CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences
Authors	Ivanov, Ivaylo P.;Firth, Andrew E.;Michel, Audrey M.;Atkins, John F.;Baranov, Pavel V.
Publication date	2011
Original Citation	Ivanov, I. P., Firth, A. E., Michel, A. M., Atkins, J. F. and Baranov, P. V. (2011) 'Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences', Nucleic Acids Research, 39(10), pp. 4220-4234. doi: 10.1093/nar/gkr007
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr007 - 10.1093/nar/gkr007
Rights	© 2011, the Authors. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited. - http://creativecommons.org/licenses/by-nc/2.5
Download date	2026-11-05 20:28:23
Item downloaded from	https://hdl.handle.net/10468/5024



University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Supplementary Dataset 1
Nucleotide and amino acid sequences of human genes of candidate or know 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

CGACGTCAAGTCGAGGCCGCCGCCGCGGGGCCTGGTTATCGCCGGTTCAGCGCAGCCCGGAGTCGCCCAGG
CCTGAACTCCTACCCAGCCTAGACTCAAGTCTGGGTTTCAGCTGCCGCCAGCCCTATTGCTGCTGTTGCTG
TTCTCTGTCCTTGGCCCAGGGGCTGGAGGCCTTTTCCTGACTGATTACTCCACCTGCTCACCCCGCAAGCT
GAGTCCTTTCCGCTCCTTTGCCAGCACCGAGCTCTTCCACTTCCATGTTCCTGAGGACACATTCCTGGCTG
TTTGGAACCTCATCATCTTCAAGGAGCAAGGGGGAACTTTTGGGGACCACTGCCCAGACCAAAGTGTGACT
GTGTATTTCCGGTCCGGGGCACCCCCTGTCATCAATCCCCTGCATACACACTTCCCAGGGGACACAGCTGT
GCCTGGGGTTTTCTCACTGACCCTCAGCTGGACACTGCCCAACCGCACCTCAGGCATCTTTAACGTCAGCA
GCCCCTTACCTGGGGACTGGTTCTTGGCTGCCCACCTTCCCCAGGCCCACGGCCACATCTCTGTCAAGGGT
CTCCAGGATGAGTGTCAGTACCTCCTTCAGCCGCAGCTGATTGTCCGGCGTTTGCTGGACGTCGCTGTGCT
GGTTCCTGGCCGGCCCTCAGAGCAAACCCTCTCCCCACACAATCGCTCAGCCCTGTACAAGGTCTTTGTGC
CCAGCTTCACTTACAGGGTTTCAGCACAGCTGGTGTGTGTGGGGGGCCGTGGGGTATCTGCCTGCCCCCTG
TCACTGCGTCTGCGTCCCAAAGCCCCACCCCTGCACAACTCAAGCTCTGTGGCCTGTGGAGGTGCCTCAGG
ATGCCAGCTGGAGCTGGCACTGCCCCCCTGGGGGCACTGGGTCTACGTGCGTGTGGAAACATCATCCCGGG

