCCCCCGCCGCCGCCCATATGTACAGATGATAATAGGG
TGTGGAATGTCGTGAGTGGCAAACATTTACAGATTTTTATTTTGTGTTCTGTCTTCAACATTTTTGACACT
GTGCTAATAGTTATATTAGTACATGAAAAGATACTACTGTGTTGAAAGCTTTTTAGGAAATTTTTGACAGT
ATTTTTGTACAAAACATTTTTTTGAAAAATACTTGTAAATTTATTCTATTTTAAATTTGCCAATGTCAATA
AAAAGTTAAGAAA

LDATSPPSVFGAEIAKTEGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPHHHQQQPQHAFNA
LMGEHIHYGAGNMNATSGIRHAMGPGTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQKLNN
QYFNHHPYPHNHYMPDLHPAAGHQMNGTNQHFRDCNPKHSGSSTPGGSGSSTPGGSGSSSGGGAGSSNS
GGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMGLDRIKELPELWLQNEFDFTDFVCKQ
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

GCGACGGCCCTGCTGCCTCTCCAGCCAAGTGGCTGGAGTCGGGAGGCTGGAAAGAGACTCCGAGAAAGTAC
CAGCGGAAGGCGGCCCGCTACGGCGATTTCGAGGGAGTAGCAGACGAAGACGGTGGCCGCCACTAGC
CACCACGTGTGGAGGATAAACGGTCTACACGGCCATTCCGGCGCCGAGTCTAGGGAAAGAGTTAGCGACGA
CGGGGAAAGAAAATGTGAAGAGAGCGACCCCGCTCCAGGGTCGCTGCAGGAAGCCTAAGTGCAGACGCC
GCTTCTCCGCACTGACTTGAGAAGGGTCAGTAAAACCTCGGCCACTGCCGACGCTCTCTAGGGAGAGA
GTTAGGGGAGATAGTGGCCACAGTACAGCTGCTCTTGGGAGAGAGTTAGGGGAGACAGCACCTTCTGCAG
CAGCGACGTGAATTTTGTGAAGTTGGAGGCCACCAAACCTACCGACTCCAGGGGAACAGCCAGAGAAGACC
GAGGCCTCCGCCTCAGTGGTCCTTGGGAGGGAGTCACTGACATTCCGGACCCGCGAAGTGTGACTTCGGG
GATAGAGTCACTGACGATCGCAGTCCCGCTTTCAGTGGCTCCTGGGAGGGAGGGAGTGTGCAAGGCGGCCA
CAGCGTTGGTAGTTCTTGGGAGGAAGTAAGTGGAGACCGCGCTACGCAGCCAGCGACTCCTCTGGTGTGA
GCGGCAGTGAAGACGCCAGCTACCGCTTTCAGTGGCTTTTGGGAGAGAGAAAGTGAAGACGAAGTTTCCGC
TGCAGCTTCTGGGAGAGAGCAAGAGAGGACCTTGGGCCCCGTCTAGTGACGACGGAGAAGAGGGCCGCTG
CCGCTGCAGTGGCTCGTGGGTGAGAGCAAGTGAAGACCGCCGACGATCAGGGGCCTGGACTCAACTCCTC
CCCAGAGTCCGAGGTGTTGCGCCATGCCCCGGGTGGCCAATTCAGGCCCTCCACTTCTCTAGGGAGACT
GCAAACCCCTGTTCCAGGAAGAAGGTGCATTTTTGGCAGCATAATGATGCAGTACGAGCTGGAGATGTA
GCAGCTTTCAGAAATAGTGTGCTTTCATTGGCTGCTCTGGCATGGAGCTGATATCACACACGTAACAACGA
GAGGTTGGACAGCATCTCACATAGCTGCAATCAGGGGTGAGGATGCTTGTGTACAGGCTCTTATAATGAAT
GGAGCAAATCTGACAGCCCAGGATGACCGGGGATGCACTCCTTTACATCTTGTGCAACTCATGGACATTC
TTTCACTTTACAAATAATGCTCCGAAGTGGAGTGGATCCCAGTGTGACTGATAAGAGAGAATGGAGACCTG
TGCATTATGCAGCTTTTCATGGGCGGCTTGGCTGCTTGCACCTTCTTGTAAATGGGGTGTAGCATAGAA
GATGTGGACTACAATGGAACCTTCCAGTTCCTTAGCAGCCATGGAAGGCCACCTTCACTGTTTCAAATT
CCTAGTCAGTAGAATGAGCAGTGCAGCAGCAAGTTTTAAAAGCTTCAATGATAATGGAGAAAATGTACTGG
ATTTGGCCCAGAGGTTCTTCAAGCAGAACATTTACAGTTTATCCAGGGGGCTGAGTATGGAAGAAAAGAC
CTAGAGGATCAGGAAACTTTAGCATTTCCAGTTCATGTGGCTGCCTTTAAGGGTGATTTGGGGATGCTTAA
GAAATTAGTGAAGATGGAGTAATCAATATTAATGAGCGTGTGATAATGGATCAACTCCTATGCATAAAG

CTGCTGGACAAGGCCACATAGAGTGTGGCAGTGGTTAATTAAAATGGGAGCAGACAGTAATATTACCAAC
AAAGCAGGGGAGAGACCCAGTGTGGCAAAGAGGTTTGGCCATTTGGCAGCAGTGAAGCTGTTAGAGGA
GCTACAGAAATATGATATAGATGACGAAAATGAAATTGATGAAAATGATGTGAAATATTTTATAAGACATG
GTGTTGAGGGAAGCACTGATGCCAAGGATGATTTATGTCTGAGTGAAGTGGATAAAAACAGATGCCAGAAGA
CCATCAAAGAAGTGCAGGGCCAGCTGGAGTATGAACGACTACGTAGAGAAAAATTAGAATGTCAGCTTGAT
GAATATCGAGCAGAAGTTGATCAACTCAGGGAAACACTGGAAAAAATTCAAGTCCCAAACCTTTGTGGCTAT
GGTTGGTGTCTCTTTTAAATACATTTATTTTCTCAAGAAGTATATACAAGAGTGGCCAAGAGTACAAGCTT
TGGGCTGGGGCTCTTTGGACTTGAATCCTGGCTATAGTCTTTTTTGTATTTGGGCTACTCATTTAACCTTTT
TGTGCCTCAACTATAAACTGGAGATGATAACTTATACTTTCTAGAGCTGATGACAGGATTAAGGAATACA
CACACACACACACACACACACACACACACACTCTATATGTAACTATAACTTTCTAGATACATACA
CACATCTGTCTATCTTCAAACAGCAGCCTAGTTCTATAAGTATTATTGTTAACATTTGTACCATAGAAG
GAGAAAATGTTTTCTATAAGTAATCAATCAGAATTCTAAGACAGCTAGAGGATATGTATATTTTAAATTTCT
AAATAACCAAATAAAGTGCTTTGAATGGAATCCATATTTTCTTTCCATAGGGAAGTTTCTTCATCAATCA
TCATGGATGAATTAATTCTGTTTGGAGCTTGTGTACCTCAAATCTGATCTATTAATATTATAGAATCTTC
CTGGCTTCTTTTTATTTTGCACAACTAGTCACTGGTCTCTTTAATAACCACTTTAAAAATATTTAAATAA
TAAACATGTTTGTGTATAGTGTAAAAACAAGATAGATGTTCTTCTGACACTAAGTTCCACTCTCCAG
AGGAAGCTATTGTTAACAATTTAGTACATACCATATTAGACAAAATTCTATTTATCTGCAAACATGTATGT
GTATTTCTATGCATATGCAAATATAACCACTTCTCAGTTTTCTCTGGGTTTTGGTGTCTCACCAGTGGTTG
ATTTTCCAAGAAATATAATTAACTATAATAATT

LERDSEKVP AEGGRRY GDSQGV ADEDGGRRTSHHVWRINGLHGHSGAESRERVSD DGERKCEESDRRSRVA
AGSLADAGFSRSDLRRVSEN LGHCRSVSRERV RGD SGHSHSCSWERV RGDSTFCSSDVNFSEVGGHQTTD
SRGTAREDRGLRLSGPWEVSDIRDPRTSDFGDRVSDDR SRRFSGSWE GGSVEGGH SVGSSWEEVSGDRGY
AASDSSGVSGSEDASYRFSGFWERESEDEGFRC SFWERAREDLGPRPSDDGEEGRCSRCSGSWVRASEDRRS
IRGLDSTPPQSRCCAMPGVANS GPSTSSRETANPCSRKKVHF GSIHDAVRAGDVKQLSEIVCLHWLLWHG
ADITHVTTTRGWTASHIAAIRGQDACVQALIMNGANLTAQDDR GCTPLHLAATHGHSFTLQIMLRSGVDPSV
TDKREWRPVHYAAFHGRLGCLQLLVKWC SIEDVDYNGNLPVHLAAMEGHLHCFKFLVSRMSSATQVLKAF
NDNGENVLDLAQRFFKQNILQFIQGA EYEGKDLEDQETLAFPGHVAAFKGDLGMLKKLVEDGVININERAD
NGSTPMHKAAGQGHIECLQWL IKMGADSNITNKAGERPSDVAKRFAHLA AVKLLLEELQKYDIDDENEIDEN
DVKYFIRHGVEGSTDAKDDLCLSDLDKTDARRPSKNCRASWSMNDYVEKN

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

TAGGAGCTGCTCCTTCATCACTGGAGTCGCCCCCTACCTCTCTGCCCCAGCCCCGAGAGCCCCAGGCGGGGA
CCCCGGATCGGACGTCCCCAAGCCTCCGGGCACCTGGCTCAGCAGGAGGCCCGGCTCGGGGCAGGGCA
GGGCCGGCGGGCGAGCCGGAGCCCGCCCCCTGCCGGGCCCGCCGAGCCCTCGGAGCCACCCATGGG
GCACCTGCCCTTGGCCTCCTTGCCGGCCGCGCCAGCCCGGCGTCCCGAGCAGCGCAGGGGAGGATCC
CCGCGCAGTGACCCGGGAGCCACCACAGACTCTGGGAGGCTCGGCGGCTGGAGCAGCAGCCAGCTCCCCG
AGCTCCCGCTTCCAGGCAGCTCTCTGAGCCTGCGCAGAGGCCCGGCCATTCCAGCCCCGAGCC
ATGATGAAGACTTTGTCCAGCGGGAAGTGCACGCTCAGTGTGCCCGCCAAAACTCATACCGCATGGTGGT
GCTGGGTGCCTCTCGGGTGGGCAAGAGCTCCATCGTGTCTCGCTTCTCAATGGCCGCTTTGAGGACCAGT
ACACACCCACCATCGAGGACTTCCACCATAAGGTATACAACATCCGCGGCGACATGTACCAGCTCGACATC
CTGGATACCTCTGGCAACCACCCCTTCCCCGCCATGCGCAGGCTGTCCATCCTCACAGGGGATGTCTTCAT
CCTGGTGTTCAGCCTGGATAACCGGGAGTCTTCGATGAGGTCAAGCGCCTTCAGAAGCAGATCCTGGAGG
TCAAGTCTGCCTGAAGAACAAGACCAAGGAGGCGGCGGAGCTGCCCATGGTCATCTGTGGCAACAAGAAC
GACCACGGCGAGCTGTGCCGCCAGGTGCCACCACCGAGGCCGAGCTGCTGGTGTCCGGGCGACGAGAAGT

CGCCTACTTCGAGGTGTGCGCCAAGAAGAACACCAACGTGGACGAGATGTTCTACGTGCTCTTCAGCATGG
CCAAGCTGCCACACGAGATGAGCCCCGCCCTGCATCGCAAGATCTCCGTGCAGTACGGTGACGCCTTCCAC
CCCAGGCCCTTCTGCATGCGCCGCGTCAAGGAGATGGACGCCTATGGCATGGTCTCGCCCTTCGCCCCCG
CCCCAGCGTCAACAGTGACCTCAAGTACATCAAGGCCAAGGTCTTTCGGGAAGGCCAGGCCCGTGAGAGGG
ACAAGTGCACCATCCAGTGA GCGAGGGATGCTGGGGCGGGGCTTGGCCAGTGCCTTCAGGGAGGTGGCCCC
AGATGCCCCTGTGCGCATCTCCCCACCGAGGCCCGGCAGCAGTCTTGTTACAGACCTTAGGCACCAGA
CTGGAGGCCCGGGCGCTGGCCTCCGCACATTCGTCTGCCTTCTCACAGCTTTCCTGAGTCCGCTTGTCC
ACAGCTCCTTGGTGGTTTCATCTCCTCTGTGGGAGGACACATCTCTGCAGCCTCAAGAGTTAGGCAGAGAC
TCAAGTTACACCTTCTCTCCTGGGGTTGGAAGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCAT
ATGGAGCCTCCTGGGACAAGCCTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAAGGTCTCCTCTCTC
CTGGCATAAACACCCAGCTTGGTTTGGGTGGCAGCTGGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGC
TCTGGTTACAGTGCCTCTGCACCCCTCCACCCCTCCACCCACACACACAAGTTGGCCCCCAGCTGCGCCTGA
CATTGAGCCAGTGGACTCTGTGTCTGAAGGGGGCGTGGCCACACCTCCTAGACCACGCCACCCTTAGAC
CACGCCACCTCCTGACCGGTTCTCAGCCTCCTCTCCTAGGTCCCTCCGCCGACAGTTGTGCTTTGTT
GTGGTTGACAGTGTTCGTGTCTGTATAGTAGTAAATGGAATCATTGTACTGTAAAAGCCTAGTGA
CTCCCTCCTTGGCCAGGCCCTCACCCAGTTCAGATCCACGGCCTCCACCCGGGACGCCTTCTCCTCTGCT
CCCAAACAGGGTTTCCGTGGCCTGTTTGCAGCTAGACATTGACCTCCGCCATTGAGCTCCACGGTTTACAG
ACAATTGCACAAGCGTGGGGTGGGCAGGCCAGGACTGCTTTTTTTAATGCTCCCATTTACAGAGGATAC
CACCGAGACTCGGAGGGGACACGATGAGCACCAGGCCACCTTTGTCCCCTAGCAAATTCAGGGTACAGC
TCCACCTAGAACCAGGCTGCCCTCTACTGTGCTCGTTCCTCAAGCATTTATTAAGCACCTACTGGGTGCTG
GGTTCCTGTGCTCTAGGAAACCAAGAGGGTCCCAGTCTGCGCCTCTGCCCGCCCTGCTGCCCCACCAC
CTTCTGCACACACAGCGGTGGGGAGGCGGGGAGGAGCAGCTGGGACCCAGAAGTGGAGCTGGGAGGGATCC
GACAGAAAAGCTCAGGGCGGGTCTTCTCCTTGTGCCCGGGATTGGGCTATGCTGGGTACCACCATGTACTC
AGGCATGGTGGGTTTTGAACCCATAAACCAAAGGCCCTTGTCTCAGCTCTTAACAAGTATATTTTTGTATT
TTAATCTCTCTAAACATATTGAAGTTTTAGGGCCCTAAGGAACCTTAGTGATCTTCTATTGGGTCTTTCTG
AGGTTTCAGAGAGGGTAAGTAACTTCTCCAGGTACACAGCAAGTCTGTGGGTGGCAGAAGCAAGCTAGCG
CTGGGCATTTCAGTACATAACCAGATGTGCTCCCTCTCTTGTGCTTGGCCCCCTGGGGCCTTCAGGGCTTTG
GGACATCTTGTCTCAACCCTCTCCCTAGATCAGTCTGTGAGGGTCCCTGTAGATATTGTGTACACCATGC
CCATGTATATAACAAGTACACACAGATGTACACACAGATGTACACATGCTCCAGCCCCAGCTCTGCATACCT
GCACCTGCACCCAGCCTTGGCCCCCTGCCTGCGTCTGTGCTCAAAGCAGCAGCTCCAACCCTGCCTCTGTC
CCCTTCCCCACCCACTGCCTGAGCCTTCTGAGCAGACCAGGTACCTTGGCTGCACCGGTGTGTGGCCCGCT
CTCACCCAGGCACAGCCCCGCCACCATGGATCTCCGTGTACACTATCAATAAAAAGTGGGTTTTGTTACAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LAQQEAPGSGQGRAGGGEPEPAPCPGPAEPSEPTHGAPAPCASLPGRAQPGVPSSAGEDPRAVTPREPPQTL
GGSAAGAAGSSPQLPALPGSSLRARGPARHSQPRAMMKTLSSGNCTLSVPAKNSYRMMVVLGASRVGKSSI
VSRFLNRFEDQYPTIEDFHRKVYNIRGDMYQLDILDTSGNHPFPAMRRLSILTDGVFILVFLDNRESF
DEVKRLQKQILEVKSCLKNKTKAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENCAFYEVSAKNT
NVDEMIFYVLFMAKLPHEMSPALHRKISVQYGDFAHPRPFCMRRVKEMDAYGMVSPFARRPSVNSDLKYIK
AKVLRGQARERDKCTIQ

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

CGGACGCGTGGGGCGCTGGGAGCCGGAGGGGAGCGCAGCGAGTTTGGCCAGTGGTTCGTGCAGTCCAAGGGG
CTGGATGGCATGCTGGACCCAAGCTCAGCTCAGCGTCCGGACCCAATAACAGTTTTACCAAGGGAGCAGCT
TTCTATCCTGGCCACACTGAGGTGCATAGCGTAATGTCCATGTTGTTCTACACTCTGATCACAGCTTTTTCT

GATCGGCATACAGGCGGAACCACACTCAGAGAGCAATGTCCCTGCAGGACACACCATCCCCAAGTCCACT
GGACTAAACTTTCAGCATTCCCTTGACACTGCCCTTCGCAGAGCCCCGAGCGCCCCGGCAGCGGCGATAGCT
GCACGCGTGGCGGGGAGACCCGCAACATTACTGTGGACCCAGGCTGTTTAAAAAGCGGCGACTCCGTTC
ACCCCGTGTGCTGTTTAGCACCCAGCCTCCCCGTGAAGCTGCAGACACTCAGGATCTGGACTTCGAGGTCG
GTGGTGTGCCCCCTTCAACAGGACTCACAGGAGCAAGCGGTTCATCATCCCATCCCATCTTCCACAGGGC
GAATTCTCGGTGTGTGACAGTGTGAGCGTGTGGGTTGGGGATAAGACCACCGCCACAGACATCAAGGGCAA
GGAGGTGATGGTGGTTGGGAGAAGTGAACATTAACAACAGTGTATTCAAACAGTACTTTTTTGGAGACCAAG
TGGCGGGGACCCAAATCCCCTTGACAGCGGGTGCCTGGGGCATTGACTCAAAGCACTGGAACATATTGTA
CCACGACTCACACCTTTGTCAAGGCGCTTTACCATGGGATGGCAAGCAGGGATAACCCTGGCCGGGTTATCC
CGGATGAGATACGGGCCGGGAAAAGTGGGGCTCAACAAGGGAGGATGAAGAGAAGGAGGCTAAAGGCTGG
CGGAAGGACATCCCCTTCCCCTGGGGCCATTCAAAAACCTTCTGCGGGGGGCCCTCCCTTAACGGGAAAC
CTGGTAAAATGGATTTTTTAAGGGTGGTAA

LASGRAVQAGWHAGPKLSSASGPNNSFTKGAIFYPGHTEVHSVMSMLFYTLITAFLLIGIQAEPHSESNP
AGHTIPQVHWTKLQHSLDTALRRARSAPAAAIAARVAGQTRNITVDPRLFKKRRLRSRVLVSTQPPREAA
DTQDLDFEVGGAAPFNRTHRSKRSSSHPIFHRGEFSVCDSSVSVWVGDKTTATDIKKEVMVVGRSEH

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

GTGAGTGATGATGATAAGCTCCCCTGCAGGTGTGAATAGAGACCCCGGAGGCGCGTCCCTAGCCCTCATCT
GGGGAAGCGCACCTGCATACAGACGGGTGCACCGGGGAGGAGGCGATCTGCCGCGTGTTCCTGCAAGCAGA
AAAGGAGTAACTAAGTGTACATTTGAAGACGAGCACTGAGGATGAGGAACCAACTGAAGAATATGAAAAT
GTTGAAATGCAGCATCTAAGTGGCCAAAAGTGGAGGATCCCTATCCCTGAATCTAAGGTTGGTGACACATG
TGTGGGATAGCAAGGTAGAGAATCAACAGAAAAGCCTGTGGAAAACAGGATGAAGGAGGACAAAAGCA
GCATCAGGGAAGCAATCAGCAAAGCCAAAGAGTACAGCAAATATAAAGACAGAACAGGAAGGTGAGGCATCT
GAGAAGAGCTTGCATCTGAGCCCACAGCATATCACACACCAGACTATGCCTATAGGACAGAGAGGCGATGA
GCAAGGCAAACGTGTGGAGAACATTAATGGAACCTCCTACCCTAGTCTACAGCAGAAAACCAATGCTGTTA
AGAAATTACATAAATGTGATGAATGTGGGAAATCCTTCAAATATAAATCCCCTTGTTC AACATAAAATT
ATGCACACTGGGAAAAGCGCTATGAATGTGATGACTGTGGAGGGACTTTCGGAGCAGCTCGAGCCTTCG
GGTCCACAAAACGGATCCACACTGGGGAGAAGCCGTACAAGTGTGAGGAATGTGGGAAAGCCTACATGTCCT
ACTCCAGCCTTATAAACCACAAAAGCACCCATTCTGGGGAGAAGAACTGTAAATGTGATGAATGTGGAAA
TCCTTCAATTATAGCTCTGTTCTGGACCAGCATAAAAGGATCCACACTGGGGAGAAGCCCTATGAATGTGG
TGAGTGTGGGAAGGCCTTCAGGAACAGCTCTGGGCTCAGAGTCCACAAAAGGATCCACACGGGGGAGAAGC
CCTATGAATGCGACATCTGTGGGAAAACCTTCAGTAACAGCTCTGGCCTTAGGGTCCATAAAAGGATCCAC
ACAGGTGAGAAACCTTACGAATGTGATGAGTGTGGGAAGGCCTTCATTACTTGTAGAACACTTCTCAACCA
TAAAAGCATCCACTTTGGAGATAAACCTATAAATGTGATGAGTGTGAGAAATCTTTAATTATAGCTCTC
TTCTCATTACAGCATAAAGTCATCCACACTGGAGAGAAACCTTATGAATGTGATGAATGTGGGAAGGCTTTC
AGGAACAGCTCAGGCCTCATAGTGCATAAAAGGATCCACACAGGAGAGAAACCTTACAAGTGTGATGTCTG
TGGCAAAGCATTACAGCTATAGCTCAGGCCTCGCAGTCCATAAAAGCATTACCCTGGGAAGAAAAGCCCATG
AATGTAAGGAGTGTGGGAAATCCTTTAGTTATAACTCACTACTTCTTCAAACACAGAACTATTACATACCGGA
GAGAGACCTTATGTATGTGATGTGTGTGGGAAAACGTTTCAAGAAACAAATGCAGGCCTCAAAGTCCACAGGAG

GCTCCATACTGGGGAAAAACCATATAAGTGTGATGTGTGTGGGAAAGCCTATATCTCACGCTCTAGCCTTA
AAAATCACAAAGGAATCCACCTTGGGGAGAAGCCCTATAAATGTAGCTATTGTGAGAAATCCTTCAACTAC
AGCTCTGCCCTTGAACAGCATAAAAGGATTACATACCAGGGAAAAACCCCTTTGGGTGTGATGAGTGTGGTAA
AGCTTTCAGAAATAATTCTGGCCTTAAAGTACATAAAACGAATCCACACTGGGGAACGACCTTACAAATGTG
AAGAATGTGGGAAAGCATAACATCTCTCTCGAGCCTTATAAATCATAAAAGTGTACACCCTGGGGAGAAG
CCCTTTAAGTGTGACGAGTGTGAGAAGGCCTTCATCACATACCGAACCCCTTACAAACCACAAAAAAGTTCA
TCTTGGGGAGAAGCCCTACAAATGTGATGTGTGTGAGAAATCTTTTAATTACACATCGCTCCTTTCTCAGC
ACAGAAGGTCCACACTAGAGAGAAAACCCCTATGAATGTGACAGGTGTGAGAAGGTCTTCAGAAACAACCTCA
AGCCTTAAAGTTCATAAAAGAATCCATACTGGGGAGAGGCCCTATGAATGTGATGTGTGTGGAAAAGCCTA
CATCTCACACTCAAGCCTTATTAACCATAAGAGTACCCACCCTGGCAAGACACCCCATACATGTGATGAAT
GTGGAAAAGCTTTTTCTCAAGCAGAATCTTATAAGCCATAAAAGAGTCCATCTTGGGGAGAAAACCCCTTC
AAGTGTGTTGAGTGTGGGAAATCTTTCAGTTACAGCTCTCTCCTTTCTCAGCACAAGAGGATCCACACAGG
GGAGAAAACCCCTATGTGTGTGATAGGTGTGGGAAGGCCTTCAGGAACAGCTCAGGCCTCACAGTGCATAAAA
GGATCCACACAGGTGAGAAAACCCCTATGAATGTGATGAGTGTGGGAAGGCATACATCTCACACTCAAGTCTT
ATCAATCATAAAAGTGTCCACCAGGGGAAGCAGCCCTATAAATTGTGAGTGTGGGAAATCCTTCAATTATAG
ATCAGTCCTTGACCAGCACAAAAGGATCCACACTGGAAAAGAAGCCATACCGATGTAATGAGTGTGGTAAGG
CTTTTAATATCAGATCAAATCTCACCAAGCATAAAAAGAACCCTACTGGAGAGGAATCTTTAAATGTGATA
TATGTGGGAAGTTATAGTGGCACATCCAGAAAGAGAACCCTATGAGGGAGGGAATGCCCTGGATGGGGGCAG
GATGAGGATGCCTCTGTAGCAGGCAGAGCTTACCAAGTCTCTCCGAACCTCAAATGGAAGAAATACCTTATG
AATGTAAGAATGTAGGGGGTTCATGGCTTGTAATTTACACAGTGTAAATGAAACCATCCTAGAGGATTATGA
GGAATCCTTTCTATGTGATTTTCAATCATAGCAAGCAAGAAAGGCTCCAGTGTCAAGGTAGTTCAGCTCTT
ACAGGATATAAAACAGTCCATACTTGAGAGAAAAACTTAGATCTGAGTGTGGAATGTGAAGCAAATCTTC
AAAATCAGTAGACATTTCTGGACATAAAACACAGATGAGGAAAGGGCTTCAATTAGAAGTTACGTAATCAC
CATCAGAAAGTTCATGTTTGGTAAATTTCTGTTACTAGAAATGTAGGAAATTCAGGTATAGCTTTGAATCCC
AATTACACATTTGGTCAAGTGGGAAAACCTAAGGGCCTCCAACAGGCAAATTCAGGGAGGATAGGTTTCAGGGA
ATATAAATTTATTTAATATTAGTGGTCTTTAAGTATAAACTTGATGTAATTTGGTTTGGGAGGGGGCAGTGA
TGATGACTTCTGAAACAAAATTTGGATTTCTTTTAGGAAAAGTAGAAAGCATAGACTTACAAGTCTAACA
GGAGATAGGAGAGAGTCACTCATAAAAAATGCAAATGATGAACGTACTATTGTGATACATTAGTTGAATG
GATGAAACTTTTTAAAGTTTTAGATGAATCCCAATGAAATGATGAATTTGTGATGAGGATAAATGCTGG
AAGTGGTATTTACACATTATGCTACAATAAAAGTTTCTACCGTGGAGAGGATTTTGCACATTAGTAAGT
AATGGAACACACCGTCAACATGAATTCGCACCTTACATGACAGAAGTGATTGAGGGATTCTATGAATAGA
AATGCTGAGAAGGAACGCATTTTTATTGCAGAAGCTAAAAAGCTAAAGTACCAGTCATCTAGAGAGAAGGAA
ATTAATGTTTCTTAATAATCCTGTTAAATGTTTGGATTGTTTTTGGAAATGTGTTATTGTAAAGATGTCATGC
AGGACATGTATATGTTGTCTGTTGTAATAATGTTAACGAATACTTTGTTTCCAGGGCTCACTCTCTTTTGTCA
TGAAAGCCAGCTCCTTGTGGCGAGGTAAAGTGAATTCGAATAAAGAAATTCCTTAAATCAAAA

VSHLKTSTEEDEEPTEEYENVGNAASKWPKVEDPIPIESKVGDTCVWDSKVENQKKPVENRMKEDKSSIREA
ISKAKSTANIKTEQEGEASEKSLHLSPQHITHQTMPIGQRGSEQGKRVENINGTSYPSLQKKTNAVKKLHK
CDECGKSFKYNSRLVQHKIMHTGEKRYECDDCGGTFRSSSSLRVHKRIHTGEKPYKCEECGKAYMSYSSLI
NHKSTHSGEKNCKCDECGKSFNYSSVLDQHKRIHTGEKPYECGECGKAFRNSSGLRVHKRIHTGEKPYECD
ICGKTFSNSSLRVHKRIHTGEKPYECDCEGKAFITCRTLLNHKSIHFGDKPYKCECEKSFNYSSLLIQH
KVIHTGEKPYECDCEGKAFRNSSGLIVHKRIHTGEKPYKCDVCGKAFSYSSGLAVHKS IHGPKKAHECKEC
GKSFSYNSLLQLHRTIHTGERPYVCDVCGKTFRNNAGLKVHRRLLHTGEKPYKCDVCGKAYISRSSLNKHG
IHLGEKPYKCSYCEKSFNYSSALEQHKRIHTREKPFGCDECGKAFRNNSGLKVHKRIHTGERPYKCEECGK
AYISLSSLINHKSVHPGEKPFKCECEKAFITYRTLINHKVHLGEKPYKCDVCEKSFNYTSLLSQHRRVH
TREKPYECDRCEKVFRRNSSLKVHKRIHTGERPYECDVCGKAYISHSSLINHKSTHPGKTPHTCDECGKAF
FSSRTLISHKRVLGEKPFKCECGKSFYSLLSQHKRIHTGEKPYVCDRCGKAFRNSSGLTVHKRIHTG
EKPYECDCEGKAYISHSSLINHKSVHQKQPYNCECGKSFNYRSVLDQHKRIHTGKKPYRCNECGKAFNIR
SNLTKHKRTHTEEESLNVIVGYSYSGTSQKRTYEGGNALDGGRRMPL

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

AGTTGAAGCTCCTCTCCCTTCCGAACAGCTCCGCGCACCGCCCCGCGACGCAGCCCGGGCGCAACTACTTTTC
 TTCTCTCTCCTTTCTTTCTTCTCTCCTTTTTTCCCCTGCTGGGTAGTGGCTGCGGCGGGGTGGGGGAGACT
 TTGAATGACCGAGCTCGCGTCCACCTTTCTCTTCATGTCGACGTCCCTGGAAACGGCCACACGGATGCCAT
 GGTACTTTTTGCCACGATCTTACAGGTGAACAAGGTGATGTCCATCTTGTGTTTATGTGATATTTCTCGCTT
 ATCTCCGTGGCATCCAAGGTAACAACATGGATCAAAGGAGTTTGCCAGAAGACTCGCTCAATTCCTCATT
 ATTAAGCTGATCCAGGCAGATATTTGAAAAACAAGCTCTCCAAGCAGATGGTGGACGTTAAGGAAAATTA
 CCAGAGCACCTGCCCCAAGCTGAGGCTCCCCGAGAGCCGGAGCGGGGAGGGCCCGCCAAGTCAGCATTCC
 AGCCGGTGATTGCAATGGACACCGAAGTGTGCGACAACAGAGACGCTACAACCTACCCGCGGGTCTGTGTG
 AGCGACAGACCTCCCTTGGAGCCCCCGCCCTTGTATCTCATGGAGGATTACGTGGGCAGCCCCGTGGTGCC
 GAACAGAACATCACGGCGGAAACGGTACGCGGAGCATAAGAGTACCCGAGGGGAGTACTCGGTATGTGACA
 GTGAGAGTCTGTGGGTGACCGACAAGTTCATCGGCCATCGACATTTCGGGGACACCAGGTACCGGTGCTGGGG
 GAGATCAAAACGGGCAACTCTCCCGTCAAACAATATTTTTATGAAACGCGATGTAAGGAAGCCAGGCCGGT
 CAAAAACGGTTGCAGGGGTATTGATGATAAACACTGGAACCTCTCAGTGCAAAACATCCCAAACCTACGTCC
 GAGCACTGACTTCAGAGAACAATAAACTCGTGGGCTGGCGGTGGATACGGATAGACACGTCCTGTGTGTGT
 GCCTTGTGAGAAAAATCGGAAGAACA TGA ATTGGCATCTCTCCCCATATATAAAATTATTACTTTAAATTA
 TATGATATGCATGTAGCATATAAAATGTTTATATTGTTTTTATATATTATAAGTTGACCTTTATTTATTTAAA
 CTTTCAGCAACCCTACAGTATATAAGCTTTTTTCTCAATAAAATCAGTGTGCTTGCCTTCCCTCAGGCCTCT
 CCCATCTGTTAAAACCTTGTGTTTGTGATCCGGCTCTCAGGAGTCACTCTGTAAAATCTGTGTACACCAGTAT
 TTTGCATTCAAGTATTGTCAAGGCCATGACTGTTGTTTTAGTAAACTTGTAAAATCAAAAAAAAAAAAAA
 LNDRARVHLSLHVDVPGNGHTDAMVTFATILQVNKVMSSILFYVIFLAYLRGIQGNMMDQRSLPEDSLNSLI
 IKLIQADILKNKLSKQMVVDVKENYQSTLPKAEAPREPERGGPAKSAFQPVIAMDTELLRQRRYNSPRVLL
 SDSTPLEPPPLYLMEDYVGGSPVVANRTSRRKRYAEHKSHRGEYSVCDSESLWVTDKSSAIDIRGHQVTVLG
 EIKTGNPQVYFYETRCKEARPVKNGCRGIDDKHWSQCKTSQTYVRALTSENNKLVGWRWIRIDTSCVC
 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

TCTTCACGTCCCAGCGCGGGTGGGCGCCGGCGGCTCCTCTTAACCACAGGTTCCAGAAGTCTCTGCAGAA
 GTGCTTCCCTCTCTCATTTCCAGGACCACAATTCCCAGAGACTTCGGCTTACGACGTTTCTCTTTTTGCC
 CGATCTCTCCCGAGCTGGCTGGGCTTCGGCCGGCCAGAGGCCACAGCGACGACGTGATCCGTGCTGAGC
 GGGTCCCAGGGTTTCTCGGCGGCCCTTTTTTCTCTCCCTCGGTCGTCCTCCCTTGCAGGCTGTGCGGGCT
 GGCTTGAGCGGTGACCTGGCGGGTTCGCGCCTGCGCTCTGCCCTGTTTCTGCTGGCTGGTGGCGGGCC
 ATTTTGTTCATCTCCTCCTCCTCCTGCTCCTCCTGGTTGGAGCGCAGTGTCCGGAGCGGGCTGGGGGGAG
 AGAGCCCAGAGCAGGGTTCGGTGCCTTTTCTCTGTCCCAGCCGGTGCAGAGCCCCCTCCCCTTCC
 TCCCCACCCCTCCCCTCCCCAACCTGCCCTCCCCCTTGTCCCGGATCGCTCCGTGCGACCCACCATGA
 TGGAAGACGACGGGACGCCCGGACTCTATACGTAGGTAACCTTCCAGAGATGTGACAGAAGTCTTATA
 CTTTCAGTTGTTTCAGTCAAGTGGACCCTGTAAAAGCTGTAAAATGATAACAGAGCATAACAAGCAATGACCC
 ATATTGCTTTGTGGAATTTTATGAACACAGAGATGCAGCTGCTGCATTAGCTGCTATGAATGGGAGAAAAA
 TTTTGGGAAAGGAGGTCAAAGTAAACTGGGCAACCACCAAGTAGCCAGAAAAAGATACTTCCAATCAC

TTCCATGTGTTTGTGGGGATTTGAGTCCAGAAATTACAACAGAAGATATCAAATCAGCATTGCCCCCTT
TGGTAAAATATCGGATGCCCCGGGTAGTTAAAGACATGGCAACTGGAAAATCCAAAGGCTATGGTTTTGTAT
CTTTTTATAACAACTGGATGCAGAAAATGCGATTGTGCATATGGGCGGTGAGTGGTTGGGTGGTTCGTC
ATCCGAACCAATTGGGCCACTCGTAAACCACCTGCACCTAAAAGTACACAAGAAAAACAACACTAAGCAGTT
GAGATTTGAAGATGTAGTAAACCAGTCAAGTCCAAAAAATTGACTGTGTACTGTGGAGGAATTGCGTCTG
GGTTAACAGATCAGCTTATGAGACAGACATTCTCACCATTTGGACAAATTATGGAAATAAGAGTTTTCCCA
GAAAAGGGCTATTCAATTTGTGAGATTTTCAACCCATGAAAAGTGCAGCCCATGCCATTGTTTCGGTGAACGG
TACTACGATTGAAGGACATGTGGTTAAATGCTATTGGGGTAAAGAATCTCCTGATATGACTAAAACTTCC
AACAGGTTGACTATAGTCAATGGGGCCAATGGAGCCAAGTGTATGGAAACCCACAACAGTATGGACAGTAT
ATGGCAAATGGGTGGCAAGTACCGCCTTATGGAGTATACGGGCAACCATGGAATCAACAAGGATTTGGAGT
AGATCAATCACTTCTGCTGCTTGGATGGGTGGATTTGGTGTCTCAGCCTCCCCAAGGACAAGCTCTCCCC
CTGTAATACCTCTCCTAACAAGCCGATATGGTATGGCAAGTTACCAAAACAGTGA