GCCCTGGTAGGACCATCCGCTTCCAGCTGTGTGTGCGGTTGCAAGAGTGCCCACAGCCCGGCCTGCTCCGA
GCCCTGGTCCCTGGAGCTGCCATGAACATGCCCCAGTCCCTGGGCAACCAGCCACTGCCCCCAGAACCGCC
ATCCCTTGGAACCCCTGCGGAGGGGCCTGGGACCACGTCCCCACCCGAGCACTGCTGGCCAGTGCGCCCGA
CTCTGCGCAACGAGCTGGACACCTTCTCTGTCCACTTCTACATCTTCTTTGGCCCAAGTGTGGCCCTTCCC
CCTGAGCGCCCAGCCGTGTTCGCCATGAGGCTGTTGCCAGTGCTGGACAGTGGAGGCGTCCTCAGCCTGGA
GCTCCAGCTCAATGCGAGCTCCGTGCGCCAGGAAAACGTGACGGTGTTTGGATGCTTGACTCACGAGGTGC
CCTTGAGCCTGGGGGATGCAGCAGTGACCTGTTCCAAAGAGTCCCTGGCCGGCTTCCTCCTCTCTGTCAGT
GCCACCACCAGGGTTGCCAGGCTGCGAATCCCATTCCCGCAGACGGGGACCTGGTTCCTGGCCCTCCGCTC
CCTGTGCGGGGTGGGGCCTCGGTTCGTGCGGTGCCGCAACGCGACGGCCGAGGTGCGGATGCGCACCTTCC
TGTCCCCATGCGTGGACGACTGCGGGCCCTACGGCCAGTGCAAGCTGCTGCGCACACACAATTATCTGTAC
GCAGCCTGCGAGTGCAAGGCCGGGTGGAGAGGCTGGGGCTGCACCGACAGTGCAGATGCGCTCACCTATGG
ATTCCAGCTGCTGTCCACACTCCTGCTCTGCCTGAGCAACCTCATGTTTCTGCCACCTGTGGTCCTGGCCA
TTCGGAGTCGATATGTGCTGGAAGCTGCAGTCTACACCTTCACCATGTTCTTCTCCACGTTCTATCATGCC
TGTGACCAGCCAGGCATCGTGGTTTTCTGCATCATGGACTACGATGTGCTGCAGTTCTGTGATTTCCTGGG
CTCCTTAATGTCCGTGTGGGTCACTGTCATTGCCATGGCTCGTTTACAGCCCGTGGTCAAGCAGGTGCTGT
ATTTGCTGGGAGCTATGCTGCTGTCCATGGCTCTGCAGCTTGACCGACATGGACTCTGGAACCTGCTTGGA
CCCAGTCTCTTCGCCCTGGGGATCTTGGCCACAGCCTGGACAGTACGCAGCGTCCGCCGCCGGCACTGCTA
CCCACCCACGTGGCGCCGCTGGCTTTTCTACTTGTGCCCTGGCAGCCTTATTGCAGGCAGTGCCGTCCTGC
TTTATGCTTTTGTGGAGACCCGGGACAACTACTTCTACATTCACAGCATTTGGCATATGCTCATTGCGGGC
AGTGTGGGCTTCCTGCTGCCCCCTCGTGCCAAGACTGACCACGGGGTCCCATCTGGAGCCCGGGCCCGGGG
CTGTGGTTACCAGCTATGCATCAACGAGCAGGAGGAGCTGGGCCTCGTGGGCCCAGGAGGGGCCACTGTCA
GCAGCATCTGTGCCAGCTGAGAGGGGCTTTGGGCCTGGCCCTGAGGGGATATGAATGCTTCCTAGAGTTCT
TTCTGGGGGTGTGGAGCCCTCTTAGAAGGAGACAGGCTGTATTTCTTGAGGACATGGAGTCTTTCTCAAGG
ACACAAAACTCTTCCAGGGACCTGGAGCCCTTCCCAGGACATGGAGAACTTCCTGAGGGCCTGGAGTCCCC
CTGCATCATGGAGTCCTTCTTAAGGACTGGAGCCTATGCAGGCACAGAGTCCCTCAGGACCAAGGAGTCCC
TCCTGCAGGTGTGGAGCCTTTCCTGGGATGCAGAGCCTTCCCAAGACATGGATTCCTTCCCAGGGAGACAA
AGCCCTGTCAGGAGCACAGCATCTTTCCAGAGGAGGTGGAGTCTATCTTGGGGAAACCAAATTTCCAGATT
TTCCCAGAGGCTCAGCAACTCTGGCCTCAGGCTTCCTTCCCAGAGGCAGCGTCTGGGCTGTGCTGTGCTGT
GGAGGAGGGATTGCAGGATGGATGGAGCTGGGACTGGGGCTGTCTGGGTGGCTGGTATCCTCGTTTGATAC
AGGTGGAGTCTGTGTGTCTCCAGTGATTGATTGGTTCAGAATGGTTCTGTGATGCCTTTTTTCCCCCTGGG
GTCAGGGGTGTGGGGAAGTGGGGAAAGAGGGCCCCTCAGGGAATCAGCAGGGCTGATGGGAGCTACTGCCG
GAGGCTTTTGTCACACCCTGTACACAGTCTGATCCCGCTCATCTGGGCCCTGCATTCATTTCTAAACAGTT
TCTAATGCCTATTCCCCAATTCCTATTGAGCCCGATTTGCAGTATCTGAGGGGTGTGTGTGTGTGTGTGTG
TGTTTATGTATGTATACGTATGCTGAGATGATTTAAATCAGTGAGACAGACTTTCCTTATGATGCCCACAC
CCAACACAGAAGGAAGCCGAGGTCCCAGGAAATTGGAATAGCAGGTACACGTCTCAGGTGTGCAAGAAATA
TCACAAGAATGTAACTTCTGTTTTGAGCCCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LLLLFSVLGPGAGGLFLTDYSTCSPRKLSPFRSFASTELFHFHVPEDTFLAVWNLIIFKEQGGTFGDHCPD
QSVTVYFRSGAPPVINPLHTHFPGDTAVPGVFSLTLSWTLPNRTSGIFNVSSPLPGDWFLAAHLPQAHGHI
SVKGLQDECQYLLQPQLIVRRLLDVAVLVPGRPSEQTLSPHNRSALYKVFVPSFTYRVSAQLVCVGGRGVS
ACPLSLRLRPKAPPLHNSSSVACGGASGCQLELALPPWGHWVYVRVETSSRGPGRTIRFQLCVRLQECPQP
GLLRALVPGAAMNMPQSLGNQPLPPEPPSLGTPAEGPGTTSPPEHCWPVRPTLRNELDTFSVHFYIFFGPS
VALPPERPAVFAMRLLPVLDSGGVLSLELQLNASSVRQENVTVFGCLTHEVPLSLGDAAVTCSKESLAGFL
LSVSATTRVARLRIPFPQTGTWFLALRSLCGVGPRFVRCRNATAEVRMRTFLSPCVDDCGPYGQCKLLRTH
NYLYAACECKAGWRGWGCTDSADALTYGFQLLSTLLLCLSNLMFLPPVVLAIRSRYVLEAAVYTFTMFFST
FYHACDQPGIVVFCIMDYDVLQFCDFLGSLMSVWVTVIAMARLQPVVKQVLYLLGAMLLSMALQLDRHGLW
NLLGPSLFALGILATAWTVRSVRRRHCYPPTWRRWLFYLCPGSLIAGSAVLLYAFVETRDNYFYIHSIWHM
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
CAACAGTTCCAGTGGACTGCGTGGACCCGTGAGCTCAGGAGCCTCAGACGCCTCCCTGGAGAGCCAAGCTG
GTGTTCGAGGTGGCGCCTCCAGGGTCCACCCTGCTGCCCAACAGCCCCGCGGCCACCAGAGGGCCGTCCCT
GGCCCGGCTGTGTGCCCTGGTGGACCTGTGTCTGGGCTGCTCCCGCTGCACCCAGCGGCTCAATGAAAGCA
CCTACGTCCTCCGTAGGGTGGAGCATGACTGCTCCCGCGAGATCCTGCTGGCCCGCTTTAAGCAGGCCACC
AAGAGCAAGGTCTGGCGCGTGGTGGGCTGCCGGCCCACCTTCCCAAGGCCCCTGTGCTACCAAGTCTGCCA
CTACTACAGCCCTGGGCTCGGCTGCCGGCGCCACCGAAACCGGTGCACCTTTGCCCGCAGTCGCGAGGAGG
CCCTGGTCTGGACCTTCGAGCGTCAGCACAACCTCCAGCGCCTATGGCTGAAGGCGGAGGTGCAGGGCAGC
GGGGCCCAGGGAGGGGCAGGCCGGGCGGCCGACGCCATCCTTACGGAGTTTGGCGGCCGCTTCGAGCTGCT
TTGCTCCCTCTGCTTCAGGCGCTGTCCCCCATGCATCTGTCGCGTGGACCCCCAGGGGCAGTGCCCTGAGC
ACGGAGCCTGCCCCTCCCTCCTGGCCCACGTGAGCGCCGAGGGCCGCCGCAAGCAACAGTTTGTGGTGGTG
AGGCCGCGGCCCCGGGCCGGCCAGCCTCCTGCCTACTGCAGGTTTGTGGGGCGTGGGCAGCCGTGCTGGCG
TGGGGAGTCCCGCTGCCAGTTTGCACACAGCGCCGTGGAGATGGCTGTGTGGGAGGCCGAGCAGCTGGGTG
GCCTCCAGCGGGGGGACCTGCTCACACCCCCTGCCCCTGATGGCGACGGGCGCACGGCCCCCCTTGGCCAG
CCCCCTGGGGCCCAGCTGTACTGCCCGGCCTGCTTGGTCACCTGCCACTCTCAGGAGGCCTTCGAGAACCA
CTGCGCATCCTCGGAGCACGCACAGATGGTGGCCTTCGACCAGGCCCTGCCCTGGGAGCACCGTTCCCCAC
CCCCGGGACTCTCCAAGTTCGAGCTCTGCCCAAAGCCTGACCTCTGTGAGTATGGGGACGCCTGCACCAAG
GCACACTCAGCACAGGAGCTGCAGGAGTGGGTCCGGCGCACGCAGGCTGTGGAGCTGCGGGGGCAGGCGGC
CTGGCAGGACGGGCTGGTGCCCTACCAGGAGCGGCTGCTGGCCGAGTACCAGCGCAGCAGCAGTGAGGTCC
TTGTGCTGGCAGAGACCCTTGATGGAGTGCGTGTCACCTGCAACCAGCCCCTGATGTACCAGGCCCAGGAG
AGGAAGACCCAGTACAGCTGGACGTTTGCCGTCCACTCTGAGGAGCCCCTGCTACACGTGGCCCTGCTGAA
GCAGGAGCCAGGAGCCGACTTCTCTCTGGTGGCTCCCGGCCTCCCGCCAGGCCGGCTCTACGCACGGGGTG
AGCGCTTCCGTGTGCCCAGCTCCACTGCCGACTTCCAGGTGGGAGTGCGTGTGCAGGCTGCCTCGTTCGGC
ACCTTTGAGCAATGGGTGGTCTTCGACTTTGGCCGCCGGCCGGTGCTGCTACAAAAGCTGGGGCTGCAGCT
GGGCCAGGGGCGTCGCCCAGGACCCTGCAGGAATCTGGCGCTCGGCCACCCTGAGGAGATGGAGCGCTGGC
ACACTGGCAACCGCCACGTGGTGCCTGGCGTGGAGCGGACGGCCGAGCAGACGGCCCTGATGGCCAAGTAC
AAGGGCCCTGCCCTGGCCCTGGAGTTCAACCGCAGCAGCGTGGCCTCGGGCCCCATCTCACCAACCAACTA
TCGGCAGAGGATGCACCAGTTTCTCTATGAGGAGGAGGCGGCTCAGCAGCAGCTGGTGGCCAAGCTGACCC
TGCGGGGCCAGGTGTTCCTGAAGACGGCATTGCAGACGCCAGCGCTGAACATGCTCTTCGCGCCTCCGGGA
GCACTGTACGCAGAGGTCCCCGTCCCCTCCTCCCTGATGCCAGACACAGACCAGGGCTTCCTGCTGGGCCG
GGCGGTCAGCACAGCCCTGGTGGCCCCTGTACCTGCACCCGACAATACGGTGTTCGAGGTGCGGCTGGAGA
GGCGGGCCAGCTCAGAGCAGGCGCTGTGGCTGCTGCTTCCGGCCCGCTGCTGCTTGGCCCTGGGGCTGCAG
CCTGAGGCCCGCCTGGTCCTGGAGGTGCAGTTCCAGATTGACCCGATGACCTTCCGCCTCTGGCACCAGGC
AGTGGACACACTGCCTGAGGAGCAGCTGGTGGTGCCCGACTTGCCCACCTGCGCCCTGCCCAGACCTTGGT
CTGTCCCACCCTTGCGGCGTGGCAACCGCAAGCAGGAGCTGGCCGTGGCGCTCATCGCGGGCTGGGGCCCT
GGGGATGGGAGGCGTGTCCCCCCGCTACTCATCTATGGCCCCTTTGGCACCGGCAAGACCTACACGCTGGC
CATGGCCTCCCTGGAGGTCATCCGGAGGCCTGAAACCAAGGTGCTCATCTGCACACACACCAACAGTGCCG
CCGACATCTACATCCGGGAGTATTTCCACAGCCACGTCAGCGGCGGCCACCCCGAGGCCACTCCTCTCCGT
GTGATGTACACGGACCGGCCGCTGAGCCAGACGGACCCAGTCACGCTGCAGTACTGTTGCCTGACCGACGA
CCGCCAGGCTTTCCGCCCGCCCACACGGGCAGAGCTGGCGCGGCACCGCGTGGTGGTCACCACCACCTCCC
AGGCCCGTGAGCTCAGGGTGCCGGTCGGCTTCTTCTCCCACATTCTCATCGATGAGGCGGCCCAGATGCTG
GAGTGCGAGGCCCTCACCCCGCTGGCCTATGCCTCGCACGGCACCCGCCTCGTGCTGGCGGGCGACCACAT
GCAGGTCACACCCCGGCTGTTCAGTGTGGCCAGGGCCCGGGCGGCCGAGCACACGCTGCTGCACCGCCTCT
TCCTGTGCTACCAGCAGGAGACTCACGAGGTGGCGCGGCAGAGCCGCCTGGTCTTCCACGAGAACTACCGC
TGCACGGACGCCATTGTCAGCTTCATCTCGCGGCACTTCTACGTGGCCAAGGGCAACCCCATCCACGCCAG
GGGCAAGGTTCCGCCCCACCCCCGGCACTACCCGCTCATGTTCTGCCACGTGGCGGGCAGCCCAGACCGGG
ACATGTCCATGGCGTCCTGGCTGAATCTGGCTGAGATTGCGCAGGTCGTCGAGAAGGTGCAGGAGGCCTAC
AACACCTGGCCCAGCTGCTGGGGCGGCCGCGAGCAGAGGTGCATCTGTGTCGTTTCCCACGGTGCCCAGGT
CAGTGCACTGAGGCAGGAGCTGAGGAGGCGGGACCTAGGCCAGGTGTCTGTCGGCAGTTTTGAGATCCTGC
CAGGGCGGCAGTTCCGGGTCGTGGTGCTCAGCACTGTGCACACCTGCCAGAGCCTGCTCAGCCCTGGGGCA

CTGGCCCCTGAGTTCTTCACCGACGCCCGCGTGCTCAACACCGTCCTGACCCGCGCCCAGTCCCAGCTGGT
GGTAGTGGGGGACGCCGTGGCCCTCTGCTCCTTCGGGGCCTGCGGCAAGCTCTGGGAGAGCTTCATCCGTG
AGTGCGTGGAGCGGCACAGTGTCTGCCCCGAGGGCCTGTCCATGGAGCAGGTCGAGCAGGGTGTGGCGCAG
AGACGGCGCTGGCCTCCCCGAGGCACACAGGCTGGGGCAGCGGGGAACTGGGAGGCTGCCCCAGAGCCAGT
AGGGGACCTGGCCGAGGAGCAGGCGGCTGTGGTGACGGCCATGGTGAAGGCAGAGCCGGGAGATGAGGCTC
TGAGCCCAGCATCCCGTGACATCACGGCAACCACAGCGCAGACGGAGGCTGCGGCAGCACCAGCAGGAGAC
GCAGTGAAGGAAGACGTGGTGCCCGGGGCCTGTGCGGCAGGAGCGGCTGCTGCAGCGGGCGTGGAGTCCAC
GGAGGCTGAGGATGCAGAGGCTGACTTCTGGCCGTGGGACGGGGAGCTCAACGCTGACGACGCCATCCTAC
GGGAGCTTCTGGACGAGAGCCAGAAGGTGATGGTGACCGTCGGGGAGGACGGGCTGCTGGACACTGTCGCC
AGGCCCGAGTCCCTGCAGCAGGCCCGGCTGTACGAGAACCTGCCCCCGGCTGCGCTACGGAAGCTGCTGCA
CGCGGAGCCCGAGCGGTACCGCCACTGCTCTTTCGTGCCAGAGACCTTCGAGCGGGCGTCAGCCATCCCGC
TGGACGATGCCTCCTCGGGCCCCATCCAGGTCAGGGGCCGCCTGGACTGTGGGATGGCCTTCGCCGGGGAT
GAGGTGCTGGTGCAGCTCCTTTCGGGAGACAAGGCGCCCGAGGGGCGGCTTCGGGGCCGCGTGCTGGGCGT
GCTGAAGAGGAAGAGGCACGAGCTGGCGTTTGTGTGCCGCATGGACACGTGGGACCCGCGCATCATGGTCC
CCATCAATGGCTCCGTGACCAAGATCTTCGTGGCCGAGCTGAAGGACCCATCGCAGGTCCCCATCTACAGC
CTCCGGAAGGGCCGGCTGCAGCGTGTGGGGCTTGAGAGGCTCACCGCCGAGGCCCGGCACAGCCGGCTCTT
CTGGGTCCAAATTGTCCTGTGGCGGCAAGGCTTCTACTACCCGCTGGGCATCGTCCGGGAGGTGCTGCCTG
AGGCCAGCACCTGGGAGCAGGGCCTCCGCATCCTCGGCCTGGAGTACAGCTTGAGGGTGCCCCCGTCGGAC
CAGGCCACCATCACCAAGGTGCTGCAGAAATACCACACGGAGCTTGGCCGGGTTGCCGGCCGCCGAGAGGA
CTGCCGCGCCTTCTTGACCTTCACTGTGGACCCCCAGGGCGCCTGCAACCTCGATGATGCCCTCAGTGTCC
GAGACCTGGGTCCCAGGTGCGAGGTGGCTGTGCACATCACTGATGTGGCCAGCTTCGTGCCCAGGGACGGG
GTGCTGGACGTGGAGGCGCGAAGGCAGGGCGCTGCGTTCTATGCCCCCGGCAGGGAGCCAGTGCCCATGCT
GCCGGCCAGCCTCTGCCAGGACGTCCTCAGCCTCCTGCCTGGCCGGGACCGCCTGGCCATCTCCCTGTTCC
TCACCATGGAGAAGGCCAGTGGCCAGCTGAAGAGCCTGCGCTTTGCACCCTCCGTGGTCCAGTCTGACCGC
CAGCTGTCCTACGAGGAGGCGGAGGAGGTGATCAGGCAGCACCCGGGTGCCGGCCGTGAGCTGCCGGCCCG
CCTGGACTCCGTGGACGCCTGCGTCGTGGCCGCGTGCTACTTCTCTCGGCTGCTGCGCCGGCACCGCCTGC
GGTCCGACTGCTTCTATGAGCAGCCGGACGAGGACGGCACCCTGGGCTTCCGCGCGGCCCACATCATGGTG
AAGGAGTACATGATTCAGTTTAATAGGCTCGTGGCTGAGTTCCTGGTGGGCAGCGAGTGCACGCGGACGGT
CACGCCTCTGCGGTGGCAGCCAGCACCCCGCAGCCAGCAGCTCAAGGCCCTGTGTGAGAAGCATGGGGACC
GGGTGCCCCTGTCACTGCACCTCGGCCACCACCTGCACGGCGGCGGGGGCAGTCCCCCCGACACGCGGCTG
CACCTCCTGGCCTCCCTCTGGAAGCAGGTCCAGTTTGCTGCCCGCACCCAGGACTACGAACAGATGGTGGA
CTTGGTCACCACGGACGACATGCACCCATTCCTGGCTCCTGCAGGCCGCGACCTCCGCAAGGCCCTGGAGC
GCTCGGCGTTCGGCCGCTGCGCCCGGGGCCACCAGCAGCAGGGCGGCCACTACTCGCTGCAGGTGGACTGG
TACACGTGGGCCACCTCGCCCATCCGCAGGTACCTGGACGTGGTGTTGCAGCGGCAGATCCTGCTGGCGCT
GGGCCATGGGGGCTCTGCCTACTCTGCCAGGGACATCGATGGGCTCTGCCAGGCCTTCAGCCTCCAGCACG
CACTTGCCCAGAGCTATCAGCGGCGGGCGCGCAGCCTGCACCTGGCCGTGCAGCTCAAGGCCCAGCCTCTG
GACAAGCTGGGCTTCGTGGTGGACGTGGAGGCGGGCTCCCGCTGCTTCCGGCTGCTCTTCCCCAGCAACCG
GGAGACGCTGCCTGACCCCTGCCCCGTCCCCTACGGCTCCCTGCAGCTGGCCGAGCACCCCCACGCCCTGG
CAGGCCGGCCGGGCCTGCGGCTCCTGTGGCGGCGCCGTGTCTACTCAGCGCAGGGATCCAGCCCGCCCCTG
CCACTGCCTGGCACTGTGCCGGACCCACACACCCTGGCCGTGGAGACGGCCCTGTGGAAGCAGCTGCTGGA
GCTGGTGGAGCTGCAGCGCTGGCCGGAGGCGGCTGCTCTCATCCAGGAGAAGGGCGAGGCGTCCCAGCGGC
GGGAGCTGGTGCAGGTGCAGCGGAGCCACTGTGGCCATTTCCTGGAGGTGGCCCGGGAGCTGGGCAGTGGG
GACACCCTGCAGGTGCAGCTCGGCACCAGCCTGCAGCACGGCTTCCTGGTACCGAGCCCTCAGCTCTGGAC
GGTGGCACCGGGCTTCAGCCTCTGCCTGGAGCACGTGGAGCGGCCCGGAGACTGCTTCTCAGGCCGTGTGT
ACCGGGCCCCGAGGGACCGGTACCGCGACGTGGATGAGTACGCCTGCGTGTGGGAACCATTCTGCGCCCTG
GAGTCGGCCACCGGCGCGGTTGCCGAGAATGACTCCGTCACACTTCAGCACCTGAGTGTCTCCTGGGAGGC
GTCACGGACGCCGCAGGGGCAGCTGCAGGGCGCCTTCCGCCTGGAGGCCGCCTTCCTCGAGGAGAACTGTG
CCGACATCAACTTCAGCTGCTGCTACCTCTGCATCCGGCTCGAGGGGCTGCCGGCTCCCACGGCCAGCCCA
CGCCCTGGGCCCAGCAGCCTCGGCCCTGGCCTGAATGTTGACCCCGGCACGTATACCTGGGTGGCCCACGG
GCAGACGCAGGACTGGGACCAGGAGCGCCGGGCAGACCGGCAGGAGGCTCCCAGACGGGTGCACCTCTTCG
TCCACCACATGGGCATGGAGAAGGTTCCGGAAGAGGTGCTGAGGCCGGGCACCCTGTTCACCGTTGAGCTG
CTGCCCAAGCAGCTTCCTGATCTCCGCAAGGAGGAAGCCGTGCGTGGACTAGAGGAGGCGTCCCCGCTGGT
CACCAGCATCGCACTGGGCCGGCCTGTCCCGCAGCCCCTCTGCAGAGTGATCCCCAGCAGGTTCCTGGAGC
GGCAGACCTACAACATCCCCGGAGGCCGCCACAAGCTGAACCCCAGCCAGAACGTGGCGGTCAGGGAGGCT
CTGGAGAAGCCTTTCACGGTCATTCAGGGCCCACCAGGTACAGGGAAGACGATCGTGGGCCTCCACATCGT
ATTCTGGTTTCATAAATCAAACCAGGAGCAGGTGCAGCCCGGAGGCCCCCCCCGTGGGGAGAAGCGGCTGG
GGGGTCCCTGCATCTTGTACTGCGGCCCCTCCAACAAGTCGGTGGATGTCCTGGCAGGACTGCTCCTGAGA

AGGATGGAGCTGAAGCCCCTCCGTGTGTACAGTGAGCAGGCTGAGGCCAGCGAGTTCCCAGTGCCGCGTGT
GGGCAGCAGGAAGCTGCTCAGGAAGAGCCCCCGGGAGGGGAGGCCGAACCAGAGCCTCAGGAGCATCACCC
TGCACCACCGGATCCGGCAGGCCCCCAACCCTTACTCGTCGGAAATCAAGGCCTTTGACACCCGGCTGCAG
AGAGGGGAGCTCTTCTCCAGGGAGGACCTGGTCTGGTACAAGAAGGTCTTGTGGGAGGCTCGGAAGTTCGA
GCTGGACCGGCATGAGGTCATCCTCTGCACCTGCTCCTGTGCAGCCTCTGCCAGCCTCAAAATCCTGGACG
TGAGGCAGATCCTTGTTGACGAGGCAGGCATGGCCACGGAACCTGAAACCCTCATCCCCCTGGTGCAGTTC
CCACAGGCCGAGAAGGTGGTTCTTCTCGGAGACCACAAGCAGCTGCGGCCTGTGGTCAAGAATGAGCGGCT
GCAAAACCTGGGTCTGGACCGGTCTCTGTTCGAGCGGTACCACGAGGACGCACATATGCTGGACACTCAGT
ACCGCATGCATGAGGGCATCTGTGCCTTCCCCTCTGTGGCGTTCTACAAGAGCAAGCTGAAGACGTGGCAG
GGCCTGAGGAGGCCGCCCAGTGTCCTGGGCCACGCTGGCAAGGAGAGCTGCCCTGTCATCTTTGGCCACGT
GCAGGGCCACGAGCGGAGCCTGCTGGTGTCCACGGACGAAGGGAATGAGAACTCCAAGGCCAACCTGGAGG
AGGTGGCTGAGGTGGTCCGTATCACCAAGCAGCTGACCCTGGGGAGGACCGTAGAGCCCCAGGACATCGCC
GTCCTCACGCCCTACAACGCGCAGGCCTCTGAGATCAGCAAGGCCCTTCGGCGAGAGGGCATCGCCGGGGT
GGCCGTGTCCTCCATCACCAAGAGCCAGGGGAGCGAGTGGCGCTATGTGCTGGTGAGCACCGTCCGCACCT
GTGCCAAGAGCGACCTGGACCAGCGGCCCACCAAGAGCTGGCTCAAGAAGTTTCTGGGCTTCGTTGTGGAC
CCCAACCAAGTGAATGTGGCTGTCACGCGGGCCCAGGAGGGGCTCTGCCTGATCGGAGACCACCTCCTTCT
GCGCTGCTGCCCCCTCTGGCGTAGCCTCCTGGACTTCTGCGAGGCTCAGCAGACCCTCGTGCCTGCCGGCC
AGGTGCGCGTCTGCAGGAGGCCAACTATGCCTTCCTGAAGAGCCCTCTCCACCTGCAAGGTGCCAGGACTG
GGAGGGAAAGTCCAGGGCCCCCCCACCCTCCCCACCACCCTCCGCCATCAGCCTGCCTGGGCCAGCCTGCC
TCCTGGCCTCAGGATCTCCCTGGGGCAGGAAGTGGGTGGACTTGGAGGAGGAACGGCAAGGGCCGGTCCCC
ACACTCTTGCTCCTCCTGGGCATCTGGGGACAGGAACACCTCTCAGGGTGGCCAAGAGGGGCTTCCTTGGG
ACCAGTGCAAGGAAGTTCTGCTGTCCCTGGGAAGTCTCAGAACCAGGAAGGAAACCCTGTGGAGGCTGCCC
TTTGACCTTTGGACGCCACCTGTAATCACTCTGGATGGTGTCTGTGGAGTTCAGGTTTATTTTCTATTTAA
ATGAACTAAAAGCCAGCATCCTGCCGGGCGTGGTGGCTCACGCCCATAATCCAGCACTTTGGGAAGCCGAG
GCGGGCAGATCACCTGAGGTCAGGAGTTCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCCACTAA
AAATACAAAATTAGCTGGGCGTGGTGGCGGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGA
ATCGCTTGACTCCAGGAGGCGCAGGCTGCAGTGAGCTCTGATGGTACCACTGCACTCCAGCCTGGGTGACA
GAGCGAGACCCTGTCTTAAAAAAAAAAACAAAACAAAACCAGCATCTGTTACTGGAACAGAGACCTCAGCC
CAAGCTCAGGACAAGGAGGCCTCCCTGGGAGAGGGGGCCTTTCTCCCACCATCACCGCCTGTGTCCTTCAG
GGGCTGCTGGAGGGCAGCCCAAAGCCAGGGGCCCACCCCGCCCCACCCCACCCCACCCCACCCTCTGAAAT
GTGAAAAGCCTGCACTCTTCCTGTGGGCTGCGCAGTGGGCTCGGGGGTTGGGGGGTGCCGGGGCGATGCTT
CCACCCTGGACTGAGTGTTCGGGGCAGCCTGCTCGGCTGTATGTTTCTCGCACGCTCACCCGCTTCAGTGC
GGCTCTGTCCCTGCTGCCGCCTCCAGTCCTGAGGGAGGGGCTTGCCCCACCTTCTACATTCCAATTTTTAT
ATCTTTGAATTATGTGATTAGATATTAATTTAATGATAAAACCACTCTGAAAGCTCTTCTCA