AAAAAAATTTGTAATTCATGATAGGCTTCGATTTTCTGTGACACTCTGAAGACATGAAAAGTAGACATCGGA
AAATGAAAATATTTATTTTAAAAATTTGAAATGTTTGGAACTTTAGCACAGATTTGCTTTGGTGAAGGACA
CGTGTCTTCTAGTTCTGCCTTTTTAAGTTTTTGTTCATGATGGATATGAACATGATTTTTCTTTATGTACA
AAAATAAAATAAAGTCAATAAAGACAATTCTGACTACAAATTTTGATATAATAGGAAAAATGGCTAATAC
ATTTTGATTCTTAGATACTATTCCATTTTTATCTTGCTGTTGAGTATTTTAACTCACTGTGTTTTTAAAG
AGCAAAAAGGGAGGATCGTGA AACCTGGGAATCACATATAAGTTCATCCTGAATCCTGATACTCCCCTC
CCCTTCCCTGAGGTGGACCACATTTGAAGTGCAGCAGAGAAAAAGTGTGATATTGAGAAGAAATGCGTGATT
TTGGAGTCGCTTTGGAGGAAATATTTTCTTCTATGCCTAAAGAACTGAAGCCAGACTGAAGTTTTGC
ACCCTAAAAAGGAACAGCATTGTTTGGAGTACTTGGCAAAATGTTGGTGGTCCACGTTAAGACATATTTT
TAAAATTTCCAAAAGTGTGATATTAAAATTTAGTATTTTACATTTTATTTTGGGGGAAATCCAAGTA
TGGTGTGTTGATTGAAGTGCAGACAGTCATACTTGTGCTTTTACATGAAGTTTAAATGATACATATTGAAA
TATTCAATAACTACAGTGTAAAAAGCATGCTTCAACATAGAAGTAGCAGCAATGTAATTTTGAAGTA
ACACTTAACACACTCCGCTGCATTGAATGCAGTGGATTGATCAGAATGTTAAGACTGACATTTCCAAGGTT
GGCTACTATGTA AAAATTA AAAATTAACAAAATTTGTGCAGAAAAAGCCTTAATTTTAAATTTTATACAATCTTT
GATGCATTAGTATGTTCTAAAATGTCATTGGGAATTAGTTTTTTGTTTTTGTTTTTTTTTTTTTTTTTTTT
TGCTTTACATACTTGGTATGTAATACTTTGATTTAAACCTTGTAAACCAATTTCAAGGTTACTATAAGT
TGTATAGTACAAGTGTTTTTTAAAAATCTTGGGGTGTTTTTTAAAAATTAAGATATATTTTGGCCAAGAATT
TTTTTAAACAAGATTGCTAAAAACATCTTATTTAGACACTTCAATGTACCAATTTATAATTTGGATTCAGT
TTAAATAGTACACAGAGTTGTGGCTTTTTATTTTCAATTAATTTTTTTTCTTGTGGGCAGTGTGCATGGTAT
AATAAGCCTGAGCAGAGGCTTAAAGTTGTATGTGTGCAGAGTTTGTAAAGGAATCAATTGGAAGATGCAGAA
GACCGAGGTTTTGCTTTCAAGGTATTTTTTCAAGGCTGTGTGGGTAAAATTTGCCTCAAATTTCTATCAAACAG
GAATGTA AAAATAGATAAAAATCCTATGATTTGAATTGTGAGAGCTAGGGAGTGCAAATGTTTTGGCAATGT
ATTCAAAATGCTGGCCTGGGCACCAAAGAGAAAATAGCCTTTTACAGTTACATAGTAAGATGCGATTAGTA
CCCACAAATTAAGTTTTTCTAAACATTTGAAGTTTTTACGATTAGCTTTAAAATAATGATTTTATAAATTGG
TGGTCACAATAATTTGGTATTACTTTCTCCTTTTCCCCTTAGCAATATAGCCAAATGTATTCAACATA
AAAATTCATAGGCTGTGAAATTCATAGCTGGGCCAAATTTTTTATGGCACCTTAGTTTTTACCATAATGGTC
ATCTATTACACTCTTCTGTTATAAAAATATACCCTTATTTCTTTTGTATAGTATCTTTGAGGAATGTTTT
TGGAAAAGTTAATTTATATTTTATAGGGAGAACACTCAATAAATTATGTTAACTGTGCCCCGAGTTAAAA
ATTTTATGAGTATATGTGAACTTGAACAACCTGAAGACTTTTTTTAATTGATAAAAATGCTTAGTATGCCT
GTTTTGGTCTGCCAGTAAATTAAGTAGCTTATTGAGATAACTAACAGCTAAATATAGCTGTAGTGTTCCT
GACTGTATATTCTATGATTTAATAAAAATTAATCCAGACTAGTTATATTGCCACAGTAAAAA

VTWRVAPALCPVSLAGGGGHFVHPPPPAPPGWSAVSAGWGERAREQGSVPFPLSPAGAQSPPPLP
LPSPTLPSPLSRDRSVAPTMEDDGGQPRTLVGNLSRDVTEVLILQLFSQIGPCKSCKMITEHTSNDPYCF
VEFYEHRDAAAALAAMNGRKLIGKEVKVNWATTPSSQKKTDSNHFHVFGDLSPFITTEDIKSAFAPFGKI
SDARVVKDMATGKSKGYGFVSFYNKLEAENAIVHMGQWLGGRRQIRTNWATRKPAPKSTQENNTKQLRFE
DVVNQSSPKNCTVYCGGIASGLTDQLMRQTFSPFGQIMEIRVFPEKGYSFVRFSTHESAHAIVSVNGTTI
EGHVVKCYWGKESPDMTKNFQQVDYSQWGWQSQVYGNPQQYQYQYMANGWQVPPYGVYQVWPNQQQFGVDQS
PSAAWMGGFGAQQPPQQAPPPIPPPNQAGYGMASYQTQ

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

CCATTTATTTCCCCCTTTCCAACCTCTTCTCCCCAGCCTCCGCACCCTACCCTTGTTTCCTGTCCCTGT
CGCGCCCAGGTGTTTACCTGGCACTCAGGTGAGTGGTGCGCTCTGGCTGTTTTCTGTGCGAGCCGCCCGCC
TCTTCCTTCAGCGCGTCCCACAAATCCCGACGGC**ACGGAGGGGGCCCCAGGCCAAGGGCGATGGGCCCTGA**
GCCCTGACACCGCTTCGCCGCTGCTGCAGGTGCCCTGGCCGGCAGCGCCGCCGTGGTCCCGGAGCGCGGC
GACATGCCGGAGCTGGTGGTGACCGCGCTGCTGGCGCCGTGCGGCCTGTGCTGAAGCTGCTGCGCGCCTT
CATGTGGAGCCTGGTGTTCCTGGTGGCGCTGGTGGCCGCGGGCGGTCTACGGCTGCATAGCGCTCACGCACG
TGCTGTGCCGGCCCCGGCGCGGCTGCTGCGGGCGCCGTGCGGAGCGCGTCCCCCGCCTGCCTGAGCGACCCC
TCGCTGGGTGAGCACGGTTTCTGAACCTCAAGAGCTCGGGCCTGCGTCTGCACTATGTCTCGGCTGGACG
AGGTAACGGACCCCTCATGCTGTTTCTGCACGGCTTCCCTGAGAACTGGTTCTCCTGGCGTTACCAGCTCC
GGGAGTTCAGAGCCGCTTCCATGTTGTGGCTGTGGACTTGCAGAGGCTATGGCCCCCTCGGATGCACCTCGG
GATGTGGACTGCTACACAATCGACCTGCTGCTGGTGGACATCAAAGATGTCATCCTAGGCCTGGGTTACTC
GAAGTGCATCCTTGTGGCCATGACTGGGGTGCCTCCTTGCCTGGCATTCTCCATCTACTACCCATCCC
TGGTCGAGCGGATGGTTGTGGTCAGTGGTGCCCCCATGTCGGTGTACCAAGACTATTCCTGCACCACATC
AGCCAGTTCTTCCGTTCCCACTACATGTTCTGTTCCAGCTGCCCTGGCTGCCCGAGAAGCTGCTGTCTAT
GTCTGACTTTCAGATTCTGAAGACCACCCTCACCCACCGCAAGACAGGCATCCCATGCTTGACCCCCAGCG
AGCTCGAGGCCTTCTTTATAACTTCTCACAGCCTGGTGGCCTCACTGGGCCCCCTCAACTACTACCGAAAC
CTCTTCAGGAACTTCCCCCTGGAACCCAGGAGCTGACCACACCCACATTGCTGCTGTGGGGGGAGAAGGA
CACTTACTTGGAGCTGGGGCTGGTGGAAAGCCATCGGCAGCCGCTTTGTGCCGGGCCGCTTGGAGGCCACA
TCCTGCCAGGCATAGGGCATTGGATCCACAGAGCAACCCCCAGGAGATGCACCAGTACATGTGGGCCTTC
TTGCAAGACCTGCTGGACTAGTGGTCTTGTGCTGGCCTGCCAGGAGGCATGGATACTCAGGAAGGAACACA
CACCCATTATCTGTGTGTGCCTGGGAGTCCATAAATGTCCATACATGGGTGAACTCTTGAATCGCTCATA
GGCATGGGACTCCTGGATCCACACAAGCGCACCTATGGGTGCCTCGGGACACACCAACCCCTATACTCAC
ACACAGGCATGGATGCATGTGTGTGTAACAAACACTTTGACCCTGGGAACTGGGTATACCTCTCTTCCAGT
GGAGCCAGATGCCGAGACCGAGTGTCCACCTCCTCCCTTCCCTGGGGCCTCACTCTGCGCTTTGCCAAAGT
CGCTTCTCTGCCATAGCTGCACAGACCTTAAACCCTGACCTTCCCTGTCCCTGCCTTCGACCTCCTGCCTGG
GTCTTCAGCTCAGTGCTACTCTGAATACAATACTGCTGACCACATTTGGCTACTTCAGCTTAAATGTATGT
TAATTGAAATGAAATACAATGAAAATGCAGTGCTCAGTCCCA

TEGPQAKGDGPLSPDTASPLLQVPLAGSAAVVPERGDMPELVVTTALLAPSRLSLKLLRAFMWLSLVFSVALV
AAVYGCIALTHVLCRPRRGCCGRRRSASPACLSDP SLGEHGFLNLKSSGLRLHYVSAGRNGPLMLFLHG
FPENWFSWRYQLREFQSRFHVAVDLRGYGPSDAPRDVDCYTIDL LLDIKDVLGLGYSKILVAHDWGA
LLAWHFSIYYPSLVERMVVSGAPMSVYQDYSLHHISQFFRSHYMFLFQLPWLPEKLLSMSDFQILKTTLT
HRKTGIPCLTPSELEAFLYNFSQPGGLTGPLNYRNLFRNFLEPQELTTP TLLLWGEKDTYLELGLVEAI
GSRFVPGRLEAHILPGIGHWIPQSNPQEMHQYMWAFLQDLLD

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

AGAGTCCTGGCTGGCTCTGCCAAGTGTAAACAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAG
ACTCCAAGGAACCTCGTCAGGAAGGCAGGAGACAGGAGACGGGACCTCTACAGGGAGACGGTGGGCCGGCC
CTTGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCAGGCCCCCAAGGA
GCCGGCCCTACACCCCATGGGTTTGTCACTGCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGT
TCTGCAGATGGTTCCAGAGACGGGAGTCTGGGCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAG
AGGATCTGGGAGTCTCCTCTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAA
GTATGAGGATTGCAAGGTGCACCAGAGAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATG
ACAACCTGGAGGCCGCCATGGTGTCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCATGACATCTGAG
CTCTATGAGGGTCACTGCACATCGCTGTTGTGAACAGAACATGAACCTGGTGCAGGCCCTGCT
TGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCTGCAACCTCATCT
ACTTTGGGGAGCACCTTTGTCTTTGCTGCCTGTGTGAACAGTGAGGAGATCGTGCGGCTGCTCATTGAG
CATGGAGCTGACATCCGGGCCAGGACTCCCTGGGAAACACAGTGTTACACATCCTCATCTCCAGCCCAA
CAAAACCTTTGCCTGCCAGATGTACAACCTGTTGCTGTCTACGACAGACATGGGGACCACCTGCAGCCCC
TGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTGTGATG
TTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGACTCTCTATGA
CCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACCTTATCATCACCACCAAGAAGCGGG
AGGCTCGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGTGGACCTCAAGTGGAAAGCGGTACGGGCGG
CCGTACTTCTGCATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTACCATGTGCTGCATCTACCG
CCCCCTCAAGCCCAGGACCAATAACCGCACGAGCCCCGGGACAACACCTCTTACAGCAGAAGCTACTTCC
AGGAAGCTACATGACCCCTAAGGACGATATCCGGCTGGTGGGGAGCTGGTACTGTCTATTGGGGCTATC
ATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAAGTGGGGGTCACTCGCTTCTTTGGACAGACCATCCT
TGGGGGCCCATTCATGTCTCATCATCACCTATGCCTTTCATGGTGTGGTACCATGGTGATGCGGCTCA
TCAGTGCCAGCGGGGAGGTGGTACCCATGTCTTTGCACTCGTGTGGGCTGGTGGCAACGTGATGACTTTC
GCCCGAGGATTCCAGATGCTAGGCCCTTACCATCATGATTGAGAAGATGATTTTTGGCGACCTGATGCG
ATTCTGCTGGCTGATGGCTGTGGTTCATCCTGGGCTTTGCTTTCAGCCTTCTATATCATCTTCCAGACAGAGG
ACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTTCAGCACCTTCGAGCTGTTTCTTACC
ATCATCGATGGCCAGCCAACTACAACGTGGACCTGCCCTTTCATGTACAGCATCACCTATGCTGCCTTTGC
CATCATCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCC
ATGAGCGGGATGAGCTGTGGAGGGCCCAGATTGTGGCCACCACGGTATGCTGGAGCGGAAGCTGCCTCGC
TGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGA
AGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCTTCCACACCCGGGGCTCTGAGG
ATTTGGACAAAGACTCAGTGGAAAACTAGAGCTGGGCTGTCCCTTTCAGCCCCACCTGTCCCTTCTTATG
CCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCTTGAGGAG
AGACCTGCGTGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTG
TTCTCACTTCGCTTCTGGAACCTGCTCTCATTTTCTGGGTGCATCAAAACAAAACAAAACAAAACACC
AGAGGTCTCATCTCCAGGCCCCAGGGGAGAAAGAGGAGTAGCATGAACGCCAAGGAATGTACGTTGAGAA
TCACTGTCCAGGCCTGCATTACTCCTTTCAGCTCTGGGGCAGAGGAAGCCCAGCCCAAGCACGGGGCTGGC
AGGGCGTGAGGAACCTCCTGTGGCCTGCTCATCACCCCTTCCGACAGGAGCACTGCATGTGAGAGACTTT
AAAAACAGCCAGCCTGCTTGGGCGCTCGGTCTCCACCCAGGGTATAAGTGGGGAGAGAGCCCTTCCCA
GGGCACCCAGGCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCT
CTGGGGGTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTTCGGATCCCTGAAAA
AAAAAAAAAAAAAAAAAAAAA

TGPLQDGGPALGGADVAPRLSPVRVWPRPQAPKEPALHPMGLSLPKEKGLILCLWSKFCRWFQRRESWAQ
SRDEQNLLQKRIWESPLLLAAKDNDVQALNKLKLYEDCKVHQRGAMGETALHIAALYDNLLEAMVLMEEAA
PELVFEPMTSELYEGQTALHIAVNVQNMNLVRLALLARRASVSARATGTAFRRSPCNLIYFGEHPLSFAACV
NSEEIVRLLIEHGADIRAQDSLGNLTVLHILILQPNKTFACQMYNLLLSDRHRGDHLQPLDLVPHQGLTPF
KLAGVEGNTVMFQHLMQKRKHTQWTYGPLTSTLYDLTEIDSSGDEQSLELEIITTKKREARQILDQTPVKE
LVSLKWKRYGRPYFCMLGAIYLLYIICFTMCCIYRPLKPRNTNNRTPRDNTLLQKLLQEAYMTPKDDIRL
VGELVTVIGAIILLVEVPDIFRMGVTRFFGQITLGGPFHVLIIITYAFMVLVTMVMRLISASGEVVPMSFA
LVLGWCNVMYFARGFQMLGPFTIMI QKMI FGDLMRFCWLMAVVILGFASAFYIIFQTEDPEELGHFYDYPM
ALFSTFELFLTIIDGPANYVDL PFMYSITYAAF AIIATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVA

TTVMLERKLPRLWPRSGICGREYGLGDRWFLRVEDRQDLNRQRIQRYAQAFHTRGSEDLDKDSVEKLELG
CPFSPHLSLPMPSVSRSTSRSSANWERLRQGTLLRRDLRGIINRGLEDGESWEYQI

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

GCGCTCCGGAGGGAGAGCTGGGGCTGGAGGTTCTACCCCTCGGGCGCCCGCATCTGCCCCGCGCGCCC
GCCC**TGA**GCCCCGCCCCGACTGGGCAGGCGGGGAGCCCTACTTCTCTCCCCCGGGCGGGGAGCCGGG
GGCAGCGCCGGAGCCCGGGGGAGCTCAGCCCCGCCGACCGGCCAGGGCAGGGGGCAGCTAGGACGG
CCCCGGTCCAGGTGGAGGCCGAGAGGGCCAGGGCAAGCAGAGGCAGCAATGG**TTGGTCTGACGGTGGC**
TGAGCCCCAGCCCCTGGAATATGCAGCCCGGGGAGCCCCAGACAGCGGCAAGGACGAGGTGGCGGAGTG
GGGCGGGAGGCATGGTCTCCACCTACCGGGTGGCCGTGCTGGGGGCGCGAGGTGTGGGCAAGAGTGCCATC
GTGCGCCAGTTCTTGTACAACGAGTTCAGCGAGGTCTGCGTCCCCACCACCGCCCGCCCTTTACCTGCC
TGCTGTGCTCATGAACGGCCACGTGCACGACCTCCAGATCCTCGACTTCCACCCATCAGCGCCTTCCCTG
TCAATACGCTCCAGGAGTGGGCAGACACCTGCTGCAGGGGACTCCGGAGTGTCCACGCCTACATCCTGGTC
TACGACATCTGCTGCTTTGACAGCTTTGAGTACGTCAAGACCATCCGCCAGCAGATCCTGGAGACGAGGGT
GATCGGAACCTCAGAGACGCCCATCATCATCGTGGGCAACAAGCGGGACCTGCAGCGCGGACGCGTGATCC
CGCGCTGGAACGTTGTCACCTGGTACGCAAGACCTGGAAGTGCGGCTACGTGGAATGCTCGGCCAAGTAC
AACTGGCACATCCTGTCTCTTTCAGCGAGTCTCAAGAGCGTCCGGCTGCGCCCGTTGCAAGCACGTGCA
CGCTGCCCTGCGCTTCCAGGGCGCGCTGCGCCGCAACCGCTGCGCCATCATGTGACGCCTGCGCGCCCTC
GGGCTGCACCGGCACTGGCCGAGCGGAGGGCGGGGCCGTAAGTGCAGGGGCTGGGGCGGGGAGCGGGCGGGAA
ATGGAAGTGTGACGGTCCCGCCCTGAGGCCCTGCAGCCACGCACCTCCCGGTGAGAAGCAGAGCGCGAGA
GGGAGCCCTCCGTAAGTGCAGCCCTGCCCCCTGCCCCCGTGGCTTCTTGGGACAGCCGCCTTCCAGTGCT
GTATTTAGTGCAGTGCAGCGCCGCCCCGCGGGGTGCCACAGCCTTTTGGGATGGGGGTGAGCGTGCAAT
GGAGGCTGGGGGTGGCGAGGTGCCGCCTTGGCCGGGCCCCACGTGTCTTCTCCAGAATGTGTCTGTCTTT
GCCTGGTGTCTTCTTTCCCGTGTCCGCCCCACCCAGCGTCTGTTGGTACTTACCTGTCTCACCTACCCTC
CAGTCCCCTCCAGCTCCGCTCACAGGGCTCTCATTTTCGTCATCCCCTTGTGCGAGATCCTGGCAGCTTC
TTTGTGAGGCCAGGCCTTCTGACTGTGAGCACCACCGGCACAGGGCAGAGATGCGGGTGGCCCAAGGACCA
CGATCAAGGGGTCCGGGGGACCGAGGTCCAGATCAGTGAGGGGAGAAGGTTGAGCTCTCCGGCTTCCAGG
GAGACCTCCCCGCCCAGCAGCCCCAGAGACACAACAACCTACCTTCCAGCCTTAACTCGATGGTCCGTCC
CTGCCAGGTGCCCTCACTCTTCCCTGACCCCAAAGCCAGATCACCCCTGGGTTAAAACCTTTTTTTCTTTT
TTTTTTTTGGACAGAGTGTGGAAAGGGAGCCCCCAAAGGATAGCTTCTTTTTTCATGATGCCAGGCTCCAG
TCCTTTATTCCCTTCTGCATACTGCAATCTGATCTGTGAGACTGGGGAATGTTGGGTTCTGGGGTCTGGTC
GTGGGCAGGATGGTGCCAGAAGGGGTTAGGTTGTCCCAGTGAATACTGTTGCCCGCTCTCAACCCCA
TCTGACTACCCAGACTCTGCCTGCCTCAGATCTCAGACTATCCTGATTAATCTGGGGAAGAACAGAGCCA
GGGAAAGAATGGTGGGGACCCCTGTACTTGGGGGAGACACACCTGCATCTTCTCTGCCAGATGGAGG
CCCTCAGGATCTGACACCTCTTGTCCCAACACCAGTCAAGCCCTATACCCTAACTCACTCCACCCATTTT
CTCCGGCTGCCTGGCCGGGTTTCTACCTCTCGTACCAGGCTGATCACTGTGAGTTTTGTACCGATTTAG
AAATAACAATAATAATGAAGATTCTAGGAATGGCATGAGGGATTGATGGGGACTTGGAGGGAGGGACAAG
TGGTGCCCTGTCCCTGCTCCCTGGCCAAAGAAAGCTGTCTTGGAGGCTGAGCCCTCAGCCCTGGCCTGG
TGGGGGACAGCAAGGTCCCTTGTATAAGAGGGGAGAGAGGACAACCTCCGCTTGGCCAACCTAGCCAA
GGCTGCAGCATATAGACCAGGAAATCAGGTAGCCAGACTGGTGTGAGGAGAGAGTCTGGGGGAAGGGTCG
TGGGTGGGGAATTTATACCAACATCCATTGTAGGGGGAATCTATGATTCTGCTTCCCCAGCGGATTCCCA
CTCTGTCCACCAAGTGGGGGTAGCACAGCCTCACAGCAACCGCCCTGACCTTGGGCAGTCTAGTGTTCCT
GCATTCTAGTCCCTGCTGTGCTGCAGGACTTTGGGCAAGTACCTGCCCTCTGTGAGCCTCCCTCTGACAC
AGAGGAGGTGGCTCCCCTTCCCCACACCTTAGAGTGGCTGGGAGGGTAACAAAGAGGGCCTGCCCTTTAG
TCTCCTGCACCCCTGCCCTGGTTACCAGAGGGAGCGGATGAAGGATGGCAGCATCTCACATGCCCCAT

CACCAACTCTGAGGCACCTGGGGTGGGGGGGGCGGAGCCCAGGCCTCTGGCTGCTCCCCTGTGGGAGCCATT
GGAATGTATCCCCTGACAGGCCCCCTTCCGCCTCCACCTCAACCCAGGTCTTGGATTTTCAGGTCCCTCCAC
CCCCATTCTGAGTCTCTGTCTTCTCCTTCCACCCGCTCCCAGGGTTTCCCACCACAGGGTCTGGAAGTGT
GTGTGACGCCATTGAGCTGTTACCCGAAGTCAGATTAATAATCAGGGAGTGTTCCTCGTTTCTGTAA
AAAAAAAAAAAAAAAAAAAA

LVLTVAEPPAPGICSPGEPQTAARTRWRSGAGGMVSTYRVAVLGARGVVGKSAIVRQFLYNEFSEVCPPTTA
RRLYLPAVVMNGHVHDLQILDFFPISAFPVNTLQEWADTCCRGLRSVHAYILVYDICCFFDSFEYVKTIRQQ
ILETRVIGTSETPIIIVGNKRDLRGRVIPRWNVSHLVRKTKWCGYVECSAKYNWHILLFSELLKSVGCA
RCKHVHAALRFQALRRNRCAIM

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*, *Procyon capensis*, *Rattus norvegicus* and *Sus scrofa*.

C17orf95 chromosome 17 open reading frame 95

(homology to AdoMet-dependent methyltransferases)

Synonyms: C17orf95

AGTTCTGCGCGTGTGAGTCTCTTTCGCCTTGCTCCGGGCTTTCTTCGCTCGCAGCGCGGCAGGGTTATCAC
CAGATCTGGGCTTTCCCCTTCTTGCCGTCAGGTGCTACGGCCACGTGGCCCCGGGCTTCCCCTCGCGCAG
TCTGGCAGCCCCGAGCCTTCCGCGTCCCCACCCGCGGGGCCAACGACGCCCTACTGGGCGAGCACG
ATTCCGAGGACAGGGGGTCCGGGCCAGCGCTTTCGATTCTCGGAGGAGCCGGGTCCGGGGCCGACGGG
GCTGTCTGAGGTTCCAGTCCCAGGTCCTGCATCTCCAGTATGGAATGTATGTTTGGCCCTGTGCTGT
GGTCTGGCCAGTACCTTTGGTTTTACAGAAGATCTCTGCCAGGCAAGGCCATCTTAGAGATTGGAGCTG
GAGTGAGCCTTCCAGGAATTTTGGCTGCCAAATGTGGTGCAGAAGTAATACTGTCAGACAGCTCAGAACTG
CCTCACTGTCTGGAAGTCTGTGCGCAAAGCTGCCAAATGAATAACCTGCCACATCTGCAGGTGGTAGGACT
AACATGGGGTCAATATCTTGGGATCTTCTGGCTCTACCACCACAAGATATTATCCTTGCATCTGATGTGT
TCTTTGAACCAGAAGATTTTGAAGACATTTTGGCTACAATATATTTTTTGGATGCACAAGAATCCCAAGGTC
CAATTGTGGTCTACTTATCAAGTTAGGAGTGCTGGCTGGTCACTTGAAGCTTTACTCTACAAATGGGATAT
GAAATGTGTCCACATTCCTTCTGAGTCTTTTGGATGCAGACAAAGAAGATATAGCAGAATCTACCCTTCCAG
GAAGACATACAGTTGAAATGCTGGTCATTTCTTTGCAAAGGACAGTCTCTGAATTATACCTACAACCTGT
TCTGGGACAGTATCAATACTGATGAGCAACCTGGCACACAACTATGAGCAGACCACTTCAGCTTGAGAAT
GCAGTGGGTCTGAAGATGGTCAAGTCTGTTTGCCTTAGATTTTGGATGTCACCTAGACAACACTTAAACTCA
TATGAAACAAAAATTAATAATACGTATTACAAGTAAAAAAAAAAAAAAAAAAAA

LAARSLPRSPTRPGPNDALLGEHDFRQGVRAQRFRFSEEPGPGADGAVLEVHVPQVLHLQYGMVYVWPCAV
VLAQYLWFHRRSLPGKAILEIGAGVSLPGILAACKGAEVILSDSSELPHCLEVCRQSCQMNNLPHLQVVGL
TWGHISWDLALPPQDIILASDVFFEPEDFEDILATIYFLMHKNPKVQLWSTYQVRSAGWSLEALLYKWD
M KCVHIPLESFDADKEDIAESTLPGRHTVEMLVISFAKDSL

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

AGATGCAGGGGCGCAAACGCCAAAGGAGACCAGGCTGTAGGAAGAGAAGGGCAGAGCGCCGGACAGCTCGG
 CCCGCTCCCCGTCCTTTGGGGCCCGGGCTGGGGAACACTACAAGGCCAGCAGGCAGCTGCAGGGGGCGGAGG
 CGGAGGAGGGACCAGCGCGGGTGGGAGTGAGAGAGCGAGCCCTCGCGCCCCGCGGCATAGCGCTCGGA
 GCGCTCTTGCGGCCACAGGCGCGGCTCCTCGGCGGGCGGCGGAGCTAGCGGGAGCCGGGACGCCGGTGC
 AGCCGCAGCGCGCGGAGGAACCCGGGTGTGCCGGGAGCTGGGCGGCCACGTCCGGACGGGACCGAGACCCC
 TCGTAGCGCATTGCGGCGACCTCGCCTTCCCCGGCCGCGAGCGCGCCGCTGCTTGA^{AAAGCCGCGGAACCC}
 AAGGACTTTTCTCCGGTCCGAGCTCGGGGCGCCCCGAGGGCGCACGGTACCCGTGCTGCAGTCGGGCACG
 CCGCGGCGCCGGGGCCTCCGCAGGGCGATGGAGCCCGGTCTGCAAGGAAAGTGAGGCGCCGCCGCTGCGTT
 CTGGAGGAGGGGGGCACAAGGTCTGGAGACCCCGGGTGGCGGACGGGAGCCCTCCCCCGCCCCGCTCCG
 GGGACCAGCTCCGGTCCATTGTTCCCGCCCGGGCTGGAGGCGCCGAGCACCGAGCGCCCGGGGAGTCG
 AGCGCCGGCCGCGGAGCTCTTGCACCCCGCCAGGACCCGAACAGAGCCCGGGGGCGGCGGGCCGGAGCCG
 GGGACGCGGGCACACGCCCGCTCGCACAAAGCC^{ACGGCGGACTCTCCCGAGGCGGAACCTCCACGCCGAGCG}
^{AGGGTCAGTTTGA^{AAAGGAGGATCGAGCTCACTGTGGAGTATCCATGGAGATGTGGAGCCTTGTCA^{CAAC}}}
^{CTCTA^{ACTG}CTGAGACTGGGATGTGGAGCTGGAAGTGCCTCCTCTTCTGGGCTGTGCTGGTCA^{CAGCCACAC}}
^{CTGCACCCGCTAGGCCGTC^{CCCCGACCTTGCCTGAACAAGCCAGCCCTGGGGAGCCCCGTG^{GGAAGTGGAG}}}
^{TCCTTCTGGTCCACCCCGGTGACCTGCTGCAGCTTCTGTGCGGCTGCGGGACGATGTGCAGAGCATCAA}
^{CTGGCTGCGGGACGGGGTGCAGCTGGCGGAAAGCAACCGCACCCGCATCACAGGGGAGGAGGTGGAGGTGC}
^{AGGACTCCGTGCCCGCAGACTCCGGCCTCTATGCTTGCCTAACCAGCAGCCCTCGGGCAGTGACACCACC}
^{TACTTCTCCGTCAATGTTTTCAGATGCTCTCCCTCCTCGGAGGATGATGATGATGATGATGACTCCTCTTC}
^{AGAGGAGAAAGAAACAGATAACACCAAACCAAACCGTATGCCCGTAGCTCCATATTGGACATCCCCAGAAA}
^{AGATGGAAAAGAAATTGCATGCAGTGCCGGCTGCCAAGACAGTGAAGTTCAAATGCCCTTCCAGTGGGACC}
^{CCAAACCCACACTGCGCTGGTTGAAAAATGGCAAAGAATTCAAACCTGACCACAGAATTGGAGGCTACAA}
^{GGTCCGTTATGCCACCTGGAGCATCATAATGGACTCTGTGGTGCCTCTGACAAGGGCAACTACACCTGCA}
^{TTGTGGAGAATGAGTACGGCAGCATCAACCACACATAACCAGCTGGATGTCGTGGAGCGGTCCCCTCACCGG}
^{CCCATCCTGCAAGCAGGGTTGCCCGCAACAAAACAGTGGCCCTGGGTAGCAACGTGGAGTTTATGTGTAA}
^{GGTGTACAGTGACCCGCAGCCGCACATCCAGTGGCTAAAGCACATCGAGGTGAATGGGAGCAAGATTGGCC}
^{CAGACAACCTGCCTTATGTCCAGATCTTGAAGACTGCTGGAGTTAATACCACCGACAAAGAGATGGAGGTG}
^{CTTCACTTAAGAAATGTCTCCTTTGAGGACGCAGGGGAGTATACGTGCTTGGCGGGTAACTCTATCGGACT}
^{CTCCCATCACTCTGCATGGTTGACCGTTCTGGAAGCCCTGGAAGAGAGGCCGGCAGTGATGACCTCGCCCC}
^{TGTACCTGGAGATCATCATCTATTGCACAGGGGCCTTCTCATCTCCTGCATGGTGGGGTCCGTATCGTC}
^{TACAAGATGAAGAGTGGTACCAAGAAGAGTGA^{CTTCCACAGCCAGATGGCTGTGCACAAGCTGGCCAAGAG}}
^{CATCCCTCTGCGCAGACAGGTAACAGTGTCTGCTGACTCCAGTGCATCCATGAACTCTGGGGTTCTTCTGG}
^{TTCCGGCCATCACGGCTCTCCTCCAGTGGGACTCCCATGCTAGCAGGGGTCTCTGAGTATGAGCTTCCCGAA}
^{GACCCCTCGTGGGAGCTGCCTCGGGACAGACTGGTCTTAGGCAAACCCCTGGGAGAGGGTCTCTTGGGCA}
^{GGTGGTGTGGCAGAGGCTATCGGGCTGGACAAGGACAAACCAACCGTGTGACCAAAGTGGCTGTGAAGA}
^{TGTTGAAGTCGGACGCAACAGAGAAAGACTTGTGAGACCTGATCTCAGAAATGGAGATGATGAAGATGATC}
^{GGGAAGCATAAGAATATCATCAACCTGCTGGGGGCCTGCACGCAGGATGGTCCCTTGTATGTCATCGTGA}
^{GTATGCCTCCAAGGGCAACCTGCGGGAGTACCTGCAGGCCCGGAGGCCCCAGGGCTGGAATACTGCTACA}
^{ACCCCAGCCACAACCCAGAGGAGCAGCTCTCCTCCAAGGACCTGGTGTCTGCGCCTACCAGGTGGCCCGA}
^{GGCATGGAGTATCTGGCCTCCAAGAAGTGCATACACCGAGACCTGGCAGCCAGGAATGTCTGGTGCAGAG}
^{GGACAATGTGATGAAGATAGCAGACTTTGGCCTCGCACGGGACATTACCACATCGACTACTATAAAAAAGA}
^{CAACCAACGGCCGACTGCCTGTGAAGTGGATGGCACCCGAGGCATTATTTGACCGGATCTACACCCACCAG}
^{AGTGTGTGTGGTCTTTCCGGGGTGTCTCTGTGGGAGATCTTCACTCTGGGCGGCTCCCATAACCCGGTGT}
^{GCCTGTGGAGGAACCTTTCAAGCTGCTGAAGGAGGGTCACCCGCATGGACAAGCCCAGTAACTGCACCAACG}

AGCTGTACATGATGATGCGGGACTGCTGGCATGCAGTGCCCTCACAGAGACCCACCTTCAAGCAGCTGGTG
GAAGACCTGGACCGCATCGTGGCCTTGACCTCCAACCAGGAGTACCTGGACCTGTCCATGCCCTGGACCA
GTACTCCCCAGCTTTCCCGACACCCGGAGCTCTACGTGCTCCTCAGGGGAGGATTCCGTCTTCTCTCATG
AGCCGCTGCCCCGAGGAGCCCTGCCTGCCCCGACACCCAGCCAGCTTGCCAATGGCGGACTCAAACGCCGC
TGA

CTGCCACCCACACGCCCTCCCCAGACTCCACCGTCAGCTGTAACCCTCACCCACAGCCCCTGCTGGGC
CCACCACCTGTCCGTCCCTGTCCCCTTTTCTGCTGGCAGGAGCCGGCTGCCTACCAGGGGCCTTCTGTGT
GGCCTGCCTTACCCCACTCAGCTCACCTCTCCCTCCACCTCCTCTCCACCTGCTGGTGAGAGGTGCAAAG
AGGCAGATCTTTGCTGCCAGCCACTTCATCCCCTCCCAGATGTTGGACCAACACCCCTCCCTGCCACCAGG
CACTGCCTGGAGGGCAGGGAGTGGGAGCCAATGAACAGGCATGCAAGTGAGAGCTTCTGAGCTTTCTCCT
GTCGGTTTGGTCTGTTTTGCCTTACCCATAAGCCCTCGACTCTGGTGGCAGGTGCCTTGTCTCAGGG
CTACAGCAGTAGGAGGTCACTGCTTCTGTCCTGATTGAAGGTGACCTCTGCCCCAGATAGGTGGTGCCA
GTGCGCTTATTAATTCGATACTAGTTTGTCTTGTGCTGACCAAATGCCTGGTACCAGAGGATGGTGAGGCGAA
GGCCAGGTTGGGGCAGTGTGTGGCCCTGGGGCCAGCCCCAACTGGGGGCTCTGTATATAGCTATGAA
GAAAACACAAAGTGTATAAATCTGAGTATATATTTACATGTCTTTTTTAAAAGGGTCGTTACCAGAGATTTA
CCCATCGGGTAAGATGCTCCTGGTGGCTGGGAGGCATCAGTTGCTATATATATATAAAAAACAAAAAGAAAAA
AAGGAAAATGTTTTTAAAAGGTCATATATTTTTTGTACTTTTTGCTGTTTTATTTTTTAAATTATGTTT
TAAACCTATTTTTCAGTTTAGGTCCCTCAATAAAAAATTGCTGCTGCTTCATTTATCTATGGGCTGTATGAAA
AGGGTGGGAATGTCCACTGGAAAGAAGGGACACCCACGGGCCCTGGGGCTAGGTCTGTCCCAGGGCACC
CATGCTCCCGGCGCAGGTTCTTGTAACTCTTCTTCTTAGGTCCTGCACCCAGACCTCACGACGCACCTC
CTGCCTCTCCGCTGCTTTTTGGAAAGTCAGAAAAAGAAGATGTCTGCTTCCAGGGCAGGAACCCCATCCATG
CAGTAGAGGCGCTGGGCAGAGAGTCAAGGCCCAGCAGCCATCGACCATGGATGGTTTTCTCCAAGGAAACC
GGTGGGGTTGGGCTGGGAGGGGGCACCTACCTAGGAATAGCCACGGGGTAGAGCTACAGTGATTAAGAGG
AAAGCAAGGGCGCGGTTGCTCACGCCTGTAATCCCAGCACTTTGGGACACCGAGGTGGGCAGATCACTTCA
GGTCAGGAGTTTGGAGACCAGCCTGGCCAACCTTAGTGAAACCCCATCTCTACTAAAAATGCAAAAATTATCC
AGGCATGGTGGCACACGCCTGTAATCCCAGCTCCACAGGAGGCTGAGGCAGAATCCCTTGAAGCTGGGAGG
CGGAGGTTGCAGTGAGCCGAGATTGCGCCATTGCACTCCAGCCTGGGCAACAGAGAAAAACAAAAAGGAAAA
CAAATGATGAAGGTCTGCAGAACTGAAACCCAGACATGTGTCTGCCCCCTCTATGTGGGCATGGTTTTGC
CAGTGTCTTAAGTGCAGGAGAACATGTCACTGAGGCTAGTTTTTGCATTGAGTCCCTGGCTTCTGTTTTCT
TGTTGGTATGCCCTCCCAGATCGTCTTCTCTGTATCCATGTGACCAGACTGTATTTGTTGGGACTGTCCGA
GATCTTGGCTTCTTACAGTTCTTCTGTCCAAACTCCATCCTGTCCCTCAGGAACGGGGGGAAAAATTCTCC
GAATGTTTTTGGTTTTTGGCTGCTTGGAAATTTACTTCTGCCACCTGCTGGTCATCACTGCTCCTCACTAAG
TGGATTCTGGCTCCCCCGTACCTCATGGCTCAAACCTACCCTCCTCAGTCGCTATATATAAGCTTATATTT
TGCTGGATTACTGCTAAATACAAAAGAAAGTTCAATATGTTTTTCACTTCTGTAGGGAAAATGGGATTGCTG
CTTTAAATTTCTGAGCTAGGGATTTTTTGGCAGCTGCAGTGTGGCGACTATTGTAATAATTCTCTTTGTTT
CTCTCTGTAATAGCACCTGCTAACATTACAATTTGTATTTATGTTTTAAAGAAGGCATCATTTGGTGAACA
GAACTAGGAAATGAATTTTTAGCTCTTAAAAGCATTGCTTTGAGACCGCACAGGAGTGTCTTTCTTTGTA
AAACAGTGATGATAATTTCTGCCTTGGCCCTACCTTGAAGCAATGTTGTGTGAAGGGATGAAGAATCTAAA
AGTCTTCATAAGTCTTTGGGAGAGGTGCTAGAAAAATATAAGGCACTATCATAATTACAGTGATGTCCTTG
CTGTTACTACTCAAATCACCCACAAATTTCCCCAAAGACTGCGCTAGCTGTCAAATAAAAGACAGTGAAT
TGACCTGAAAAAAAAAAAAAAAAA

TADSPEAEPRRARVSLKRRIELTVEYPWRCGALSPTSNCRTGMWSWKCLLFWAVLVTATLCTARPSPTLP
EQAQPWGAPEVEVESFLVHPGDLQLRCLRDDVQSINWLRDGVQLAESNRTRITGEEVEVQDSVPADSGLY
ACVTSSPSGSDTTYFSVNVSDALPSSDDDDDDSSSEEKETDNTKPNRMPVAPYWTSPEKMEKKLHAVPA
AKTVKFKCPSGTPNPTLRWLKNGKEFKPDHRIGGYKVRATWSIIMDSVVP SDKGNYTCIVENEYGSINH
TYQLDVVERSPHRPILQAGLPANKTVALGNSVEFMCKVYSDPQPHIQWLKHIEVNGSKIGPDLNPYVQILK
TAGVNTTDKEMEVLHLRNVSFEDAGEYTCLAGNSIGLSHSAWLTVLEALEERPAMVTSPLYLEIIIIYCTG
AFLISCMVGSVIVYKMKSGTKKSDFHSMQMAVHKLAKSIPLRRQVTVSADSSASMNSGVLLVRPSRLSSSGT
PMLAGVSEYELPEDPRWELPRDRLVLGKPLGEGCFGQVVLAEAIGLDKDKPNRVTKVAVKMLKSDATEKDL
SDLISEMEMMKMIGKHKNIINLLGACTQDGPLYVIVEYASKGNLREYLQARRPPGLECYNPSHNPEEQLS
SKDLVSCAYQVARGMEVYLASKKCIHRDLAARNVLTEDNVMKIADFGGLARDIHHIDYKKTNTNGLRVPVKWM
APEALFDRIYTHQSDVVSFVGLLWEIFTLGGSPYGPVVEELFKLLKEGHRMDKPSNCTNELYMMMRDCWH
AVPSQRPTFKQLVEDLDRIVALTSNQEYLDLSMPLDQYSPSFPDTRSSTCSSGSDSVFSHEPLPEEPCLPR
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