VAPPGSTLLPNSPAATRGPSLARLCALVDLCLGCSRCTQRLNESTYVLRRVEHDCSREILLARFKQATKSK
VWRVVGCRPTFPRPLCYQVCHYYSPGLGCRRHRNRCTFARSREEALVWTFERQHNLQRLWLKAEVQGSGAQ
GGAGRAADAILTEFGGRFELLCSLCFRRCPPCICRVDPQGQCPEHGACPSLLAHVSAEGRRKQQFVVVRPR
PRAGQPPAYCRFVGRGQPCWRGESRCQFAHSAVEMAVWEAEQLGGLQRGDLLTPPAPDGDGRTAPLGQPPG
AQLYCPACLVTCHSQEAFENHCASSEHAQMVAFDQALPWEHRSPPPGLSKFELCPKPDLCEYGDACTKAHS
AQELQEWVRRTQAVELRGQAAWQDGLVPYQERLLAEYQRSSSEVLVLAETLDGVRVTCNQPLMYQAQERKT
QYSWTFAVHSEEPLLHVALLKQEPGADFSLVAPGLPPGRLYARGERFRVPSSTADFQVGVRVQAASFGTFE
QWVVFDFGRRPVLLQKLGLQLGQGRRPGPCRNLALGHPEEMERWHTGNRHVVPGVERTAEQTALMAKYKGP
ALALEFNRSSVASGPISPTNYRQRMHQFLYEEEAAQQQLVAKLTLRGQVFLKTALQTPALNMLFAPPGALY
AEVPVPSSLMPDTDQGFLLGRAVSTALVAPVPAPDNTVFEVRLERRASSEQALWLLLPARCCLALGLQPEA
RLVLEVQFQIDPMTFRLWHQAVDTLPEEQLVVPDLPTCALPRPWSVPPLRRGNRKQELAVALIAGWGPGDG
RRVPPLLIYGPFGTGKTYTLAMASLEVIRRPETKVLICTHTNSAADIYIREYFHSHVSGGHPEATPLRVMY
TDRPLSQTDPVTLQYCCLTDDRQAFRPPTRAELARHRVVVTTTSQARELRVPVGFFSHILIDEAAQMLECE
ALTPLAYASHGTRLVLAGDHMQVTPRLFSVARARAAEHTLLHRLFLCYQQETHEVARQSRLVFHENYRCTD
AIVSFISRHFYVAKGNPIHARGKVPPHPRHYPLMFCHVAGSPDRDMSMASWLNLAEIAQVVEKVQEAYNTW
PSCWGGREQRCICVVSHGAQVSALRQELRRRDLGQVSVGSFEILPGRQFRVVVLSTVHTCQSLLSPGALAP
EFFTDARVLNTVLTRAQSQLVVVGDAVALCSFGACGKLWESFIRECVERHSVCPEGLSMEQVEQGVAQRRR
WPPRGTQAGAAGNWEAAPEPVGDLAEEQAAVVTAMVKAEPGDEALSPASRDITATTAQTEAAAAPAGDAVK
EDVVPGACAAGAAAAAGVESTEAEDAEADFWPWDGELNADDAILRELLDESQKVMVTVGEDGLLDTVARPE
SLQQARLYENLPPAALRKLLHAEPERYRHCSFVPETFERASAIPLDDASSGPIQVRGRLDCGMAFAGDEVL
VQLLSGDKAPEGRLRGRVLGVLKRKRHELAFVCRMDTWDPRIMVPINGSVTKIFVAELKDPSQVPIYSLRK
GRLQRVGLERLTAEARHSRLFWVQIVLWRQGFYYPLGIVREVLPEASTWEQGLRILGLEYSLRVPPSDQAT

ITKVLQKYHTELGRVAGRREDCRAFLTFTVDPQGACNLDDALSVRDLGPRCEVAVHITDVASFVPRDGVLD
VEARRQGAAFYAPGREPVPMLPASLCQDVLSLLPGRDRLAISLFLTMEKASGQLKSLRFAPSVVQSDRQLS
YEEAEEVIRQHPGAGRELPARLDSVDACVVAACYFSRLLRRHRLRSDCFYEQPDEDGTLGFRAAHIMVKEY
MIQFNRLVAEFLVGSECTRTVTPLRWQPAPRSQQLKALCEKHGDRVPLSLHLGHHLHGGGGSPPDTRLHLL
ASLWKQVQFAARTQDYEQMVDLVTTDDMHPFLAPAGRDLRKALERSAFGRCARGHQQQGGHYSLQVDWYTW
ATSPIRRYLDVVLQRQILLALGHGGSAYSARDIDGLCQAFSLQHALAQSYQRRARSLHLAVQLKAQPLDKL
GFVVDVEAGSRCFRLLFPSNRETLPDPCPVPYGSLQLAEHPHALAGRPGLRLLWRRRVYSAQGSSPPLPLP
GTVPDPHTLAVETALWKQLLELVELQRWPEAAALIQEKGEASQRRELVQVQRSHCGHFLEVARELGSGDTL
QVQLGTSLQHGFLVPSPQLWTVAPGFSLCLEHVERPGDCFSGRVYRAPRDRYRDVDEYACVWEPFCALESA
TGAVAENDSVTLQHLSVSWEASRTPQGQLQGAFRLEAAFLEENCADINFSCCYLCIRLEGLPAPTASPRPG
PSSLGPGLNVDPGTYTWVAHGQTQDWDQERRADRQEAPRRVHLFVHHMGMEKVPEEVLRPGTLFTVELLPK
QLPDLRKEEAVRGLEEASPLVTSIALGRPVPQPLCRVIPSRFLERQTYNIPGGRHKLNPSQNVAVREALEK
PFTVIQGPPGTGKTIVGLHIVFWFHKSNQEQVQPGGPPRGEKRLGGPCILYCGPSNKSVDVLAGLLLRRME
LKPLRVYSEQAEASEFPVPRVGSRKLLRKSPREGRPNQSLRSITLHHRIRQAPNPYSSEIKAFDTRLQRGE
LFSREDLVWYKKVLWEARKFELDRHEVILCTCSCAASASLKILDVRQILVDEAGMATEPETLIPLVQFPQA
EKVVLLGDHKQLRPVVKNERLQNLGLDRSLFERYHEDAHMLDTQYRMHEGICAFPSVAFYKSKLKTWQGLR
RPPSVLGHAGKESCPVIFGHVQGHERSLLVSTDEGNENSKANLEEVAEVVRITKQLTLGRTVEPQDIAVLT
PYNAQASEISKALRREGIAGVAVSSITKSQGSEWRYVLVSTVRTCAKSDLDQRPTKSWLKKFLGFVVDPNQ
VNVAVTRAQEGLCLIGDHLLLRCCPLWRSLLDFCEAQQTLVPAGQVRVCRRPTMPS

AUG initiation = 294.6 kDa
GUG initiation = 322.3 kDa

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

#3

NM_001010858

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

RNF187 ring finger protein 187

Synonyms: RNF187

TCTAGCGAGGTGACAGCGTAGAACCAGGTGCGCGTCCCCGGCGTTGGCGTCTTCGTCCTGTTGCTGGTCTC
CGTCCGGTCGCCGGCCGTCTAGGTCTCCGGCCCTCCCCAGCCGCTCCTGCGCCCTTGCCGGCCCCGCCGCC
CGCAGCCCTGGCGCTCCCTGCGGGCCCCGCCGAGGCCGCCTGCGCCCTGTGCCAGCGCGCGCCCCGGGAAC
CGGTGCGCGCCGACTGCGGCCACCGCTTCTGTCGGGCGTGCGTGGTGCGCTTCTGGGCCGAGGAGGACGGG
CCCTTCCCGTGCCCCGAGTGCGCCGACGACTGCTGGCAGCGCGCCGTGGAGCCCGGCAGGCCCCCGCTCAG
CCGCCGCCTTCTGGCGCTCGAGGAGGCGGCCGCGGCGCCCGCGCGCGACGGCCCGGCCAGCGAGGCCGCGC
TGCAGCTGCTGTGCCGCGCCGACGCCGGCCCGCTCTGCGCCGCCTGCCGTATGGCTGCGGGCCCCGAGCCG
CCCGAGTGGGAACCGCGCTGGAGGAAGGCGCTGCGCGGCAAGGAGAACAAGGGGTCTGTGGAAATCATGAG
AAAGGACTTGAATGACGCCCGGGACCTGCATGGCCAGGCAGAGTCAGCAGCTGCAGTGTGGAAGGGACACG
TGATGGACCGTAGGAAGAAGGCACTGACCGACTACAAGAAGCTGCGGGCCTTCTTTGTGGAGGAGGAGGAG
CATTTCCTGCAGGAGGCTGAGAAGGAGGAGGGGCTCCCTGAGGACGAGCTGGCTGACCCCACTGAGCGGTT
CAGGTCACTGCTGCAGGCGGTCTCGGAGCTGGAGAAGAAGCATCGCAACCTGGGCCTCAGCATGCTGCTGC
AGTGATGGCGCCAACCCGTGGCAGTCCCAGAGCTGGAGGCAGGAGGATGGATCCTCATCTCCATGGGAAGT
GTCAGCGTGTGGCTGCCAGGGAAGCGTGGCAGGCGCCTGGCCTTGGGTCCATCTACATAGTTGCGTGTTTC
AACAATGTCCATTTATCCTTCACCCCGAGGCGTGTTTTGGGGGCTGCAAACACCTCCCGGTAGAGGCTGGA
CCTGAGGACCCTTCCCACCTGTGCCCGTCCCTTCCTGAAGTCCTAGCCACAGCCCATCCTCCATGAGTCCC
GGCAGCTCTGGGTCATGCCCTTCCCTGGTCACCCATCTGCCCCTCACCTCGTCATCCAGGGACCCAGACCC
TGCACCTTCCATGTGGGCCCACAGATCCTTGGCAGGTACCTGAGGTGCACCATTGAGTGTCGGATTTGGGG
TTAGCATCCAGAAAGAAGAATGCGCATGACGCTCTGTGAAGGCTGGAACTCAGGTCTTCAGGGAGAGAAAG
GAAGACTGGATTGCACCTTGATGCCTCCTGAGGAGGCGGCCCCCCTCTTGAGGTGGGCGTGGGCCCGGCCC
AGCCTTATCCAAGTCGCTCTGTCCACCTCCCCCTTCCTGGCCCCCACCCCACTCCTGTGCCTCCCAGGAGC

CCTCCCTGTGCTCCACCTGCCTCCGCAGAAGGAAGCCTCTTTCTCTGTTTCCCTGGGTGAGGGGGCTGGCA
GGTGGCTAACCCCATTTAGCATCTCCAGGCCCTGCCATCGTGTCTCATCTTGCTGTTATCTCTAGCTCTTT
CCCTCCTCCCATTTCCTTTAGTAGTTGAATTTTGCAAAGCTTGTAGCAGTAGCTCAGTTGCCTGCAGCATC
CTTGTGTGTAGATAAATTAGTCGACAGAAACTCAGCACTGGGGACAGGATTGCAAAGTCGGGGACATAGAT
GCAGACAGTTGTTGAGATTTGGGGATAGCCGGGCTTGTGAGCGGTGCCCATTTCCAGATGAAGCCTTTCAG
CCCTTCTGAGTCCCCGGCCCTTGGTGCGATGTCTGTGAGTTTGACCTGCCCAGCGTGTGGGCTGGCTCAAT
GCTGAATAAAGTGGGTTTGTGTCAGCTCGTTTGCTTCGTCTCCGTGTGTCCACCTGGCCTCTTCCCCCTGC
CCTGGCCACCCTCCAGTGTCAAAGGAAACTTCCTCGTGACACGTGCTAAAGCATGGTGAGGAGGACTTTGA
TTGGGACCATTGAGATGGGTGTGGGACCCTTTCCTTGGGGCCTGGGGGGAGATGGGGCTCCACCCCGACGT
AGCAGGGCAGGGGTTGGAGGAGCGAGGAGCAGTATAGGGTCCATGGGTGGGAATGACTGTGAGGAGACATC
AGGGCTGAGGGGGCTCTGGCTAAACCCACCTCACAGAGTCCTTGCTGCAGGCAGGCAGGGCGATCAGACAT
TGGCTGCAAACGGTCAGAGAGGAACCCAGTCAGGTACCATTGAGGGTGGTCAGATATTATGGTTAACCAAA
TTAGGGTTCTTGCTAAAACTGGATTTCATAAGAAAGGGCAAAGAGGGCCCTAGGAGAAGATTCCAGAGCCT
GGCCAGAGTTTGGCCAAGTAGAGAATCTTTGTCAGCACGCCAACAACATCCCGACCCTGAGACCTCCAGTT
TGTCTTTCTCACTGTCTCCGCCTGCTGCAGTCTGCTGTCATCCCTGAGCATCCCTGCCCCTGCCCTGCACA
CCTGTGATGCTTGCCCGGACAGGTCCTGATGGCAGAGTCTCCCACAACATCAGTGTCTCCACATCACCAGG
TCCGACAGTGGCTTCACCATCCTCACCTAACCTAGCTGACCAGCAACATCCCACCCTGTCAATCACAACCT
CTTTCTATTTAAGAAAATTATATATTTATGGGGCACAGTG
CGTCCTGTTGCTGGTCTCCGTCCGGTCGCCGGCCGTCTAGGTCTCCGGCCCTCCCCAGCC

LALPAGPAEAACALCQRAPREPVRADCGHRFCRACVVRFWAEEDGPFPCPECADDCWQRAVEPGRPPLSRR
LLALEEAAAAPARDGPASEAALQLLCRADAGPLCAACRMAAGPEPPEWEPRWRKALRGKENKGSVEIMRKD
LNDARDLHGQAESAAAVWKGHVMDRRKKALTDYKKLRAFFVEEEEHFLQEAEKEEGLPEDELADPTERFRS
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

CTCGGGCGCGCTGGCCCCTGGGGACGCCGAGGGCGGCTGCGACGCGCCGAGAGGCCGCGGCTCTCCCACCT
GTCACCCTGGCCCTTCTCTGCTTGGATGGTGTCTTCCTCTCCTCAGCCGAGAATGACTTCGTCCACCGGAT
CCAGGAGGAACTGGACCGCTTTCTGCTGCAGAAGCAGCTGTCAAAGGTTCTTCTTTTCCCCCCACTCTCCA
GTCGCCTCCGGTACCTGATCCATAGAACAGCAGAGAATTTTGATCTCTTGAGCAGCTTCTCCGTTGGGGAG
GGCTGGAAGAGGAGGACGGTCATCTGTCACCAGGACATCAGGGTACCCAGTTCGGATGGCCTCTCTGGCCC
CTGCCGCGCTCCTGCCTCCTGCCCCAGCAGGTACCACGGTCCTCGGCCCATCTCCAACCAAGGAGCAGCTG
CGGTTCCCCGAGGTGCCCGGGCTGGCCGGTGGTATCGTGGACGCAAGCCTGACCAGCCTTTGTATGTGCCC
CGGGTGCTGCGCAGGCAGGAAGAATGGGGGCTGACCTCTACCTCGGTGCTCAAGAGAGAGGCCCCAGCTGG
CAGGGACCCAGAAGAGCCTGGAGATGTTGGTGCTGGAGACCCCAACTCTGATCAGGGACTCCCTGTGCTGA
TGACTCAGGGAACAGAGGACCTAAAGGGCCCAGGACAAAGGTGTGAGAATGAGCCACTGCTGGACCCTGTT
GGCCCTGAGCCTCTGGGGCCTGAGAGTCAGTCAGGGAAGGGAGACATGGTGGAGATGGCCACACGGTTTGG
GTCCACCCTGCAGCTAGACCTGGAAAAGGGGAAGGAGAGTCTGTTGGAGAAGAGGCTGGTGGCAGAGGAGG
AAGAGGACGAAGAGGAGGTGGAAGAGGATGGCCCCAGCAGCTGCTCGGAGGACGATTACAGTGAGCTGCTG
CAGGAGATCACAGACAACCTGACGAAGAAGGAGATTCAGATAGAGAAGATCCATTTGGACACATCCTCCTT
CGTGGAGGAGCTGCCTGGAGAGAAGGACCTTGCCCACGTGGTAGAGATCTATGACTTTGAACCAGCGCTCA
AGACGGAGGACCTGCTGGCAACGTTTTCTGAGTTCCAAGAGAAGGGGTTCAGGATTCAGTGGGTGGATGAT
ACTCACGCACTCGGCATCTTTCCCTGCCTGGCCTCAGCTGCGGAAGCCCTGACCCGGGAGTTCTCGGTGCT

CAAGATCCGGCCCCTCACACAGGGAACCAAGCAGTCAAAGCTCAAAGCCTTGCAGAGGCCAAAACTCCTGC
GTCTGGTGAAGGAGAGGCCACAGACAAATGCGACTGTGGCCCGGCGGCTGGTGGCCCGGGCCCTGGGACTC
CAACACAAAAAGAAAGAGCGGCCTGCTGTCCGGGGTCCGCTGCCGCCCTGAGGCCTGGAGACCCAACTGGC
CTGGATCTGCGTCCCGACGTAGCTGGCGCCCCCAACACCATAAGCCTTCACAGACGCCAGAGCAGCCCCGC
ACCACCCTCGAGCTTCACCATGGGGTGTGGTGGGCTTTAGTTTAGTCCCAGAAATGGAGAAAAAATAAAAA
CTCACGTTGTTCTAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAA

LALLCLDGVFLSSAENDFVHRIQEELDRFLLQKQLSKVLLFPPLSSRLRYLIHRTAENFDLLSSFSVGEGW
KRRTVICHQDIRVPSSDGLSGPCRAPASCPSRYHGPRPISNQGAAAVPRGARAGRWYRGRKPDQPLYVPRV
LRRQEEWGLTSTSVLKREAPAGRDPEEPGDVGAGDPNSDQGLPVLMTQGTEDLKGPGQRCENEPLLDPVGP
EPLGPESQSGKGDMVEMATRFGSTLQLDLEKGKESLLEKRLVAEEEEDEEEVEEDGPSSCSEDDYSELLQE
ITDNLTKKEIQIEKIHLDTSSFVEELPGEKDLAHVVEIYDFEPALKTEDLLATFSEFQEKGFRIQWVDDTH
ALGIFPCLASAAEALTREFSVLKIRPLTQGTKQSKLKALQRPKLLRLVKERPQTNATVARRLVARALGLQH
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