GATAGCGGTTGGCGAGAGGCCACATTTCCGCCACGTGACCCGCGCATGCGCCTGCTTGCTGGAGAGCGAGC
 GTCTTTTGGCCACGAACACCTGCGGCGTGCCGAAGTTCCTCTTCTGCCCCGGTTCGGGGCGGTCCGCTGG
 GAGTGGCGCCCGAGGGACACCCGTGCCTGGGGTCGGGGCCAGTGTCTTCTCGGAGCTCCAGAAGTTCT
 GCTGAAGAAGCGCGGCGGAGCAAGACGACTTCTCCGAGCCGCCGAGCTGGAGTTAGAGGTGGAGCTCCGT
 GGGCCGGGCCCCGGCTGCGGGCAGCGGCTCCTGCAGGCGGAGGCCCGGCGGAGAAATGAGCCGGAGCC
 GGAGGTGGTGGTGGTCTCCTGGCAGAGCGGTGGACCGGGAGCAAGCTGCGGTGGTTACCACCTTGATGC
 TGTGTGCCTCCTTCTGGGGCTGGGATTGAGTGTTGCTATAGTGGGACCCACGTTTCAAGATTTGGCAACA
 AACGTGAACCGAAATATCAGTAGTCTGTCTTTTCAATTTTTGTGGGTGCTGCCTTGGGATATTTGAGTGGCTC
 TGTGATTGGTGGATTTCTTGTGATGTGATGAATTATTTTTACTTTTGGGAATCTCAATGTCGGCTACCA
 CCGTTGGTCTTTATCTTGTTCCTTTTTGCAAGACAGCAATATTACTCACTGTCATGATGTCTATCTTCGGT
 GTTTCAATTGGCATTCTGGATACAGGTGGTAACGTCTTATCTTGGCTATTTGGGGGGACAAAGGAGCCCC
 ACATATGCAGGCCTTACACTTCTCTTTGCTTGGGTGCCTTTTTGGCTCCACTGCTAGCTAAACTGGCTT
 TGGGTCCGACAGCGTCTGCTGAAAACACACAGAGTCTGACTTCCATCCTGCACTCAACCAATCATCTGAT
 GCTGACTCAGAAGCTCTGTTTGGAGTACCTAATGATAAGAATTTACTGTGGGCTTATGCTGTTATCGGTAC
 TTACATGTTCTTAGTTTTCTGTCAATTTTTTTGTCTGTTTTTAAAGAATAGCTCAAAGCAAGAAAAAGCAA
 GAGCATCTGCTGAGACATTTTCAAGAGCAAAATATCACAACGCCCTTCTTTGTCTCCTTTTTCTGTTCTTC
 TTTTTTATGTTGGAGCTGAGGTAACATATGGCTCTTATGTTTTCTCATTTGCAACCACCCATGCTGGCAT
 GAAAGAAAGTGAAGCTGCTGGGTTGAACTCCATCTTCTGGGGGACATTTGCAGCCTGCAGGGGCTGGCAA
 TCTTTTTTGCTACCTGTTTACAGCCTGGAACCATGATTGTGTTGAGCAACATTGGCAGCCTGACTTCATCT
 TTATTTCTGGTGTCTTTTGAACAAGAACCAATTTGTCTCTGGATAGCAACTTCAGTGTATGGGGCTTCAAT
 GGCAACCACATTTCCGAGTGGTGTCTTGGATTGAGCAGTACACGACCATCCATGGGAAATCTTCAGCAGCAT
 TTTTGTAAATGGTGTCTTCCCTGGGAGAAATGGCTATTTCTGCAGTCATTGGAATCTTCAAGGAAATAC
 CCTGATTTGCCTGTAGTTCTGTATACCTCTTTGGGAGCATCAATAGCTACTGGTATTTTATTTCTGTGCT
 ATATAAATTAGCCACTTACCTCTTGATCGCCAGCGAAAAGAAGACAGAAAGAGTGAGGACCAGAAAGCTC
 TGCTCTCTAGCTCCGGGCTAAATGAATATGAGGAAGAGAATGAAGAGGAGGATGCAGAAAAATGGAATGAA
 ATGGATTTTGAATGATTGAAACGAATGATACAATGAGGCATTCTATAATAGAGACATCTAGAAGTAGTCT
 GACGGAGCCACAGCTGAAGTCTATAATCAATACCCATCAAATGCACTGGTGTGTTGAGTCTTCTCCTTTTA
 ATACTGGCAGTGCCCATGTGAAGCACTTGCCAGAAACCAGGACAAAAGGGACTAACGTTTAGAGAAGATGG
 ATTACTCACTGACATCTTTGAATAACTGCCACTTCTAAGGACGCCATTTCAGAGCAGAGCATTAGCAGAATT
 TCACCATGCTTTGGGGATTAATAATTTTAGGTCCATATTATAGCAAATGCTAAAAATTTTTGAAATGTTCT
 GTAATTTCCAGAGTCTTCCATAAATAACCAAATGGTCTCATAACAGTACAATAGGATCTTGTATTAACGCT
 ATCATTTGACATGTGACTGATTAGGAAATATTTTATGGTCTCCACCCCTCAGAGCACCCAAAGAAGGCATT
 GTTTTTGAGAAGGGGGGATTCTCAGAATGAGACCATGTAGGAGTTGAATTGATTGGTTGAACAAATTATC
 AGGTATCAATATTTTTCCAGGAAAGTTGAGGAGAATAGTTTTATAAAAGAGAAATGGGCCGGGTGCAGTGG
 CTAAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACGGGCAGATCACGAGGTGAGGAGATCGAGACCAT
 CCTGGCTAACACAGTGAACCCCTGTCTCTACTAAAAATACAAAAAATTAGCCAGGCGTGGTGGCAGGCACC
 TGTAGTCCCAGCTACTTTGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGTGGAGCTTGCAGTGAG
 CCGAGATCATGCCACACCGTACTCCAGCCTGGGCGACATCTAGACTCCGTCTCAAAAAAGAAATGGAGAGG
 TATTTTATTAATGCTTTACCTTATCTGTAGGGTGGATCAAAGAATGTAATGACTCATTGTATTTCTAA
 TTCTCAATGAAGAAGACAGACAAGGGCCCAAGGGTTCGGGCTTCAAGCCTTGACTTCTTATGCCCTGCA
 CTCTGTCACTTAACGTAGGCCTTGTATCTCGCCTCCTGATCTGTGATTCTAGAAATGAAAATGAGATGA
 TTTCCAAATTGGGAATCTGTAAGTTTTAAAGTTTGTGCAAAATGTCTGCTAGAAATTAGTACAAGTGACAG

CATGTTCTTCATTTCAAATGAGAAGTGAAACCAAATGATAAATAGCTTTAAGAATGTGCTAATGATAAAT
GATTACATGTCAATTTAATGTACTTAATGTTTAAATACCTTATTTGAATAATTACCTGAAGAATATATTTTT
TAGTACTGCATTTTATTGATTCTAAGTTGCACTTTTTACCCCATACTGTAAACATATCTGAAATCAGAAT
GTGTCTTACAATCAGTGATCGTTTAAACATTGTGACAAAGTTTAAATGGACAGTTTTTTCCCATATGTATATA
TAAAATAATGTGTTTTACAATCAGTGGCTTAGATTGAGTGAATAACAGTAATTCATTCAATTATGATAGTA
TCTTTACAGACATTTTAAAAATAAGTTATTTTTATATGCTAATATTCTATGTTCAAGTGGAAATTTGGAGAC
ACTATGTTCTAATTTAAGAAGTGTTCAGTGTATTAGATTAGTAGTAGGGCTGTAGATGGGGTCTTTTCG
AGGTCCTTTGCCTACCTCAGTAACAGAGGCTGGGTCTGAGCCTTGCTCGAGATTAGCTCTTCACCAGGTCT
GTGTGGTAATCATATCGTAGTTACCGTAGGTCACGCTGCAAGTGTAAATTTACAGTGAGGCATAGTCTTTTC
CAGTGATTGGAATACTTGTGTACCGATCCAGGGTTTCTCTTTCCTGAGTCTGAAATACCTGGGAAACCAG
TTAGAAGGTGACTGGGTAGCCTAGGAGTCAAGGATCAGAAAGGGTACTTCATATATATGTGGCAGAATTGG
TGCAGGTGATCCTTTGAAAAATGACAAGAATTGATGAAATTTTGTACTTCTCCAATTTTTTGTGCTGAAT
AACATCTCCTTTCTACCTGATTGGAGTAGGCTTGTCTTCTTACTTAATTTCTATAAATTTGCCATTGTT
TTGTCTCAAGTAAAAAAAAAAAAA

LELEVELRGAGPPAAGQRLQLQAEAPAENEPEPEVVVSWQSGGTGSKLRWFTTLLMLCASFLGLGLSVAIVG
PTFQDLATNVNRNIISSLSFIFVGRALGYLSGSVIGGLVDVMNYFLLLGISMSATTVGLYLVPFCKTAILL
TVMMSIFGVSIGILDTGGNVLILAIWGDGKAPHMQALHFSFALGAFLLAKLALGPTASAENHTESDFH
PALNQSSDADSEALFGVPNDKNLLWAYAVIGTYMFLVSVIFFCLFLKNSSKQEKARASAETFRRAKYHNAL
LCLLFLFFFYVGAEVTYGSYVFSFATTHAGMKESEAAGLNSIFWGTFAACRGLAIFATCLQPGTMIVLS
NIGSLTSSLFLVLFDKNPICLWIATSVYGASMATTFPSGVSWIEQYTTIHGKSAAFFVIGASLGEMAIIPAV
IGILQGKYPDLVVLYTSLGASIAATGILFPVLYKLATSPLDRQRKEDRKSEDQKALLSSSGLNEYEEENE
EDAENWEMDFEMIEETNDTMRHSIIETSRSSLEPTAEVYNQYPSNALVFESSPFNTGSAHVKHLPETRTK
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

AGAAGGGGCACCGCTGAGCTGCCGAGAGGAAGCTCGCTCTGCCCGGCTGCCCTCTTGTAGTCCGCCGGCGA
GGGGCAGTTCTCGGTGAGGAGGAAGAGAGCAGCGGACGGCACAGCACCCGCGCGGGCCCTCCACAACAGC
TCCAGCTGGCAGCATCACTTCCCGCCAATTTATCCAATTTCTGCCAAGGCTCTGAAATGCCAACACGTCG
AGGCCTGCACTTGATGTCAAGGGTGGCACCTCACCTGCGAAGGAGGATGCCAACCAAGAGATGAGCTCCGT
GGCCTACTCCAACCTTGCAGGTAAGATCGCAAAGCAGTGGCCATTCTGCACTACCCTGGGGTAGCCTCAA
ATGGAACCAAGGCCAGTGGGGCTCCCCTAGTTCTCGGGATCTCCAATAGGCTCTCCTACAACCACCCCT
CCCCTAAACCCCATCCTTCAACCTGCACCCCGCCCCTCACTTGCTGGCTAGTATGCACCTGCAGAACT
TAATAGCCAGTATCAGGGGATGGCTGCTGCCACTCCAGGCCAACCCGGGGAGGCAGGACCCCTGCAAACT
GGGACTTTGGGGCCCAGGCGGGAGGGGCAGAATCACTCTCTCTCTCTGCTGGTGGCCAGAGCCCTGCTATC
ATCGATTCCGACCCAGTGGATGAGGAAGTGTGATGTGCTGGTGGTGGAACTGGGGTTGGACCGAGCCAA
TGAGCTTCCGGAGCTGTGGCTGGGGCAGAATGAGTTTGACTTCACTGCGGACTTTCCATCTAGCTGCTAAAT
GCCAAGTGTCCCTAAAGATGGAGGAATAAAGCCACCAATTCTGTTGTAATAAAAAATAAAGTTACTTACAA
AGAGACGGGCCAAAAA

LAASLPANLSNFCQGSEMPPTSRPALDVKGGTSPAKEDANQEMSSVAYSNLAVKDRKAVAILHYPGVASNG
TKASGAPTSSSGSPIGSPTTTPPTKPPSFNLHPAPHLLASMLHQKLNQYQGMMAATPGQPGEAGPLQNW
FGAQAGGAESLSPSAGAQPAAIIDSPPVDEEVLMSLVVELGLDRANELPELWLGQNEFDFTADFPSSC

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

CAAGCCGAGAGTCTCGGGATCGACACGTGGGGGCGCCTGAGCGAAGATAACCGTAATAAATAGTAACCTAA
CGGTCCAGTCATCGTTCTGTGGTCCCTTTCTTTTATGATTACAAGGAATGACCCTCTTCATCGCCTCTCCT
AATTCAGTCCTCACAACAGTCCTTTTACAAATGGGACAACAGGTTAGAGGAAGTCAGGCAGATTTCCAGCA
TCATAGAGAGTAAAGGACCAGGGAAGGATCAGGATTCAAGGACTGCACCCAGGCTCTGCTTCCAGCTTGCT
GTGTGACTTTGGGTAATTTTGTTCCTTAGGGAACTGAGCTTTCTCATTTGTAAATGCAAACAGGCTGTTG
GGAGGATCAAATGAGATCCAGGGGTGAAAACAGCTTAGTTTACTTTTCAGGAATTTACCCACGCGGTATATA
AAGGCAAAATATTATTATAGTCAGGTGATTGTAGATTGAGGAACCCATTTCCCTCATTCTGCAAATTGCAA
CCTGAGGGCCCAAAGAGGGACAGGGGCTTGCCCCAGGTCTCAGCAGGCTGTGAGCAAGAGCTAAAGCCTAA
TCCTCCTGCCTTTGGGCCTGGAGCCCTTCCCTGTACCCCAGGGGTGAGTGTCTTTGTTGGATACAGGCTTA
GATTGACTGACTGTACCTTGAGAACCTAGGGGAGTCCCTGTTCCCAATTCTTCTCCTACCCACCTTGCC
CTGATGGAGGAAGACCCTGCTGTGTTGAGATGAGCACCAGAGCCAAGAAGCTGAGGAGGATCTGGAGAATT
CTGGAGGAAGAGGAGAGTGTGCTGGAGCTGTACAGACCCTGCTTCTCAGGTCCAGGAAGGTGGCGTCAG
CATCTGCAGCCGCGTCGACGTTGTGCGGAGCCTCCGCGGAGGCCAGGAGAGCCGGACTAGGACCAGGGCC
CTGGGCCTCCCCACACTCCCCATGGAGAAGCTGGCGGCCTCTACAGAGCCCCAAGGGCCTCGGCCGGTCTCT
GGGCCGTGAGAGTGTCCAGGTGCCCGATGACCAAGACTTTTCGAGCTTCCGGTCAGAGTGTGAGGCTGAGG
TGGGCTGGAACCTGACCTATAGCAGGGCTGGGGTGTCTGTCTGGGTGCAGGCTGTGGAGATGGATCGGACG
CTGCACAAGATCAAGTGCCGGATGGAGTGTCTGTGATGTGCCAGCCGAGACACTCTACGACGTCCTACACGA
CATTGAGTACC GCAAGAAATGGGACAGCAACGTCAATTGAGACTTTTGACATCGCCCGCTTGACAGTCAACG
CTGACGTGGGCTATTACTCCTGGAGGTGTCCCAAGCCCCTGAAGAACCGTGATGTCATCACCCCTCCGCTCC
TGGCTCCCCATGGGCGCTGATTACATCATTATGAACTACTCAGTCAAACATCCCAAATACCCACCTCGGAA
AGACTTGGTCCGAGCTGTGTCCATCCAGACGGGCTACCTCATCCAGAGCACAGGGCCCAAGAGCTGCGTCA
TCACCTACCTGGCCAGGTGGACCCCAAAGGCTCCTTACCCAAGTGGGTGGTGAATAAATCTTCTCAGTTC
CTGGCTCCCAAGGCCATGAAGAAGATGTACAAGGCGTGCCTCAAGTACCCCGAGTGGAAACAGAAGCACCT
GCCTCACTTCAAGCCGTGGCTGCACCCGGAGCAGAGCCCCTTGCCGAGCCTGGCGCTGTGCGGAGCTGTGCG
TGCAGCATGCGGACTCACTGGAGAACATCGACGAGAGCGCGGTGGCCGAGAGCAGAGAGGAGCGGATGGGC
GGCGCGGGCGGCGAGGGCAGCGACGACACCTCGCTCACC TGA GCGCCGCACCGCTTCAGGGACGGAGA
CAGGACCGGGCGAGCCCTGGGGCGGCGGCCGCTCCTGCACCTTCTCCCCTCCCCACCCGGCACCTGGTGG
CACCGGGCCAGGCCAGGCGGGTGTGTCAGCCTGGCTGGACAGAGCCCCAATAAACGATCCCACAGCCTCA

VASASAAASTLSEPPRRQTESRTRTRALGLPTLPM EKLAASTE PQGPRPVLGRESVQVPDDQDFRSFRSEC
EAEVGNLTYSRAGVSVWQAVEMDRTLHKIKRMECCDVPAETLYDVLHDI EYRKKWDSNVIETFDIARL
TVNADVGYYSWRCPKPLKNRDVITLRSWLPMGADYIIMNYSVKHPKYPPRKDLVRAVSIQTGYLIQSTGPK
SCVITYLAQVDPKGLPKWVVKSSQFLAPKAMKKMYKACLKYPEWKQKHLPHFKPWLHPEQSPLPSLALS
ELSVQHADSLENIDESAVAESREERMGGAGGEGSDDDTSLT

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

TGCGCGATCTCCCGGAGC**ATG**CGCAGCAGCGGCGCCGACGCGGGGCGGTGCCTGGTGACCGCGCGCTCC
CGGAAGTGTGCCGGCGTCGCGCGAAGGTT**CAGCAGGGAGCCGTGGGCCGGGCGCGCCGGTTCCCGGCACGT**
GTCTCGGCACGTGGCAGCGCGCCTGGCCCTGGGCTTGGAGGCGCCGGCGCC**CTGGATCCGCCGGCCGTGGT**
CGCCGAGTCGGTGTGTCCTTGACCATCGCCGACGCGTTTATTGCAGCCGGCGAGAGCTCAGCTCCGACCC
CGCCGCGCCCCGCGCTTCCAGGAGGTTTTCATCTGCTCCTTCCCTGACTGCAGCGCCAATTACAGCAAAGCC
TGGAAGCTTGACGCGCACCTGTGCAAGCACACGGGGGAGAGACCATTTGTTTGTGACTATGAAGGGTGTGG
CAAGGCCTTCATCAGGGACTACCATCTGAGCCGCCACATTCTGACTCACACAGGAGAAAAGCCGTTTGT
GTGCAGCCAATGGCTGTGATCAAAAATTCAACACAAAATCAAACCTGAAGAAACATTTTGAACGCAAACAT
GAAAATCAACAAAAACAATATATATGCAGTTTTGAAGACTGTAAGAAGACCTTTAAGAAACATCAGCAGCT
GAAAATCCATCAGTGCCAGCATA**CCAATGA**ACCTCTATTCAAGTGTACCCAGGAAGG**ATGTGGGAAACACT**
TTGCATACCCAGCAAGCTGAAACGAC**ATGCCAAGGCCACGAGGGCTATGTATGTCAAAAAGGATGTTCC**
TTTGTGGCAAAAAC**ATGG**ACGGA**ACTTCTGAAACATGTG**GAGAGAAACCATAAAGAGGAAATACTATGTGA
AGT**ATGCCG**AAAAACATTTAAACGCAAAGATTACCTTAAGCAACAC**ATG**AAAACTCATGCCCCAGAAAGGG
ATGTATGTGCGTGTCCAAGAGAAGGCTGTGGAAGA**ACCTATA**CAACTGTGTTTAATCTCCAAAGCCATATC
CTCTCCTTCCATGAGGAAAGCCGCCCTTTTGTGTGTGAACATGCTGGCTGTGGCAAAAACATTTGCAATGAA
ACAAAGTCTCACTAGGCATGCTGTTGTACATGATCCTGACAAGAAGAAAATGAAGCTCAAAGTCAAAAAAT
CTCGTGAAAAACGGAGTTTGGCCTCTCATCTCAGTGGATATATCCCTCCCAAAGGAAACAAGGGCAAGGC
TTATCTTTGTGTCAAAACGGAGAGTCA**CCCAACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTACT**
TACCCTTGGCTAAGA**ACTGCACTGCTTTGTTTAAAGGACTGCAGACCAAGGAGCGAGCTTTCTCTCAGAGC**
ATGCTTTTCTTTATTA**AAAATTACTGATGCAGAACATTTGATT**CCTTATCATTTC

LDPPAVVAESVSLTIADAFIAAGESSAPTPPRPALPRRFICSPDCSANYSKAWKLD AHLCKHTGERPFV
CDYEGCGKAFIRDYHLRHLTHTGEKPFVCAANGCDQKFNTKSNLKKHFERKHENQKQYICSFEDCKKT
FKKHQQLKIHQCHTNEPLFKCTQEGCGKHFASPSKLRHAKAHEGYVCQKGC SFVAKTWTPELLKHVRETH
KEEILCEVCRKTFKRKDYLKQHM**KTH**APERDVCRCPREGCGRTYTTVFNLQSHILSFHEESRPVCEHAGC
GKTFAMKQSLTRHAVVHDPDKKKMKLKVKK**SREK**SLASHLSGYIPPKRKQGLSLCQNGESPNCVEDKM
LSTVAVLTLG

+++++

#2

NM_001418

Region conserved from insects to human.

EIF4G2 eukaryotic translation initiation factor 4 gamma

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

GCAGACGGCAACCGGGCCGCTGATTGGGCGGCGAAGGAGCCATT**CGGGGAGACTCTGGTGGGTTCCGGCTGC**
CCCAAGAGTGATAAGTT**CGGCTT**CAGACACGCCTTAGCGCCAGCAGT**GAGTCGGAGCTCTATG**GAGGTGGC
AGCGGGTACCGAGTGGCGGCTGCAGCAGCGACTCCTCTGAGCTGAGTTT**GAGGCCGTCCCCGACTCCTTCC**
TCCCCCTTCCCTCCCCCTTTTTTTTTGTTTTCCGTTCCCTTTCCCTTCCCTTCCCTATCCCCGACGACCGG
ATCCTGAGGAGGCAGCTGCGGTGGCAGCTGCTGAGTTCTCGGTGAAGGTATTT**CATTTCTCCTGTCCCTC**

CCCTCCCCACCCCATCTATTAATATTATTCTTTTGAAGATTCTTCGTTGTCAAGCCGCCAAAAGTGGAGAGT
GCGATTGCAGAAGGGGGTGCTTCTCGTTTCAGTGCTTCTTCGGGCGGAGGAGGAAGTAGGGGTGCACCTCA
GCACTATCCCAAGACTGCTGGCAACAGCGAGTTCCTGGGGAAAACCCAGGGCAAACGCTCAGAAATGGA
TTCCTGCACGAAGCACTAGACGAGATGACAACCTCCGCAGCAAACAACCTCCGCAAACGAAAAAGAACGACAT
GATGCAATCTTCAGGAAAGTAAGAGGCATACTAAATAAGCTTACTCCTGAAAAGTTTTGACAAGCTATGCCT
TGAGCTCCTCAATGTGGGTGTAGAGTCTAAACTCATCCTTAAAGGGGTGCATACTGCTGATTGTGGACAAAG
CCCTAGAAGAGCCAAAGTATAGCTCACTGTATGCTCAGCTATGCTCTGCGATTGGCAGAAGATGCACCAAAC
TTTGATGGCCCAGCAGCAGAGGGTCAACCAGGACAGAAGCAAAGCACCACATTGACACGCCTCCTAATTTCC
CAAATTACAAGATGAATTTGAAAACCGAACTAGAAATGTTGATGCTATGATAAGCGTGAAAATCCCCTCC
TCCCCGAGGAGGAGGAACAGAGAGCCATTGCTAAGATCAAGATGTTGGGAAACATCAAATTCATTGGAGAG
CTTGGCAAGCTTGATCTTATTACGAATCTATCCTTCATAAGTGCATCAAAACACTTTTGGAAAAGAAGAA
GAGATCCAACCTCAAAGATATGGGAGAGGATTTGGAGTGCCTCTGTGAGATAATGAGGACAGTGGGACCTA
GATTAGACCATGAACGAGCCAAGTCTTAATGGATCAGTACTTTGCCGAATGTGCTCCTTGATGTTAAGT
AAGGAATTGCCAGCAAGGATTTCGTTTCTGCTGCAGGATACCGTAGAGTTGCGAGAACACCATTGGGTTCC
TCGCAAGGCTTTTTCTTGACAATGGACCAAAGACGATCAATCAAATTCGTCAAGATGCAGTAAAAGATCTAG
GGGTGTTTATTCTGCTCCTATGGCTCAAGGGATGAGAAGTGACTTCTTTCTGGAGGGACCGTTTCATGCCA
CCCAGGATGAAAATGGATAGGGACCCACTTGGAGGACTTGTGATATGTTTGGACAAATGCCAGGTAGCGG
AATTGGTACTGGTCCAGGAGTTATCCAGGATAGATTTTACCCACCATGGGACGTCATCGTTCAAATCAAC
TCTTCAATGGCCATGGGGGACACATCATGCCTCCACACAATCGCAGTTTGGAGAGATGGGAGGCAAGTTT
ATGAAAAGCCAGATTAGCCTGAGGCCTGCTCAGTCTGTTTCTAATGAATAAAAATCAAGTGCCAAAGCTTCA
GCCCCAGATAACTATGATTCTCCTAGTGCACAACCACCACGCACTCAAACACCACCTCTGGGACAGACAC
CTCAGCTTGGTCTCAAACCTAATCCACCGCTTATCCAGGAAAAGCCTGCCAAGACCAGCAAAAAGCCACCA
CCGTCAAAGGAAGAACTCCTTAAACTAACTGAAACTGTTGTGACTGAATATCTAAATAGTGGAAATGCAAA
TGAGGCTGTCAATGGTGTAAAGAGAAATGAGGGCTCCTAAACACTTTCTTCTGAGATGTTAAGCAAAGTAA
TCATCCTGTCACTAGATAGAAGCGATGAAGATAAAGAAAAAGCAAGTTCTTTGATCAGTTTACTCAAACAG
GAAGGGATAGCCACAAGTGACAACCTCATGCAGGCTTTCTGAATGTATTGGACCAGTGTCCCAAACCTGGA
GGTTGACATCCCTTTGGTGAATCCTATTTAGCACAGTTTGCAGCTCGTGCCATCATTTCAGAGCTGGTGA
GCATTTCAGAACTAGCTCAACCCTAGAAAAGTGGCACCCTTTTCTCTCTTCTACTTTTGTCTTCAGCAG
TTAGCTAAATTAAGATCGAGAATGGTTAACAGAACTTTTTCAACAAAGCAAGGTCAATATGCAGAAAAT
GCTCCCAGAAATGATCAGAATAAGGACCGCATGTTGGAGATTTTGAAGGAAAGGGGACTGAGTTTCTTAT
TCCCCTCCTCAAATGGAGAAGGAAGTGTGAAGCAAATAAAGTTGGATCCATCCCCTCAAACCATATAT
AAATGGATTAAGATAACATCTCTCCCAAACCTCATGTAGATAAAGGATTTGTGAACATCTTAATGACTAG
CTTCTTACAGTACATTTCTAGTGAAGTAAACCCCCCAGCGATGAAACAGATTTCATCCTCTGCTCCTTCCA
AAGAACAGTTAGAGCAGGAAAAACAACCTACTACTATCTTTCAAGCCAGTAATGCAGAAAATTTCTTCATGAT
CACGTTGATCTACAAGTCAGTGCCCTGTATGCTCTCCAGGTGCACTGCTATAACAGCAACTTCCCAAAGG
CATGTTACTTTCGCTTTTTTGTGCACTTCTATGACATGGAAATATTGAAGAAGAAGCTTTCTTGGCTTGG
AAGAAGATATAACCCAAGAGTTTCCGGGAAAAGGCAAGGCTTTGTTCCAGGTGAATCAGTGGCTAACCTGG
TTAGAAACTGCTGAAGAAGAAGAATCAGAGGAAGAAGCTGACTAAAGAACCAGCCAAAGCCTTAAATTGTG
CAAACATACTGTTGCTATGATGTAAGTGCATTTGACCTAACCACTGCGAAAATTCATTCCGCTGTAATGT
TTTCACAATATTTAAAGCAGAAGCAGTCAGTTAGGATTTCTTCTGCATAAGGTTTTTTTTGTAGTGTAAT
GTCTTAATCATAGTCTACCATCAAATATTTTAGGAGTATCTTTAATGTTTAGATAGTATATTAGCAGCATG
CAATAATTACATCATAAGTTCTCAAGCAGAGGCAGTCTATTGCAAGGACCTTCTTTGCTGCCAGTTATCAT
AGGCTGTTTTAAGTTAGAAAACCTGAATAGCAACACTGAATACTGTAGAAATGCACTTTGCTCAGTAATACT
TGAGTTGTTGCAATATTTGATTATCCATTTGGTTGTTACAGAAAAATTTAACTGTAATTGATGGTTGTT
GCCGTAATAGTATATTGCCTGTATTTCTACCTCTAGTAATGGGCTTTATGTGCTAGATTTTAAATATCCTTG
AGCCTGGGCAAGTGCACAAGTCTTTTTAAAAGAAACATGGTTTACTTGCACAAAACCTGATCAGTTTTGAGA
GATCGTTAATGCCCTTGAAGTGGTTTTTGTGGGTGTGAAACAAATGGTGAGAATTTGAATTGGTCCCTCCT
ATTATAGTATTGAAATTAAGTCTACTTAATTTATCAAGTCATGTTTCATGCCCTGATTTTTATATACTTGTAT
CTATCAATAAACATTGTGATACTTGTATGATGTA

VESAI AEGGASRFSASSGGGSRGAPQHYPKTAGNSEFLGKTPGQNAQKWI PARSTRRDDNSAANNSANEK
ERHDAIFRKVRGILNKLTPKFDKLCLELLNVGVESKLILKGVILLIVDKALEEPKYSSLYAQLCLRLAED
APNFDGPAAEGQPQKQSTTFRLLI SKLQDEFENRTRNVVDYDKRENPLLPEEEEEQRAIAKIKMLGNIK
IGELGKLDL IHESILHKCIKTLLEKKKRVQLKDMGEDLECLCQIMRTVGPRLDHERAKSLMDQYFARMCSL
MLSKELPARIRFLLQDTVELREHHWVPRKAFLDNPGPKTINQIRQDAVKDLGVFIPAPMAQGMRSDFFLGEP
FMPPRMKMDRDLPLGLADMFGQMPGSGIGTGPVGIQDRFSPTMGRHRSNQLFNHGGHIMPPTQSQFGEMG

KGFMKSQISLRPAQSFLMNKNQVPKLQPQITMIPPSAQPPRTQTPPLGQTPQLGLKTNPPLIQEKPAKTSK
KPPPSKEELLKLTETVVTEYLNLSGNANEAVNGVREMRAPKHFLPEMLSKVIIISLDRSDEDKEKASSLISL
LKQEGIATSDNFMQAFNLVLDQCPKLEVDIPLVKSYLEAQAARAIISELVSISELAQPLESGTHFPLFLLC
LQQLAKLQDREWLTELQSKVNMQKMLPEIDQNKDRMLEILEGKLSFLFPLLKLEKELLKQIKLDPSQ
TIYKWIKNISPKLHVDKGFVNILMTSFLQYISSEVNPPSDETDSSSAPSKEQLEQEKQLLLSFKPVMQKF
LHDHVDLQVSALYALQVHCYNSNFPKGMLLRFFVHFYDMEIEEEAFLAWKEDITQEFPGKGKALFQVNQW
LTWLETAEEEESEEEAD

+++++

#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

CCTCGCCCGCCTGCCGCCTTTTTGTGCGCGTGTGAGTGTGGGCCCCAGCGTGCCTCCCGGGGGTGGGTTCC
CGGGCGGAAGGCGGAGGCCCGGCGCGCAGCCCGCCCGCCGCTGCCCGCGGACCGGGGAGCCGGGGTGTCTT
GGAGCGGGGGACGCCAGGCGTGGGCTGGCGGGCGGGACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCGCGG
GCTGGCGCTTGCCCGGGCGCAGTCGGCGGGGACCGAGTCGTACTTCCTGTGCGAAAGGCGGCCCGACCCTA
ACCGCCACCCCTCCCTGTCTCCCTCTCTGAACCCGCCATTGGGGGTAGGACACTCAGCCGTACCAGC
TCGCTCTGCTGGCCGCTACCTGCAGCAAGATAGGGCCGCCATCGCCGGGCGACGACGAGGAGGAGGCGGCC
GCCGCGCCGGGCCCCCGCCGCGCCGGAGCAGAGTGAATTTGGCTTCTGCACAGTTAGGAGGAGGACACC
AAACCGATGGGAGGTTTTGTGACCCACACCTCAACATAAAAAGATGAAGCTGGTAATCTAGTCCAGATTC
CAAGTGTCTACTTCAAGTGGGCAGTATGTTCTTCCCTTTCAGAAATTTGCAGAATCAACAAATATTTTCC
GTTGCACCAGGATCAGATTCATCAAATGGTACAGTGTCCAGTGTCAATATCAAGTGATACCACAGATCCA
GTCAGCAGATGGTTCAGCAGGTTCAAATGGTTTTACAGGCTCTTCAGATAATGGGGTATAAATCAAGAAA
GCAGTCAAATTCAGATCATTCTGGCTCTAATCAAACCTTACTTGCCTCTGGAACACCTTCTGCTAACATC
CAGAATCTCATACCACAGACTGGTCAAGTCCAGGTTCCAGGAGTTGCAATTGGTGGTTCATCTTTTCTGG
TCAAACCCAAGTAGTTGCTAATGTGCCTCTTGGTCTGCCAGGAAATATTACGTTTGTACCAATCAATAGTG
TCGATCTAGATTCTTTGGGACTCTCGGGCAGTTCTCAGACAATGACTGCAGGCATTAATGCCGACGGACAT
TTGATAAACACAGGACAAGCTATGGATAGTTCCAGACAATTCAGAAAGGACTGGTGAGCGGGTTTCTCCTGA
TATTAATGAAACTAATACTGATACAGATTTATTTGTGCCAACATCCTCTTCATCACAGTTGCCTGTTACGA
TAGATAGTACAGGTATATTACAACAAAACACAAATAGCTTGACTACATCTAGTGGGCAGGTTTATTCTTCA
GATCTTCAGGGAAATTATATCCAGTCGCCTGTTTCTGAAGAGACACAGGCACAGAATATTAGGTTTCTAC
AGCACAGCCTGTTGTACAGCATCTACAACCTCAAGAGTCTCAGCAGCCAACAGTCAAGCCAAATTTGTGC
AAGGTATTACACCACAGACAATCCATGGTGTGCAAGCCAGTGGTCAAATATATCACAACAGGCTTTGCAA
AATCTTCAGTTGCAGCTGAATCCTGGAACCTTTTTAATTCAGGCACAGACAGTGACCCCTTCTGGACAGGT
AACTTGGCAAACGTTTCAAGTACAAGGGGTCCAGAATTTGCAGAATTTGCAAATACAGAATACTGCTGCC
AACAAATAACTTTGACGCCTGTTCAAACCTCACACTTGGTCAAGTTGCGGCAGGTGGAGCCTTCACTTCA
ACTCCAGTTAGTCTAAGCACTGGTCAAGTGTCCAAATCTACAAACAGTTACAGTGAATCTATAGATTCTGC
TGGTATACAGCTACATCCAGGAGAGAATGCTGACAGTCCTGCAGATATTAGGATCAAGGAAGAAGAACCTG
ATCCTGAAGAGTGGCAGCTCAGTGGTGAATCTACCTTGAATACCAATGACCTAACACACTTAAGAGTACAG
GTGGTAGATGAAGAAGGGGACCAACAACATCAAGAAGGAAAAAGACTTCGGAGGGTAGCTTGCACCTGTCC
CAACTGTAAAGAAGGTGGTGGAAAGAGGTACCAATCTTGGGAAAAAGAAGCAACACATTTGTATATACCAG
GATGTGGTAAAGTCTATGGGAAGACCTCACATCTGAGAGCTCATCTGCGTTGGCATTCTGGAGAACGCCCT
TTTGTGTTGTAAGTGGATGACTGTGGTAAAAGATTTACTCGAAGTGAATTAACAGAGGCACAGAAGAAC
ACATACAGGTGAGAAGAAATTTGTTTGTCCAGAATGTTCAAACGCTTTATGAGAAGTGACCACCTTGCCA
AACATATTAACACACCAGAATAAAAAAGGTATTCACTCTAGCAGTACAGTGCTGGCATCTGTGGAAGCT
GCGCGAGATGATACTTTGATTACTGCAGGAGGAACAACGCTTATCCTTGCAAATATTCAACAAGGTTCTGT
TTCAGGGATAGGAACTGTTAATACTTCCGCCACCAGCAATCAAGATATCCTTACCAACACTGAAATACCTT
TACAGCTTGTACAGTTTCTGGAAATGAGACAATGGAGTAAATATTACACAAATACTTATTATTGTTGGTT
ATTTTTTATACAGTAGTGAGAAGAATATTGTTCTTAAGTCTTAGATATCTTTTTATTGATGTGCAAAAAT

TTTGGATTGACAGTAACTTGGTTATAACATGACACTGAAATGCCTTACTTTGTATGATATTCCATAGTATAT
TAAAAATGGTAAAATTGCATGGGTTTTGTAGGTACTTTTTGGAATCTAGAAGAAATGAAATTTTACCAAGTT
ATATAAAGAGAAAATTGAATTTAACAAATGCGAATGGTAGTCTAACCAAATGCATCAATCCTGTGTGGTTTA
GTGTA AAAATGAGAACATGTTGGTATTTATCTATTGTAAGATAAAAAAGCTGGTGGGTGAAAGAAATCATG
TTATGATAAAAAATTTTGTAAATTTCTTGATGACTGGAATTTTTATTATGCATAACTGACAAATCAAGTTT
CCAAGCAAATGTTACATAGTGTAGGCTTTACTTAGCTTATCAATTTGTCATTTTGAAGCTAATTATTTTAA
TTAGGTAACTATGTACAATATTTTAAAGCATTACTCTTGTAAAGATTTTGA AAACTACATTTTAAACATGGAA
CTCTAGGGATAGTCACCTTTTAAATCCTGTTGAAAAGCCATGTTTAAAGATTTAATTTGCCAAAATAATGTC
TTGTTAATATTCTTTCAATAACGAAGTTGGGCAATATAACCAATGTTTAAAAAAGTTTAAAAATGTATAAGT
TGAGGCATTTGGGTGGTAAGAGAATGTTATAGTGAATTATCCCTTTTCTTGACTATTGGAGGACCAAAAA
ATAAGGTGTATTGCGTCTTAGCAGTGATTTTTATCCAATCTTGTTC AAAAACCATGTCTCCAGGCCT
TAAAGCCATCATGTAAATACCAGTAAAGTGAACATATGCAAACATAACAAAATCACTTCCATAGTGAC
GATACTCCAACCATATGGATATTAGTCATAGAAGAACTAGAGGTTTTATGATATTTTTTTAAGTCTTTTTT
TTTTTTGTCTAGGTAGTCAGTCTGCACTTAAATATCAATCATTTTCTTTTTTGTCTTCTCCCTTAAAT
TATATGTATCCAGTACATTTAATTGAGAAGCGTATGTTTTTTATTATGCTGTATTTTTCTTTTTATTTTTTA
ATTATTGTTTATATTTTCAATTC AAAATGTACAAAATAAAGTTACATTGCTGGTCTTGTAAAGAGCTATAC
AGTTTTCTTAAATGTATACCTGTA ACTGCAGCAGTTCACCTATTTCAAAAATTTGGAATTTCTGTTCAATTTG
TTATTCTTAAAGACCACCTCAAATTTAAAGGCTACCTTATTGTACGTTTTAAAGTGTATTATAACAGTGTGGT
AGTTAATAAAACACTATTTTTTTTTCTTTTGA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAA

IGPPSPGDDEEEAAAAGAPAAAGATGDLASAQLGGAPNRWEVLSATPPTTIKDEAGNLVQIPSAATSSGQY
VLPLQNLQNLQIFSVAPGSDSSNGTVSSVQYQV I PQIQSADGQQVQIGFTGSSDNGGINQESSQIQIIPGS
NQTLASGTPSANIQNLI PQTGQVQVQVQVAIGGSSFPQGTQVVANVPLGLPGNITFVPINSVDLDSLGLSG
SSQMTAGINADGHLINTGQAMDSSDNERTGERVSPDINETNTD TDLFVPTSSSSQLPVTIDSTGILQQN
TNSLTSSGQVHSSDLQGNYIQSPVSEETQAQNIQVSTAQPVVQHLQLQESQQPTSQAQIVQG ITPQTIHG
VQASGQNISQQALQNLQLQNLNPGTFLIQAQTVTPSGQVTWQTFQVQGVQNLQNLQIQNTAAQQITLTPVQT
LTLGQVAAGGAFTSTPVSLSTGQLPNLQTVTVNSIDSAGIQLHPGENADSPADIRIKEEEPDPPEEWQLSGD
STLNTNDLTHLRVQVVDDEEGDQHQEGKRLRRVACTCPNCKEGGGRGTNLGKKKQHICHIPGCGKVKYKTS
HLRAHLRWHSGERPFVCNWMYCGKRFTRSDQLRHRRTHTGEKKFVCPECKRFMRSDHLAKHIKTHQNKK
GIHSSSTVLASVEAARDTLITAGGTTLILANIQQGSVSGIGTVNTSATSNDILNTNTEIPLQLVTVSGNE
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

ATTAGAGTCTGTGCTTCACTTCCGTTCCAGCCTCAGCGGCAGCTGGATCGCTCGACGGAGTGCCTCTGGTA
GTTGGCCAAAGACGCCGAATATCAAATCTT CAGCGGCAGCTCCCACCAGGACTTATCCAGAAAATTGCTG
ACCGCCTGGGCCTGGAGCTAGGCAAGGTGGTACTAAGAAATTCAGCAACCAGGAGACCTGCGTGGAAAT
GATGAGAGTGTGCGTGGAGAGGATGCTACATCGTTCAGAGTGGTTGTGGCGAAATCAACGACAGTCTAAT
CGAGCTTTTTGATCATGATTAATGCCTGCAAGATTGCTT CAGCTAGCCGAGTTACTGCAGTCATCCCATGCT
TCCCTTATGCCCGACAGGATAAGAAGGATAAGAGCCGGTCCCAATCTCTGCCAAGCTTGTGCAAAATATG
CTCTCTATAGCAGGTGCGGATCATATCATCACCATGGACCTACATGCTTCTCAAATTCAGGGCTTTTTTGA
TATCCAGTAGACA ACTTGTATGCAGAGCCA ACTGTCTGAAGTGGATAAGGGAGAATATCCCTGAGTGGGA
AGA ACTGCATTATTGTCTCGCCAGATGCTGGTGGAGCTAAAAGAGTGACCTCCATTGCAGACCAGTTGAAT

GTGGACTTTGCTTTGATTTCATAAAGAACGGAAGAAGGCCAATGAAGTGGACTGCATAGTGTCTAGTGGGAGA
TGTGAATGATCGTGTGGCTATCCTTGTAGATGACATGGCAGACACTTGTGTTACAATCTGCCTCGCAGCTG
ACAAACTTCTCTCAGCTGGAGCAACCAGAGTTTATGCTATCTTGACTCATGGAATCTTTTCTGGCCCAGCC
ATTTCTCGCATCAACACTGCATGCTTTGAAGCAGTGGTAGTCACCAATACCATACCTCAAGATGAGAAGAT
GAAGCATTGCTCCAAAATACGAGTAATTGACATCTCCATGATCCTTGCAGAAGCCATAAGGAGAAGCTCATA
ATGGGGAATCTGTTTCTACCTGTTTCAGCCATGTTCTTTTAAACAGAAATAACTTCTAGGTTATGCTATTT
TAAAATAAATAAGATTAATAAAAAA

TPNIKIFSGSSHQDLSQKIADRLGLELGLKVVTKKFSNQETCVEIDESVRGEDVYIVQSGCGEINDSLMELL
IMINACKIASASRVTAVIPCFPYARQDKKDKSRSPISAKLVANMLSIAGADHIITMDLHASQIQGFFDIPV
DNLYAEPTVLKWIWIRENIPEWKNCIIVSPDAGGAKRVTSIADQLNVDFALIHKERKKANEVDCIVLVGDVND
RVAILVDDMADTCVTICLAADKLLSAGATRVYAILTHGIFSGPAISRINTACFEAVVVNTNIPQDEKMKHC
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

GTGGCCCCGCCGCGCCACTCCCTCCGGCTCCCTCCCTCCCGCCGCGGGCGCGCATCTCATTCCAGCCCTCATT
CGCGCATTCCAGCGTCTCCTCGCACACTCGAGGCCAGGGGGCGGGAGGGCCGCGAGCTCCGGCGCCGCGC
GTCCC GCCAGAACGATCGCCGCGGCCGGAAGAGTTGGCGCTCGGGGCGGACTCCTTGGAACTGGCTTAGCG
CACCCATCCCACCTTCCCGCACCTGGGACCGGTCCAACGAGCGCTCCTCCAAGCGGAGCCTTGGAGGGCA
CGGCCGCCACCATTACCTCCAACGAGTGGAGCTCTCCACCTCCCTGAGGGGAGCACCGCCTCTGGGGGC
AGTCAGGCACTGGACAAGCCCATCGACAATGACGCAGAGGGCGTGTGGAGCCCGGATATTGAGCAGAGTTT
CCAGGAGGCCCTCGCCATCTACCCGCCCTGTGGCAGGCGCAAATCATCCTGTCTGGACGAGGGCAAGATGT
ATGGTCGGAACGAGCTGATTGCCCGCTACATCAAGCTCCGGACAGGGAAGACCCGCACCAGGAAGCAGGTC
TCCAGCCACATCCAGGTGCTGGCTCGTCGCAAAGCTCGCGAGATCCAGGCCAAGCTAAAGGACCAGGCAGC
TAAGGACAAGGCCCTGCAGAGCATGGCTGCCATGTCGTCTGCACAGATCATCTCCGCCACGGCCTTCCACA
GTAGCATGGCCCTCGCCCGGGCCCCGCGCCAGCAGTCTCAGGGTTTTGGCAAGGAGCTTTGCCAGGC
CAAGCCGGAACGTCCCATGATGTGAAGCCTTCTCTCAGCAAACCTATGCTGTCCAGCCTCCGCTGCCTCT
GCCAGGTTTGGAGTCTCCTGCAGGGCCCCCAGTCCGCTCTCGCCCCCGGCACCCCCATGGCAGGGCC
GCAGCGTGGCCAGCTCCAAGCTCTGGATGTTGGAGTCTCTCGCTTCTGGAGCAGCAGCAGGACCCGGAC
ACGTACAACAAGCACCTGTTCTGTCACATTTGGCAGTCCAGCCCAAGCTACAGCGACCCCTACCTCGAAGC
CGTGGACATCCGCCAAATCTATGACAAATCCCGGAGAAAAAGGGTGGACTCAAGGATCTCTTGAACGGG
GACCCTCCAATGCCTTTTTTCTTGTGAAGTCTGGGCAGACCTCAACACCAACATCGAGGATGAAGGCAGC
TCCTTCTATGGGGTCTCCAGCCAGTATGAGAGCCCCGAGAACATGATCATCACCTGCTCCACGAAGGTCTG
CTCTTTCCGCAAGCAGGTGGTGGAGAAAGTTGAGACAGAGTATGCTCGCTATGAGAATGGACACTACTCTT
ACCGCATCCACCGGTCCCCGCTCTGTGAGTACATGATCAACTTCATCCACAAGCTCAAGCACCTCCCTGAG
AAGTACATGATGAACAGCGTGTGGAGAACTTACCATCCTGCAGGTGGTCAACAACAGAGACACACAGGA
GACCTTGCTGTGCATTGCCTATGTCTTTGAGGTGTGAGCCAGTGTGACACGGGGCTCAGCACCATCTACA
GGCTGGTGAAGAATGAGAGACTCGGGGAGCAGGGAGGGGGGAAGAGACGTGTGTGCAGGAAACGGGGACG
TGGGGAGGGGACCTGCAGGGGCGACCCCTGAAGTGCCAAGAGAGCTGAGAGGAGCAGTTGTGACTCTACC
CAGGAACAAACTGTGCCTGAACCTGAGGTGCCCAACCCCAAATAAACCCAAAGATGCTGTGTATTTTCAGAG
GAAAA

LEGTAGTITSNEWSPTSPEGSTASGGSQALDKPIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDE
GKMYGRNELIARYIKLRTGKTRTRKQVSSHIVLARRKAREIQAKLKDQAAKDKALQSMAAMSSAQIISAT
AFHSSMALARGPGRPAVSGFWQALPGQAGTSHDVKPFSQQTYAVQPPLPLPGFESPAGPAPSPSAPPAPP

WQGRSVASSKLMLEFSAFLEQQQDPDITYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDKFPEKKGGLKDL
FERGSPNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPENMIITCSTKVCSTFGKQVVEKVETERYENG
HYSYRIHRSPLCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTNRDTQETLLCIAVFEVSASEHGAQH
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; ETRF-1; TEAD3

TCCTCAACACAAACTTTCCGTCCTCCGCTCCCTCCTCCGCGCTCGGCGCCTCCCGCTCCAGCCCGGCTC
ATTCGCGACATTCCGGCCAGCCCCCTCCCCACGACCCCCCTTCCCGGCCCCCTTGCAGGCTCCCTCGGGC
CCGGCCGAGCGGGCCCGGCGGAGCGCCCCGCGAGCTCGGACCAGGCTCAGCCGCCAGTGGGCTCAGGCC
CAGAGCCAGAGCAACCAGCACAATAGCGTCCAACAGCTGGAACGCCAGCAGCAGCCCCGGGGAGGCCCGG
GAGGATGGGCCCCGAGGGCCTGGACAAGGGGCTGGACAACGATGCGGAGGGCGTGTGGAGCCCGACATCGA
GCAGAGCTTCCAGGAGGCCCTGGCCATCTACCCGCCCTGCGGCCGGCGGAAGATCATCCTGTGACAGCAGG
GCAAGATGTACGGCCGAAATGAGTTGATTGCACGCTATATTAACCTGAGGACGGGGAAGACTCGGACGAGA
AAACAGGTGTCCAGCCACATACAGGTTCTAGCTCGGAAGAAGGTGCGGGAGTACCAGGTTGGCATCAAGGC
CATGAACCTGGACAGGCTCCAAGGACAAAGCCCTTCAGAGCATGGCGTCCATGTCTCTGCCAGATCG
TCTCTGCCAGTGTCTGCAGAACAAAGTTTCCAGCCACCTTCCCCTCTGCCCCAGGCCGTCTTCTCCACTTCC
TCGCGGTTCTGGAGCAGCCCCCTCTCCTGGGACAGCAGCCTGGACCCTCTCAGGACATCAAGCCCTTTCG
ACAGCCAGCCTACCCCATCCAGCCGCCCTGCGCCGACGCTCAGCAGTTATGAGCCCTGGCCCCGCTCC
CCTCAGCTGCTGCCTCTGTGCCTGTGTGGCAGGACCGTACCATTGCCTCCTCCCGGCTGCGGCTCCTGGAG
TATTCAGCCTTCATGGAGGTGCAGCGAGACCCTGACACGTACAGCAAACACCTGTTTGTGCACATCGGCCA
GACGAACCCCGCTTCTCAGACCCACCCCTGGAGGCAGTAGATGTGCGCCAGATCTATGACAAATTCCCCG
AGAAAAAGGGAGGATTGAAGGAGCTCTATGAGAAGGGGCCCTAATGCCTTCTTCTTGTCAAGTTCTGG
GCCGACCTCAACAGCACCATCCAGGAGGGCCCGGAGCCTTCTATGGGGTCACTCTCAGTACAGCTCTGC
TGATAGCATGACCATCAGCGTCTCCACCAAGGTGTGCTCCTTTGGCAAACAGGTGGTAGAGAAGGTGGAGA
CTGAGTATGCCAGGCTGGAGAACGGGCGCTTTGTGTACCGTATCCACCGCTCGCCCATGTGCGAGTACATG
ATCAACTTCATCCACAAGCTGAAGCACCTGCCCGAGAAGTACATGATGAACAGCGTGTGGAGAAGTTCCAC
CATCCTGCAGGTGGTCCAGAGCCGGGACTCCCAGGAGACCCTGCTTGTGCTTTTGTCTTTCGAAGTCT
CCACCAGTGAGCACGGGGCCAGCACCATGTCTACAAGCTCGTCAAAGACTAGGGTGCCTCTGCGCCTCC
TTAAGGATGCAGGGTGCAGCATCTCCTCTCCACACCTGCCTGGCACCCTGGGGGGTCCAGGATTGAGGAT
TCATCTACCTGCCAGGCCTCAGGCCAGGACCAGGAGGCCTCCCCACCTACCCAGCACACACTCCCT
GCCACTGTTCTGCGCTTTAATTGTGGGAGAAGAGAGGAGAGGAGGGCTCAGCGGTGGGGCAGCCTGTCCGG
GGCGCTGACCCACCATACCCTGCTCTGCCAGCCTCGCGTGACCTCAGAGAGGTGGGGATAGGGGACACC
TTCAGCCTCCAGCATGTGTGGCCACTGTACCCCAACCCCTTGGGGGAGCATGATGGGCAGGTGAGGGC
AGGATGGAGACCAAGGGAGTCACTGAGCAGAGGCCCTGGGAGTGTCCGGTTGGGGTTGGACTGAGGACAGA
GGGGCCACACTTCTTGGCCCTTTGTGTCCAGGCCTGGTGCCAGACTCCTTGCATGGCTTGTGTGGTC
CTCAGACTCCGCACAGCGAGCGTAGGTCTCTGGGTTTTAGATGAAGTGCCAGGCTCCAGGAAGTTGAGGG
ACCCACAGGAGAGGTGGGCAGAGCTGGAGTTCTCATCCAGGGCTGCTTGTCCCCAGAGCCCAGGTTTATAC
TACCTCCCTGGGGCGGGGCTGGCCGAGGGTAGGGGAGAGGCTCTGCAGTGTGGAGTGGAGCCTCATCGA
GGGGCGCTGGGTTAGGGGAGCACCTGTTTTCAGACTGGGCATGAAGAAGGGAGCACAGCAGCTACTAGACCC
CATTAGCACCTCATTAGCCACAAGCCAGCCAGGGGCCCCAGGAAGATGGGGCACCCCCAGCACCCCTCCA
GATTGAGAGCAAGGTAGAGGAAGGAGTCCCAGCCTCTGGGCAGACCAGAGGCCCCAGAGGGAGAGAGTAGCA
GAAGGCTTTTGTATTTTCTCTTGCCTGAGGCTTGAATCTGACAAACCCTTGGTGGGCACTGCTCCCTTAGG
TTCTTCCCCACCTCAATCTACCTGCCTAGAGTAGCAGCTCCCAGACCCAGTTCTGGGACTGAAGGTTAACC
CTTACCTGCTGTCCCTTCTTAACACCCAGGCCCCAGAGCCAGCTGGGCCTGTCCAGCAGCCACCTGTGG
GTATTTATGAGTTTTCATATGAAGTACTGTGCCCTTCCCTTCCCTCATCCCGACCCTGCCCGAGCTTCCCTGA
AGGTCCCTCACTGTTTGCATATCGCTCAGGCCACCTCCAAACCCACCTAGGTTTTATAATGTATATTATAT
ATTTTTTTGTGTATTTTTTAAAATCCAGCTGTGATGGGTTATATCATAAATGCAGCTTGGGGTTGGAGCAGG

GGCCCTCAAAGGCCAGCTCCTGCTCAAAAAAAAAAAAAAAAAAAAAAAAAAATTAAAGTTATTTGTTTGTGGGTCA
GTCATGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

IASNSWNASSSPGEAREDPGLDKGLDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRNE
LIARYIKLRTGKTRTRKQVSSHIQVLARKKVREYQVGIKAMNLDQVSKDKALQSMASMSAQIVSASVLQN
KFSPPSPLPQAVFSTSSRFWSSPPLLGQQPGPSQDIKPFAPAYPIQPPLPPTLSSYEPLAPLPSAAASVP
VWQDRTIASSRLRLLEYSAFMEVQRDPDTYSKHLFVHIGQTNPAFSDPPEAVDVRQIYDKFPEKKGGLKE
LYEKGPPNAFFLVKFWADLNSTIQEGPGAFYGVSSQYSSADSMTISVSTKVCSEFGKQVVEKQVETEARLEN
GRFVYRIHRSPMCEYMINFIHKLKHLPEKYMNSVLENFTILQVVTSRDSQETLLVIAFVFEVSTSEHGAQ
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

CCCCGCTTCTGCCTGCGCTGTGAACCCCCCCCCAGCCGCCGGCACGGCCCCGCCCGCTGCCCGGTGGT
GGCCACGGCCCCCGGCTGCCCGTGGTCAAACTGGAGTCGCTGAAGCGCTGGAACGAAGAGCGGGGCCTC
TGGTGCAGAGAAGGGGGTGCAGGTGCTGCTGACGACGGTGGGCGCCTTCGCCCTTCGGCCTCATGACCAT
CGCCATCAGCACTGACTACTGGCTCTACACGCGCGCCCTCATCTGCAACACCACCAACCTCACGGCCGGCG
GCGACGACGGGACCCCCACCGCGGGGGCGGCGGCGCCTCGGAGAAGAAGGACCCCGGCGGCCTCACGCAC
TCGGGCCTCTGGAGGATCTGCTGCCTGGAAGGGTTGAAAAGAGGCGTCTGCGTGAAGATCAATCATTCCC
GGAGGACACGGACTACGACCACGACAGCGCGGAGTATCTACTCCGAGTTGTCCGGGCCTCCAGCATCTTCC
CCATCCTTAGCGCCATCCTGCTGCTGCTCGGGGGTGTGTGCGTGGCGGCCTCCCGCTTACAAGTCCAAG
AGGAACATCATTCTGGGCGCAGGGATCCTGTTTCGTGGCAGCAGGCCTGAGCAACATCATCGGCGTGATCGT
GTACATCTCCGCCAACGCGGGCGAGCCGGGCCGAAGCGGGACGAGGAGAAGAAAAACACTACTCGTACG
GCTGGTCCTTCTACTTCGGCGGGCTGTCTGTTTCTGCTGCGCCGAGGTGATAGGCGTGTGGCCGTCAACATC
TACATCGAGCGCAGCCGCGAGGCGCACTGCCAGTCTGCTCGGACCTGCTCAAGGCCGGCGGGGGCGCGGG
CGGCAGTGGCGGGAGCGGCCCTCGGCCATCCTCGTCTGCCAGTTACCGCTTCCGCTACCGCCCGCT
CCCCTCTAGCTCCCCTCCAGCGAGCCGTGCGCGTCCGCGGACGCGTCTCCCGCGGCCCCGGGGGCCCG
GGCTTTGCCTCCACGGACATCTCCATGTACACGCTCAGCCGCGACCCCTCCAAGGGCAGCGTGGCCGCGGG
GCTGGCGGGGGCCGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
GCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
CACAACGCCTTCCCCAAGGAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
GCCCCGCGCCACCGCGCCCTCTGCGCCCGCCCCGGGACCCTGGCCAAGGAGGCGGCGGCGGCGGCGGCGG
ACACGCTCAACAGGAAAACCACGCCTGTGTAGGGCGCGGCGGGGAGCCGAGGGGCGTGTCCGGGGC

LES�KRWNEERGLWCEKGVQVLLTTVGAFAAFGLMTIAISTDYWLYTRALICNTTNLTAGGDDGTPHRGGG
GASEKKDPGGLTHSGLWRICCLEGLKRGVCVKINHFPEDTDYDHDSA EYLLRVVRASSIFPILSAILLLLG
GVCVAASRVYKSKRNIILGAGILFVAAGLSNIIGVIVYISANAGEPGPKRDEEKKNHYSYGWSFYFGGLSF
ILAEVIGVLAVNIYIERSREAHQSRSDLLKAGGGAGSGSGPSAILRLPSYRFRYRRRSRSSSRSSSEPS
PSRDASPGGPGGPGFASTDISMYTLSRDPSKGSVAAGLAGAGGGGGGAVGAFGGAAGGAGGGGGGGGAGA
ERDRGGASGFLTLHNAFPKEAGGGVTVTVTGPPAPPAPAPPAPSAPAPGTLAKEAAAANTNTLNKRTTPV

+++++

#8

NM_016178

CUG initiation well conserved in mammals.

OAZ3 ornithine decarboxylase antizyme 3

Synonyms: AZ3; OAZ-t; TISP15; OAZ3

GTTGCCTAAACCTCTGCCACCCACCTGTGAACCTCACTTTGCCACAGGGAGGCACCTGAACTGAAAACTGC
CTTGTAAGAGGTGTGCGCCCTCTGTCTACTCCCTTTCTTATATCAAGAGGGGAAAAACACGTAACCTACCTC
TACCCGATCTGGTCACCATACGCCTATTACCTTTACTGTTACAAGTACCGGATCACTCTCCGGGAGAAGAT
GCTGCCTCGTTGTTATAAAAGCATCACTTATAAGGAAGAGGAGGACTTGACACTCCAGCCCCGTTCCCTGCC
TCCAGTGCTCCTGAGTCCCTAGTAGGCCTCCAGGAGGGCAAAGCACCGAGCAGGGTAACCACGACCAGCT
TAAAGAACTGTATTCCGGCTGGGAACTTGACGGTGCTGGCTACTGACCCCTGCTCCACCAGGACCCAGTAC
AGTTAGACTTTCACTTCCGCCTTACCTCCCAGACCTCTGCCATTGGCACGGCCTTCTCTGTGACCGTCA
CTCTTCCCTGGATATCCCATATCAGGCCTTGGATCAAGGCAACCGGGAAAGTTTGACTGCAACCCTGGAGTA
CGTGGAAAGAGAAGACAAATGTGGACTCTGTGTTGTGAACTTCCAGAATGATCGGAACGACAGAGGTGCC
TGCTGCGGGCCTTCAGCTACATGGGCTTTGAGGTGGTCAGACCAGATCACCTGCCCTCCCTCCCTTGGAC
AATGTCATCTTTATGGTGTATCCCTTGAAAGGGATGTTGGCCACCTGCCCAGTGAGCCTCCTTGAACATG
CTTATTCCAACGCTTTGAGGGGCTGGAAGCCTTGACACATGGAATCAGGGGCCCGGGATGTGATTCAGGAC
ACTTTCCATCCTAGGAATAAAGGGTAGTGCAATCAAAAAAAAAAAAAA

LPCKRCRPSVYLSYIKRGKTRNYLYPIWSPYAYLYCYKYRITLREKMLPRCYKSITYKEEDLTLQPRS
CLQCSESLVGLQEGKSTEQGNHDQLKELYSAGNLTVLATDPLLHQDPVQLDFHFRLTSQTSAHWHGLLCDR
RLFLDIPYQALDQGNRESLTATLEYVEEKTNVDSVFNFNQDRNDRGALLRAFSYMGFEVVRPDHPALPLP
DNVIFMVYPLERDVGHLPSPEP

+++++

#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

TTCCGAACATTCTTAGCATCGCTCGCGCCGCGCCGCGCCGCTGAGCCGAGCCGAGCCTCTGCTGCCGCCG
CCGCGCCCCCGCCCGCCCGCGGGCGCCACCAAGCACTTTGCAGACTCGCTTCCACCCTGCGGGCCATT
CCGCGCGGGGGCCCGGGCCCGGGCGGCGCGTCCAGGCACAGGCCATGCAGTGACGCCCCCACC
TCCACCTTTGCCCGGAGCGCGGGCAGCAGCCAGCGCGCCAGCCGGCCCCGGGGCAGGAGCGGTGCTAGGC
AGGGGTGGGGTGGCCGGGCCAGGGACCGGGAGCCGGGGAGGGAGCCGGGCACCGAGCAGAGGGCGGGGA
AGCGGCGCCGAAGTTTGCCTCGGACTCGCCGGGCGCTGCGGTGGCTCCCTGGGCCGAGGACTGTTGCTGCC
GCTGCCGCCCGCTTCAATTGCACATTCAAGTGGAAAATTTTTCAGGAGTCAGCAGAAAACATTGTGTCCAAA
AAAGACTGAGTGCAGTTACCACCAAACCCAGGAGGAGACTCTCCCTGGAAAACCTCCCTTCCCTTTCCGGT
TTATTTTCTTGAAGAGGCTCCAGGCTTCCGGCTTGAAAATCCCACCGCCAAAATTGAGCCCAGCAGCTGGA
GCGGCAGTGAGAGCCCTGCCGAAAACATGGAAAGGATGAGTGACTCTGCAGATAAGCCAATTGACAATGAT
GCAGAAGGGGTCTGGAGCCCCGACATCGAGCAAAGCTTTTCAGGAGGCCCTGGCTATCTATCCACCATGTGG
GAGGAGGAAAATCATCTTATCAGACGAAGGCAAAATGTATGGTAGGAATGAATTGATAGCCAGATACATCA
AACTCAGGACAGGCAAGACGAGGACCAGAAAACAGGTGTCTAGTCACATTACAGGTTCTTGCCAGAAGGAAA
TCTCGTGATTTTCAATCCAAGCTAAAGGATCAGACTGCAAAGGATAAGGCCCTGCAGCACATGGCGGCCAT
GTCCTCAGCCCAGATCGTCTCGGCCACTGCCATTATAACAAGCTGGGGCTGCCTGGGATTCCACGCCCGA
CCTTCCCAGGGGCGCCGGGGTTCTGGCCGGGAATGATTCAAACAGGGCAGCCAGGATCCTCACAAGACGTC
AAGCCTTTTGTGCAGCAGGCCTACCCCATCCAGCCAGCGGTACAGCCCCCATTCCAGGGTTTGGACCTGC
ATCGGCCCCAGCTCCCTCAGTCCCTGCCTGGCAAAGGTGCTCCATTGGCACAACCAAGCTTCGCTGGTGG
AATTTTCAGCTTTTCTCGAGCAGCAGCGAGACCCAGACTCGTACAACAACACCTCTTCTGTGCACATTGGG

CATGCCAACCAATTCTTACAGTGACCCATTGCTTGAATCAGTGGACATTCGTCAGATTTATGACAAAATTTCC
TGAAAAGAAAGGTGGCTTAAAGGAACTGTTTGGAAAGGGCCCTCAAATGCCTTCTTCTCGTAAAATTTCT
GGGCTGATTTAAACTGCAATATTTCAAGATGATGCTGGGGCTTTTTATGGTGTAAACCAGTCAGTACGAGAGT
TCTGAAAATATGACAGTCACCTGTTCCACCAAAGTTTGTCTCTTTGGGAAGCAAGTAGTAGAAAAAGTAGA
GACGGAGTATGCAAGTTTTGAGAATGGCCGATTTGTATACCGAATAAACCGCTCCCAATGTGTGAATATA
TGATCAACTTCATCCACAAGCTCAAACACTTACCAGAGAAATATATGATGAACAGTGTTTTTGGAAAATTTCC
ACAATTTTATTGGTGGTAAACAAACAGGGATACACAAGAACTCTACTCTGCATGGCCTGTGTGTTTGAAGT
TTCAAATAGTGAACACGGAGCACAAATCATATTTACAGGCTTGTAAAGGACTGAACATGGTTATTTATAT
ATATAGATATCTGTATATACACACACACATATGTGCACACACACACTCTCTCTCCATTATCGAACGACTGA
CTGTAAACCTCACCACACAGGGTGGTGCCTGGCCCCGAGGTACCCCCGACTTTTCTAAATCTTGTTTGAG
TGAAGTCATTTTTTTCATGTGTTTACACTATCATTGTAGCTGTGAAGTTCTGGTACAGTTGTAAGAAAGAGAA
ATTTAGTTGTTTTCTCTATGTTCTTTCAGATGTGCAGCCACAATTCCTCGGAAAGGTGAACCTGAACAAACC
CAAGTCTCTCTGTCAGAGCCCTGTTTTCTAATTGTGGTAGAAAATATTGAGACAGAGCATTGGCCATGGGA
CATTTACAGCCTTTATACAAATGTATTTAGTTCTCTTTTTTCCAACATAAAAATTTCTGTTTTAAGATACAA
GTAAAATTAATCTTTAAATATAAATGTAAATTAGTACACAAAATAAGAATCTTTAGACTTATCTTTGTAA
CTAATTAGGGTGGAAAGTTATGAAAGAATGTAATTCATAAATTTTTTTTTAAATGAAACCTTTTTTTTTTCT
TTTTGAAACCAAATGTTAAACTATAGCCTTAAAGAAATGCTTGGTAGAAGTGTCTAATGAGACAAATTTGT
ACTTTTTATCCTCAAGGTTAAACACTAATCTCCTAATCCATTAACCTCTTGAACAGGTATTACAAAGGAAGAA
AACTTCACCCCTTATCCTTAAACATATATAGTATATTTAAAAAATATAAAAATTGTATTGTACTAATGTGATG
ATGGATTATTTAATGAAAAAGAAAAATGGCTCTTTTTGCAATAAGTAGATACATACTGAAAAAATCTAAA
CTTACAATGTTTATAGTCTTGTGTGTGCAGTTATATTTTTATATGGACGACCAAATTTTTTTATTAAGATGAG
TAAATATTTGAACCACTGAATTTAATAACAAAATTTAAAATTTGGCATGAATACGGAATACTGCACTGTG
AGATGCAAAGTATACAGAATCTGTGGCTGGGAGAAAATTTTCATCAAATAGACAAGTAAAAGGCTCATCAGT
TTTAGCATCTCTGCTCCCCAGAAAATTTGTAAGCATCCTCACCAGCCTGTGGATACATTCTTTATTTCTAGT
GACCAATATGCATATTAACCTGCTATAACTAGGGCTATATGTGTAGGTATGTGTATACATATACACAAAT
GCACATATAGAGTTAACACATTTAGTGAACACTTGTTTAGTGTCACTCAGTTTGTAGGTGCTGATATGTA
CGTATATCTCAATGTGTCTGTAGACTTAGATAACATCCTCTTGAAGCACATCCATTTCTTTAGCGTCTCTCA
GTAAGTTACAGTACTTGTTTGACTTAGGTTTAAAGAGGCCAGCTACCTATCTCTGACCTTTTCAAATAGGC
TCATTTGGGAGATTTCTTTGGCAGGAGAGATTTCAACTTTCCAATCTAAGTATTCCAGAGCATGCCAGGC
AGAGTTGGTTGATGTGGCCAGATGTTTTGAGTTATTTCCCTTAAAGTGTCTTACTGGGGAGAGAACAGGGA
GTGCTCCTCCAGCTTCCCAAAGAAATATGTTTTTGTAAAGTGGTAGGAACATGTGCACACAATAGAACATGA
AATAAGTTTTTTAACTTGTAAAACATGTCAAGATTTTTTCCACCAAGCTAGAAAATAAAAAACTTAGTTCTA
CCACATCCAATTAACCTTACACACCCCTTCCCTGTCTCAACACCTGCTTTGACCCTGCTTTTTCTATTATTA
CATCAGTCAGCATCTTGTGGTCCCTAACATGAGGATGTGGCTGGCTCGTGGGAAACAGCAAAACACTAAGC
CTGACCTCTCCCAAATTTGGGAAGACCAGAGGAGAAAGTGCAAAACTGTCCCCATTTGGAATGCCATTCTCT
TCTAGAAACCAGTTGGACAGTGTCTCTGCCCCTTATAAACAGACTACTGTTGGGTCCCTGATTCCAGGC
TGGCCTGTGAAGGATTGCCCCAGGTGTCCCCTTTCACGGTTGTACATTTACAGTGACTTCTGTTGAACAC
CCCTCTTAGGGATGTTTTCTTTGCTCTTATTTCTGCATCTTTCTTAAAGGAAGCCCCATCCTCTCCCAG
GACCAGGAGTTTATGACCAGGCGAGCACAAATGGCTAAAAGCCAAGCTGTCTTAGAATTTCAAGTGGGAGAG
CTGTCTGGTTTCAATTTTACCCAGGAATGGTACTTTTTCAGTGCAGCCAGGAGGGCTCTTGGGATTTCTTTT
CCAAAGCACAAAATACTGGGACCCAAGAAGAACAGCTAGAGGACAACCTCTGTTGGCACAGAGACGGGGAC
AGCCCAGTCTGCTGACCTCACAGGGTCAGCTGGGCCCCCTGGTGCTTACCACCTGCATCCTCTTGCTCA
GAATGCCTTTGCAGTTGAGTTTTCTGGGTTTTCTATGATTGACCTTGAGGTTTACTCCTTGCTCTTACAACA
TTTTCTAAGGATTTTTAAAAGTTTACTTCTTGTCTTGTCTTCTTAAAGCTTTCTCCAGGACAGATATTTTCC
CTGTCTTAAACACTGGTCCAGTCATCCCAGTGGGCTTCTCTTTGTCTCTCCAGATTAGACCTTTGGGTGA
GATTGGCATCACAACATCTAATCTGAGTCTGTCTTTTGTCTTCTTAAAGCTTTCTCCAGGACAGATATTTTCC
TAAAAGCTTTCTAAAGCATACTAAAGAAGCCTTCCAGAGCCCCGTCTTGCTTCTCTCCAGGTGCTCTAT
CCCCCTCGAGACCTCTGTTGCCAGGCTTGTCTTACGGCCATCTTGTGTTGTCACTGCAGAGTTTGGAGGCC
AGTTTTCCACAGCCTAAACAGGGAGGAGCTGCAGAATGGGGCTCTGGTCTCTGGGCATTCAATTTCCCTCAT
AGAGGCTGAGAATAAAAACAAGGACTTATTACACATGTTCTAGAACCCCAGAATGGCCCAAGTTACCTGAG
ACCAGGGTTTTCTCAACCTTGACACCATTGACATTTTGGACTGGGTAATTCTTTGTTCTGCAGAGCTGTCTT
TTGCACTGTAGGAGATTTACTAATATCCCTGGCCTTACCCAGTAGTACCCTAGCACCTATTCCCCACCC
AGCGTGTCTCCAGATATTGTCAAATATCCCATCGGGTGCAAAATGATCCCTGGTCAAGATCTGTTGCCCAA
GATGTTACAGGTCACAATGACCACATTTGAAATTGTTTTCCCTTTCAATTTACCCTGTGAAAGCATCTCTC
CTAGAGCCTTGCAAGAGGCAGGTGACATTTGTGTCCATATTTCTTCTGTTTCCAGAACTTCTGTTTCCAAAC
AATTTCTCTCTCGCTACAAGTATTCTTTCACTCAGCACTGGGGAAGTTGGGAACAGCTGGTCACCATCATC
CCTTTAATCAACTCACACCTGTTTTAAAGAGTGTCTGATTTGACCTTCATCCCTTAGTTTTACTGGCGTTA