TCACGCGTCCGCTGCTCGTTGAGTGAAGAAAATCCACCGGCATCGCCTGAGCCCCGCTACCGAGAAGGGCG
CCGCTTCCTCCGGGGAGGGGGATAAAGATCCCCCGCCGCCGGCCCATGAGGATATTGCCGTGAAAGGCACA
GCGACTGCAGCAGGAACCGGACCCGGCACCGGAGCGGCGGCGGCGGCGGCAGCAGCGGTACCGCCTCCTCA
CCCGGCGGCGGCAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCAGCGGTCCCCCCTCCTCACC
CGAACATCAGGGCCCTCCAGACTCAGGCGCCCCAACAAATTCCTAGAGGACCTGTGCAACAACCTCTTGAG
GATCGAATCTTCACTCCCGCTGTCTCAGCAGTCTACAGCACGGTAACACAAGTGGCAAGACAGCCGGGAAC
CCCTACCCCATCCCCTTATTCAGCACATGAAATAAACAAGGGGCATCCAAATCTTGCGGCAACGCCCCCGG
GACATGCATCGTCCCCTGGACTCTCTCAAACCCCTTATCCCTCTGGACAGAATGCAGGTCCAACCACGCTG
GTATACCCTCAAACCCCTCAGACAATGAATTCACAACCTCAAACCCGTTCTCCGTTTTTCCAGAGGCCTCA
AATACAGCCTCCTAGAGCTACCATCCCGAACAGCAGTCCTTCCATTCGTCCTGGTGCACAGACACCCACTG
CAGTGTACCAGGCTAATCAGCACATCATGATGGTTAACCATCTGCCCATGCCGTACCCAGTGCCCCAGGGG
CCTCAGTACTGTATACCACAGTACCGTCATAGTGGCCCTCCTTATGTTGGGCCCCCCCAACAATATCCAGT
TCAACCACCGGGGCCAGGTCCTTTTTATCCTGGACCAGGACCTGGGGACTTCCCCAATGCTTATGGAACGC
CTTTTTACCCAAGTCAGCCGGTGTATCAGTCAGCACCTATCATAGTGCCTACGCAGCAACAGCCGCCTCCA
GCCAAGAGAGAGAAAAAAACTATAAGAATTCGGGATCCAAACCAGGGAGGTAAAGACATAACAGAGGAGAT
TATGTCTGGAGGTGGCAGCAGAAATCCTACTCCACCCATAGGAAGACCCACGTCCACACCTACTCCTCCTC
AGCAGCTGCCCAGCCAGGTCCCCGAGCACAGCCCTGTGGTTTATGGGACTGTGGAGAGCGCTCATCTTGCT
GCCAGCACCCCTGTCACTGCAGCTAGCGACCAGAAGCAAGAGGAGAAGCCAAAACCAGATCCAGTGTTAAA
GTCTCCTTCCCCAGTCCTTAGGCTAGTCCTCAGTGGAGAGAAGAAAGAACAAGAAGGCCAGACATCTGAAA
CTACTGCAATAGTATCCATAGCAGAGCTTCCTCTGCCTCCATCACCTACCACTGTTTCTTCTGTTGCTCGA
AGTACAATTGCAGCCCCCACCTCTTCTGCTCTTAGTAGCCAACCAATATTCACCACTGCTATAGATGACAG
ATGTGAACTCTCATCCCCAAGAGAAGACACAATTCCTATACCCAGCCTCACATCTTGCACAGAAACATCAG
ACCCTTTACCAACAAATGAAAATGATGATGATATATGCAAGAAACCCTGTAGTGTAGCACCTAATGATATT
CCACTGGTTTCTAGTACTAACCTAATTAATGAAATAAATGGAGTTAGCGAAAAATTATCAGCCACGGAGAG
CATTGTGGAAATAGTAAAACAGGAAGTATTGCCATTGACTCTTGAATTGGAGATTCTCGAAAATCCCCCAG
AAGAAATGAAACTGGAGTGTATCCCAGCTCCCATCACCCCTTCCACAGTTCCTTCCTTTCCTCCAACTCCT
CCAACTCCTCCAGCTTCTCCTCCTCACACTCCAGTCATTGTTCCTGCTGCTGCCACTACTGTTAGTTCTCC
GAGTGCTGCCATCACAGTCCAGAGAGTCCTAGAGGAGGACGAGAGCATAAGAACTTGCCTTAGTGAAGATG
CAAAAGAGATTCAGAACAAAATAGAGGTAGAAGCAGATGGGCAAACAGAAGAGATTTTGGATTCTCAAAAC

TTAAATTCAAGAAGGAGCCCTGTCCCAGCTCAAATAGCTATAACTGTACCAAAGACATGGAAGAAACCAAA
AGATCGGACCCGAACCACTGAAGAGATGTTAGAGGCAGAATTGGAGCTTAAAGCTGAAGAGGAGCTTTCCA
TTGACAAAGTACTTGAATCTGAACAAGATAAAATGAGCCAGGGGTTTCATCCTGAAAGAGACCCCTCTGAC
CTAAAAAAAGTGAAAGCTGTGGAAGAAAATGGAGAAGAAGCTGAGCCAGTACGTAATGGTGCTGAGAGTGT
TTCTGAGGGTGAAGGAATAGATGCTAATTCAGGCTCCACAGATAGTTCTGGTGATGGGGTTACATTTCCAT
TTAAACCAGAATCCTGGAAGCCTACTGATACTGAAGGTAAGAAGCAGTATGACAGGGAGTTTCTGCTGGAC
TTCCAGTTCATGCCTGCCTGTATACAAAAACCAGAGGGCCTGCCTCCTATCAGTGATGTGGTTCTTGACAA
GATCAACCAACCCAAATTGCCAATGCGAACTCTGGATCCTCGAATTTTGCCTCGAGGACCAGACTTTACAC
CAGCCTTTGCTGATTTTGGAAGGCAGACACCTGGTGGAAGAGGCGTACCTTTGTTGAATGTTGGGTCACGA
AGATCTCAACCTGGCCAAAGAAGAGAACCCAGAAAGATCATCACAGTTTCTGTAAAAGAAGATGTACACCT
GAAAAAGGCAGAAAATGCCTGGAAGCCAAGCCAAAAACGAGACAGCCAAGCCGATGATCCCGAAAACATTA
AAACCCAGGAGCTTTTTAGAAAAGTTCGAAGTATCTTAAATAAATTGACACCACAGATGTTCAATCAACTG
ATGAAGCAAGTGTCAGGACTTACTGTTGACACAGAGGAGCGGCTGAAAGGAGTTATTGACCTGGTCTTTGA
GAAGGCTATTGATGAACCCAGTTTCTCTGTGGCTTACGCAAACATGTGTCGATGTCTAGTAACGCTGAAAG
TACCCATGGCAGACAAGCCTGGTAACACAGTGAATTTCCGGAAGCTGCTACTGAACCGTTGCCAGAAGGAG
TTTGAAAAAGATAAAGCAGATGATGATGTCTTTGAGAAGAAGCAGAAAGAACTTGAGGCTGCCAGTGCTCC
AGAGGAGAGGACAAGGCTTCATGATGAACTGGAAGAAGCCAAGGACAAAGCCCGGCGGAGATCCATTGGCA
ACATCAAGTTTATTGGAGAACTCTTTAAACTCAAAATGCTGACTGAAGCCATCATGCATGACTGTGTGGTG
AAGCTGCTAAAGAACCATGATGAAGAATCCCTGGAGTGCCTGTGTCGCCTGCTCACCACCATTGGCAAAGA
CTTGGACTTTGAAAAAGCAAAGCCACGTATGGACCAGTACTTTAATCAGATGGAGAAAATTGTGAAAGAAA
GAAAAACCTCATCTAGGATTCGGTTCATGCTTCAAGATGTTATAGACCTAAGGCTGTGCAATTGGGTATCT
CGAAGAGCAGATCAAGGGCCTAAAACTATCGAACAGATTCACAAAGAGGCTAAAATAGAAGAACAAGAAGA
GCAAAGGAAGGTCCAGCAACTCATGACCAAAGAGAAGAGAAGACCAGGTGTCCAGAGAGTGGACGAAGGTG
GGTGGAACACTGTACAAGGGGCCAAGAACAGTCGGGTACTGGACCCCTCAAAATTCCTAAAAATCACTAAG
CCTACAATTGATGAAAAAATTCAGCTGGTACCTAAAGCACAGCTAGGCAGCTGGGGAAAAGGCAGCAGTGG
TGGAGCAAAGGCAAGTGAGACTGATGCCTTACGGTCAAGTGCTTCCAGTTTAAACAGATTCTCTGCCCTGC
AACCTCCAGCACCCTCAGGGTCCACGCCATCCACGCCTGTAGAGTTTGATTCCCGAAGGACCTTAACTAGT
CGTGGAAGTATGGGCAGGGAGAAGAATGACAAGCCCCTTCCATCTGCAACAGCTCGGCCAAATACTTTCAT
GAGGGGTGGCAGCAGTAAAGACCTGCTAGACAATCAGTCTCAAGAAGAGCAGCGGAGAGAGATGCTGGAGA
CCGTGAAGCAGCTCACAGGAGGTGTGGATGTGGAGAGGAACAGCACTGAGGCTGAGCGAAATAAAACAAGG
GAGTCAGCAAAACCAGAAATTTCAGCAATGTCAGCTCATGACAAGGCTGCATTATCAGAAGAGGAACTGGA
GAGGAAGTCGAAATCTATCATTGATGAATTTCTACACATTAATGATTTTAAGGAAGCCATGCAGTGTGTGG
AAGAGCTGAATGCCCAGGGCCTACTACATGTTTTTGTGAGAGTGGGAGTGGAGTCCACCCTGGAAAGGAGC
CAGATCACCAGGGATCACATGGGCCAATTACTCTATCAGCTGGTACAGTCAGAAAAACTCAGCAAACAGGA
CTTTTTCAAAGGTTTTTCAGAAACTTTGGAATTGGCAGATGACATGGCCATTGATATTCCCCATATTTGGT
TGTACCTTGCTGAACTGGTGACCCCCATGTTAAAAGAAGGTGGAATCTCCATGAGAGAACTTACCATAGAA
TTTAGCAAACCTTTACTTCCTGTTGGAAGAGCTGGGGTCTTGCTATCTGAAATATTGCACCTACTATGCAA
ACAAATGAGCCATAAGAAAGTGGGAGCCTTATGGAGGGAGGCTGACCTCAGCTGGAAGGACTTTTTACCAG
AAGGAGAAGATGTACATAATTTTCTTTTGGAGCAGAAGTTGGACTTCATAGAGTCTGACAGTCCCTGTTCC
TCTGAAGCACTTTCAAAGAAAGAACTGTCTGCCGAAGAGCTGTATAAGCGACTCGAGAAACTCATTATTGA
GGACAAAGCGAATGATGAACAGATCTTTGACTGGGTAGAGGCTAATCTAGACGAAATCCAGATGAGTTCAC
CTACATTCCTTAGAGCTTTAATGACTGCTGTTTGTAAAGCAGCTATTATAGCCGACTCTTCTACCTTCAGA
GTGGACACTGCTGTTATCAAGCAGAGAGTGCCGATCTTACTCAAGTACCTAGACTCAGATACAGAGAAGGA
ACTGCAAGCACTTTATGCACTACAAGCATCGATAGTAAAACTTGATCAACCTGCCAATTTGCTGCGGATGT
TTTTTGATTGTCTATATGACGAGGAGGTGATCTCCGAGGATGCCTTCTACAAATGGGAGAGCAGCAAGGAC
CCTGCAGAGCAGAATGGGAAGGGCGTGGCTCTGAAATCTGTCACGGCATTCTTCACGTGGCTGCGGGAAGC
AGAAGAGGAGTCTGAGGATAACTAAAACTTCAAATACACAAAATGAAACAAAAGAAACAATTTAAGTATTT
TTTTAAAAAGTTTCACGTCTTCGCCAATCACAGTGCAGCAAGGCCAATTCTCGCAGAAACCCCCACGTGTG
CACGAGTGGGAGAGGGGAAAGAGAAAAAAAGGTGATCATGGAGGAAAAAGGTACTGGATAAAAGTAAACTT
CAAACCTTAGGGCGGGAGCACTAAAACCAAAATACATGTATTATTTATAGAAAATATTTTCTGTTTTAATC
TTTTCTTTTTAAACAAGGACTCATACTTAAAAAAATGTTTAGCAAAAAAAAAAAAAAGTTGAGAACTTTTAA
TTTATTTTAAGGACTGCAAATGCCAGTGTAATTTTTTAATTTGCAGTTTCTGTAAACAACTTGTATAATAG
AAAAGCAGAGAAATAAATTTCCCTCCCCTTCAAGATGCACCTCATGTTTGTTTTAAGGTATAGCATTTAGT
CCAGATTTGAGAAAGTTTGGGGTGAACAAGGTAAGAAAGATTTTTTTTTTTTTGGCATCAAATCTTTCTGC
CTGCCTCTCAGCTTGCTTCAGAAAATTTAAAAAATCACAATAGTAATCAAAACATACATAACATTGAAACA
GAAGGAAATGCTGTGGACCACAGAACTCCAAGAATTGTTTAAAAAAAAAAAAGTGCTACCCTGAGAAAAGT
ACTCTTAATACTCTTGAAATCTTTAGAGCAACTTTAAGGCTTGTAAATACATAGAACAAATATTTAAAAAA

ACAAAAAGAAATTGACTCAGTACTATTTCTTTTCACTTTGAAAATATAAAGAACAAAATAAAGACAAACAT
TGCAAGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

IHRHRLSPATEKGAASSGEGDKDPPPPAHEDIAVKGTATAAGTGPGTGAAAAAAAAVPPPHPAAAAAAAAA
AAAAAAAAVPPPHPNIRALQTQAPQQIPRGPVQQPLEDRIFTPAVSAVYSTVTQVARQPGTPTPSPYSAHE
INKGHPNLAATPPGHASSPGLSQTPYPSGQNAGPTTLVYPQTPQTMNSQPQTRSPFFQRPQIQPPRATIPN
SSPSIRPGAQTPTAVYQANQHIMMVNHLPMPYPVPQGPQYCIPQYRHSGPPYVGPPQQYPVQPPGPGPFYP
GPGPGDFPNAYGTPFYPSQPVYQSAPIIVPTQQQPPPAKREKKTIRIRDPNQGGKDITEEIMSGGGSRNPT
PPIGRPTSTPTPPQQLPSQVPEHSPVVYGTVESAHLAASTPVTAASDQKQEEKPKPDPVLKSPSPVLRLVL
SGEKKEQEGQTSETTAIVSIAELPLPPSPTTVSSVARSTIAAPTSSALSSQPIFTTAIDDRCELSSPREDT
IPIPSLTSCTETSDPLPTNENDDDICKKPCSVAPNDIPLVSSTNLINEINGVSEKLSATESIVEIVKQEVL
PLTLELEILENPPEEMKLECIPAPITPSTVPSFPPTPPTPPASPPHTPVIVPAAATTVSSPSAAITVQRVL
EEDESIRTCLSEDAKEIQNKIEVEADGQTEEILDSQNLNSRRSPVPAQIAITVPKTWKKPKDRTRTTEEML
EAELELKAEEELSIDKVLESEQDKMSQGFHPERDPSDLKKVKAVEENGEEAEPVRNGAESVSEGEGIDANS
GSTDSSGDGVTFPFKPESWKPTDTEGKKQYDREFLLDFQFMPACIQKPEGLPPISDVVLDKINQPKLPMRT
LDPRILPRGPDFTPAFADFGRQTPGGRGVPLLNVGSRRSQPGQRREPRKIITVSVKEDVHLKKAENAWKPS
QKRDSQADDPENIKTQELFRKVRSILNKLTPQMFNQLMKQVSGLTVDTEERLKGVIDLVFEKAIDEPSFSV
AYANMCRCLVTLKVPMADKPGNTVNFRKLLLNRCQKEFEKDKADDDVFEKKQKELEAASAPEERTRLHDEL
EEAKDKARRRSIGNIKFIGELFKLKMLTEAIMHDCVVKLLKNHDEESLECLCRLLTTIGKDLDFEKAKPRM
DQYFNQMEKIVKERKTSSRIRFMLQDVIDLRLCNWVSRRADQGPKTIEQIHKEAKIEEQEEQRKVQQLMTK
EKRRPGVQRVDEGGWNTVQGAKNSRVLDPSKFLKITKPTIDEKIQLVPKAQLGSWGKGSSGGAKASETDAL
RSSASSLNRFSALQPPAPSGSTPSTPVEFDSRRTLTSRGSMGREKNDKPLPSATARPNTFMRGGSSKDLLD
NQSQEEQRREMLETVKQLTGGVDVERNSTEAERNKTRESAKPEISAMSAHDKAALSEEELERKSKSIIDEF
LHINDFKEAMQCVEELNAQGLLHVFVRVGVESTLERSQITRDHMGQLLYQLVQSEKLSKQDFFKGFSETLE
LADDMAIDIPHIWLYLAELVTPMLKEGGISMRELTIEFSKPLLPVGRAGVLLSEILHLLCKQMSHKKVGAL
WREADLSWKDFLPEGEDVHNFLLEQKLDFIESDSPCSSEALSKKELSAEELYKRLEKLIIEDKANDEQIFD
WVEANLDEIQMSSPTFLRALMTAVCKAAIIADSSTFRVDTAVIKQRVPILLKYLDSDTEKELQALYALQAS
IVKLDQPANLLRMFFDCLYDEEVISEDAFYKWESSKDPAEQNGKGVALKSVTAFFTWLREAEEESEDN

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
GAGAGTGGCGCGGCGTGGGGCGCCGCGAGGGCACTACGAGCCCAGGAACTGCCGCGCCGGCCCGATTTTCC
GCGGCTGTATAAGGACTAGCGCTTCGGTAGCCGGGGAGCTGGAGGGAACCTGATTTGAGATCAATTCTTGG
AGAACACAAAGAAGTGTGATGGATTGAACCAGCTGCCCTTTGACCCTGCTAACAACAACGAGCCCCTGCAG
TTTGGTAGTGCCAGTGGCCCTCTGGTCACAGAAGGCCTCATTGAGAATGGAGGGGAATCAAGCAAGAAAAG
AAAGAGAACAAATACTCCTTCAGCCGATAACAGTAGAACTCTGAACGTGGATTCCACTGCAATGACACTAC
CTATGTCTGATCCAACTGCATGGGCCACAGCAATGAATAATCTTGGAATGGCACCGCTGGGAATTGCCGGA
CAACCAATTTTACCTGACTTTGATCCTGCTCTTGGAATGATGACTGGAATTCCACCAATAACTCCAATGAT
GCCTGGTTTGGGAATAGTACCTCCACCAATTCCTCCAGATATGCCAGTAGTAAAAGAGATCATACACTGTA
AAAGCTGCACGCTCTTCCCTCCAAATCCAAATCTCCCACCTCCTGCAACCCGAGAAAGACCACCAGGATGC
AAAACAGTATTTGTGGGTGGTCTGCCTGAAAATGGGACAGAGCAAATCATTGTGGAAGTTTTCGAGCAGTG
TGGAGAGATCATTGCCATTCGCAAGAGCAAGAAGAACTTCTGCCACATTCGCTTTGCTGAGGAGTACATGG
TGGACAAAGCCCTGTATCTGTCTGGTTACCGCATTCGCCTGGGCTCTAGTACTGACAAGAAGGACACAGGC