AAAAAAGTCTCAGCAATTTTTCATTATTTTCTCGTGGGTCTCATTATCAAACCTTTACTTATTTTCGGCATATT
TCCTCTGGGCTTCTTCTAGTTTTCTGCCTTACAAGCAATGCTGTTCTGTAAATTTATTGAAACCTCTGGAAC
ATTTACCTTTAGAGATGGAGGATGGAAGGATTGGTACCAGAAGAGGGCTAAGATACGTTTTTCTGTCTTGA
GCTGAAAGCACAGTCTACTCTCCTTCGTTTTGTGATGAGAAAGTTGAGGCCAGAGGGGAGGTGACATGTT
TAGAGTCACCCAGCTGGTTAGTGACAGAAAAAGCGTGAGAGTTGTCTAGGATTCCTGCCACTTTGGTCCCT
GGCCTCTCCTGGGGGAGGCTGCTGTTCTTAGGTGCTCTAAGCTTAATCCCTCAGAATGTGTGGACAGGTCA
GCTTAGAAGAGATGGGGAGATTGAGGATCCCCCTGTGCCAGAGCACAGCCTCACCGGATGCTGCTTCCCAC
ACTGAAGTGTCTGTCCGACCATTGCTATCTGAGGCATCCACAAGCAGGTAGGAAAGCTGGCGAGCCATTT
TACTTCTGAGGACAATTCCCCAGCCACAGGCTCTGAGTCAAATTTCTATTTGGTAAGCATCCTAGCAGCA
AAGTCTGCACCTCAGACCAGCCAAAAACAGCCCCATTCCAAGTACTTGGTGTCAAAAGTCCCCGAACGA
CTTTTAAACCCCAAGTCTTCTTAAGGTTTCAGTACTGTGGTGGCTTTAGCAGTTGTTTTGTGCAACTATAA
ATTTTAAATCATCTGAGATGACAGTCAATTTTACAACAGGTACATATTAATTTGTATAATTTGTAT
ATGCTCTGGTACACTACCTGAACTAACGAAGGGTAGAACTAATTCTGTTTGTGAGTGTTCACACCTGTAAC
ATTAGGAGGATATGTCTGCATTGCTTATTTCTTTATGTTGGTGTCTTCTGTGGCAAAGCCCTGCACATGGCA
TTTCTGAAAAGCCTTAAATCTTTAAGATGTTGCATGTAGGGTATGCAGTGCAAAAGGCTGCCTCAGAAGT
TGAGCCCTTTTGTAAAGCTGGAAGCATTCTCTTACTACTGTTACTTTTGTAGGAAGTTTTCAATTCAGAGC
TGCCAAAGTGTTCCTCGTAAGCAGTGCCTTAGTAATACCTTAGTCATGCCGCCAGCCTTTTCTTACACCAAT
TCCTAATGTTTCATTTACGAATTGGCCCAATATTGGAACAAAAACAAGCAAAATTTGTCTTCATTTTTGTTT
TGTAAGCCCATTTTTTCTCCAGTTCTATAGGAACTGACTGCTTGGTGTAAAATCCGAAACTGGACACAAG
TCAGTTCTTTTACCACACTCAAATGTATATACAAAACAAAAGTTGCAACTTCATAGTTTACTATGAAAA
GCAATTTGTACTTTTTAATGTTGCCTTTTTAAATTCATGACCAAATACTTAGCTATTTGTGAATCTTCTGCA
CTCTAGCATGAAAGTGCCTTTGGTTTTGAGATTCCAGCTTAGAAAAGTGTGCCATAATAACGATAATTTGT
AGAGAGACCAAAAAATTTTTGAGATCACCGTAATGCCTTTGGTTTTACCGGGATGAGTAACCAACCACAGGC
CTCTGTTTACAAGAGCACGACGTGGTCCCCGCCTGCTGCTAGTCTGTCTGCCACTGGGGGCCTCCCAACAT
CCATAGCACACTTCAGCGGAAGGACCCAGAACTGTTGTGTTTGTGTGTGCTGATGACCTAGTGTGTCAT
TTCACCTCGTCACCCAGCCCTGCGTCCGGATGAGGGGACTTCTGCACAAATGACAGAATCTCGGCTGGTGG
ACAGATACTACAGCTTTCTCCTCCTCCTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT
AATGACAGTGCACACTTATCTGGTTTACACAATGATACCAATTTTAAAAGTTGGAAGCCTCAAACCTGAGACG
ACAGTGCAGAAAAACAAAAGTGAAGTTAGGGTCTTAAAATTTGAAGTGTCTTCTTAGGGCAAAACATGTT
GACTCCGAGTATTGTGATGAATGTGCTACGAGAACTTCAAAGAGCACCATTACAATTTGGCATTTTT
AAAGAATGTTCCAGCCCTCAAAGGGCAACTCTTTAAAAGTCTTGTGGCTTTTATCCAAACCTTGTAGAA
ATTGGGAAAGCTGATAGAGGTAAGGAAGACGAGTGAAGGACAAGAAGGCCAAACACCAGCCAAAAAGAA
ACTAGGAAAAAAGATTTTTCTTTGCTAATATAGATGTAAAATAACATCAGACATCTTTGAAAATTAGCCT
CTAAACTCTTAATACATACGTTCTGTGTGTCTCTACCTGGCGTCTTTAAGAATATCCTCTCTGGGCTCTGA
AATTTTAGGAGTGATTCTTATCCACTCCAAGTTGTAAGTATTTGTAGAAATTTGTGCAAAACAAAACAAAAC
TATCAAATGAAAAGAAAATGTACTCAACCTAACTTATAGTTAGCAGCTGGAATTTCAACTCTTCCCTGCC
AGCACTATACCACAGTGTGGAAGAAATTAGTCAAATGCTTGTGTTTTCTGCTTCTCTTTTCAACTGTTACTG
TGCTTTGTTTGAAGTAGTTTTCTCTCTCAAAGCCGTTGCTTATATCGTTAAGAATGAAGGTTTGTGTTTA
AAATTTATTGCATTGCAAAGGGTAGTTTCACTGAAGTCATGCACCATTAAATAAGATGAAATATTTGTATT
TATTGTCCTACTTCTAAGCCGTAACCTCTTTTCTCTGTGAATTTGCATTGAGTCACTCATGCTACACTA
CATCGCTTTAGTATTTGAGATGGCATTATGTTTCTCTCGTTTATCATGAAATGGGGTCAGATTCCATCA
GATTCACCTCTGTGAGGTGGACTCTTGTCTGCCTTCCATGATGAGATTTTTTTTTCTCCTTCCCCTTTCTT
TAAGAGAGGCTGACAGATCTAGGTGTCAATCAATTGGAACCCAGTCTCTGATTTTTTTTTCTATTAGTTATTT
TCTATCATTAGTTTCACTGTGTAAATTAGATATCAACTGCACTTCTTTAAAAAAAATACATCTCCCTATT
ACCTCCTTGAAGATTTACTTCTGTAGGCCTTTTTCAATAGGCTCATGACTGCAGACAAGGAAAAAAAAG
TAAAAACAAAACAGTATGTGCCTGAAAATGACAAAAAAAATTTGTAACATTTAAAAAAGAAACCTGAA
TAGCCTTTAATTTCTTAATAATACACTTAAATTTTATGTAAATCGGTTTTTCGCCACGTGTGTTTGTTCACA
TTCTAAATGACTTAATGGGATTTCTCACGGTCTGTGTCTTTGTGTGACGTGTATAAAATGGGCTTGTGATGT
AAGCGTTTCTCTGGTCAAGTGGTTCTTTGATATTGACTGCTGCTGGGAGTGGGCTGTGGAACCTGCCTT
CGGGTAACTGGGTTCTTCTTGGGTAGATTGGAGAGATGGGGGTGGGCGTGGGCAAATTTCTCACACATGTTT
TCTTAACTATTTGCAGAACTTTCAAAGGCATTTGATTAAACCTCTTGGCAGTACAGTATTCTTGTATT
TGTTAACGTCTGTGTTTAGGTACTGGTACCTTTTTGTTTTAAAATGTTCTAAGTGTGGCTTTAAAGTGAA
TTTATCTTTAGTATGATAGTTATATGAAAATTATAGGATTTGTGTGCAGAGAATTTTTTTATAAAGTGCTT
TGTAAAAAAATGATATTCTAGCTTTTGGCGTACATATGTGTGATAACTTTAATACCCATGACAGTT
AAGTGCAATTTTCTACTCTAAAATGCTATTTTTGTGTGAGTTCCTGCAGGTGTTTTCTATGCTTTT
CAAAGTGACACATTTTGTATGCCTTCTTGATAAAGTGGTAGACATTTTGTAGCTTTCTAGAACTTTGTATT

CATACGGTATCAATGAAAAATAAAGAAAATGAAAGTGTGGGTACCTTTTTTATCTGCAAAAAAAAAAAAA
AAAAA

LENPTAKIEPSSWSGSESPAENMERMSDSADKPIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEG
KMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLRARRKSRDFHSKLKDQTAKDALKQHMAAMSSAQIVSATA
IHNKLGLPGIPRPTFPGAPGFWPGMIQTGQPGSSQDVKPFVQAYPIQPAVTAPIPGFEPASAPAPSVPAW
QGRSIGTTKLRLVEFSAFLEQQRDPDSYNKHLFVHIGHANHSYSDPLLESVDIRQIYDKFPEKKGGLKELF
GKGPQNAFFLVKFWADLNCNIQDDAGAFYGVTSQYESSENMTVTCSTKVCSFGKQVVEKVETEYARFENGR
FVYRINRSPMCEYMINFIHKLKHLPEKYMNSVLENFTILLVVTNRDQETLLCMACVFVSNSEHGAQHH
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

GTTAAGTTGGAGCCGACTCAGCGGCCGCCCATTTTTGTGCAGTCGCTGGGAAGGAAGGAGACGCCTAAAC
CGCGGCACTGCCCGGTTTTGAGCGTAGCCAAACCTGCCACCAGGCTTTGTAGCCCCGATTCTCTGTGTTTTG
CTCCCGTCTCCGACGAGAGAGGCGGCGACGGTGGCGTCTGCGACGGGAGACAGCGCGTCGGAGCGAGAGAG
CGCTGCGCCTGCCGCCGCCCAACAGCGGAGGCGCCGCCCATCGGTCTGTACCAGACCGGAGCCGCAGG
CCCTCCCAGCCCGGCCATCCGTGCCCGCTCCCAGATCTCTATCCTTTTTGGGACCATGCGCGGAGGAGGC
TTTGGGGACCGGGACCGGGATCGTGACCGTGGAGGATTTGGAGCAAGAGGTGGTGGTGGCCTTCCCCGAA
GAAATTTGGTAATCCTGGGGAGCGTTTTGCGTAAAAAAAGTGGGATTTGAGTGAGCTCCCCAAGTTTGAGA
AAAATTTTTATGTGGAACATCCGGAAGTAGCAAGGCTGACACCATATGAGGTTGATGAGCTACGCCGAAAG
AAGGAGATTACAGTGAGGGGGGAGATGTTTTGTCTAAACCCGTGTTTGCCTTCCATCATGCTAACTTCCC
ACAATATGTAATGGATGTGTTGATGGATCAGCACTTACAGAACCAACTCCAATTCAGTGCCAGGGATTTCC
CGTTGGCTCTTAGTGGCCGGGATATGGTGGGCATTGCTCAGACTGGCTCTGGGAAGACGTTGGCGTATCTC
CTGCCTGCAATTGTTTCATATTAACCACCAGCCATACTTGGAAAGGGGAGATGGCCCAATCTGTCTAGTTCT
GGCTCCTACCAGAGAGCTTGGCCAGCAAGTACAGCAGGTGGCCGATGACTATGGCAAATGTTCTAGATTGA
AGAGTACTTGTATTTATGGAGGTGCTCCTAAAGGTCCCCAGATTTCGAGACTTGGAAAGAGGTGTTGAGATC
TGCATAGCCACTCCTGGACGTCTGATAGATTTCTGGAGTCAGGAAAGACAAATCTTCGCCGATGTACTTA
CCTTGTATTGGACGAAGCTGACAGAATGCTTGGATGGGGTTTGAACCCAGATCCGTAAAATTGTTGACC
AAATCAGCCCTGATAGGCAGACACTGATGTGGAGTGCAACCTGGCCAAAAGAAGTAAGACAGCTTGCAGAG
GATTTCCCTTCGTGATTACACCCAGATCAACGTAGGCAATCTGGAGTTGAGTGCCAACCACAACATCCTCCA
GATAGTGGATGTCTGCATGGAAAGTAAAAAGACCACAAGTTGATCCAATAATGGAAGAAATAATGGCTG
AAAAGGAAAAACAAAACAATAATTTTTGTGGAGACAAAGAGACGCTGTGATGATCTGACTCGAAGGATGCGC
AGAGATGGTTGGCCAGCTATGTGTATCCATGGAGACAAGAGTCAACCAGAAAGAGATTGGGTACTTAATGA
GTTCCGTTCTGGAAAGGCACCCATCCTTATTGCTACAGATGTAGCCTCCCGTGGGCTAGATGTGGAAGATG
TCAAGTTTGTGATCAACTATGACTATCCAAACAGCTCAGAGGATTATGTGCACCGTATTGGCCGAACAGCC
CGTAGCACCAACAAGGGTACCGCCTATACCTTCTTCAACCCAGGGAACCTAAAACAGGCCAGAGAGCTTAT
CAAAGTGCTGGAAGAGGCCAATCAGGCTATCAATCCAAAACCTGATGCAGCTTGTGGACCACAGAGGAGGCG
GCGGAGGCGGGGTAAGGGTGGTCTGTTCTCGTTACCGGACCACTTCTTCCAGCCAACAATCCAATCTGATG
TATCAGGATGAGTGTGACCGAAGGCTTCGAGGAGTCAAGGATGGTGGCCGGAGAGACTCTGCAAGCTATCG
GGATCGTAGTGAAACCGATAGAGCTGGTTATGCTAATGGCAGTGGCTATGGAAGTCCAAATTCGCCTTTG
GAGACAAGCAGGCCAATACACCTATGGTCAAGGCACCTATGGGGCAGCTGCTTATGGCACCAGTAGCTAT
ACAGCTCAAGAATATGGTGTGGCACTTATGGAGCTAGTAGCACCACTCAACTGGGAGAAGTTCACAGAG
CTCTAGCCAGCAGTTTATGTTGGATAGGCCGGTCTGGGCAGCAGCCACAGCCACTGATGTCAACAGTTTGG
CACAGCCTCCGGGAGCTACCAATATGATAGGTTACATGGGGCAGACTGCCTACCAATACCCTCCTCCTCCT
CCCCCTCCTCCTCCTCACGTAATGAACCACTCAAGTGGTAGTGACTCCAGCAGACTTAATTACATTTT

AAGGAACACTGTCTTTTCTTTTTTTTTTCTCTTCGCCTTTTCTTTTTTTTTTCTTTTTTCTTTTTTTTTT
TTAATTTTTTCCCCCAACCATCGTGATTTGTCTTTTCATGCAGATTAGTTAGAATTCAGTCCAGGTTTCT
TCTGCCACCAAATGATCCAGTCTGGAATAACATTTTGTAAAAAATAATATATATATATATATATAT
AGCTGACTGGAAGAGATTAATTTCTTCCCCCAACTTCTTGCATGTTGAAGATATTTGAGCTATTTTTCATC
TAAAAGAGTAAGGTATTAGGCCCTTTTGTGGGAGCCCATGTTTTGTTTTTCTGAGTTGGTGGGGAGGGAG
GGAGGGGGAGGGCTGAATTGTTTTGCAGAGGAAGATGGCATCTGTGCTTTAAATTTCTCATTACTGGGTTA
GAAAACAAAGAGGGATTGCCCTGCACATTTTCTTTTGTGCTTTTAAATGTTTCTTAAGTTGGAACAGGTTT
CCTCGGGCCTGTTTTGACTGATTGCTGGAGTGCATTTGATAGTTAAAAATTACTAATTGGTTTTATTTCCC
TTCACACTCTGCCTCCCCACTTCTCCCCCGTTACTGAAAAATAACCATTTTAGTGTGAGGCTAGAAAATTG
AATTGCTGAGTTTTGTGTATCCTTTAAATTAACACACAAGTGTATTATTGTAGTGGTTAAACTGTAGCAT
CTCAGCATCTGGGTGGAAGCTGCCTATATTTCTTCCAGTTTAACTGGGGACCATCTGTGAAATTAATTTT
CCTCCAGACAGCTGTGTGAGCAAATGAACATAAATGCTCGCTGGAAATTTACTAACCAGTTTTATATT
GACCTGCAGTGAAAAAGCACATTTAATTATAAAACAATATATTCAAATGGGCAAATTTATTTTCAAATG
CAGTGTAGAGCTAGATTAAGCAACTCTTGGCACCTACTCTGCCCTTTTGGCAAAGTTACCTTGAACAA
AGAATCTTAAGGGTTTATTAAGAACTCTTTATTTTTCTTACATACCCTGTTCTCTGCAGTGTCTTCTAACAGC
TTCTGGGTGCAGATTTTCTTGGCATCCTTTTGCACCTAGCTTATTACAGGTAGGTAGTGTAAAGAAAAG
TCATGGAGGACTAAAGCCTAAGTCTTTTCACTTTTCTCCATCTGAAGGTAGGTGAGTTCATCCTCTTCA
TGGTAATGCTGTTTTACCAAGACTTTATAGCAGATGGACCCAGAAAGAATTTTCTGCTATTGTGTTCACTA
CAACAGGATAGGGACATCAGACAGCCCCAGAAACCCCTTCCAGATCTGATATGGGACTATTAATTTTTATG
CTGTTAATTGGTATTCAATCACAATGCAGTTGAAGGGGAAGGCTCCACTGCATTCTTTGGCTAAGGCCTG
AATGCTTGTCTCATCTGTAAGATCTATACTCGAGTTTTGTTTTCTTTTAAATTTCTTTAGGGAGAGAGGG
ATGTTTTCTGAGGGGTTCTGAAAGTATGATTCAATGTGCAACATACAGGTAGGTCTTTCAGCATAAGCTGAA
ATATATGCATGTA AAAACTTTGACATCTTTTTTTTTTAAATTTTCACTTTCTTCTTAACCTTTACTTCTCTT
TTGTCCCCCCCCCATCTTACAGAAGTTGAGGCCAAGGGAGAATGGTAGGCACAGAAGAAACATGGCAAAC
GCTCTGTGCTTTCAAACCAAAGTGTCCCCCAACCCCAAATTTGTCTAAGCACTGGCCAGTCTGTTGTGG
GCATTGTTTTCTACAACCAAATTTCTGGGTTTTTTTTCTTCTTTCTTTAAACATAGAGGTACCACCACAAGGG
ATGCCCTACTCTCTCGCAGCTCTTGAAGCATCTGTTTGAAGGAAAGGTCTCTGGGCAAGCAAGTGGTTAT
TTGGATTGCTTGTCTTCCCTTTTTCCACCTGGGACATTGTAATCATAAAATAACAGTAAATTTCAAACCTCA
AAAATATTATGGCTGAGCACAGCTGAAATCTAGCAGAGTTTAACTCTTCTGCCTCCATGTCTGTCACTT
ATAATTGAGTTCTGCTGTTGGCTTTCAGAACATGAGCAGAAGAATCGTTTTATGCTAGTTATTGCTATT
GGTTGAAACTCAACTTAGGGAAAGGGTTCCAATGTATTAAGCAATGGGCTGCTTCTCCCAATCCTCCCTA
ACAATTCGTTGTGTGGACTTCTCATCTAAAAGGTTAGTGGCTTTTGGCTTGGGATCAGTGTCTCTATTGAT
GTTCTTGGTGTCTCCAGACACATTTCTGTTGCATTAAGACTTGAAGACTTGTAGATGTGTGATGTTTCAG
GCACAGGATGCTGAAAGCTATGTTACTATTCTTAGTTTGTAAATTGCTCTTTTGTATACCATCATCTTGT
TCTTTTTGTAGGTATAAATAAAAAACTGTTGACAATAAAAAAAAAAAAAAAAAA

TVASATGDSASERESAAPAAAPTAEAPPPSVVTRPEPQALPSPAIRAPLPDLYPFGTMRGGGFGDRDRDRD
RGGFGARGGGGLPKKFGNPGERLRKKKWDLSELPKFEKNFYVEHPEVARLTPYEVDLRRKKEITVRGGD
VCPKPVFAFHANFPQYVMDVLMQHFTEPTPIQCQGFPLALSGRDMVGIQGTSGKTLAYLLPAIVHINH
QPYLERGDGPICLVLAPTRELAQQVQVADDYGKCSRLKSTCIYGGAPKGPQIRDLERGVEICIATPGRLI
DFLESKTNLRCTYLVLDEADRMLDMGFEPQIRKIVDQIRPDRQTLMWSATWPKEVRQLAEDFLRDYQI
NVGNLELSANHNILQIVDVCMESEKDHKLIQLMEEIMAEKENKTIIFVETKRRCDLRRMRRDGPAMCI
HGDKSQPERDWVLNEFRSGKAPILIAITDVASRGLDVEDVKFVINYDYPNSSEYVHRIGRTARSTNKG
TAYTFFTPGNLQARELIKVLEEANQAINPKMLQLVDHRGGGGGGGKGRSRYRTTSSANNPNLQYDECDRL
RGVKDGGRRDSASYRDRSETDRAGYANGSGYSPNSAFGAQAGQYTYGQTYGAAAYGTSSYTAQYEGAGT
YGASSTTSTGRSSQSSSQFSGIGRSGQPPQPLMSQQAQPPGATNMIGYMGQTAYQYPPPPPPPPSRK

+++++

#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

GGCTTGGGGCAGCCGGGTAGCTCGGAGGTCGTGGCGCTGGGGGCTAGCACCAGCGCTCTGTCTGGGAGGCGC
AGCGGTTAGGTGGACCGGTGAGCGGACTCACCAGCCAGGGCGCTCGGTGCTGGAATTTGATATTCATTGAT
CCGGGTTTTATCCCTCTTCTTTTTTCTTAAACATTTTTTTTTTAAAACGTATTGTTTCTCGTTTTAATTTA
TTTTTGCTTGCCATTCCCCTTGAATCGGGCCGACGGCTTGGGGAGATTGCTCTACTTCCCCAAATCACT
GTGGATTTTGGAAACCAGCAGAAAGAGGAAAGAGGTAGCAAGAGCTCCAGAGAGAAGTCGAGGAAGAGAGA
GACGGGGTCAGAGAGAGCGCGGGCGTGCAGCAGCGAAAGCGACAGGGGCAAAGTGAGTGACCTGCTTT
TGGGGGTGACCGCCGGAGCGCGGGCGTGAGCCCTCCCCCTGGGATCCCGCAGCTGACCAGTCGCGCTGACG
GACAGACAGACAGACACCGCCCCAGCCCCAGCTACCACCTCCTCCCCGGCCGGCGGGCAGAGTGGACGC
GGCGGCGAGCCGCGGGCAGGGGCCGGAGCCCGCGCCCGAGGCGGGGTGGAGGGGGTTCGGGGCTCGCGGCG
TCGCACTGAAACTTTTTCGTCCAATTTCTGGGCTGTTCTCGCTTCGGAGGAGCCGTGGTCCGCGCGGGGGAA
GCCGAGCCGAGCGGAGCCGCGAGAAGTGCTAGCTCGGGCCGGGAGGAGCCGAGCCGGAGGAGGGGGAGGA
GGAAGAAGAGAAGGAAGAGGAGAGGGGGCCGAGTGGCGACTCGGGCGCTCGGAAGCCGGGCTCATCGGACGG
GTGAGGCGGGCGGTGTGCGCAGACAGTGTCCAGCCGCGCGCGCTCCCCAGGCCCTGGCCCGGGCTCGGGC
CGGGGAGGAAGAGTAGCTCGCCGAGGCGCCGAGGAGAGCGGGCCGCCCCACAGCCCGAGCCGGAGAGGGAG
CGCGAGCCGCGCCGGCCCCGGTCTGGGCCTCCGAAACCATGAACTTTCTGTGTCTTGGGTGCATTGGAGCC
TTGCCTTGCTGCTCTACCTCCACCATGCCAAGTGGTCCCAGGCTGCACCCATGGCAGAAGGAGGAGGGCAG
AATCATCAGAAGTGGTGAAGTTTATGGATGTCTATCAGCGCAGCTACTGCCATCCAATCGAGACCCCTGGT
GGACATCTTCCAGGAGTACCCTGATGAGATCGAGTACATCTTCAAGCCATCCTGTGTGCCCTGATGCGAT
GCGGGGGCTGCTGCAATGACGAGGGCCTGGAGTGTGTGCCACTGAGGAGTCCAACATCACCATGCAGATT
ATGCGGATCAAACCTCACCAAGGCCAGCACATAGGAGAGATGAGCTTCTACAGCACAACAAATGTGAATG
CAGACCAAAGAAAGATAGAGCAAGACAAGAAAAAATCAGTTCGAGGAAAGGGAAAGGGGCAAAAACGAA
AGCGCAAGAAATCCCGGTATAAGTCTTGGAGCGTGTACGTTGGTGCCCGCTGCTGTCTAATGCCCTGGAGC
CTCCCTGGCCCCCATCCCTGTGGGCCTTGCTCAGAGCGGAGAAAGCATTGTGTTGTACAAGATCCGCAGAC
GTGTAATGTTCTTGCAAAAACACAGACTCGCGTTGCAAGGCGAGGCAGCTTGAGTTAAACGAACGTACTT
GCAGATGTGACAAGCCGAGGCGGTGAGCCGGGCAGGAGGAAGGAGCCTCCCTCAGGGTTTCGGGAACCAGA
TCTCTCACCAGGAAAGACTGATACAGAACGATCGATACAGAAACCACGCTGCCGCCACCACACCATCACCA
TCGACAGAACAGTCTTAATCCAGAAACCTGAAATGAAGGAAGAGGAGACTCTGCGCAGAGCACTTTGGGT
CCGGAGGGCGAGACTCCGGCGGAAGCATTCCCGGGCGGGTGACCCAGCACGGTCCCTCTTGAATTGGATT
CGCCATTTTTATTTTTCTTGCTGCTAAATCACCGAGCCCGGAAGATTAGAGAGTTTTATTTCTGGGATTCCT
GTAGACACACCCACCCACATACATATATATATATATATATATATATAAAAAATAAATATCTCT
ATTTTATATATATAAAATATATATATATCTTTTTTTTTAAATTAACAGTGCTAATGTTATTGGTGTCTTCACTG
GATGTATTTGACTGCTGTGGACTTGAGTTGGGAGGGGAATGTTCCCACTCAGATCCTGACAGGGAAGAGGA
GGAGATGAGAGACTCTGGCATGATCTTTTTTTGTCCCCTTGGTGGGGCCAGGGTCTCTCCCCTGCCCA
GGAATGTGCAAGGCCAGGGCATGGGGGCAAATATGACCCAGTTTTGGGAACACCGCAAACCCAGCCCTGG
CGCTGAGCCTCTACCCAGGTGAGCAGGACAGAAAGACAGATCACAGGTACAGGATGAGGACACCCGGC
TCTGACCAGGAGTTTGGGGAGCTTCAGGACATTGCTGTGCTTTGGGGATTCCCTCCACATGCTGCACGCGC
ATCTCGCCCCAGGGGCACTGCCTGGAAGATTGAGGAGCCTGGGCGGCCTTCGCTTACTCTCACCTGCTTC
TGAGTTGCCAGGAGACCCTGGCAGATGTCCCGGCGAAGAGAAGAGACACATTGTTGGAAGAAGCAGCCC
ATGACAGCTCCCCTTCTGGGACTCGCCCTCATCTCTTCTGCTCCCCTTCTGGGGTGCAGCCTAAAAG
GACCTATGCTCTCACACCATTGAAACCACTAGTCTGTCCCCCAGGAGACCTGGTTGTGTGTGTGTGAGT
GGTTGACCTTCCCTCCATCCCCTGGTCTTCCCTTCCCTTCCCGAGGCACAGAGAGACAGGGCAGGATCCAC
GTGCCATTGTGGAGGCAGAGAAAAGAGAAAGTGTATATACGGTACTTATTTAATATCCCTTTTTAAT
TAGAAATTAACAGTTAATTTAATTAAGAGTAGGGTTTTTTTTTTCAGTATTCTTGGTTAATATTTAATTT
CAACTATTTATGAGATGTATCTTTTGCTCTCTCTTGTCTCTTATTTGTACCGTTTTTTGTATATAAAATT
CATGTTTCCAATCTCTCTCTCCCTGATCGGTGACAGTCACTAGCTTATCTTGAACAGATATTTAATTTTGC
TAACACTCAGCTCTGCCCTCCCCGATCCCCTGGCTCCCAGCACACATTCCTTTGAAATAAGGTTTTCAATA
TACATCTACATACTATATATATATATTTGGCAACTTGTATTTGTGTGTATATATATATATATATATATATGTTTATGTA
TATATGTGATTCTGATAAAATAGACATTGCTATTCTGTTTTTATATGTAATAAAACAAAACAAGAAAAATA
GAGAATTCTACATACTAAATCTCTCTCCTTTTTTAATTTAATATTTGTTATCATTTATTTATTTGGTGCTA
CTGTTTATCCGTAATAATTGTGGGAAAAGATATTAACATCACGTCTTTGTCTCTAGTGCAGTTTTTCGAG
ATATCCGTAGTACATATTTATTTTTAAACAACGACAAAGAAATACAGATATATCTTAAAAAATAAAGC
ATTTTGTATTAAGAATTTAATTTCTGATCTCAAAAAAAAAAAAA

LTDRQTDTPSPSYHLLPGRRRRTVDAAASRGQPEPAPGGGVEGVGARGVALKLFVQLLGCSTRFGGAVVRA
GAEAPSGAARSASSGREEPQPEEGEEEEEEKEEERGPQWRLGARKPGSWTGEAAVCADSAPAARAPQALARA
SGRGGRVARRGAEESGPPHSPSRRGSASRAGPGRASET MNFLLSWVHWSLALLLLYLHHAKWSQAAPMAEGG
GQNHHEVVKFMDVYQRSYCHPIETLVDIFQEYPDEIEYIFKPSCVPLMRCGGCCNDEGLECVPTTESNITM
QIMRIKPHQGQHIGEMSFLQHNKCECRPKKDRARQEKKSVRGKGGQKRKRKKSRYKSWSVYVGARCCCLMP
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

TTCTTAACCCTTTCCAGCTTTCCACCCTCTTTGGCTTTAGCCATGGCCTTCTGATCTGTGTTTCTCAGGG
GACCTGCAGGCCCCAGATATAGCCCCATGCTGTCTCCTACCCAGAGCACACTGTTCCAGGCTACTTCCAC
TGGTACTGAAATCCAGTATTTCACTTACTCTTTTTCTTTCCAATATCCTCATGACATTCAATATTTCACTT
ACTCTAGGTCTCCTGCCTAAGGCCCAAGTCAACTTTCTGTCCAGTGGGATTTGTAATCCAATACCTCCT
AGCCCTAGCAGAATCCCATGTGGATAATCAGAAATGTGACTGGAAAAAGGACAGAGCTCTATGGCTGTGGG
TCCCAGTCCCCTGCTGGCAGTAAGTCCCAGCAGTGTGTAAGCACCTTACATTCTGCGCTTGGT
TGAAAACAGCAAGGCAAGCATCCACTTGAGAAATGTCAACCCTAGGAAATCCCAGCCTCAAGTCTTTCTC
ATCCCTTGGGAAGTGCAAATTGGATAGAGAAGAAACCAATTAATAAAACAAAACAAACAAATCATACTTAGAT
ATTCTGGCTTTTCTCACCAGGGCTGGATTAAGCATGTACTTCAAATAATAACAACCTTAAGTCAATAAAT
AAATGTAAGGAAGTCAAATGTTACCTGAAGACAACCTGTGGTCATTTTTTGGCAATCCCAGGTTCTCTTT
TCTACCTGTTTGTCTCAATCGTGGTCTCCCTCTCCCTCTCTTGTGGGGCCCATGCCCTGCTTTACTGTTG
CCAGAGGCTTGTACTTGTGGCTTTTAGGTAGGAGCAGTTACTTCCACTCCCCTCACCTGCCATAAAGCA
TCTTTATAAAACAAAGCAAGTAGAAGAAACACATCCTGGTATCCACCACATTCGGCTTTTGTGATTCTGTT
CACTTGGGAGCACCTGCTGCTAGGGAATAAGAAGGTTGAGGCTGAAGAGTGAGGACTCTTCAGCTCCCCTC
TGGCAGGACCCGGGAGAGGAAAGAGCCCTCAGCTGGTCCATCCTCCCCTCCTGGTCAGCCTTCTGTTCT
GAGATCAAAGTGGTGGGTACATTCTCGAGAATGTGCTCAGCCCCCTCATCTCACACTTTCCCTCTC
CCTGTGTCCCTGCCCTTTACATAACCATGCTGGTGATTGGCACCGTCAATAATCAATACTTTGCTCA
CTTTCACATCAAGTAACACTATCCAGGGAGGTGGTTTCAACAAAGGAGGAAGTATAAGGAGATCTAGGTTT
AAATTAATGTTGCCCTAGTGGTAAAGGACAGAGACCCTCAGACTGATGAAATGCACTCAGAATTACTTAG
ACAAAGCGGATATTTGCCACTCTCTTCCCCTTTTCTGTGTTTTTGTAGTGAAGAGACCTGAAAGAAAAA
GTAGGGAGAACATAATGAGAACAATAACGGTAATCTCTTCAATTTGCTAGTTCAAGTGCTGGACTTGGGACT
TAGGAGGGGCAATGGAGCCGCTTAGTGCCTACATCTGACTTGGACTGAAATATAGGTGAGAGACAAGATTG
TCTCATATCCGGGAAATCATAACCTATGACTAGGACGGGAAGAGGAAGCACTGCCTTTACTTCAGTGGGA
ATCTCGGCCCTCAGCCTGCAAGCCAAGTGTTCACAGTGAGAAAAGCAAGAGAATAAGCTAATACTCCTGTCC
TGAACAAGGCAGCGGCTCCTTGGTAAAGCTACTCTTGAATCGATCCTTTGCACCGGATTGTTCAAAGTGGA
CCCCAGGGGAGAAGTCGGAGCAAAGAATTACCACCAAGCAGTCCAAGAGGCCAGAAGCAAACCTGGAGG
TGAGACCCAAAGAAAGCTGGAACCATGCTGACTTTGTACTCTGTGAGGACACAGAGTCTGTTTCTGGAAAG
CCCAGTGTCAACGCAGATGAGGAAGTCGGAGGTCCCCAAATCTGCCGTGTATGTTGGGGACAAGGCCACTGG
CTATCACTTCAATGTCAATGACATGTGAAGGATGCAAGGGCTTTTTTTCAGGAGGGCCATGAAACGCAACGCC
GGCTGAGGTGCCCTTCCGGAAGGGCGCCTGCGAGATCACCCGGAAGACCCGGCGACAGTGCCAGGCCTGC
CGCCTGCGCAAGTGCCTGGAGAGCGGCATGAAGAAGGAGATGATCATGTCCGACGAGGCCGTGGAGGAGAG
GCGGGCCTTGTCAAGCGGAAGAAAAGTGAACGGACAGGGACTCAGCCACTGGGAGTGCAGGGGCTGACAG
AGGAGCAGCGGATGATGATCAGGGAGCTGATGGACGCTCAGATGAAAACCTTTGACACTACCTTCTCCCAT
TTCAAGAATTTCCGGCTGCCAGGGGTGCTTAGCAGTGGCTGCGAGTTGCCAGAGTCTCTGCAGGCCCCATC
GAGGGAAGAAGTGCCTGCAAGTGGAGCCAGGTCCGGAAGATCTGTGCTCTTTGAAGGTCTCTCTGCAGCTGC
GGGGGAGGATGGCAGTGTCTGGAACATAAAACCCAGCCGACAGTGGCGGGAAGAGATCTTCTCCCTG
CTGCCCCACATGGCTGACATGTCAACCTACATGTTCAAAGGCATCATCAGCTTTGCCAAAGTCATCTCCTA

CTTCAGGGACTTGCCCATCGAGGACCAGATCTCCCTGCTGAAGGGGGCCGCTTTTCGAGCTGTGTCAACTGA
GATTCAACACAGTGTTCACCGCGGAGACTGGAACCTGGGAGTGTGGCCGGCTGTCTACTGCTTGGGAAGAC
ACTGCAGGTGGCTTCCAGCAACTTCTACTGGAGCCCATGCTGAAATTCCACTACATGCTGAAGAAGCTGCA
GCTGCATGAGGAGGAGTATGTGCTGATGCAGGCCATCTCCCTCTTCTCCCCAGACCGCCAGGTGTGCTGC
AGCACC CGTGGTGGACCAGCTGCAGGAGCAATTGCGCATTACTCTGAAGTCCTACATTGAATGCAATCGG
CCCCAGCCTGCTCATAGGTTCTTGTTCCTGAAGATCATGGCTATGCTCACCGAGCTCCGCAGCATCAATGC
TCAGCACACCCAGCGGCTGCTGCGCATCCAGGACATACACCCCTTTGCTACGCCCTCATGCAGGAGTTGT
TCGGCATCACAGGTAGCTGAGCGGCTGCCCTTGGGTGACACCTCCGAGAGGCAGCCAGACCCAGAGCCCTC
TGAGCCGCCACTCCCGGGCCAAGACAGATGGACACTGCCAAGAGCCGACAATGCCCTGCTGGCCTGTCTCC
CTAGGGAATTCCTGCTATGACAGCTGGCTAGCATTCTCAGGAAGGACATGGGTGCCCCCAACCCAGTT
CAGTCTGTAGGGAGTGAAGCCACAGACTTTACGTGGAGAGTGCCTGACTGACCTGTAGGTGAGGACCATGAG
GAGCAAGGTGGCCCTTTCTTTTAAAAGGCCCTGTGGTCTGGGGAGAAATCCCTCAGATCCCACATAAAGT
GTCAAGGTGTGGAAGGGACCAAGCGACCAAGGATGGGCCATCTGGGGTCTATGCCACATACCCACGTTTG
TTCGTTTCTGAGTCTTTTTCATTGCTACCTCTAATAGTCTGTCTCCCACTTCCCACTCGTTCCCTCCTC
TTCCGAGCTGCTTTGTGGGCTCCAGGCCTGTACTCATCGGCAGGCGCATGAGTATCTGTGGGAGTCTCTA
GAGAGATGAGAAGCCAGGAGGCCTGCACCAAATGTCAGAAGCTTGGCATGACCTCATTCCGGCCACATCAT
TCTGTGTCTCTGCATCCATTTGAACACATTATTAAGCACCGATAATAGGTAGCCTGCTGTGGGGTATAACAG
CATTGACTCAGATATAGATCCTGAGCTCACAGAGTTTATAGTTAAAAAACAACAGAAACACAAACAATT
TGGATCAAAAGGAGAAATGATAAGTGACAAAAGCAGCACAAAGGAATTTCCCTGTGTGGATGCTGAGCTGTG
ATGGCGGGCACTGGGTACCCAAGTGAAGGTTCCCGAGGACATGAGTCTGTAGGAGCAAGGGCACAACTGC
AGCTGTGAGTGCCTGTGTGTGATTTGGTGTAGGTAGGTCTGTTTGCCACTTGATGGGGCCTGGGTTTGTTC
CTGGGGCTGGAATGCTGGGTATGCTCTGTGACAAGGCTACGCTGACAATCAGTTAAACACACCCGAGAAGA
ACCATTTACATGCACCTTATATTTCTGTGTACACATCTATTCTCAAAGCTAAAGGGTATGAAAGTGCCTGC
CTTGTATATAGCCACTTGTGAGTAAAAATTTTTTGCATTTTACAAATTATACTTTATATAAGGCATTCC
ACACCTAAGAAGTGTGTTTGGGAAATGTAGCCCTGGGTTTAAATGTCAAATCAAGGCAAAAGGAATTAATA
ATGTACTTTTGGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LEVRPKESWNHADVFHCEDESVPGKPSVNADEEVGGPQICRVCGDKATGYHFVMTCEGCKGFFRRAMKR
NARLRCPFRKGACEITRKRTRQCQACRLRKCLESGMKKEMIMSDEAVEERRALIKRKKSERGTQPLGVQG
LTEEQRMMIRELMDAQMKTDFDTTFSHFKNFRLPGVLSGCELPELQAPSREEAAKWSQVRKDLCSLKVSL
QLRGEDGSVWNYKPPADSGGKEIFSLPHMADMSTYMFKGIISFAKVISYFRDLPIEDQISLLKGAAFELC
QLRFNTVFNAETGTWECGRLSYCLEDTAGGFQQLLLEPMLKFHYMLKQLHEEEYVLMQAIISLSPDRPG
VLQHRVVDLQEQFAITLKSYLECNRPQPAHRFLFLKIMAMLTELRINAQHTQRLRLRIQDIHPFATPLMQ
ELFGITGS

+++++

#11

NM_001172131

Region conserved in mammals.

HCK hemopoietic cell kinase

Synonyms: JTK9; HCK

GTCCCAGCTCGGGAGCACATCAGAGGCTTAGAGGCGAGTGGGAAGGGACTCAGACAGTGCAGGACGAGAAA
CGCCCCGCGGCACCAAAGCCCTCAGAGCGTCGCCCCCGCCTCTAGTTCTAGAAAGTCAGTTTCCCGGCCT
GGCACCCCGGAACCTCAGGGGCTGCCGAGCTGGGGGGCGCTCAAGCTGCGAGGATCCGGGCTGCCCGCGA
GACGAGGAGCGGGCGCCAGGATGGGGTGCATGAAGTCCAAGTTCCTCCAGGTCCGAGGCAATACATTCTC
AAAACTGAAACCAGCGCCAGCCACACTGTCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGG
GGCCTAATAGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC
CTGTATGATTACGAGGCCATTACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATGGTGGTCTAGA
GAACTCGGGGAGTGGTGAAGGCTCGATCCCTGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATG
TCGCCCCGCTTGAATCTCTGGAGACAGAGGAGTGGTTTTTCAAGGGCATCAGCCGGAAGGACGCAGCGC
CAACTGCTGGCTCCCGGCAACATGCTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTA

CTCTTTGTCCGTGCGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGG
ACAACGGGGGCTTCTACATATCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCACTACAAG
AAGGGGAACGACGGGCTCTGCCAGAACTGTCCGGTGCCTGCATGTCTTCCAAGCCCCAGAAGCCTTGGGA
GAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAGAAGAACTTGGAGCTGGGCAGTTTGGGG
AAGTCTGGATGGCCACCTACAACAAGCACACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCC
GTGGAGGCCTTCTCGGCAGAGGCCAACGTGATGAAAACCTCTGCAGCATGACAAGCTGGTCAAACCTTCATGC
GGTGGTCACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGTCTGGACTTTCTGA
AAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACCTCATTGACTTCTCAGCCCAGATTGCAGAAGGCATG
GCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCCAACATCTTGGTCTCTGCATCCCT
GGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGTTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGG
CCAAGTTCGCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTACCATCAAGTCAGACGTC
TGGTCTTTGGTATCCTGTGATGGAGATCGTCACTACGGCCGATCCCTTACCAGGGATGTCAAACCC
TGAAGTGATCCGAGCTCTGGAGCGTGGATAACGGATGCCTCGCCAGAGAAGTCCCAGAGGAGCTCTACA
ACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCCAATACATCCAGAGTGTGCTG
GATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCAGGGCCA
GGGGGTGCCAGGTGGTGGCTGCAAGGTGGCTCCAGCACCATCCGCCAGGGCCACACCCCCCTTCTACTC
CCAGACACCCACCCTCGCTTCCAGCCACAGTTTCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTC
TTTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCAAGTTGATATTTCTATTTCTGGAA
TGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAATAGTGAATGAATATTTAAATAA
AAGATATAAATGCCAAAGTCTTTACCAAAAAAAAAAAAAAAAAAAAAA

LGGRSSCEDPGCPRDEERAPRMGCMKSKFLQVGGNTFSKTETSASPHCPVYVPDPTSTIKPGPNSHNSNTP
GIREAGSEDIIVVALYDYEAIIHEDLSFQKGDQMVVLEESGEWWKARSLATRKEGYIPSNYVARVDSLETE
EWFVKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYLSVRDYDPRQGDTVKHYKIRTLDNNGFFYISPR
STFSTLQELVDHYKKGNLGLCQKLSVPCMSKPKQKPEKDAWEIPRESLKKLEKKGAGQFGEVWMATYNKH
TKVAVKTMKPGSMSVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPL
PKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGGLARVIEDNEYTAREGAKFPIKWTAP
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

CCAGGCAGCTGGGGTAAGGAGTTCAAGGCAGCGCCACACCCGGGGGCTCTCCGCAACCCGACCGCCTGTC
CGCTCCCCACTTCCCGCCCTCCCTCCCACCTACTCATTACCCACCCACCCACCCAGAGCCGGGACGGCA
GCCAGGCGCCCGGGCCCGCCGTCTCCTCGCCGCGATCCTGACTTCTCTTGCTGAGGACCCGGCTTC
CACGTGTGTCCCGGAGCCGGCGTCTCAGCACACGCTCCGCTCCGGGCCTGGGTGCCTACAGCAGCCAGAGC
AGCAGGGAGTCCGGGACCCGGGCGGCATCTGGGCCAAGTTAGGCGCCGCGAGGCCAGCGCTGAACGTCTC
CAGGCGCGGAGGAGCGCGGGCGTCCGGTCTGAGCCGAGCAAATGGGCTCCGACGTGCGGGACCTGAA
CGCGCTGCTGCCCCGCCCTCCCTCCCTGGGTGGCGCGCGCGGCTGTGCCCTGCCTGTGAGCGCGCGCGC
AGTGGGCGCCGGTGTGACTTTGCGCCCCGGGCGCTTCGGCTTACGGGTGCTTGGGCGGCCCCGCGCG
CCACCGGCTCCGCCGCCACCCCGCCGCCCGCCTCACTCCTTCAACAACAGGAGCCGAGCTGGGGCGG
CGCGGAGCCGCACGAGGAGCAGTGCCTGAGCGCCTTCACTGTCCACTTTTCCGGCCAGTTCACTGGCACAG
CCGGAGCCTGTGCTACGGGCCCTTCCGGTCTCCTCCGCCAGCCAGGCGTCATCCGGCCAGGCCAGGATG
TTTCTAACGCGCCCTACCTGCCAGCTGCCTCGAGAGCCAGCCCGCTATTGCAATCAGGGTTACAGCAC

GGTCACCTTCGACGGGACGCCCAGCTACGGTCACACGCCCTCGCACCATGCGGGCGCAGTTCCCCAACCCT
CATTCAAGCATGAGGATCCCATGGGCCAGCAGGGCTCGTGGGTGAGCAGCAGTACTCGGTGCCGCCCCCG
GTCTATGGCTGCCACACCCCCACCGACAGCTGCACCGGCAGCCAGGCTTTGCTGCTGAGGACGCCCTACAG
CAGTGACAATTTATACCAAATGACATCCCAGCTTGAATGCATGACCTGGAATCAGATGAACTTAGGAGCCA
CCTTAAAGGGCCACAGCACAGGGTACGAGAGCGATAACCACACAACGCCCATCCTCTGCGGAGCCCAATAC
AGAATACACACGCACGGTGTCTTCAGAGGCATTAGGATGTGCGACGTGTGCCTGGAGTAGCCCCGACTCT
TGTACGGTCGGCATCTGAGACCAGTGAGAAACGCCCTTCATGTGTGCTTACCCAGGCTGCAATAAGAGAT
ATTTTAAGCTGTCCCCTTACAGATGCACAGCAGGAAGCACACTGGTGAGAAACCATAACCAGTGTGACTTC
AAGGACTGTGAACGAAGGTTTTCTCGTTTACAGACCAGCTCAAAAGACACCAAAGGAGACATACAGGTGTGAA
ACCATTCCAGTGTAAAACCTTGTGAGCGAAAGTCTCCCGGTCCGACCACCTGAAGACCCACACCCAGGACTC
ATACAGGTAATAAAGTGAAGAACCCCTTCCAGTGTGCGGTGGCCAAGTTGTGAGAAAAGTTTGCCCGGTCA
GATGAATTAGTCCGCATCAACAATGCATCAGAGAAAACATGACCAAACCTCCAGCTGGCGCTTTGAGGGGT
CTCCCTCGGGGACCGTTTCAAGTGTCCCAGGCAGCACAGTGTGTGAACTGCTTTCAAGTCTGACTCTCCACTC
CTCCTCACTAAAAGGAACTTCAAGTGTCTTCTTTCATCCAACCTTCCAAGACAAGATACCGGTGCTTCTG
GAACTACCAGGTGTGCCTGGAAGAGTTGGTCTCTGCCCTGCCTACTTTTTAGTTGACTCACAGGCCCTGGA
GAAGCAGCTAACAATGTCTGGTTAGTTAAAAGCCCATTGCCATTTGGTGTGGATTTTTCTACTGTAAGAAGA
GCCATAGCTGATCATGTCCCCCTGACCCTTCCCTTCTTTTTTTATGCTCGTTTTTCGCTGGGGATGGAATTA
TTGTACCATTTTCTATCATGGAATATTTATAGGCCAGGGCATGTGTATGTGTCTGCTAATGTAACTTTGT
CATGGTTTCCATTTACTAACAGCAACAGCAAGAAATAAATCAGAGAGCAAGGCATCGGGGTGAATCTTGT
CTAACATTTCCGAGGTGAGCCAGGCTGCTAACCTGGAAAGCAGGATGTAGTTCTGCCAGGCAACTTTTAAA
GCTCATGCATTTCAAGCAGCTGAAGAAAAAATCAGAATAACCAGTACCTCTGTATAGAAATCTAAAAGAA
TTTTACCATTAGTTAATTCAATGTGAACACTGGCACACTGCTCTTAAGAACTATGAAGATCTGAGATTT
TTTTGTGTATGTTTTTACTCTTTTTGAGTGGTAATCATATGTGTCTTTATAGATGTACATACCTCCTTGCA
CAAATGGAGGGGAATTCATTTTCACTGAGGAGTGTCTTAGTGTATAAAAACCATGCTGGTATATGGCT
TCAAGTTGTAAAATGAAAGTACTTTAAAAGAAAATAGGGGATGGTCCAGGATCTCCACTGATAAGACTG
TTTTTAAGTAACCTAAGGACCTTTGGGTCTACAAGTATATGTGAAAAAATGAGACTTACTGGGTGAGGAA
ATCCATTGTTTTAAGATGGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
TTTTAAGGGAGGGAATTTATTATTTACCGTTGCTTGAATTACTGTGTAAATATATGTCTGATAATGATTT
GCTCTTTGACAATAAATTAGGACTGTATAAGTACTAGATGCATCACTGGGTGTTGATCTTACAAGATAT
TGATGATAACACTTAAAATTGTAACCTGCATTTTTCACTTTGCTCTCAATTAAGTCTATTCAAAAAG

LDLFLLLQDPASTCVPEPASQHTLRSGPGCLQQPEQQGVDRDPGGIWAKLGAAEASAERLQRRSRGASGSEP
QQMGSDVRLNALLPAVPSLGGGGCALPVSQAAQWAPVLDFAAPPASAYGSLGGPAPPPAPPPPPPPPH
SFIKQEPSWGAEPHEEQCLSAFTVHFSGQFTGTAGACRYGPFPPPPPSQASSGQARMFPNAPYLPSCLES
QPAIRNQYSTVTFDGTTPSYGHTPSHHAQQFPNHSFKHEDPMGQQGSLGEQQYSVPPPVYGCHTPTDSTG
SQALLLRTPYSSDNLQMTSQLECMTNQMNLGATLKGHSTGYESDNHTTPILCGAQYRIHTHGVFRGIQD
VRRVPGVAPTLVRSASETSEKRPFMCAYPGCNKRYFKLSHLQMHRSRKHTEKPYQCDFKDCERRFSRSDQL
KRHQRRTGVPKPFQCKTCQRKFSRSDHLKTHTRTHTGKTSEKPFSCRWPSQKKFARSDELVRHNMHQRN
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

GCAGGCCGGGGCGGGGCTGGGAAGTAGTCGGGCGGGGTTGTGAGACGCCGCGCTCAGCTTCCATCGCTGGG
CGGTCAACAAGTGCGGGCCTGGCTCAGCGCGGGGGGGCGCGGAGACCGCGAGGCGACCGGGAGCGGCTGGG
TTCCCGGCTGCGCGCCCTTCGGCCAGGCCGGGAGCCGCGCCAGTCCGAGCCCCCGGCCAGCGTGGTCCGC

CTCCCTCTGGGCGTCCACCTGCCCGGAGTACTGCCAGCGGGCATGACCGACCCACCAGGGGCGCCGCCGCGC
GGCGCTCGCAGGCCGCGGATGAAGAAGAAAACCCGGCGCCGCTCGACCCGGAGCGAGGAGTTGACCCGGAG
CGAGGAGTTGACCCTGAGTGAGGAAGCGACCTGGAGTGAAGAGGCGACCCAGAGTGAGGAGGCGACCCAGG
GCGAAGAGATGAATCGGAGCCAGGAGGTGACCCGGGACGAGGAGTTCGACCCGGAGCGAGGAGGTGACCAGG
GAGGAAATGGCGGCAGCTGGGCTCACCGTGACTGTCAACACAGCAATGAGAAGCACGACCTTCATGTTAC
CTCCCAGCAGGGCAGCAGTGAACCAGTTGTCCAAGACCTGGCCAGGTTGTTGAAGAGGTTCATAGGGGTTT
CACAGTCTTTTCAGAAACTCATATTTAAGGGAAAATCTCTGAAGGAAATGGAAAACCCGTTGTGACACTT
GGAATACAAGATGGTTGCCGGGTTCATGTTAATTGGGAAAAAGAACAGTCCACAGGAAGAGGTTGAACTAAA
GAAGTTGAAACATTTGGAGAAGTCTGTGGAGAAGATAGCTGACCAGCTGGAAGAGTTGAATAAAGAGCTTA
CTGGAATCCAGCAGGGTTTTCTGCCAAGGATTTGCAAGCTGAAGCTCTCTGCAAACCTTGATAGGAGAGTA
AAAGCCACAATAGAGCAGTTTATGAAGATCTTGGAGGAGATTGACACACTGATCCTGCCAGAAAATTTCAA
AGACAGTAGATTGAAAAGGAAAGGCTTGGTAAAAAAGTTTCAGGCATTCTAGCCGAGTGTGACACAGTTGG
AGCAGAACATCTGCCAGGAGACTGAGCGGCTGCAGTCTACAAACTTTGCCCTGGCCGAGTGAAGGTGTAGCA
GAAAAAGGCTGTGCTGCCCTGAAGAATGGCGCCACCAGCTCTGCCGTCTCTGGAGCGGAATTTACCTGATT
TCTTCAGGGCTGCTGGGGGCAACTGGCCATTTGCCAATTTTCTACTCTCACACTGGTTCTCAATGAAAAA
TAGTGTCTTTGTGATTTTGTAGTAAAGCTCCTATCTGTTTTCTCCTTCTGTCTCTGTGGTTGTACTGTCCAG
CAATCCACCTTTTCTGGAGAGGGCCACCTCTGCCCAAATTTTCCAGCTGTTTGGACCTCTGGGTGCTTTC
TTTGGGCTGGTGAGAGCTCTAATTTGCCTTGGGCCAGTTTCAGGTTTATAGGCCCCCTCAGTCTTCAGATA
CATGAGGGCTTCTTTGCTCTTGTGATCGTGTAGTCCCATAGCTGTAAAACCAGAATCACCAGGAGGTTGCA
CCTAGTCAGGAATATTGGGAATGGCCTAGAACAAGGTGTTTGGCACATAAGTAGACCACTTATCCCTCATT
GTGACCTAATTCAGAGCATCTGGCTGGGTTGTTGGGTTCTAGACTTTGTCTCACCTCCCAGTGACCCTG
ACTAGCCACAGGCCATGAGATACCAGGGGGCCGTTCTTGGATGGAGCCTGTGGTTGATGCAAGGCTTCT
TGTCCCCAAGCAAGTCTTCAGAAGGTTAGAACCAGTGTGACTGAGTCTGTGCTTGAAACCAGGCCAGAG
CCATGGATTAGGAAGGGCAAAGAGAAGGCACCAGAATGAGTAAAGCAGGCAGGTGGTGAAGCCAACCATAA
ACTTCTCAGGAGTGACATGTGCTTCTTCAAAGGCATTTTGTAAACCATATCCTTCTGAGTTCTATGTTT
CCTTCACAGCTGTTCTATCCATTTTGTGGACTGTCCCCACCCCCACCCCATCATTGTTTTTAAAAAATTA
AGGCCTGGCGCAGCAGCTCATGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACTTGAGG
CCAGGAGTTTGGAGACCAGCCCAGGCAACATAGCAAAACCCATTCTGCTTAAAAAATAAAAAAAAAAAAA
TTAGCTTGGCGTAGTGGCATGTGCCTATAATCCCAGTACTGGGAGGCTGAGGCACAAGAATCATTTTGAA
CTGGGAGGTAGAGTTGCTGTGAGCCGAGATTACGCCCTGCCTCCAGCCTGAGTCCAGCCTGGGTCCAGAGT
CCATCTCAGAAAAAATAAATTGAGTCAGGTGCAGTAGCTCCTTCTGTAGTCCCAGCTACTTGGGAGGC
TGAGGCTAGAGGATCACTTGAGCCCAGGAGTTTGTAGTCTAGTCTGGGCAACATAGCAAGACCCCATCTCTA
AAATTTAAGTAAGTAAAAGTAGATAAATAAAGAAAAAATAAAGTGTATGTGCTCATCATAAAGTAGAA
GAGTGGTTTTGCTTTTTTTTTTTTTTTTGGATTAATGAGGAAATCATTCTGTGGCTCTAGTCATAATTTATG
CTTAATAACATTGATAGTAGCCCTTTGCGCTATAACTCTACCTAAAGACTCACATCATTGAGCAGAGAGAG
AGTCGTTGAAGTCCCAGGAATTCAGGACTGGGCAGGTTAAGACCTCAGACAAGGTAGTAGAGGTAGACTTG
TGGACAAGGCTCGGGTCCCAGCCCACCGCACCCCAACTTTAATCAGAGTGGTTCATATTGATCTATTTTT
GTGTGATAGCTGTGTGGCGTGGGCCACAACATTTAATGAGAAGTTACTGTGCACCAAACCTGCCGAACACCA
TTCTAAACTATTATATATATATTAGTCATTTAATTTTACATAACTTGAGAGGTAGACAGATATCCTTATTT
TAGAGATGAGGAAACCAAGAGAATTTAGGTCAATAGCGCAAGGTTGTAGAGTAAGCGGCAAAGCCAAGACA
CAAAGCTGGGTGGTTTTGGTTTTAGAGCCAGTGCTTTTCCCCTCTACTGTACTGCCTCTCAACCAACACAGG
GTTGCACAGGCCATTCTCTGATTTTTTCTCTTGTCTCTGCCTCTCCCTCTAGCTCCCCTTCTCTCTC
TGCTCTAGTTCATTTTCTTTAGAGCAGCCGAGTGATCATGAAGTGCAAATCTTGCCATGTCAGTCCCCTG
CTTAGAACCTCCAATGGCTCACTTTCTCTTTAGGCAAAAGTCTTTACCCCATGCCTTCTCCCATCTCATC
TCAACCCCTCATTGTTGGCTGTCTGTCTGTGAGCCACTCTTCTTTTCCAGTCCCTCAGATGCACTGCACCCT
CTCCTGCCTGGGGTCTTTGCTCCTGCTACTACCTCTGCTTGAACAGCTCCTCACCTTCTCTTCTCCAACC
CACCTTTGTATAGGTGACTTTTGTTCATCCTTCAAGATTTCAACTCACATGTCTCTTGCATGGAGAACCCT
CTACTACTGTGTTGAGACCCTGTCCAGCCCCAGGTGGGATCCTCTCTCGACTTCCCATACATTTCTTTCA
CAGCATTTACATAGTCCATGATAGTTTACTTGTGGGATTATTTGGTTAATCTTTGCCTTTAACACCAGGGT
TCCTTGGGTGAAGGAGCTTCTTTATCTTGGTAACAGCATTATTTCAAGCATAACTTGTAATATAGTTATAT
TACATATATAACATATATATATATAACATAACATATATAACATATATAACAAGCATAACTTGTTATATAGT
CTTGTATATAGTAAGACCTCAATAAATATTTGGAGAACAAAAAATAA

LAQRGGARRPRGDRERLGSRLRALRPGREPRQSEPPAQRGPPPSGRPPARSTASGHDRPTRGAAAGARRPR
MKKKTRRRSTRSEELTRSEELTLSEEATWSEEATQSEEATQGEEMNRSQEVTRDEESTRSEEVTREEMAAA
GLTVTVTHSNEKHDHLHVTSQQGSSEPVVQDLAQVVEEVIGVPQSFQKLIKFKGSLKEMETPLSALGIQDGC

RVMLIGKKNSPQEEVELKCLKHLEKSVEKIADQLEELNKELTGIQQGFLPKDLQAEALCKLDRRVKATIEQ
FMKILEEIDTLILPENFKDSRLKRKGLVKKVQAFLAECTVEQNICQETERLQSTNFALAE

+++++

#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

GTACCCACTCCCTGGCACTTCCACTCCTAGAGGGAGGAGGCTGAGCAGGCAGAGAATGGGACGTGTCCCCT
CAGAGGAGCCTCGAGCCCAGTTCCAGCCAGCGGCCACTCAGTGAAGGTGCTCAAGTACCCACGTCCCCCGC
CAGCTGCCAGGGTTCCCTCTCCTCCCTCCGTCCCTCCCCCATCTGGGGAGCCAGCGGTACTGAGGGGGC
GGAACGAGGCGGGGCCACCGAGCGGTTATAGCTGGGCCTGCAGGGGACCCACGGCTCGCCTCCAGCCTCCT
GCGCTCCGGTACCTGGGCGTCCCAACTCCACTGCGCGCCAAACCCAGCCGAGCCGGTTCGTGGCCCGCCC
CGCCGGGCGGCCGTGCAGCGGAGCGCCCTGGCGTGGCGCCAGGGGAGCGGGGGGCTCCCGCGAGCCGGCC
GCGGCTGGCACTGCTGCTGCTTCTGCTCCTGCTGCCGCTGCCCTCCGGCGCGTGGTACAAGCACGTGGCGA
GTCCCCGCTACCACACGGTGGGCCGCGCCGCTGGCCTGCTCATGGGGCTGCGTTCGCTCACCTATCTGTGG
CGCCGCGCGCTGCGCGCGGGCCGCGGGCCCTGGCCAGGGACACCCTCTCCCCCGAACCCGCAGCCCGCGA
GGCTCCTCTCCTGCTGCCCTCGTGGGTTTCCAGGAGCTGTGGGAGACGCGACGCAGGAGCTCCCAGGCAGGGA
TCCCCGTCCGTGCGCCCCGGAGCCCGCGCGCCCCAGAGCCTGCGCTGGAACCCGGAGTCCCTGGACTTCAGC
GGAGCTGGCCAGAGACTTCGGAGAGACGTCTCCCGCCAGCGGTGGACCCCGCAGCAAACCCGCTTGGCCT
GCCCTGCCTGGCCCCGGACCCTTCTGACAGCGTCCCCCGCCCGCCGTGGCGCCTCCGCGCCTGACCCAG
GAGGAGTGGCCGCGCGCTTCCAGGAGCCGCTCATAGACCCCGCCTGCCGTCCGGTCAATAAAAATCCGCGCTG
ACTCCTGCGCCCCCGCATGCGAAAAA

LAWRPGERGAPASRPLALLLLLLLPLPSGAWYKHVASPRYHTVGRAAGLLMGLRRSPYLWRRALRAAAG
PLARDTSLPEPAAREAPLLLPSWVQELWETRRLSSQAGIPVRAPRSPRAPEPALEPESLDFSGAGQLRRD
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

ACCCCCGAGCTGTGCTGCTCGCGGCCGCCACCGCCGGGCCCCGGCCGTCCTGGCTCCCCTCCTGCCTCGA
GAAGGGCAGGGCTTCTCAGAGGCTTGGCGGGAAAAAGAACGGAGGGAGGGATCGCGCTGAGTATAAAAGCC
GGTTTTCGGGGCTTTATCTAACTCGCTGTAGTAATTCCAGCGAGAGGCAGAGGGAGCGAGCGGGCGGCCGG
CTAGGGTGAAGAGCCGGGCGAGCAGAGCTGCGCTGCGGGCGTCCCTGGGAAGGGAGATCCGGAGCGAATAG
GGGGCTTCGCTCTGGCCCAGCCCTCCCGCTGATCCCCAGCCAGCGGTCCGCAACCCCTTGCCGCATCCAC
GAAACTTTGCCATAGCAGCGGGCGGGCACTTTGCACCTGGAACCTTACAACACCCGAGCAAGGACGCGACTC
TCCCAGCGGGGAGGCTATTCTGCCAATTTGGGGACACTTCCCCGCCGTGCCAGGACCCGCTTCTCTGA
AAGGCTCTCCTTGAGCTGCTTAGACGCTGGATTCTTTCGGGTAGTGGAAAACAGCAGCCTCCCGCGAC
GATGCCCCCTCAACGTTAGCTTACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCT
ACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGCGCCAGCGAG

GATATCTGGAAGAAATTCGAGCTGCTGCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTC
GCCCTCTACGTTGCGGTACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCA
CGGCCGACCAGCTGGAGATGGTGACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTTCATCTGCGAC
CCGGACGACGAGACCTTCATCAAAAACATCATCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGC
CAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCTGCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCC
GCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGATCTGAGCGCCGCCGCTCAGAGTGCATC
GACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCAAGTCTGCGCCTCGCAAGACTC
CAGCGCCTTCTCTCCGTCTCGGATTCTCTGCTCTCCTCGACGGAGTCTCCCCGAGGGCAGCCCCGAGC
CCCTGGTGTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGAAGATGAGGAA
GAAATCGATGTTGTTTCTGTGGAAAAGAGGCGAGGCTCCTGGCAAAGGTGAGAGTCTGGATCACCTTCTGC
TGGAGGCCACAGCAAACCTCCTCACAGCCACTGGTCTCAAGAGGTGCCACGTCTCCACACATCAGCACA
ACTACGACGACCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTGAGA
GTCCTGAGACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAA
GAGGCGAACACACAACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTTCGCTGCGTG
ACCAGATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATA
ATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTCTGAAGAGGACTTGTGCGGAAACGACGAGAACA
GTTGAAACACAACTTGAACAGCTACGGAAGTCTTGTGCGTAAAGGAAAAGTAAGGAAAACGATTCTTCTA
ACAGAAATGTCCTGAGCAATCACCTATGAACTTGTTCAAATGCATGATCAAATGCAACCTCACACCTTG
GCTGAGTCTTGAGACTGAAAGATTTAGCCATAATGTAAGTGCCTCAAATTGGACTTTGGGCATAAAAGAA
CTTTTTTATGCTTACCATCTTTTTTTTTTCTTTAACAGATTTGTATTTAAGAATTGTTTTTAAAAAATTTT
AAGATTTACACAATGTTTCTCTGTAAATATTGCCATTAATGTAATAACTTTAATAAAACGTTTATAGCA
GTTACACAGAATTTCAATCCTAGTATATAGTACCTAGTATTATAGGTACTATAAACCTAATTTTTTTTTAT
TTAAGTACATTTTGTCTTTTTTAAAGTTGATTTTTTTCTATTGTTTTTAGAAAAAATAAAATAACTGGCAAAT
ATATCATTGAGCCAAAAAAAAAAAAAAAAAAAAA

LDFFRVVENQPPATMPLNVSFTNRNYDLDYDSVQPYFYCDEEENFYQQQQSELOPPASEDIWKKFELL
PTPPLSPSRRSGLCSPSYVAVTFPSLRGDNDDGGGFSFSTADQLEMVTELLGGDMVNQSFICDPDDETFIKN
IIIIQDCMWSGFSAALKLVSEKLASYQAARKDSGSPNPARGHSVCSTSSLYLQDLASAAASECIDPSVVFYP
LNDSSSPKSCASQDSSAFSPSSDLSLSTESSPQGSPEPLVLHEETPPTTSSDSEEEQEDEEEDVVSVEK
RQAPGRSESGSPSAGHSKPPHSPLVLKRVSTHQHNYAAPPSTRKDYPAAKRVKLDVSRVLRQISNNR
KCTSPRSSDTEENVKRRTHNVLERQRRNELKRSFFALRDQIPELENNEKAPKVVILKKATAYILSVQAEEQ
KLISEEDLLRKRREQLKHKLEQLRNSCA

C)

#1

NM_002006

Does not appear conserved beyond humans.

FGF2 fibroblast growth factor 2 (basic)

Synonyms: BFGF; FGFB; HBGF-2; FGF2

CGGCCCCAGAAAACCCGAGCGAGTAGGGGGCGGCGCGCAGGAGGGAGGAGAACTGGGGGGCGCGGGAGGCTG
GTGGGTGTGGGGGGTGGAGATGTAGAAGATGTGACGCCGCGGCCCGGGTGCAGATTAGCGGACGCGG
TGCCCGCGGTTGCAACGGGATCCCGGGCGCTGCAGCTTGGGAGGCGGCTCTCCCAGGCGGCGTCCGCGGA
GACACCCATCCGTGAACCCAGGTCCCGGGCCGCGGCTCGCCGCGCACCAGGGGCCGGCGGACAGAAGAG
CGGCCGAGCGGCTCGAGGCTGGGGGACCGGGGCGCGGCCGCGCGCTGCCGGGCGGGAGGCTGGGGGGCCG
GGGCCGGGGCCGTGCCCCGAGCGGGTCCGAGGCCGGGGCCGGGGCCGGGGACCGCGGCTCCCCGCGCGG
CTCCAGCGGCTCGGGGATCCCGGCCGGGCCCGCAGGGACCATGGCAGCCGGGAGCATCACACGCTGCC
GCCTTGCCCCGAGGATGGCGGCAGCGGCCTTCCCGCCCGGCCACTTCAAGGACCCCAAGCGGCTGTACTG
CAAAAACGGGGGCTTCTTCTGCGCATCCACCCGACGGCCGAGTTGACGGGGTCCGGGAGAAGAGCGACC
CTCACATCAAGCTACAACCTCAAGCAGAAGAGAGAGGAGTTGTGTCTATCAAAGGAGTGTGTGCTAACCGT

TACCTGGCTATGAAGGAAGATGGAAGATTACTGGCTTCTAAATGTGTTACGGATGAGTGTTCCTTTTTTGA
ACGATTGGAATCTAATAACTACAATACTTACCGGTCAAGGAAATACACCAGTTGGTATGTGGCACTGAAAC
GAACTGGGCAGTATAAATCTGGATCCAAAACAGGACCTGGGCAGAAAAGCTATACTTTTTCTTCCAATGTCT
GCTAAGAGCTGAATTTAATGGCCACATCTAATCTCATTTTACATGAAAAGAAGAAGTATATTTTAGAAATTT
GTTAATGAGAGTAAAAGAAAATAAATGTGTATAGCTCAGTTTGGATAATTGGTCAAACAATTTTTTATCCA
GTAGTAAAATATGTAACCATTGTCCCAGTAAAGAAAATAACAAAAGTTGTAATAATGTATATTCTCCCTTT
TATATTGCATCTGCTGTTACCCAGTGAAGCTTACCTAGAGCAATGATCTTTTTTACGCATTTGCTTTATTC
GAAAAGAGGCTTTTTAAATGTGCATGTTTAGAAAACAAAATTTCTTCATGGAAATCATATACATTAGAAAAT
CACAGTCAGATGTTAATCAATCCAAAATGTCCACTATTTCTTATGTCATTTCGTTAGTCTACATGTTTCTA
AACATATAAATGTGAATTTAATCAATTCCTTTTATAGTTTTATAATTCTCTGGCAGTTCCTTATGATAGAG
TTTATAAAAACAGCTCCTGTGTAAACTGCTGGAAGTTCTCCACAGTCAGGTCAATTTTGTCAAACCTTCTC
TGTACCCATACAGCAGCAGCCTAGCAACTCTGCTGTTGATGGGAGTTGTATTTTTCAGTCTTTCGCCAGTCA
TTGAGATCCATCCACTCACATCTTAAGCATTCTTCTGGCAAAAATTTATGGTGAATGAATATGGCTTTAG
GCGGCAGATGATATACATATCTGACTTCCAAAAGCTCCAGGATTTGTGTGCTGTTGCCGAATACTCAGGA
CGGACCTGAATTTCTGATTTTTATACCAGTCTCTTCAAACCTTCTCGAACCCTGCTGCTCCTACGTAAAAA
AAGAGATGTACAAATCAATAATAATTACACTTTTAGAAAAGTATCATCAAAGATTTTTCAGTTAAAGTAGC
ATTATGTAAAGGCTCAAACATTACCCTAACAAAGTAAAGTTTTCAATACAAATTTCTTGCCTTGTGGATA
TCAAGAAATCCCAAATATTTTTCTTACCCTGTAAATTCAAGAAGCTTTTGAAATGCTGAATATTTCTTTG
GCTGCTACTTGGAGGCTTATCTACCTGTACATTTTTGGGGTTCAGCTCTTTTTAACTTCTTGGCTGCTCTTTT
TCCAAAAGGTAAAATATAGATTGAAAAGTTAAAACATTTTGCATGGCTGCAGTTCCTTTGTTTCTTGAG
ATAAGATTCCAAAGAAGTATAGATTCATTTCTTCAACACCGAAATGCTGGAGGTGTTTGTATCAGTTTTCAAG
AACTTGAATATAAATAATTTTATAATTCAACAAAGTTTTTACATTTTATAAGTTGATTTTTCAATTA
AATGCAAATTTGTGTGGCAGGATTTTTATTGCCATTAACATATTTTTGTGGCTGCTTTTTCTACACATCCA
GATGGTCCCTCTAACTGGGCTTTTCTAATTTTTGTGATGTTCTGTCTCCTCCAAAGTATTTAGGAGA
AGCCCTTTAAAAGCTGCCTTCTCTACCCTTTGCTGGAAAGCTTACAATTTGTACAGACAAAAGATTTT
TGTTCCAATACTCGTTTTGCCTCTATTTTTCTTGTGTTGTCAAATAGTAAATGATATTTGCCCTTGCAGTAA
TTCTACTGGTGAACAAATGCAAGAAGAGGAAGTACAGAAACATGTCTCAATTTCCATGTGCTGTGACT
GTAGACTGTCTTACCATAGACTGTCTTACCCTATCCCTGGATATGCTCTTGTTTTTTCCCTCTAATAGCTA
TGGAAAGATGCATAGAAAGAGTATAATGTTTTAAAACATAAGGCATTTCTGCTGCCATTTTCAATTAGCTA
CTGACTTCCCTTACAATTTGAGATTTGCCATAGTTTAAACATGGTTAGAAAACAAGTAAAGCAATAAAGAA
AAATCTAGGCCGGGTGCAGTGGCTCATGCCTATATTCCCTGCATTTGGGAGGCCAAAGCAGGAGGATCGC
TTGAGCCCAGGAGTTCAAGACCAACCTGGTGAACCCCGTCTCTACAAAAAACACAAAAAATAGCCAGGC
ATGGTGGCGTGTACATGTGGTCTCAGATACTTGGGAGGCTGAGGTGGGAGGGTTGATCACTTGAGGCTGAG
AGGTCAAGGTTGCAGTGAGCCATAATCGTGCCACTGCAGTCCAGCCTAGGCAACAGAGTGAGACTTTGTCT
CAAAAAAGAGAAATTTTCTTAAATAAGAAAAGTAATTTTTACTCTGATGTGCAATACATTTGTTATTTAAA
TTTATTATTTAAGATGGTAGCACTAGTCTTAAATTTGTATAAAATATCCCCTAACATGTTTAAATGTCCATT
TTTATTCAATTATGCTTTGAAAATAAATTATGGGGAAATACATGTTTGTATTAAATTTATTATTAAGATA
GTAGCACTAGTCTTAAATTTGATATAACATCTCCTAAGTGTAAATGTCCATTTTTATTCTTTATGTTT
GAAAATAAATTTATGGGATCCTATTTAGCTCTTAGTACCCTAATCAAAGTTCGGCATGTAGCTCATGAT
CTATGCTGTTTCTATGCTGTGGAAGCACCAGGATGGGGTAGTGAGCAAATCTGCCCTGCTCAGCAGTACC
ATAGCAGCTGACTGAAAATCAGCACTGCCTGAGTAGTTTTGATCAGTTTAACTTGAATCACTAACTGACTG
AAAATTTGAATGGGCAAATAAGTGCTTTTTGTCTCCAGAGTATGCGGGAGACCCTTCCACCTCAAGATGGATA
TTTCTTCCCAAGGATTTCAAGATGAATTTGAAATTTTTAATCAAGATAGTGTGCTTTATTCTGTTGTATTT
TTTATTATTTAATAATACTGTAAGCCAAACTGAAATAACATTTGCTGTTTTATAGGTTTGAAGAACATAGG
AAAACTAAGAGGTTTTGTTTTTATTTTTGCTGATGAAGAGATATGTTTAAATATGTTGTATTGTTTTGTT
TAGTTACAGGACAATAATGAAATGGAGTTTATATTTGTTATTTCTATTTTGTATATTTAATAATAGAATT
AGATTTGAAATAAATAATAATGGGAAATAATCTGCAGAATGTGGGTTTTCTGGTGTTTCCCTCTGACTCTA
GTGCACTGATGATCTCTGATAAGGCTCAGCTGCTTTATAGTTCTCTGGCTAATGCAGCAGATACTCTTCTC
GCCAGTGGTAATACGATTTTTTTAAGAAGGCAGTTTGTCAATTTAATCTTGTGGATACCTTTATACTCTTA
GGGTATTATTTTATACAAAAGCCTTGGAGATTGCATTCTATTTTCTATATGACCCTCTTGATATTTAAAAA
ACACTATGGATAACAATTTCTTCAATTTACCTAGTATTATGAAAAGTGAAGGAGTTCAAACAATGTGTTTC
CCAGTTAACTAGGGTTTTACTGTTTGGCCAAATATAAATGTTTAACTGTTTGTGATGGCAGTATTTCTAAAG
TACATTGCATGTTTTCTTAAATACAGAGTTTTAAATAATTTTCAAGTAAATTTTAGATGATTTCAGCTTTCATCAT
TAAGAATATCTTTTGTTTTTATGTTGAGTTAGAAATGCCTTCATATAGACATAGTCTTTTTCAGACCTCTACTG
TCAGTTTTTCAATTTCTAGCTGCTTTTCCAGGTTTTATGAATTTTCCAGGCAAAGCTTTAATTTATACTAAGCTT
AGGAAGTATGGCTAATGCCAACGGCAGTTTTTTTTCTTCTTAAATCCACATGACTGAGGCATATATGATCTC
TGGGTAGGTGAGTTGTTGTGACAACCACAAGCACTTTTTTTTTTTTTTAAAGAAAAAAGGTAGTGAATTTT

TAATCATCTGGACTTTAAGAAGGATTCTGGAGTATACTTAGGCCTGAAATTATATATATTTGGCTTGGAAA
TGTGTTTTTCTTCAATTACATCTACAAGTAAGTACAGCTGAAATTCAGAGGACCCATAAGAGTTCACATGA
AAAAATCAATTTATTTGAAAAGGCAAGATGCAGGAGAGAGGAAGCCTTGCAAACCTGCAGACTGCTTTTT
GCCCAATATAGATTGGGTAAGGCTGCAAAACATAAGCTTAATTAGCTCACATGCTCTGCTCTCACGTGGCA
CCAGTGGATAGTGTGAGAGAATTAGGCTGTAGAACAATGGCCTTCTCTTTCAGCATTACACCCACTACAA
AATCATCTTTTATATCAACAGAAGAATAAGCATAAACTAAGCAAAGGTCAATAAGTACCTGAAACCAAGA
TTGGCTAGAGATATATCTTAATGCAATCCATTTTCTGATGGATTGTTACGAGTTGGCTATATAATGTATGT
ATGGTATTTTGAATTTGTGTAAGTAAAAATCAAGCTTTAAGTACATGGACATTTTTAAATAAAATAT
TTAAAGACAATTTAGAAAATTGCCTTAATATCATTGTTGGCTAAATAGAATAGGGGACATGCATATTAAGG
AAAAGGTCATGGAGAAATAATATTGGTATCAAACAATAACATTGATTTGTGTCATGATACACATTGAATTTGA
TCCAATAGTTTAAGGAATAGGTAGGAAAATTTGGTTTCTATTTTTCGATTTTCTGTAAATCAGTACATAA
ATAATTTTAGCTTATTTTTATATTTTCTTGTCTTAAATACTGAGCTCAGTAAAGTGTGTAGGGGATTAT
TCTCAGTTGAGACTTTTCTTATATGACATTTTTACTATGTTTTGACTTCTGACTATTAATAAAATAATAGTAG
ATACAATTTTATAAAGTGAAGAATTATATAATCACTGCTTTATAACTGACTTTATTATATTTTATTTCAAA
GTTTCAATTTAAAGGCTACTATTCATCCTCTGTGATGGAATGGTCAGGAATTTGTTTTCTCATAGTTTAAATC
CAACAACAATATTAGTCGTATCCAAAATAACCTTTAATGCTAAACTTTACTGATGTATATCCAAAGCTTCT
CATTTTTCAGACAGATTAATCCAGAAGCAGTCATAAACAGAAGAATAGGTGGTATGTTCCCTAATGATATTAT
TTCTACTAATGGAATAAACTGTAATATTAGAAAATATGCTGCTAATTATATCAGCTCTGAGGTAATTTCTG
AAATGTTTCAGACTCAGTCGGAACAATTTGAAAATTTAAATTTTTATTCTTAGCTATAAAGCAAGAAAGTA
AACACATTAATTTCTCAACATTTTTAAGCCAATTAATAAATAAAGATAACACACCAATATCTTCTTCAG
GCTCTGACAGGCTCCTGGAACTTCCACATATTTTTCAACTGCAGTATAAAGTCAGAAAATAAAGTTAAC
ATAACTTTCACTAACACACACATATGTAGATTTACAAAATCCACCTATAAATGGTCAAAGTGGTTGAGAA
TATATTTTTTAGTAATTCATGCAAAAATTTTTCTAGCTTCCATCCTTTCTCCCTCGTTTTCTTTTTTTG
GGGAGCTGGTAACTGATGAAATTTTTCCCACCTTTTCTTTCAGGAAATATAAGTGGTTTTGTTTTGGTT
AACGTGATACATTTCTGTATGAATGAAACATTTGGAGGGAAACATCTACTGAATTTCTGTAATTTAAATATT
TTGCTGCTAGTTAACTATGAACAGATAGAAGAATCTTACAGATGCTGTATAAATAAGTAGAAAATATAAA
TTTTCATCACTAAAATATGCTATTTTTAAAATCTATTTTCTATATTGTATTTCTAATCAGATGTATTACTCTT
ATTATTTCTATTGTATGTGTAAATGATTTTATGTAATAATGTAATTGCTTTTCTAGATAGTATGAATAAA
ATTGATTAGTTTGTGTTTTTCTTGTCTCCC

LVGVGGGDVEDVTPRPGGCQISGRGARGCNGIPGAAWEAALPRRRPRRHPSVSNPRSRAGSPRTRGRTE
ERPSGSR**L**GDRGRGRALPGGRLGGRGRGRAPERVGGRRGRGTAAAPRAAPAARGSRPGPAGTMAAGSITTL
PALPEDGGSGAFPPGHFKDPKRLYCKNGGF^{FLR}IHPDGRVDGVREKSDPHIKLQLQAEERG**VVS**IKGV**CAN**
RYLAMKEDGRLLASKCVTDE**CFF**FERLESNNYNTYRSRKYTSWY**VAL**KRTGQYKLGSKTGP**GQ**KAIL**FLPM**
SAKS

+++++

#2

NM_014293

No conservation found.

NPTXR neuronal pentraxin receptor

Synonyms: NPR; NPTXR

CGGCCGCGGACAGCTCCAGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCT
GGCCCAGCGCGCCGGGCTCCGCGCCCCGACCCGCGCGGCGCTGCCGGGGGCCTCGGGCGCCCCCGCC
GCCCGCCTC**AC**G**CTGA**AGTT**CCTGG**CC**TGCTGCTGG**CCCG**G****CAT**GCTGGCGTT**CCT**CGGTGCCGT**CAT**
CTGCATC**ATCG**CCAGCGT**GCC**CT**GG**CG**CCAG**CCCGCGCGGGCGCTGCCGGCGGGCGCCGACA**ATG**CTT
CGGTGCCTCGGGCGCCCGCGTCCCCGGGCCCAGCGGAGCCTGAGCGCGCTGCACGGCGCGGGCGGT
TCAGCCGGGCCCCCGCGCTGCCCGGGCACCCCGCGGCCAGCGCGCACCCGCTGCCGCCCGGGCCCCTGTT
CAGCCGCTTCTGTGCACCGCGTGGCTGCTGCCTGCCCGTCGGGGGGCCAGCAGGGGGACCGGGCGGGCG
CTGCGCCGGGCGAGCGCAAGAGCTGCTGCTGCTGCAGAGCACGGCCGAGCAGCTGCGCCAGACGGCGCTG
CAGCAGGAGGCGGCATCCGCGCCGACCAGGACACCATCCGTGAGCTCACCGGCAAGCTGGGCCGCTGCGA
GAGCGGCTGCCCGCGGGCTCCAGGGCGCGGGCCCCCGCGCACACCATGGCCGACGGGCCCTGGGACT

CGCCTGCGCTCATTCTGGAGCTGGAGGACGCCGTGCGCGCCCTGCGGGACCGCATCGACCGCCTGGAGCAG
GAGCTTCCAGCCCCTGTGAACCTCTCAGCTGCCCCAGCCCCAGTCTCTGCTGTGCCACCGGCCTACACTC
CAAGATGGACCAGCTGGAGGGCAGCTGCTGGCCCAGGTGCTGGCACTGGAGAAGGAGCGTGTGGCCCTCA
GCCACAGCAGCCCGCCGGCAGAGGCAGGAAGTGGAAAAGGAGTTGGACGTCTGCAGGGTCGTGTGGCTGAG
CTGGAGCACGGGTCTCAGCCTACAGTCTCCAGATGCCTTCAAGATCAGCATCCCCATCCGTAACAACATA
CATGTACGCCCCGCGTGCAGGAAGGCTCTGCCCGAGCTCTACGCATTACCCGCCTGCATGTGGCTGCGGTCCA
GGTCCAGCGGCACCGGCCAGGGCACCCCTTCTCCTACTCAGTGCCCGGGCAGGCCAACGAGATTGTACTG
CTAGAGGCGGGCCATGAGCCCATGGAGCTGCTGATCAACGACAAGGTGGCCCAGCTGCCCCTGAGCCTGAA
GGACAATGGCTGGCACCACATCTGCATCGCCTGGACCACAAGGGATGGCCTATGGTCTGCCTACCAGGACG
GGGAGCTGCAGGGCTCCGGTGAGAACCTGGCTGCCTGGCACCCCATCAAGCCTCATGGGATCCTTATCTTG
GGCCAGGAGCAGGATACCCCTGGGTGGCCGGTTTGTGACCCAGCCCTTTGTGCGTGACATTTGCCAGTT
TAACCTGTGGGACACGCCCTGACACCAGCCAGTCTGGGCATTGCCAAGTGCCTGCGCCACTGCTGG
GCAACGTCTTCCCTGGGAAGACAAGTTGGTGGAGGCCTTTGGGGGTGCAACAAAGGCTGCCTTCGATGTC
TGCAAGGGGAGGGCCAAGGCATGAGGGGCCACCTCATCCAGGGCCCCTCCCTTGCCTGCCACTTTGGGGAC
TTGAGGGGGGTCAATTTCCCTCCTCAGCCTGCCACGCACTGGCCTTCCCTCCTGCCCCACTCCTGGCTGT
GCCTCCCATTTCCCTCACCTGTACCCACACCTCCAGAATGCCCTGCCCTGCGAGTGTGTCCCCTGTCCCC
ACCTGAGTGGGGAGGAGCGTCTCAAGTGAACAGTGGGAGCCTGCCACCTGGCACTGCCTGGAGTTGTCT
CTTACCCACCCCTCCCTGCCCATCAACTGTATCTGATTTCACTAATTTTACAGCAGCCCCCAGTAGGGTAG
GATTGTGTATGAGGGGACCCACTATCTCAGTGGTGGGGGTGGCCGCCCCGCCCTTGTCCCCCATGCAA
CAGGCCCAGTGGCTTCCCTTTCAGGGCCACAACAGGCTGTAGAAGGGGATGACGAGGACATCAGAGGTTAG
ACTTACCCTCCTCCCTCTTTCCACCAGCTGCCAGTCAAGGGCAGTGGGATCTCGATGGAGCCTCCCCCCCC
CCCCACCCATGCCTCCCTCTTCCCTCCTCTTCCCTCCTCTCTTTGTGTGTAGCGTTTTGAATGTTGGTTCCA
TGCCTGGCCCAGCCCCACCTCAGTCTCCAGGACATTCCTTTCCAGCTCCAGCCTGGAGGGAAGGGGACAA
AGACCCCAGGAGGCCAAAGGGCTGCAGTACCCCTTGTGCTCACCCATAGTGTGGCCACTGGTATAGTCA
TCGCTCTCCCTCCATGCCAAGGACAGGACTTGGACCGCTTACGCTGGGGTGGGAGCAGCCCTAAGGTAGA
GGCCTCATGGCCCAGGAGACCCACCTCTGGCAGAGCCACATTACCTACCCTGTGCATGGTCTGGGGCAG
CAAGGAAGAAGCTCAGAGGGTGGGGAGAAGCATGAAGCAGTGGCAGAGCACTGGGTGAGAGGGAGAAGAC
CTTGGTTCCCTAGCCAGCCCTGCTAATGTGCTGTGTGGCCTTCTGTAAGTCCCTGCCCTCTCTGGGCCTGGC
TTCCCTCATTTCGTGAGCTGAGCCCTCGCTTTGGTCAATTTGCTCTCCAGATTGGGTGTGAGCTTCTCTGTG
ATTCAGGTGATATGTGGGGAAGCTCTGGTGACCTGGGCTTTCGAGGGGTAGATCCCAGGACTCGGCA
GTGGATGGGATGCAGCCAGTCAAGGTTAGGGTTCAGCAGAGACTCAGAGTCCAGGGCAAGGTTCAAGGCAG
ACTAACCTCATGCATGGATTGTAATAAACCAGCTCCCTTTGGATCAACCCAGCCTGGCACCCCTTGCCTGTC
TGAGAGTGTCTCAAAGGGCTGATGGCTTCCCTGGTCCCCTTGGATCATACCAGCTTCCCCAAGAGAGTGTG
AGAATCTTAAGAGCTGAGAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGA
CAGGCAGATCACTTGGAGTTCAGGAGTTCGAAGTTCAGCCTGGCCAACGTGGTGAACCCCATCTTCACTAAA
AATACAAAACTTAGCTGGTTAGGTGGTGCATGCCTGTAGTCCCAGCTACTCGGGAGGCCGAGGCAGAAGAA
TCTCTTGAAGTGGAGGTTGGAGGTTGCAGTGGCCGAGATCACGCCATTGCACTCCAGCCTGGGCAACAG
AGCAAGACTCCATCTCAAAAAATAATAATAATCTTAAAGATGAGAAAAGCCACCCCATCTGGCACCACAG
CTGCATCTTGTCTGTGAGAAATGGGGAAGAGTTGAGGGAGGACACGTGACCTGCACAGGATCACAGAGCAT
GGGGCAGAGCCAGGACTAGAGCTCAGGGCATCTGACTCCCTCTTTCAGTGTCTTCCCCCTCCATGTTGCCT
GCCCCGTAAGACCTTTGAGTTTCAGTCTACACCTAAGCAGGTAGACATCCGCGAGGTGAGATGCTTTCCAAC
ATGACACCTGAACATCTTCCCTTATGCAACACCCAAACATCTTGGCATCCCCACCCAGGAAGTGCGGGGA
GGAGGTTATGATCCCTGGGCGCTTCCGGCAGAATGGAGAGCTGAGGTGTCCCTCCCCTGCTAGTCACCTACC
AGGTGTCTGAGCAGCTGCATGCTCCCTGGCTCAAGTGGGCACTGTACCTTTTGCCTGCCTTTTTGTTCCCT
ATCTCCACTCCCTGAGGCCACTTAGCCTGAGACATGATGCAAGAGCTGCAGGCCGGGGGGCTCAGTGCCAT
GGAAGTACTCCAAGTTGCATTGCCCTCCCGCGCCCAGATCCTGCTTTCCATTTGAGAACATAAATAGATT
GCCCAGCCCCTCCAGTACAATCCCCTGGAAGAAAAGGCAATGGCGGGCTTCAGCCAGACCTGCTGAGACC
TAGGTTGCCACGGTAACAGCCAAAGACATCAACCCAAGTGTGGTCAAGTGTCTCATCATACTGGCACTG
TTGCTGGGGTGACGGCAGAATTCAGAACTTCAATTTTCAGTGACGCCAAGCTTGATGTGTTTTCTGTTATTGT
TTTGAAGAAGGTAGCTCTTGTGGAGGACTTGGGAGAAGGATGGGGTCTTAGGAAGGAGGTGACAGCACTTG
CATGGTCACTTGGAGCCACACACAGCTCAACCCCAAGTCCCTTATGCTTTGTGACAGTGAAGATGAGACC
TCTGACGTCCAAGCCTTGTTCCTGTGCTGCATCACCCACTCAGCCTTCCAAAGGGAACAGGAACAAATTTT
CCCAGCACCACTGTTTGGGTCCCCTTTTCCCTATCTTCTGCTGCCCTGAGCACATCCAAGCAGACAGGGA
AAGAGGAGTCAGACATGGCCCAGTCACATCCTGAGCTGCTCCTGGCTGATAACCACGATGGAGCCCGTGT
TGTCTGCCATCTGGCACTGCCTGAGTGTGGCACAGGCACCGTCTGTTGATCTCACAACACAGTTCTAA
GTTAGGACGTTCTTGGCTCCGTTAGACAGGTGAGGAACTGGGGCACAGAGAGGTGATGTCTGCTGCTGG
TGTCAATCAGCTAGCAAGTGTGAGGCCAGATTTCAAACCAAGGGGGTTACGTCCAGGGGCTGAGTTCC

CACTCACCTGTGTAGAGTGCCATCTGGGCACCATTGCTCCAGACGTGTTCCGACCCCTTTCCCAGCCCACA
GGGCTTGAAGTGAAGGAACAGAGGCAGGGGTGGGCCAGCCCCAGGGCCAGGGTCCCCTTGGTGAAGCCGT
GCCAGGGGGCTCAGCTGCTTCAGGGAATGTGTCCCTCCCACCATGGGCCAGAGCTTCAGCCCTTCTTTAGC
TCAGCTAGAGTTACAGGAGAGCCAAAAAAGAAAAGGAAGCTGAGCATCTCCCGAGTCCTGGGCAGGGAAG
GGGAGGGAATTGCTGCTTCTCCAACCTTTGCTTGGGGCCAAGCCCTGCACCAGTTGCTTCCCAGCTGTTA
TCTGCCAGATCTTCCCATCTTGTGGCATGTGGTGGCCCCACCAACATCCCAAGGGGACCAATCCCCTTGGC
ACCACTTTGCATCACCTGGGACCACAGATTTGGACAGGAAGGGCTCTGAGAAGAGGGCCAAAGCCCTCATT
TACAGATGAGGAAGCTGAAGCCCGGGGAGGGGAGCGACCCTCAAGGCCACCCAGCTGGACACGGGAGACTT
GAGCCCAGCCTTCTGACTGCATTAGCCCTCTCTAGGACGCAGCAGCCTCTCCCCAGCACTGAGTCCCCC
TCCTTTGTGTGTCCCAGCACCTTGGCCTGAGTAACTTGGAAAGGGGCTCCCTCCCAGAGAAGGGACTAC
TCTCTTACCCTTTATTCCAGCTGCCTGCCACCCAGACCCACCTCCCACCTGACCCCGACCCCTG
GTTGGGGAAGGGGCTCACATGGGCCAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT
TTAGTGACCCCTACTGACAACCCCTCCCAGCCTTGGCCCTTCTCCTCTCCCTGTTTTGTACATAAATTG
ACATGAGCTGCAACATGTGTGCGTGTGTGTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
GTTTTGTTACCTTTTTTGTTTTTGTAACTTGAATGTTCAAATAAACATGCTGTTTACTCTGA

LKFLAVLLAAGMLAFLGAVICIIASVPLAASPARALPGGADNASVASGAAASPGPQRSLALHGAGGSAGP
PALPGAPAASAHPLPPGPLFSRFLCTPLAAACPSGAQQGDAAGAAPGEREELLLLQSTAEQLRQTALQOEA
RIRADQDTIRELTGKLRCEGLPRGLQGAGPRRDTMADGPWDSPALILELEDAVRALRDRIDRLEQELPA
RVNLSAAPAPVSAVPTGLHSMQDQLEGQLLAQVLALEKERVLSHSSRRQRQEVEKELDVLQGRVAELEHG
SSAYSPPDAFKISIPIRNNMYARVRKALPELYAFTACMWLRSRSGTGQGT PFSYSVPGQANEIVLLEAG
HEPMELLINDKVAQLPLSLKDNGWHHCIAWTTTRDGLWSAYQDQELQSGENLAAWHPKPHGILILGQEQ
DTLGGFRDATQAFVGDIAQFNLDHALTPAQVLGIANCTAPLLGNVLPWEDKLVFAFGGATKAADFVCKGR
AKA

+++++

#3

NM_021182

No conservation found.

HMHB1 histocompatibility (minor) HB-1

Synonyms: HB-1; HLA-HB1; HMHB1

CCACATCCCAGGAGGCCGAGGCGGCTTGCCCCGCATCTCAGAAGCCGGGCAGGCCCTGAGCCTTCTGACCT
CACATCCTCTGCCACACCACAGTGGAGAAACCAGAAAGTGGAGGAGCAGCCAGAATGCAGAGAAGAAAAAG
AGGTTCTCTGCATGTTTTGGAAGTCGGAATTGGTTGAAGTTGAAGATGATGTGTATCTGAGGCACAGCTCTT
CCCTGACTTATAGGCTTTGACTGCTGTTGAGGTTTGACTCGAAGCCCAGAGTTTTGGTGTGGATGAGCA
GGGACAAATTGCTGAGCATGAAGAAGAGTAAATTAAGCAAGTGAACATATGCCCTTTGCCTCTGCTCTG
CACAGTGAATGAAAAGTCAACCTTTGAAAAAAAAAAAAA

+++++

#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

GGCGGAGCGTGGTACTACGACCAGCGCGGGCCGGAGGGGGCGGGGGGATGCGCCGCGGGCGGGCGGGCGCG
GGAGCTGGGGTTGGTGTGGCGGCGCCAGAGCAGCGGATCCCAGTCTCGCCGAGCAGCAGCGCGGGTGT
CGTGCACCGCCTGAAGACGCCGTACCTTTCTACCCCCACCTTTTTTTTTTTTTTTTTTTTTAAATAACCGGAA
CCAATGAACGCAGCCGGGATCAGAGCTCCGGAGGCCGCCGGTGCCGATGGGACCAGGCTGGCGCCCGGGCGG
GAGCCCGTGTCTGAGGCGGGGGGGCGGCCGGAGGAGTGCAGCGCGGGCGGTGGTGGCGCCTCGCGGAGCCG
GCGAGCTGCAGGCGGGCCGGGGCGCCGCTGCGCTTTACCCCGGCTTCTCCTCGGCCTTCATCCCGCCTCG
ACTCCTGGCCCAGCGTGGGGCTGGCTGCTGCGGGCGGGCGCTGGGCTGCGTTGCTGGTGTCTGGGCTGCT
GGTAGCCGGAGCGGGACGGATGCGAGCTTGTGCCCGGCACCTCCGCGGGCGGGCGGGCGACTGGCTCTG
CCGCAACTGCCGCCTCCTCTCCCGCCGGCGGGCGGGCGATAGCCCGGCGCTCATGACAGATCCCTGCATG
TCACTGAGTCCACCATGCTTTACAGAAGAAGACAGATTTAGTCTGGAAGCTCTTCAAACAATACATAAAACA
AATGGATGATACAAAGATGGAATTGAAGTAGAGGAAAGTATGAATTCATCAGAGAAGATATGAAAT
ATAAAGATGCTACTAATAAACACAGCCATCTGCACAGAGAAGATAAACATATAACGATTTAGGATTTATGG
AAACGATGGAAGAACATCAGAAGTTTATAATTGGACCTTGAAGACACTCTTCAGTGGTTGATAGAGTTTGT
TGAACTACCCCAATATGAGAAGAATTTAGAGACAACAATGTCAAAGGAACGACACTTCCCAGGATAGCAG
TGCACGAACCTTCATTTATGATCTCCAGTTGAAAATCAGTGACCGGAGTACACAGACAAAACTTCAGCTC
AAGGCATTGGATGTGGTTTTGTTGGACCTCTAACACGCCACCTCATAACTGGATGAAAAGATTTTATCCT
CACAGTTTCTATAGTAATTGGTGTGGAGGCTGCTGGTTTGCTTATACGCAGAATAAGACATCAAAAAGAAC
ATGTTGCAAAAATGATGAAAGATTTAGAGAGCTTACAACTGCAGAGCAAAGTCTAATGGACTTACAAGAG
AGGCTTGAAAAGGCACAGGAAGAAAACAGAAATGTTGCTGTAGAAAAGCAAAATTTAGAGCGCAAAATGAT
GGATGAAATCAATTATGCAAAGGAGGAGGCTTGTGCGCTGAGAGAGCTAAGGGAGGGAGCTGAATGTGAAT
TGAGTAGACGTGATATGCAGAACAGGAATTGGAACAGGTTGCGATGGCTCTGAAAAGGCCGAAAAGAA
TTTGAAGTGAAGCAGTTGGTCTGTTCCAGATGCACTTCAGAAATGGCTTCAGTTAACACATGAAGTAGA
AGTGCAATACTACAATATTAAGAACAAAACGCTGAAATGCAGCTAGCTATTGCTAAAGATGAGGCAGAAA
AAATTAAGAAAGAGAAGCAGCTTTTGGGACTCTGCACGTTGCACACAGCTCCTCCCTAGATGAGGTA
GACCACAAAATCTGGAAGCAAAGAAAGCTCTCTCTGAGTTGACAACTTGTTTACGAGAACGACTTTTTTCG
CTGGCAACAAAATTGAGAAGATCTGTGGCTTTAGATAGCCATAACTCAGGACTCCCCAGCCTGACCTCTT
CCCTTTATTCTGATCACAGCTGGGTGGTGTGATGCCAGAGTCTCCATTCCACCCTATCCAATTGCTGGAGGA
GTTGATGACTTAGATGAAGACACACCCCAATAGTGTCAAAATTTCCCGGGACCATTGGCTAAACCTCTGG
ATCATTAGCCAGAAGCAGCAGCTGTGCCGTTACGCCCGCAGCATTGTGCCGTCTCGCTCAGCTCAGC
GAGCTCAGCTTGTCCACACGCCCCCCCCCGTCAACCCCTCGGCACCCTCACCACCCGCAACACACACCA
CACTCCTTGCTTCCCCTGATCCAGATATCCTCTCAGTGTCAAGTTGCCCTGCGCTTTATCGAAATGAAGA
GGAGGAAGAGGCCATTTACTTCTCTGCTGAAAAGCAATGGGAAGTGCCAGACACAGCTTCAGAATGTGACT
CCTTAAATCTTCCATTGGAAGGAAACAGTCTCCTCCTTTAAGCCTCGAGATATACCAAACATTATCTCCG
CGAAAGATATCAAGAGATGAGGTGTCCCTAGAGGATTCTCCCGAGGGGATTGCGCTGTAACCTGTGGATGT
GTCTTGGGGTTCTCCCGACTGTGTAGGTCTGACAGAACTAAGAGTATGATCTTCAGTCTGCAAGCAAAG
TGTACAATGGCATTGTTGGAGAAATCCTGTAGCATGAACCAGCTTTCCAGTGGCATCCCGGTGCCTAAACCT
CGCCACACATCATGTTCCCTCAGCTGGCAACGACAGTAAACCAGTTGAGGAAGCCCCAAGTGTGCCAGAAT
AAGCAGCATCCACATGACCTTTGTATAATGGAGAGAAAAGCAAAAAGCCATCAAAAATCAAAAGCCTTT
TTAAGAAGAAATCTAAGTGAAGTGGCTGACTTGGTGAATCATGTTCAAGTGGCATCTGTAAACTATTATC
CCCCACCCTCCACTCCCCACCTTTTTTTGGTTAATTTTAGGAATGTAACCTCATTGGGGCTTTCCAGGC
CGGATGCCATAGTGAACATCCAGAAGGGCAACTGTCTACTGTCTGCTTATTTAAGTGAATATATATAATC
AATTCATCAAGCCAGTTATTACTGAAAAATCATTGAAATGAGACAGTTTACAGTCATTTCTGCCTATTTAT
TTCTGCTTTGTTCTCAGTGTATATATGCAACATTTTGTGAAAGCCACGATGGACTTACAAGCTTTAATG
GACTCGTAAGCCAGCATGGGCTTGCAAAAATTTCTTGTTTACCAGAGCATCTTCTTATCTTTCCACAGAGC
TATTTACATCCTGGACTATATAACTTAAAAGAAGTAAAACGTAATTGCACTACTGTTTTCCAGACTGGAAA
AAAAAAAATCTCTGCAAGTGAAGTGTATAGAGTTTATAAAAATGACTATGGATAGGGGACTGTTTTCACT
TTTAGATCAAAATGGGTTTTTAAAGTAGAACCTAGGGTTTTCTAATTGACTTGATTTCTGGAAATGAAAACCC
TCGCTTTTTATTATGGGAAGCTTCTTGAAGTGCATTTACTATTGTGAAGTTTCAAGTCCCCTGTAAAGATC
ATGTTGTTTTGTTTTCCCAGGGCTTTCACTGTGATTTACTGCATTGCAGGCTGTATGATAAAACACACAT
AATTTAAAGAGAGAAGGCTCTTGATTCTTATGCAAGTGAAGAGTTGAAACTTGATTGAAGGACTTAAAA
CATTACAACTTAAAGCCGAGGTGGGGGGATATGGGGATTGAGCAATTGTTTACACACTTTGAATAACTG
CAAAGGATTTACGGTTTTGTGAAAATGTGTACTGTGGAAGATAATAAATGAAAGACATTATTGTGTGGG
ATTGTGCTGATTTTTGTTGATAACACAAAAAACTATGTTTTCTGGAGAGCTGTGTAAGCTGTCTTGTG
CTTAGTTGCAATATAAGAAATAGTGTATTTTTGGACGTAAGTTGTCAACAAATTTCTATTTTATATTGTTA
TATTTTTATGTAGTTTGAATGTAATAATGTTCTAATATCAAGATTAACAAATATAAATTTATGGTGCATT
TAGAAAAA

+++++

#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

AGGAGCCGCGCCAGTGGAGGGCCGGGCGCTGCGGCCGCGGCCGGGGCGGGCGCAGGGCCGAGCGGACGGG
GGGGCGCGGGCCCCCGGGAGGCCGCGGCCACTCCCCCGGGCCGCGCGCGGGGGAGGCGGAGGATGG
AAACACCCTTCTACGGCGATGAGGCGCTGAGCGGCCTGGGCGGCGGCCAGTGGCAGCGGCGGCAGCTTC
GCGTCCCCGGGCGCTTGTTCGCCGGGGCGCCCCGACGCGCCGCGGCCGAGCATGATGAAGAAGGACGC
GCTGACGCTGAGCCTGAGTGAGCAGGTGGCGGCAGCGCTCAAGCCTGCGGCCGCGCCGCTCCTACCCCC
TGCGCGCCGACGGCGCCCCAGCGCGGCACCCCCGACGGCCTGCTCGCCTCTCCCGACCTGGGGCTGCTG
AAGCTGGCCTCCCCGAGCTCGAGCGCCTCATCATCCAGTCCAACGGGCTGGTCACCACCACGCCGACGAG
CTCACAGTTCCTCTACCCCAAGGTGGCGGCCAGCGAGGAGCAGGAGTTCGCCGAGGGCTTCGTCAAGGCC
TGGAGGATTTACACAAGCAGAACCAGCTCGGCGCGGGCGCGGCCGCTGCCGCCCGCCGCCCGCCCGGG
GGGCCCTCGGGCACGGCCACGGGCTCCGCGCCCCCGGCGAGCTGGCCCCGGCGGGCGGCCCGCCGAAGC
GCCTGTCTACGCGAACCTGAGCAGCTACGCGGGCGGCGCCGGGGGGCGCGGGGGCGCCGCGACGGTGCCT
TCGCTGCCGAACCTGTGCCCTTCCCGCCGCCACCCCCAGGCGCGTTGGGGCCCGCGCCTGGCTGCG
CTCAAGGACGAGCCACAGACGGTGCCTGACGCTGCCGAGCTTCGGCGAGAGCCCGCCGTTGTGCGCCATCGA
CATGGACACGCAGGAGCGCATCAAGGCGGAGCGCAAGCGGCTGCGCAACCGCATCGCCGCTCCAAGTGCC
GCAAGCGCAAGCTGGAGCGCATCTCGCGCCTGGAAGAGAAAGTGAAGACCCTCAAGAGTCAGAACACGGAG
CTGGCGTCCACGGCGAGCCTGCTGCGCGAGCAGGTGGCGCAGCTCAAGCAGAAAGTCTCAGCCACGTCAA
CAGCGGCTGCCAGCTGCTGCCCCAGCACCCAGGTGCCCGCGTACTGAGTCCGCGCGGGGGCGCATGCGCGG
CCACCCTCCCCAAGGGGCGGGCTCGCGGGGGGTGTGCTGGGCGCCCCGACTTGGAGAGGGTGGCGCCCT
GGGACCCCCCTCCCCGAGTGTGCCAGGAACCTCAGAGAGGGCGCGCCCCGGGATTCCCCCCCCCTG
AGGTGCCAGGACTCGACAAGCTGGACCCCCCTGCTCCCGGGGGCGAGCGCATGACCCCCCGCCCTG
CGCTGCCTCTTTCCCCCGCGGGCCGCCCGTGTTCACAAACCCGCGGTCTCGGCTGCCCTTTGTACA
CCGCGCCGGAAGGGGGCTCCGAGGGGGCGAGCCTCAAACCCTGCCTTCTTTACTTTTACTTTTTTT
TTTTTTTCTTTGGAAGAGAGAAGAACAGAGTGTTCGATTCTGCCCTATTTATGTTTCTACTCGGGAACAAA
CGTTGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
CCGCCCTTTTCTCGATCTCGCTCCCCCTTCGGTTCTTTTCGACCGGTCCCCCTCCCTTTTTTTGTTCTGT
TTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
AAAGTCTCGTTACGCCAGCTCGGCACAAAAAAAAAAAAAAAAAAAAA

+++++

#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.

MRV11 murine retrovirus integration site 1 homolog

Synonyms: IRAG; JAW1L; MRV11

GGCGCTGGCCCAGGGTCTTCTCCAGGCTCACATGCTTGCTGGAGGCTCCAGGCGGGTGGGGAGATCTGCAG
GGTCCCTTTGTTCCCGTCCGGCCAGGACAGTGTCCAGTGGCACGAGTTGCCATGGGGACTGCCATCTGCTG
TGTGGATAACCCATCGGAGGGGCCAGGGTGTAGCGACTGCCCCAAAGGCCAGCCTGCCCTGAGCCCTGAGG
CCAGCTGGTTCGGCTTTCCCTGCCATTATCCCCGAGCCAGGCCGGAGTGGGGCTCCCAGAGGAGCTCTCGG
CTCTAAGTCCCAGCCCCTCTCAGAGGAGGCGTGCAGGCGAGGTTCTGCAGCCAGAGGCTTCTCGGCTCTGA
GCTCCGGAGCCAGATGTAACATTGACCTTAAATGGTAAAAGCTCCCCAGAGTGAAGAGAGGCTGGCCAGAG
GAGGAAAGGAGAATAACTCAGTTTTAGACAGGGTCTCGCACTGTCACCCAGGCTGGAGTGCAGTGGTGCAA
CCCCACCTCCCGGGTTCAAGCAATTCTCATGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGGTCTCTCTA
CCGAGGAGTATCTTGCAGTCCAACCTCCCACGATTGTCCTGACTGGGGATGCCACTTCACCAGAAGGAGAA
ACCGACAAAAACCTGGCCAACAGAGTTCACAGTCCCCACAAGAGGCTTTCTCACCGACACTTGAAGGTGTC
CACTGCCTCCCTGACATCTGTGGACCCCGCGGGCACATCATTGACCTGGTGAATGACCAGCTGCCAGACA
TCAGCATCTCAGAGGAGGACAAGAAGAAAAACCTGGCGCTGCTGGAAGAAGCCAAGTTGGTGTAGTGAGCGA
TTCTGTACCCCGCGTGGGAGGAAGTCCAGGAGCAGCCCCGGAGACTCCCCATCAGTGTTCCTCCGAACTC
CAGCCCCAGCGCTTCTCTACATCTCTCGGAGCAACTCACTTACAGTCCCCACCCCGCAGGTTTGGATG
TGTGCAGTGGCCCGCCATCCCCTCTGCCTGGAGCACCACCACAGAAGGGGGATGAGGCCGACGTCTCTTCA
CCTCACCTGGCGAGCCTAACGTCCCCAAAGGGCTAGCTGACAGGAAGCAGAATGACCAGAGGAAAGTGTG
TCAGGGCAGGCTGGCTCCTCGTCTCTCCAGTTGAGAAGTCCAAAGAGATTGCAATAGAACAAAAGGAAA
ACTTCGATCCCCTCCAGTACCCCGAGACCACACCCAAAGGCCTAGCTCCTGTTACAAACAGCAGTGGGAAA
ATGGCCCTGAACAGCCCTCAGCCTGGCCCCGTGGAGAGCGAGCTGGGGAAGCAGCTCTTGAAAACGGGCTG
GGAGGGCAGCCCTCTGCCGAGAAGTCCAACCCAGGATGCGGCAGGAGTGGGTCCCCCAGCCTCCCAGGGGA
GAGGCCAGCTGGAGAGCCGATGGGGCCCCGAGGCTGGCTCCAAAGCTGAGCTTCCACCCACTGTGTCCCGG
CCCCCGCTGCTGCGAGGGCTCTCCTGGGACAGTGGCCCTGAAGAACCTGGCCCCCGGCTGCAGAAAGTGT
TGCCAAGCTGCCACTGGCAGAGGAAGAAAAGCGTTTTGCAGGCAAGGCCGGCGCAAGCTGGCCAAGGCC
CTGGTCTCAAAGACTTTTCAGATACAAGTGCAGCCCGTGCAGGATGCAGAACTGACCAAGCTCCGAGAGGAG
CACATCCTGATGAGAAATCAGAACTTAGTGGGGCTCAAGCTTCAGACCTTAGTGAAGCAGCTGAGCAGGA
AAAAGGGCTTCTTCTGAACTCTCCCCAGCTATTGAGGAAGAAGAGTCAAAGAGTGGCTTAGATGTCATGC
CTAATATTTCTGATGTGCTGCTGCGCAAACCTGCGGGTCCACAGGAGTCTCCCTGGAAGTGGCCCTCCACTC
ACTGAAAAGGAAGTTGAGAAGCTGTTTGTGCAACTGTCTTGGCCTTTAGAAATGACAGCTACACTCTGGA
ATCTAGAATTAACCAGGCTGAAAGGGAAACGCAACCTGACAGAGGAGAACTGAGAAAAGAACTGGAAAAC
TCAAAGCTTCCATTACGTCTCAGCTTCACTCTGGCACCCTGTGAGCACCGGGAAACCTACCAGAAGTTG
CTGGAGGACATCGCTGTCTGCAACCGCTGGCTGCCCCCTCTCCAGCCGAGCTGAGGTGGTAGGCGCCGT
CCGCCAGGAAAAGCGCATGTGAAAAGCAACGGAAGTGTGATGCAGTATGTGGAGAATCTAAAGAGACGT
ATGAGAAGGACCATGCGGAGCTCATGGAGTTTAAAAAGCTTGCAAATCAGAATTCAAGCCGAGCTGTGGC
CCCTCTGAAGATGGGGTCCCTCGCACGGCACGGTCCATGTCCCTCACGCTGGGAAAAGAATATGCTCGCCG
GAGGGTCAGCGTTGCTGTGGTTTCTAAGTTTAAATGCCCTGAATCTGCCTGGCCAAACTCCCAGCTCATCAT
CCATTCCCTCCTTACCAGCCTTGTGGAATCACCCAATGGGAAAGGCAGCCTACCTGTCACTTCAGCACTG
CCTGCACTTTTGGAAAATGGAAAGACAAATGGGGACCCAGATTGTGAAGCCTCTGCTCCTGCGCTGACCCT
GAGCTGCCTGGAGGAGCTTAGTCAAGGAGACCAAGGCCAGGATGGAGGAAGAAGCCTACAGCAAGGGATTCC
AAGAAGGTCTAAAGAAGACCAAAGAACTTCAAGACCTGAAGGAGGAGGAGGAAGAACAGAAGAGTGAGAGT
CCTGAGGAACCTGAAGAGGTAGAAGAACTGAGGAAGAGGAAAAGGGCCCAAGAAGCAGCAAACCTGAAGA
ATTGGTCCATTTCTTACAAGTCATGTATCCCAAACCTGTGTGAGCACTGGCAAGTGTCTGGATGATGGCTG
CAGTGTGCTGGTCTTGACTGTTGTGCTGGGGCTCTACAATTCCTATAACTCTTGTGCAGAGCAGGCTGAT
GGGCCCTTGGAAAGATCCACTTGCTCGGCAGCCAGAGGGACTCCTGGTGGAGCTCAGGACTCCAGCATGA
GCAGCCTACAGAGCAGTAGGAAACCTCACACCTAGCCAGTGCCTGCTCTGAGACACTCAGACTACCACCC
TTTCCCCAAGTATAACGTGAGGCCCAAGTGTGGACACACTGCCGCCCATCCCATCAGGTCATGAGGAAGGG
TTCTTTTAAACTCGGCACTTCTGTGGGAGCTATTATACACAGTGACTTGATGTTCTTGGAGGATCAACA
AAACTGCCCTGGGAAAGCATCCAGTGGATGAAGAAGTCACTTCCCAAGGAACCTTATTGGAAGGGAAGG
TCTCCTGCCCCTAGCTCAGGTGGCTGGGGAGAACTAAAACACCTTCACTGGTGGTTGGGGGTAAGGAGCGG
GAGCAAGGAGGAGGAGGTAGGGGGCAGTAAAAAACTTACTCTTTTTTCTCTCTGTAATTTGTTATC
AGGAAGAATTTGCTTAATGACTAACACCCTAAGCACTCAGCACTGGAATTTGGAGTTGCAAGTACTACT
TCCCATTTCCCATCTCATTTTTCAATAACTTACGCCTCCCATTCTTTCTTTTGGAAATGAGAGTTTCTTTTA
CAGAAGTAGGAAAGGCTTCTCAGAAAAAAGTATAGGCTGAATTTAGCTCAGTGCTTGAATGG
GAAGATATGAATTATTATATACGCATCTGTCCACACATACACACATACTGTTGTGTACACACACACAACAT
GCCTGTGCACAGAGCCAACAACCTTCAAAGTGTGCTCTGGGTGTGTACCTCTGGATAAATAAGATGCAT
GCCAAGCCAACCCACAGATTTTACCAGTGTGGGGCAGTACCAGGCACCTGTTCAATGAGCTGTCCACAT

GGATTGAAGATGTTTTAAAAACACAGAAAACCTCATGGCTTCAATGGCAGACTTACTAGTCTCCATTTCAA
TGCCAACTCTGAGCTGCTGTACAGCACAATCTATTCCCTATTCTCTCTTTGAAAACAGTTAACCCACCTCA
CAGGTGAATGAGGAGAGAAGATGTGCTTTCTGCTTCAGTCTCTTACTCTGTGTGTGACCACATGCAAGAGT
AAACTTGCACCTCAGTGCTTCAGTTCAAATGGGGTTTTCCAACCCCAAGTATAAATTAGGGGTGTTTCAGAGCA
TCCCCAGTTATTTAGCACAACACTGAAGGAGCACATCCCCTCTCCATTTTGACTTCTCTCCCCACTTTTTAC
AGCCACTGCCTTCATCAGTTTTGTAGAGGTTTTGATTTCCATGTGGGTTTTGTTGTCAATTGTTTTGCATTTT
TGTTTTGTTATTGATATTGTTTGCTTTTCATTGCTAAAACCTCATATACGACTTACTATGAGCCAAGCACTGT
TCTCAGTATTACATAGGTATGAATTCATTTAAGTCCTGAAGAAAAAGAAAAAATACGAAGTGGATATT
ACCCTTCCCATTTTCAAATAAGGAACTGAAGCACAAAAAGAACAAGTAACTTGACAAGGACACCCCGGTA
GTAAATCATGGGGCTGGAGCTCAACCCAGGGTAGGCTGGCTCCAGAGCTGTGCTCTCCTTGACTCTTCTG
ATGGTCTCCTAGCTGGAAGCCTCACATTTTCAGTCTCATTCCCCCAAGTGGCCCATCAGCTTCCATCTCT
GGCTCCCCAACTAAACAGTTTTCTCTCATAGTGTGACCTCCACTACTAGTTTTTTTTTCCAGCTGTTCTT
CTCTTTTTCTTCAGGTCCTCTCTCGACCGAGTGCAAAAATTATCCCCTCCATACCAGCTTTGATGACCTT
CCTTCCATACTCCTCACCAGACACAACATAATAGGTACACACTCCTCTGTGCTTTCTGGCAGTTTTTAAA
CATTATTATTATTGACCTTTACCTATAGTATAACCATGGCCTATTTATGTATCCATCTCCCCTAGCATTTTT
CCTCAAAGACAAGAACCATGTCTTACCCATCTCTTGGGTAAGTGCCTAGCATGGTGGCTGACGCTTGGGAG
GGTGTCAATTAATGTTGCTCAAAGAACAAGCAAACATTTAAGGTGGTGGAGAGCAGCCTGGGGACAGCTG
ACATGCTGCATGCTTCTCAGTACCAGCACCATCACAATGCAAAAAGCAACATCTTTCTTAACTCAGCTTA
TTCTGTTTTTTCAGTCTACTCTGTGAGAGAGCAGGAATGAGACCAGACTAGCAACACCATTGCCAAGCTCAA
GGACTGGGCTCAATGCAGTCACTCCTTCAGAGAGACCCCCACCCCAAGCATGCCCCACTTTAAAATAGCA
TGTTTTATTGAAGGGGGCATCCTTTACAGTAGCTAGAAAATGACTGAGGCCCAAGCCAGGGTTGATCAAGGA
TGTGCCATTAAGGTAAAGAGTTACAGAGCAGGGCAGAGGGACTCTGGGGGCAGAAGTGGATGATTTGCCCG
GCCTCTTCCAGGGGGTCTGGATACAACCTGAAGGAGCTTTAGCTACATGAGGCCCTCAGAGCCAAAGACAGG
ATGCAAAATAGAGTTCTAGAGAGTGGCCGTGGAAGCAGAACTCCAGGTGGGGAATGTTCAATCTCTGCCTCC
CTTAAAGCAGGGCCAGGCTCAGCTGGCCCCATTGTTCACTTGGTCCACAAGTTTTCTACCTTTGTTTCTGGA
TGAGTCAAAGGCCAGGAAGGCAGTTATGGAGAGCTCCTGCACCTCCAGCTGCCCCACAGAAAAGCCTGCAA
GAGTACTTCCAGGCACAGGCCCTCTCCCACCCTATTCCATTTGTAAGCAAGGGAGGTGAGGAAAAGGACA
TCTCCAAAAGGGAGCATAAGAATAGCCATATATACAGGGCTGAAAAAGTGTCTCATGAGTCCATCTTTTCT
GGAAAGCAAAGACCAGCCTGAAGCAGTGGGAGCTGCTGCCCAAGCGGTAGTGAAGTGGAGAGAAAACAGGCC
CTGGATTTTTCAAGTGGATGTGGATCTGAAGAGTCCCCAAATGCCTCTGAAGTCTGACATCTCTGTCTAG
CCCTAGGAGTCTGTTTTCCCTGCCTTTCAGTTGAGAGTGTGTTTTGTGTCATTGTTGATGTCACCTCCTA
AAAAGACCTTCACTTTCTGGCTGCCACAAAGCCATATGTGTTGCTCCCCATATACAGCCTGACAGAGTAAA
TGGAGAGGAAGTGTGATTTGTGTATCACTGGCTATCAGTTTCTCATGTTGTTAAGCCTCACACAGGTGT
GCTAGCATTGAACTGTAGAGTGTACATACCTGAGTTTGAAAATAAAAGCACATTTCCAAACCTCAAAAAA
AAAA

LTGDATSPEGETDKNLANRVHSPHKRLSHRHLKVSTASLTSVDPAGHIIDLVDQLPDISISEEDKKNLA
LLEEAKLVSERFLTRRGRKSRSSPGDPSAVSPNLSPSASPTSSRSNSLTVPTPPGLDVCSGPPSPLPGAP
PQKGDEADVSSPHPGEPNVPKGLADRQNDQRKVSQGRAPRPPVVEKSKEIAIEQKENFDPLQYPETTPK
GLAPVTNSSGKMALNSPQPGPVESELGKQLLKTGWEGSPLPRSPTQDAAGVGPASPQGRGPAGEPMGPEAG
SKAELPPTVSRPPLLRLGLSWDSGPEEPGPRQLQKVLAKLPLAEEKRFAGKAGGKLAKAPGLKDFQIQVQPV
RMQKLTKLREEHILMRNQNLVGLKLPDLSEAAEQEKGLPSELSPAIEEESKSGLDVMPNISDVLLRKLRLV
HRSLPGSAPPLTEKEVENVFVQLSLAFRNDYSYTLERINQAERERNLTEENTEKELENFKASITSSASLWH
HCEHRETYQKLLIEDIAVLHRLAARLSSRAEVVGAVRQEKRMSKATEVMMQYVENLKRTEYKDHAELEMEFKK
LANQNSSRSCGPPSEDGVPRTARSMSLTLGKNMPRRRVSVAVVPKFNALNLPQTPSSSSIPSLPALSESPN
GKGLPVTSPALPALLENGKTNGDPDCEASAPALTLSCLEELSQETKARMEEEAYSKGFQEGKKTKELDL
KEEEEEEQKSESPPEEPEVEETEETEEEEKGPRSSKLEELVHFLQVMYPKLCQHWQVIWMMAAVMLVLTVVGLY
NSYNSCAEQADGPLGRSTCSAAQRDSWSSGLQHEQPTAQ

+++++

#7

NM_001085386

No conservation found.

NF-E4 transcription factor NF-E4

Synonyms: NF-E4

GACAATTCCTGTTTACGGGAAGACTATAAAACCCCTGCCCCCTACTCATTGGTGCTGACGCCATTTTAGGCC
TTAGCCTGCCTGCACCCAGGCGCTCATTAAAAACAGCAGGTTGCTCCACACTGCCTCGTGTTGTCTGTTGGC
ACACTCTCAAGAGTTTGAACGGATACAAGAATCTTTCATCTGGTGCCGAAACCCGGGAGGGGCTCCGGTCT
TCGTCCCCCGTGGACCTACCCCTCCGCCCCAGAAAGCAGGCCACAGCAGCCGGACAAAGGAAGCTCCTCAG
CCTCCAGTTGCTTCTCTGTGCATGCACATCAGTCACTGATCTCACCTACTGGGGCCCTGCAGGCCATGGGG
CCACAGCTCCACACAGAAGCCTCCTAGCAATCCACCTCCACCTGGTGCCTGCTTCAAGTGCGGCAATGAAG
GCCACTGGCCCCACACAATGCCCAAACCCAGGTAAACCCACGAGGCCATGCCCCCTCTGCGGAGGACCCCCAC
TGGAAGTTGGACTGTGAGCGGCCCTGCAAGGACCACCCCATCCCTTCCTGAGCCAATCAAACCCCTCCTA
CTCGGATCTCGTCAGCCTTGCCGCTGAAGACTGATAGTGCCTTGGAACAGACACCCAGCAACTACCATCG
CTTCATCTGAGCCAAGGGTAAACCTGATGGTGGCAGGCCAGTCCCAAGGCGCAAGGCCACTTGTGCCAGC
AGTGTGAGTCAGCAAGATAGCAGAAGCAGGAAGAGAGCCGGCCAGAAGACACCTACTCTGACTGGGAGACA
CGTACCCCTGAAGATTGAGAAAGAGGCCATCCAGGTACCACATAGCAGTTACATCAGACTGGGACATTTCC
TGTTTACAGGAGACTATAAAACCCCTGCCCCCTACTCATTGGTGCTGACGCCATTTTAGGCCCTCAGCCTG
CCTGCACCCAGGTACTCATTAAAAACAGCGTGTGGTCCAAAAA

LPRVVCWHTLTKSLNGYKNLSSGAETREGLRSSSPVDLPLRPRKQATAAGQRKLLSLQLLLCACTSVDLTY
WGPAGHGATAPHRSLLAIHHLVLPASSAAMKATGPHNAQTQVNPRGHAPSAEDPTGSWTVSGPCKDHPHPF
LSQSNPPTRISSALPLKTDLSALEQTPQQLPSLHLSQG

+++++

#8

NM_199072

No conservation found.

MDFIC MyoD family inhibitor domain containing

Synonyms: HIC; MDFIC

CCCAGGCCGGCTCTGGCCTCCTGACCCAGACAGCGCAGGGCGCGAGGGATCGCGCGGCCGAGCCCGGGTGC
CGCCGCTCCCAGCATCGGGGCCGCTAGCCAAGAGTTTCGAGGCCCTTCCCGATCCGGATGTGATGAAAAAGAG
CAACAGAGGGGAGAAGTGTTCAGGATTGTAGAGTGGAAAGAGGGGAAAGAGAGGCAGAGAGGGGGAAGGCC
CCCTCGCAGGGGAGCCGGCTGGAGTGAGCTGGCTGGAAAGAGGGGGCGGAGTCCGCGGAGTCAAGCCGCC
ACCGCTGCCGCAGTTGCCGCCACTGCGGCGTCTGGGCTGAGCCGGAGGGAGGCGGGAGGACCGCGCAGGGGC
GGCCGCCCGCTCGTCAGGCCACCGGGGCGAAAATGCGGCCGCTGCCGGAGGCTCGCTAACTTTCCGGGGC
GGAAGAGGAGGAGGAGGAGGAGGAAGGGGCTTGAGCGACTACGGGGGATGCGGAGAAGCAGTCAGTTCC
CTGCACCCAGCACCTCACAGCCCTTCTCCGTGCGCCCTGCCGGGCGGCGAGCTAGGCGGCAGCGGCGCGG
CGCGGGCTCGGCGGAGCGGCCATGTCGCGGCGGGGCGAAGCCCTCGCTCCCGGGCCCGTGGGGCCGCGAGC
GCGTGGCCGAGGCGGGCGGGCGGCCAGCTGGGCTCCACAGCCAGGGAAAATGTGATAAAGACAATACTGAG
AAAGATATAACTCAAGCTACCAATAGCCACTTACACATGGAGAGATGCAAGACCAGTCCATTTGGGGAAA
TCCTTCGGATGGTGAATCATTAGAACCAACCTCAGCGCTTGCCTCAGCTTACAGCTTACAGCCAGGTGC
CAAGTGGTGAGGAAATAGGCAAGATAAAGAACGGCCACACAGGTCTGAGCAATGGAAATGGAATTCACCAC
GGGGCCAAACACGGATCCGAGATAATCGAAACTTTCAGCACCTGTTTCTCAAAAAATGCATAGAAAAAT
TCAGTCCAGCTTGTCTGTAACAGCGATATCAGTAAGAAGAGCAAAGTAAATGCTGTCTTTTCCAAAAAGA
CAGGCTTTTACCTGAAGATTGTTGTGTCCTGATCCTGGCTTGTGTTCTGCGAATTCCTGACCCCTT
TGCAACATTGCTCTGGGACAAGCGTCATGTGGCATCTGCACCTCAGAAGCCTGCTGCTGTTGCTGTGGTGA
CGAGATGGGGGATGATTGTAACCTGCCCTTGTGATATGGACTGTGGCATCATGGATGCCTGTTGTGAATCAT
CAGACTGCTTGAAATCTGTATGGAATGCTGTGGAATTTGTTTTCTTCAATATTTTATCTTTTGTGTTG
TGTTAAAACCTGGAGAGTGTAAAAATTTCTTTTTGGGGGGAAGAAAAGCACATTGTAAGATTCTCATGAA
ACAACATGGAATTTGCACTGTTAACTCATTATTGTAAGTAATCTCTGAAAGCCTTTTTACTTTAACCAAAT
CTACATGGTTTTAATATGTGAAATTTAACTACTTTAACTAGTTTTATAAATTTCTTAATATGTTACAATAA
CTTAGGGACATTTTACACCCCTTCCCAAATGTTAAATGCCTTCTCCTTTTTACCGATATTTCTGTTTC

TTTTAACCGTTCTCAGGAGCACTTTGCTCCAAATATATTATTTTTTCAGTGTGTATTTAAACGAGGCAGTTT
ATTTTTGATATGTATCTATTTCATGATTGAAAGGAAGCAGTCTTGGCCAGGCACGGTGGCTTACACCTGTAAC
CCTGGCATTGTTGGGAGGCCAAGGTGGGCAGATTGCCTGAGCTCAGGAGTTTCGAGACCAGCCAGGGCAACAT
GGTGAACCCCATCTCTACTAAAATACAAAAAGTTAGCTGGGCTTGGCGGTGTGCGCCTGTAGTCCCAGCT
ACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGAGAGGCGGAAGTTGCAGTGAGCCGAGATTGTGCC
ACTGAACTCCAACCTGCACTCCAGCCTGGGCAACAGAGCGGAGACTCCATCTCTAAATAAATAAATAAATAA
ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
TAATAGTTGGGAACAGTGTGATCTCTCTCCCTTACCTTCTCCACTTGTTC AACAGACTCTGAATGCCGA
CTGTGTGGACTCTCTTCCCTCAGACTGTGGGGACAGATAACAATCCACTCCTGTCCACAGGAACATGAGATT
TAGCAGACTAAGGAGATCTGTAAAGAATGAACCATAACCACAAGGCATACTGAAGTGAGGATTTATAAGAGAA
ATAAACTCAAAATGCTGTGGAAATATGCAGAGAATTGCCTACCAGAATATTCAGTAAGGTTTCAGGGAGAAT
GTGGCATTGAGGACTCTCTTAGAATGAGTGATTACCTGCTATTTAAATGAATTTTAGATTTTTTGACA
AAGATTTAGGTGGACACCCTAAACTGTGTGTGCCTTTAACAGTTAAAAGAACAGTGCCTTCAGCATACTT
TTTTATTAGTTGTAGGAATACAGCTTTTTGAAAAAGCTATAAAGTTTTAAATTAACATAAATAATATGCATTTT
CTTACACATAATTTAAATGTTATCATACTTTTTTGGATGAAAAATAATGCCTTAGTAAAATAGCTCTATTT
AATAAAGAAGATTGAGTACTCTGACACATTTCAATTTAAATTAGGAAATTTTTAATATTAATAATCCCAGTGT
TCTGAGTTATTGAAAGGCTTTCTTTTATTTTGGAGCTTTAGGTCTTTTTGGGATGAGAACATTTTAGTTG
TTTAGTTTGTCTTAAAGCAGTGCTATTTTTTGTAAACACAGATAAATGGAAACCATTCTTTTCAATGCAG
AAGAAATCTAGATATCCCCTACTGTGACCAAATTTCTGTATTACGATTTTATGTTAAATTAACATAATATG
GCAGTTTATAATGATCCTTAAGTGTAAAGAAATCAGTCAATTACAAGAGTAATTGTATAGTTATTGAGACC
TATAGTGTGTGGCTTAGATGAAAGGGAGAGTAAATTTTATACCATGCTCTCTCCTACTCAGTTTGTATCTC
TCTAAAATTGTAGTTTGGTTTGAATTAATAAATCTTAGTAGAAATTTTGAAGTATGCTTTGGGATTAA
TAATTTTTTTAATTTTTCTGGCTGAATATCAAATTGATAGTAAACAACAGAAGCATAATTTTAGGAAGGCT
TTGCAAACCTAGCCTTTTTAAGAGAGGTTTTTAACTGAAGCATGAGAATATATCACCTGTGGTTTTTCT
TTGAGATGAAACGTAGTTTCTAGTTATATCATTACTTAAAGGGCTTAAAAAGAAAAAACTTAGCAAACTTT
TGAATCTTTCTTTTATTGCTATTTACACATACATACACACATACAAAACCTTTAAATTTTGGGATCTGAAT
ATAATTCTGGTAAACAGCTGTCTTCATTTTTCTCCTCTAAAGAACTTAATTCATTTGTTACATAAAATATA
AGGAAATCTTTATACTATTTTACAGTAACCACAATCTAAATATTTACATATACCCAAAATTAACCTTATGCT
CATATATTAGGATGTGAGAATATCATCTGTTTTATGGACACATGAAACCTCCTAATGACCTGGAATTTGTAG
AATATTTGACTTTTTTATGCAAAAGTTTTTCAACCAAGTGGTTTTGTCTAATATTTAAACATGTACTGGCAC
AATTTGTGATGAAAATATTAGCACATTTGCAATAATGTTTCTCCATAACAGAGAATGTTAATGGATACCAG
AATTTTTATTTTTGTATTTATGTTTATAGTACTTTTTCTCCTTGTCTACTCCAGACAGTTATTCCATAAAGCA
TTTGTATAATTTAAAGGAAAACAGAAAAAGGAAAAGTAGGCAAATGTGAAAATAGTTTCAATATATCTTAT
GATTTCTTAATGTAAATGTTTTGTTGAAGTATATGGCTATCATGACTAAGTGCTAGAATTTATAGTTACA
GGCGGTGTCCTTTTTAAATGTGGAAAGGCTTTTTAAATATTTTAAACTGGACCTGTATTATCCTGAATACA
CTATTTTGAATTTTTTAAATGACTTCTTTATTTTGTCTTTACCGTATGTTTATATCTAATTGACATATT
GACTAATGTTTGAAGAATTCAACCATAAGTTAAAATCTGAAGTTATCTTTATCATGTTTTCATCCCTGTC
TGAAGATTTCTAGTCTTCTTATGTAAATCACATGACTCATGTCCGTAAATGAACTATGAAAGATATCGAT
CAGTTTATGATCATTGACATGTGATTTCAAAACACAGTGTCTTTTTAAATCTATAATATGTCAAATAC
AAGTTTTTTTTTTTTACATCGTTTTAGTAAGTTAATTTCAATTTTACTTTGGAGCTATATTTCCACTTA
GAAAACTAAGGTAATTTTACAATATATGCTGAGATTAATAAACCAGGTAATAATGATCAAACATATATGA
AATTGAGTCTTAGATTTAATGAATTTCACTCGAAAATAAATGATCAGAAGAATTTTTCATCTAA

VRGVRAATAAAVAATAASGLSRREAGGRAGAAA VVRPPGRKCGRCRRLANFPGRKRRRRRRKGLGATTGG
CGEAVSSLHPHSPSSVRPAGRRRARRRRRQRRGAGSAERPMSGAGEALAPGPVGPQRVAEAGGGQLGSTAQ GK
CDKDNTEKDITQATNSHFTHGEMQDQSIWGNPDSGELIRTQPQRLPQLQTSQVPSGEEIGKIKNGHTGLS
NGNGIHHGAKHGSADNRKLSAPVSVQKMRKIQSSLSVNSDISKKSKVNAVFSQKTGSSPEDCCVHCILACL
FCEFLTLCNIVLGQASCGICTSEACCCCGDEMDDCNCPCMDMDCGIMDACCESSDCLEICMECCGICFPS

+++++

#9
NM_005583

Not well conserved beyond humans.

LYL1 lymphoblastic leukemia derived sequence 1

Synonyms: bHLHa18; LYL1

CAGCCTGGCCCTTATCTGCACTGGGCCAGCATCCTCCGGCCGCTGCGCCGCCAGGGGTGAGAGGGAGGAAA
CCGGGCCCGCCGGGGCGGGGAGAAGGCGGGCCGGCCCGGGAGCCGCTCACTTTCCCTGGGGGGACCTACG
CGGAGACCTCGGCTATCCTGGCCTTCCGAGGCCACGAGGAGGCGCGCCCAACGCCGGGCCTGGAGCAT
TGAGGCCGGACCCTCGCGAGACAGCAGAGCCTGGCCTGACGCTGGAAACCACACCCTGGCCAGACTGCCA
GCCCTGACGGGACAGAGCCAGGGCACTCACCAGGCTGCAAGAACAGTGCTGGGGTGAGTACCCCCACGTG
GGGTCCATGTGCCCGCCTCAGGCACAGGCAGAGGTGGGCCCCACCATGACTGAGAAGGCAGAGATGGTGTG
TGCCCCCAGCCCAGCGCCTGCCCCACCCCTAAGCCTGCCTCGCCTGGGCCCCCGCAGGTGGAGGAGGTGG
GCCACCGAGGAGGCTCCTCGCCCCCAGGCTGCCACCTGGTGTACCAGTGATCAGCCTGGGCCACAGCAGG
CCCCCAGGGGTAGCCATGCCACCACAGAGCTGGGCACTCTGCGGCCCCCGCTGCTGCAACTCTCCACCCT
GGGAACTGCCCCGCCCACTTTGGCCCTGCACTACCACCCTCACCCCTTCTCAACAGTGTCTACATTGGGC
CAGCAGGACCTTTTAGCATCTTCCCTAGCAGCCGTTGAAGCGGAGACCAAGCCACTGTGAGCTGGACCTG
GCTGAGGGGCACCAGCCCCAGAAGGTGGCCCGCGCGTGTTCACCAACAGCCGGGAGCGCTGGCGGCAGCA
GAACGTTAACGGCGCCTTCCGCGAGCTGAGGAAGCTGCTGCCGACGCACCCGCCCGACCAGGAACTGAGCA
AGAACGAGGTGCTCCGCCTAGCCATGAAGTACATCGGCTTCCCTGGTGCAGGCTGCTGCGCGACCAAGCCGCA
GCTCTGGCCGCAGGCCCCACCCCTCCCGGGCCTCGCAAACGGCCGGTGCACCCGGGTCCCAGACGACGGCGC
CCGCCGGGGATCCGGACGCAGGGCCGAGGCGGCAGCGCGCTCGCAGCCCGCGCCCCCGGCCGACCCCGACG
GCAGCCCCGGTGGAGCGGCCCGGCCCATCAAGATGGAGCAAACCGCTTTGAGCCCAGAGGTGCGGTGA
CACGCGGCAGCACCTCTGAGCCGGAGGGCACCAGGGACTCGGCCCAGGGCCGTCAAGGAAAGGGCAGTGGA
CGTGCTCGCATGTTCCGGGAGCGAACTCCCCGAAGAAGGACCAGTGAAGACGTCAGGGCAAGGTCTCGG
GGTCCGGAAGGGTGATCATCGACCCCCAAGGGACCCGAGACCCCTTAAAAAATCACCCACAACCCTCTG
GAAGTGGCCTTGCCCGGTCCCCTTCCAGGGCGAGGTGGCAAAGCAACATGGCAGAGCAGTCATAGGAA
AAAAAAAAAAAAAAAAAAAA

LAQTASPDGTEPGHSPGCKNSAGVSTPTSMSMCPQQAQAEVGPMTTEKAEMVCAPSPAPAPPPKPASPGPP
QVEEVGHRGGSSPRLPPGVVIVISLGHSRPPGVAMPTTELGLTRPPLLQLSTLGTAPPTLALHYHPHPFLN
SVYIGPAGPFSIFPSSRLKRRPSSHCELDLAEGHQPKVARRVFTNSRERWRQNVNGAFAELRKLPLTHPP
DRKLSKNEVRLAMKYIGFLVRLLRDQAAALAAGPTPPGPRKRPVHRVPDDGARRGSGRRAEAAARSQPAP
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

GGCTCCCCACTCTGCCAGAGCGAGGCGGGGCAGTGAAGGACTCCGCGACGCGTCCGCACCCTGCGGCCAGAG
CGGCTTTGAGCTCGGCTGCGTCCGCGCTAGGCGCTTTTTCCAGAAGCAATCCAGGCGCGCCCGTGGTTC
TTGAGCGCCAGGAAAAGCCCGGAGCTAACGACCCGCGCTCGGCCACTGCACGGGGCCCCAAGCCGAGAA
GGACGACGGGAGGGTAATGAAGCTGAGCCAGGTCTCCTAGGAAGGAGAGAGTGCGCCGGAGCAGCGTGGG
AAAGAAGGGAAGAGTGTCTTAAGTTTACGGCCAACGGTGGATTATCCGGGCCGCTGCGCGTCTGGGGGCT
CGGGAATGCGCGAGGAGAACAAGGGCATGCCAGTGGGGGCGGCAGCGATGAGGGTCTGGCCAGCGCCGCG
GCGCGGGGACTAGTGGAGAAGGTGCGACAGCTCCTGGAAGCCGGCGCGGATCCCAACGGAGTCAACCGTTTT
CGGGAGGCGCGGATCCAGGTCATGATGATGGGCAGCGCCCGCGTGGCGGAGCTGCTGCTGCCACGGCG
CGGAGCCCAACTGCGCAGACCCTGCCACTCTCACCCGACCGGTGCATGATGCTGCCCGGGAGGGCTTCTG
GACACGCTGGTGGTGTGTCACCGGGCCGGGGCGCGGCTGGACGTGCGCGATGCCTGGGGTCTGCTGCCCGT
GGACTTGGCCGAGGAGCGGGGCCACCGCGACGTTGCAGGGTACCTGCGCACAGCCACGGGGGACTGACGCC
AGGTTCCCCAGCCGCCACAACGACTTTATTTTCTTACCCAATTTCCACCCCCACCCACCTAATTCGATG
AAGGCTGCCAACGGGGAGCGGGCGAAAGCCTGTAAGCCTGCAAGCCTGTCTGAGACTCACAGGAAGGAGGA

GCCGACCGGGAATAACCTTCCATACATTTTTTTCTTTGTCTTATCTGGCCCTCGACACTCACCATGAAGCG
AAACACAGAGAAGCGGATTTCCAGGGATATTTAGGAGTGTGTGACATTCCAGGGGTCGTTTTGCTTTTCAGG
GTTTTCTGAGGGAAAGTGCATATGAAATCCTTGACTGGACCTGGTGGCTACGAATCTTCCGATGGATGAAT
CTCCCCTCCAGCGCTGAGTGGGAGAAGGCAGTGATTAGCACTTGGGTGACGGCAGTCGATGCGTTCACTC
CAATGTCTGCTGAGGAGTTATGGTGAACCCACAACCTTAGGCCCTAGCGGCAGAAAGGAAAACCTGAAGACT
GAGGACAAAGTGGAGGAGGGCCGAGGTGGGCTTCAGTAAGTCCCCGGCGGCGCTTTAGTTTTGAGCGCATGG
CAAGTCACATGCGTAAACGACACTCTCTGGAAGCCCTGGAGACCCTCGCCCAACTCCACCAGATAGCAGAG
GGGTAAGAGAGGATGTGCAAGCGACGACAGATGCTAAAATCCCTGGATCACGACGCTGCAGAGCACCTTTG
CACAGGATGCTGGCCTTTGCTCTTACTACACTGAGGAGAGATTCCCGCGGGTCCGCAGGCAGACTACACA
GGATGAGGTGGTGGAGTGGAGTGAAGCAATTGTAACGGTTAACTGTAACGTTTTCTTTACACACACACA
CACACACACACACACATGCTAGGATGCGGAAATCCCCTTATGACTTGTACTTTTTGATTTTGTGATAT
TTTTGTACTTTTTAGTTGTTTTCAGCAACTGTCTTATTTAATGGGGAGATTTTAAGTAACATAAAGTGGCTC
TCAGTTAAAATGTGAGGAAGAAGTACAGCTCTTAAATGTAGCAATGGCACTGTTGCAAACTCAGTGCAAAC
GCCTAGATTGCTTTCTTCTTAACTATTTATTTCTTTGTTAAATTTTTCTGATTGTTTCTTTATAGAGTG
TCTCAGGGTGCAGAGGTGAGACTAAGAAATATTCCAAATGTCTTTTAGAAGATAGATGCACTTATGCAGTA
AATTATCTTGGGATAGTTCCCAAAAGATTGCTGAAAAAGTAGATTGAGTATAAAAACTTGAAAATATATGA
TGGCTCGTGGGATGCTCTACTATCACTGAACAACTAAAGGTGCACTGCTTTGGGATTTAATTTCCAGGGT
TGCTTGATCATTATATCATTGGAACAACCTGATACTTCACTACTTTAATAAAGAATTAACAGAGATTGAACT
CCAAGAGGTGGGTAATTTGGTTTTAAAAATACATGTTTCATGGGTTTACCCTAACTCCTGAGAAATGTTAAA
GGTTCACAGGGGTTCCCTTCTCTCAATGTTTGTAAATAATTGCTCATAAGCAATACCAGCAATTCATAAAAA
CTGCTTACTTATGCCATAGAAAATTAACACAAAGTGTATACATGTATTATGCTTCTAAATGCTCATTCTA
CCAGATACACATTTAAAAGAGAAAAAAGGAACAGAAACAAGTCATTTGAGAGTGGAGACTTATAAGAAGGA
GTACATTTGAGTTGAATACACAAATCTTTACTTCTCTACCAATTCCTATTCCCAAAATGAACATATTACTG
GGGAAAGTTAGTTGAGAATCAGAGCATATGTTATTGGGGAAAGGATATGTTTATTGACACATAATCTGTAC
CAGGTATGCATTAATAATATTTGTTAATTTAATTTAAACCTGAGAGATAGGTATTGTTTCCAGATGA
GGACAATGAGGCAAAGAAATATCAAGTAACTTGCCAAAGGTTACAAGATATTCAATCCATGGATGCACAAA
GAAGTGCATCTAGTTCCACAGCTGATTATGGTTGTCTTGCTTTTCTTCCATTGCACCAGCTTGTCTCTCA
AAATCATGAATGATACACATGAAGATAACTTTTTTTAAAAAAGCAGAAATACACAATGATCTCCCTTGT
AAGTCCTAAGTGGCTTTTTCTTCTCTAATCTTAGTAAATATAAACGGTTTTGTTTTGAAAACATTTTTAA
AATGTCAACAATATGGAGAATAACCCCCCAACACACCTATAAAAAACCAAAATTTTTTGAACAAAGATAA
TGGAACCTCCATTTTCAAACCTGAAGCACAGGGACAGAAAATATATTTCTAGTTATCACTTAAGCACTCAAT
CATTAGAGGCTACAAGAATAATATTTTTTAAAGTTACAGTATTTTACAATTATTAGAAAACATTCTATATAA
AAGAAGTCAGTTGATACTTTAAAATCTCCATTTGGTTTTATAAAATCCCTTAATTTGACCTCTATATCTTA
AATTCCAAGATGTTTAAATTTGCTAGTTGCATTATACTGGGTCATGAAAAATTATCCCTTGAAATAGATAT
GAAACATGTTACTTCAATTTCTGGTTTTAAATAACTTGTGGAATCTTTCCTAATGACAACCTGATATTAAGGG
AACTAAAGAAAATGTTATTGTGGATCCCACAGTACTATATTACACTGTTTTTTTTTTGTTTTGTTTTGTTAGT
TTTTTTTTATTTAAAGCAAACCTCAAACATTATTGGGTATCAATTACCACCTGGTTGTATTGAAATAGTAAC
TTATCAATGCCATGTAATAAATTAATTCATTTTGAAGCCACCTGGCAGACAGGTTTAGCTGTTTCATCAG
CAGCCTAATATATACTGTTAAATTTGTTAAGGATTTCACTTTGAAGGATACATGCAAAACATATAGTTACT
ATTTTCATGAGTCCTGCTTCTAGCTCCATTGTGGAATACAGAAAATTAATATACCTGTTAAGTTCGTATC
TAAACCTAAGACATTACCAAGGTTTGTACAAATTTACTACCTGACATTTATTCCAAGAAGATCTGGAAAG
TTAAATAAATTTATAAATTTAATAACAAAAAAAAAAAAAAAAAAAA

TVDYPGRCASGGCGMREENKGMPSGGGSDEGLASAAARGLVEKVRQLLEAGADPNGVNRFRRAIQVMMMG
SARVAELLLLHGAEPNCADPATLTRPVHDAAREGFLDTLVVLHRAGARLDVRDAWGRLPVDLAEERGRD
AGYLRATGD

+++++

#11

NM_001085386

Apparently conserved in primates only.

NFE4 transcription factor NF-E4

Synonyms: NF-E4

GACAATTCCTGTTTACGGAAGACTATAAAAACCCCTGCCCCCTACTCATTGGGTGCTGACGCCATTTTAGGCC
TTAGCCTGCCTGCACCCAGGCGCTCATTAAAACAGCAGGTTGCTCCACACTGCCTCGTGTGTCTGTTGGC
ACACTCTCAAGAGTTTGAACGGATACAAGAATCTTTCATCTGGTGCCGAAACCCGGGAGGGGCTCCGGTCT
TCGTCCCCCGTGGACCTACCCCTCCGCCCCAGAAAGCAGGCCACAGCAGCCGGACAAAGGAAGCTCCTCAG
CCTCCAGTTGCTTCTCTGTGCATGCACATCAGTCACTGATCTCACCTACTGGGGCCCTGCAGGCCATGGGG
CCACAGCTCCACACAGAAGCCTCCTAGCAATCCACCTCCACCTGGTGCCTGCTTCAAGTGCGGCAATGAAG
GCCACTGGCCACACAATGCCCAAACCCAGGTAAACCCACGAGGCCATGCCCCCTCTGCGGAGGACCCAC
TGGAAGTTGGACTGTGAGCGGCCCTGCAAGGACCACCCCATCCCTTCTGAGCCAATCAAACCCCTCCTA
CTCGGATCTCGTCAGCCTTGCCGCTGAAGACTGATAGTGCCTTGGAACAGACACCCCAAGCAACTACCATCG
CTTCATCTGAGCCAAGGGTAAACCCCTGATGGTGGCAGGCCAGTCCCAAGGCGCAAGGCCACTTGTGCCAGC
AGTGTGAGTCAGCAAGATAGCAGAAGCAGGAAGAGAGCCGGCCAGAAGACACCTACTCTGACTGGGAGACA
CGTACCCCTGAAGATTGAGAAAAGAGGCCATCCAGGTACCACATAGCAGTTACATCAGACTGGGACATTTCC
TGTTTACAGGAGACTATAAAAACCCCTGCCCCCTACTCATTGGGTGCTGACGCCATTTTAGGCCTCAGCCTG
CCTGCACCCAGGTACTCATTAAAACAGCGTGTGCTCCAAAA

LPRVVCWHTLKSLNGYKNLSSGAETREGLRSSSPVDLPLRPRKQATAAGQRKLLSLQLLLCACTSVTDLTY
WGPAGHGATAPHRSLLAHLHLVPASSAAMKATGPHNAQTQVNPRGHAPSAEDPTGSWTVSGPCKDHPHPF
LSQSNPPTRISSALPLKTDSALEQTPQQLPSLHLSQG

+++++

#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*, *Dasyus 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

CGGAGCTGCGCGGGCGCGGGCGAGCAGGGTCTCCGGGTGGGCGGGCGACGCCCGCGCAGGCTGGAGGC
CGCCGAGGCTCGCCATGCCGGGAGAACTCTAACTCCCCATGGAGTTCGGCCGACTTCTACGAGGCGGAGCC
GCGGCCCGGATGAGCAGCCACCTGCAGAGCCCCCGCACGCGCCAGCAGCGCCGCTTCCGGCTTTCCCC
GGGGCGGGCCCCCGCGCAGCCTCCCCCCCCACCTGCCGCCCGGAGCCGCTGGGCGGCATCTGCGAGCAC
GAGACGTCATCGACATCAGCGCCTACATCGACCCGGCCGCTTCAACGACGAGTTTCTGGCCGACTGTT
CCAGCACAGCCGGCAGCAGGAGAAGGCCAAGGCGGCCGTGGGCCCCACGGGCGGCGGGCGGCGGCGACT
TTGACTACCCGGGCGCGCCCGGGCCCCGGCGGCGCCGTCATGCCGGGGGAGCGCACGGGCCCGCC
GGCTACGGCTGCGCGGCCGCGGCTACCTGGACGGCAGGCTGGAGCCCTGTACGAGCGCGTGGGGCGCC
GGCGCTGCGGCCGCTGGTGATCAAGCAGGAGCCCCGCGAGGAGGATGAAGCCAAGCAGCTGGCGCTGGCCG
GCCTCTTCCCTTACCAGCCGCCGCCCGCCGCCCTCGCACCCGCACCCGCACCCGCCCGCCCGCGCAC
CTGGCCGCCCCGCACCTGCAGTTCAGATCGCGCACTGCGGCCAGACCACCTGCACCTGCAGCCCGGTCA
CCCCACGCCGCCGCCACGCCGTGCCAGCCCGCACCCCGCGCCCGCTCGGTGCCGCCGGCCTGCCGG
GCCCTGGCAGCGCGCTCAAGGGGCTGGGCGCCGCGCACCCCGACCTCCGCGGAGTGGCGGCAGCGGCGCG
GGCAAGGCCAAGAAGTCCGTGGACAAGAACAGCAACGAGTACCGGGTGGCGCGGAGCGCAACAACATCGC
GGTGCGAAGAGCCGCGACAAGGCCAAGCAGCGCAACGTGGAGACGCAGCAGAAGGTGCTGGAGCTGACCA
GTGACAATGACCGCCTGCGCAAGCGGGTGAACAGCTGAGCCGCGAACTGGACACGCTGCGGGGCATCTTC
CGCCAGCTGCCAGAGAGCTCCTTGGTCAAGGCCATGGGCAACTGCGCGTGAAGGCGCGGGCTGTGGGACCG
CCCTGGGCCAGCCTCCGGCGGGGACCCAGGGAGTGGTTTTGGGGTGGCCGATCTCGAGGCTTGGCCGAGCC
GTGCGAGCCAGGACTAGGAGATTCCGGTGCCTCCTGAAAGCCTGGCCTGCTCCGCGTGTCCCCTCCCTTCC

TCTGCGCCGGACTTGGTGCCTAAGATGAGGGGGCCAGGCGGTGGCTTCTCCCTGCGAGGAGGGGAGAAT
TCTTGGGGCTGAGCTGGGAGCCCGGCAACTCTAGTATTTAGGATAACCTTGTGCCTTGGAAATGCAAACCTC
ACCGCTCCAATGCCTACTGAGTAGGGGGAGCAAATCGTGCCTTGTCAATTTATTTGGAGGTTTCCTGCCTC
CTTCCCGAGGCTACAGCAGACCCCATGAGAGAAGGAGGGGAGCAGGCCCGTGGCAGGAGGAGGGGCTCAGG
GAGCTGAGATCCCGACAAGCCCGCCAGCCCCAGCCGCTCCTCCACGCCTGTCCTTAGAAAAGGGGTGGAAAC
ATAGGGACTTGGGGCTTGGAACTAAGGTTGTTCCCTAGTTCTACATGAAGGTGGAGGGTCTCTAGTTCC
ACGCCTCTCCACCTCCCTCCGCACACACCCCAACCCAGCCTGCTATAGGCTGGGCTTCCCTTGGGGCGG
AACTCACTGCGATGGGGGTACCAGGTGACCAGTGGGAGCCCCACCCCGAGTCACACCAGAAAGCTAGGT
CGTGGGTGAGCTCTGAGGATGTATAACCCCTGGTGGGAGAGGGGAGACCTAGAGATCTGGCTGTGGGGCGGGC
ATGGGGGGTGAAGGGCCACTGGGACCCTCAGCCTTGTGTACTGTATGCCTTCAGCATTGCCTAGGAACA
CGAAGCACGATCAGTCCATCCAGAGGGACCGGAGTTATGACAAGCTTTCCAAATATTTTGCCTTTATCAGC
CGATATCAACACTTGTATCTGGCCTCTGTGCCCCAGCAGTGCCTTGTGCAATGTGAATGTGCGCGTCTCTG
CTAAACCACATTTTATTTGGTTTTTGTGTTTTGTTTTGGTTTTGCTCGGATACTTGCCAAAATGAGACTCTC
CGTCGGCAGCTGGGGGAAGGGTCTGAGACTCCCTTTCTTTTTGGTTTTGGGATTACTTTTATCTGGGGG
ACCAATGAGGTGAGGGGGGTTCTCCTTTGCCCTCAGCTTTCCCGAGCCCTCCGGCCTGGGCTGCCACAA
GGCTTGTCCCCAGAGGCCCTGGCTCCTGGTCGGAAGGGAGGTGGCCTCCCGCCAACGCATCACTGGGGC
TGGGAGCAGGAAGGACGGCTTGGTTCTTCTTTTTGGGGAGAACGTAGAGTCTCACTCTAGATGTTTTAT
GTATTATATCTATAATATAAACATATCAAAGTCAA

VRGRGRAGSPGRRRRPAQAGRRGSPCRENSNSPMEADFYEAEP RPPMSSHLQSPPHAPSSAAF GFPRG
AGPAQPPAPPAPEPLGGICEHETSIDI SAYIDPAAFNDEF LADLFQHSRQQEKAKAAVGPTGGGGGGDFD
YPGAPAGPGGAVMPGGAHGPPPGYGCAAAGYLDGRLEPLYERVGAPALRPLVIKQEPREDEAKQLALAGL
FPYQPPPPPPSHPHPHPPAHLAAPHLQFQIAHCGQTTMHLQPGHPTPPPTPVPSHPAPALGAAGLPGP
GSALKGLGAAHPDLRASGGSGAGKAKKSVDKNSNEYRVR RERNNI AVRKSRDKAKQRNVETQQKVLELTS
NDRLRKRVEQLSRELDTLRGIFRQLPESSLVKAMGNCA