AGACTCCACGTTGATTTCGCACAGGCTCGAGATGACCTGTATGAGTGGGAGTGTAAACAGCGTATGCTAGC
CAGAGAGGAGCGCCATCGTAGAAGAATGGAAGAAGAAAGATTGCGTCCACCATCTCCACCCCCAGTGGTCC
ACTATTCAGATCATGAATGCAGCATTGTTGCTGAAAAATTAAAAGATGATTCCAAATTCTCAGAAGCTGTA
CAGACCTTGCTTACCTGGATAGAGCGAGGAGAGGTCAACCGTCGTAGCGCCAATAACTTCTACTCCATGAT
CCAGTCGGCCAACAGCCATGTCCGCCGCCTGGTGAACGAGAAAGCTGCCCATGAGAAAGATATGGAAGAAG
CAAAGGAGAAGTTCAAGCAGGCCCTTTCTGGAATTCTCATTCAATTTGAGCAGATAGTGGCTGTGTACCAT
TCCGCCTCCAAGCAGAAGGCATGGGACCACTTCACAAAAGCCCAGCGGAAGAACATCAGCGTGTGGTGCAA
ACAAGCTGAGGAAATTCGCAACATTCATAATGATGAATTAATGGGAATCAGGCGAGAAGAAGAAATGGAAA
TGTCTGATGATGAAATAGAAGAAATGACAGAAACAAAAGAAACTGAGGAATCAGCCTTAGTATCACAGGCA
GAAGCTCTGAAGGAAGAAAATGACAGCCTCCGTTGGCAGCTCGATGCCTACCGGAATGAAGTAGAACTGCT
CAAGCAAGAACAAGGCAAAGTCCACAGAGAAGATGACCCTAACAAAGAACAGCAGCTGAAACTCCTGCAAC
AAGCCCTGCAAGGAATGCAACAGCATCTACTCAAAGTCCAAGAGGAATACAAAAAGAAAGAAGCTGAACTT
GAAAAACTCAAAGATGACAAGTTACAGGTGGAAAAAATGTTGGAAAATCTTAAAGAAAAGGAAAGCTGTGC
TTCTAGGCTGTGTGCCTCAAACCAGGATAGCGAATACCCTCTTGAGAAGACCATGAACAGCAGTCCTATCA
AATCTGAACGTGAAGCACTGCTAGTGGGGATTATCTCCACATTCCTTCATGTTCACCCATTTGGAGCAAGC
ATTGAATACATCTGTTCCTACTTGCACCGTCTTGATAATAAGATCTGCACCAGCGATGTGGAGTGTCTCAT
GGGTAGACTCCAGCATACCTTCAAGCAGGAAATGACTGGAGTTGGAGCCAGCCTGGAAAAGAGATGGAAAT
TCTGTGGCTTCGAGGGCTTGAAGCTGACCTAAATCTCTTTGCCTAACAACTTGGGATCCTGAAGATAAATA
TGTGTTGGACAAGCATAGAAAGTGATTTATATTTTTAATGGTTTTCAAGTGGAAGTTCCTTTGAATTTGTC
AGTTCATTCCTGGAAAATCTTTTGAGTTAAAATAAGGATCCTAGGACAGCACCTCGAACTACAGGCCCTAA
AGAGAAATTGCCTCAAACCACAAGTGCTGTAACTTCCTCCCCTTTCTGTCAATTGGTTGTCTTTAAATATT
GCAAAAGTCCTGATGCTAAACAGTATTTGGAGTGTTTTCAGTGTCTGTACTACTGTTGTAGACCTTGGTAT
TTTTTTAAACACTGTTAACTGAAATGTTTTGATGATTTGTATGTGATTTGTGTTTCTAAACTTCTCTTTAC
ATTAATGTTGTTACTGGTGAAAGGCATGAGAGCAGCACTAAGTCCTCTGTGTAACTGCCATTGTCTTTCCA
ATCCCCAGTAGACCAGTAAATAAATAACACATCAGTGTCTTCTAGAAGGTGCCTGACCAGGTTCACCTTTT
AAACGACAAAGCATGGTTTGTGGCTTTTTGCAAAATTACTATGAACCAAAAGTTGACAAATGTTCCAAAGT
TATTTTCTCTAACATATCACATTAAAGATCTGTTTCAGAATTGTAAAAAGTACATCTAGATGTGTTTACAG
AAAGCAAGTATCCAGTATGACTGGCATGTGTTCATGCTATTCAGAATCACTTGTAAATAGTCTGCTTTTAA
AGGAGGGCATGTTCAGTTTTCTGTGAATTAAAATATGCTCATGTGTGGGCACACACGCACAAACACACACA
CGCACGCACACAGTGGCAGAAGGGATTTATATTAATATTCTTTCCCCTCTGGCCTTCTTACAGTCTGTTGG
TCCCTTTGCTTCTGTTGTCAGTGTGTTGAATTGCAAACCGAGTACTGCTGTAAATACTATGTTTACTTCAT
GCTGAATGTTTGCAAAGACTTGATATAAGTATTAATAGTAATGAATCAATGAATAAATAATGAGCTAGGGT
TTGTGAGGCTTTCTACAAATAGGTCAGCTCCACCTGGAGTGCGAATTGCCAGAGACACCTTGGTAGTGCCC
ATCGGCAAATCGCAATGGCAGCATGTGAGTGGACCATTCAGAAACTTCTGCTTGGTGGAAAGTAAACAGAG
AGGATGGAGGTTTGGGGCGAATGTCCTGAGGCAGAGATGGTCTTTATTGTGTGTGGTGGTGGTTGTGGTAT
TTATAATAATGCAAGCATACCCTCCCTTGAGTCTCAATTGAAGATAAAAGAATGTACTGAGCAAGCAAAGC
CAATGGAGAGTATTTCACAAAAATACTTTGTAAATGAGATGCCAGTAGTGTTCAAAGTTGTATTTTTAAAA
GATAAATATTCCTTTTTATACCTCAGTTTTGTGTCCTGTTTTTAATGACTTACGCTCTAAGTAATCCATTA
GTAGTTATCTCAGTTCCTTCCTTTGGGTTACTAGAATGTTGGAAAAAGATGCCAAGTCTGTCTTGACAACT
GGAAACAGGGTTCCACAGCAGCCCATTCGTGCTGAAAACTGGCTTCCCCCTGAAGCACCCTGCTGTGGCAC
CAGCAGGAAGCTCAGGTTAATTTTACACTAGCTTGCTCACTGATGCATCTCTCATCAATGCTACGGAAGGC
TTTGATTCATCAGTCTCGGGCTCTTGGAATACCTAATTTTAATAATATCTATGAAATCAAGGGAAACTTTC
CATTTACAGTTATTTCTTGTTTAAATAAACTAAATTAATTTTTAGGGGAGAGCAGTAGGAAAAAGAGCTAA
TGCATGCGGGGTTTAATACCTAGGTGATGGGTTGAGGTGCAGCAAAACCACCATGGCACACGTTCACCTAT
GTAACAAACCTGCACATCCTGCACATGTACCCCGGAACTTACTTAAAA

LNQLPFDPANNNEPLQFGSASGPLVTEGLIENGGESSKKRKRTNTPSADNSRTLNVDSTAMTLPMSDPTAW
ATAMNNLGMAPLGIAGQPILPDFDPALGMMTGIPPITPMMPGLGIVPPPIPPDMPVVKEIIHCKSCTLFPP
NPNLPPPATRERPPGCKTVFVGGLPENGTEQIIVEVFEQCGEIIAIRKSKKNFCHIRFAEEYMVDKALYLS
GYRIRLGSSTDKKDTGRLHVDFAQARDDLYEWECKQRMLAREERHRRRMEEERLRPPSPPPVVHYSDHECS
IVAEKLKDDSKFSEAVQTLLTWIERGEVNRRSANNFYSMIQSANSHVRRLVNEKAAHEKDMEEAKEKFKQA
LSGILIQFEQIVAVYHSASKQKAWDHFTKAQRKNISVWCKQAEEIRNIHNDELMGIRREEEMEMSDDEIEE
MTETKETEESALVSQAEALKEENDSLRWQLDAYRNEVELLKQEQGKVHREDDPNKEQQLKLLQQALQGMQQ
HLLKVQEEYKKKEAELEKLKDDKLQVEKMLENLKEKESCASRLCASNQDSEYPLEKTMNSSPIKSEREALL
VGIISTFLHVHPFGASIEYICSYLHRLDNKICTSDVECLMGRLQHTFKQEMTGVGASLEKRWKFCGFEGLK
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

CAGCCACCCAGCCCCGAAGCAGGGCCCAGCGAGACAGGGCCCAGCCCCAAGCGTGACCCAGCTGCAGAGAA
GCTGAGACGGTGATTCGGACTGACTGGACTCTGGCCTCCAGGTGTCTGCACCTCACACTGGACCCACCTGG
AGACTCCACGGGCTCCCGGCAGCTCCGGCTGCTTGGGTGGATCCGGGTGCCCGCAGGGGCTGTGGGAGGGC
CCCGGGGCCTGCACTGCTCCCCCGATGGCCTGCTCTTCCTCACGGCCGGGGCTGCACCCTGTGTCCACGTG
CTAGATCTGGAGGGACGCCCCATCTGCCTCCTGCCCTGCCGCACTCCGGGGAGCGGGGCCTTCGTTCCAGA
GGACGTGGCTGTGACAGCGTCAGGGCTTGTGGTGGTCAGCGATCCCATCCATGGGGCTGTCCATGCACTCC
AGCACACAGCCCGGGACCCCGGGGGCCACTGGGTGACAGTGGGCACCTTCCTGTCTCCCCGAGGGCTGGCT
GTGGATGCCCTCAACCGCCTCCTGGTGACGGACTACTTGCCTGGGGCTGTGCACAGCTTCTCGTTGGGTCC
TGCTTGGGAGCCCCTGGCCCCAGCCTCCATGCTGGGTCTGGAGGGCCCCTGCTGGGTGGGCCCAGGGCCTG
ATGGGGGCCTTGCTGTGAGTGAGGAGTTTGGGGATGTGAGGCTGTTTGGCAGTGCCCGCCAACCCCTGGGC
TCCCTGGGGGGCTGGACGGGGCACACTTTCGGCTGCCCAGCGGGCATCTGCTCCAACTCAGAGGGCAATGT
TATTGTGGCAGACGAGCAGAGGCGCCAGGTGACCCTGTTTCCCCGGGCTGGGCCACCCATCTGCCTGGTGT
CAGAGGGGCTTGGGCAGCCCTTGGGAGTGGCCTGTGCACCCCAGGGCCAGCTCCTGGTGGCTGATGCCAAG
GACAACTCCATCAAGGTGTACCAGGGCCTCAAGGAGCTGGCCTGACCTGAGGCTGGGTTGGAGCAGCCCTC
CTGTGCCTGAGGCCAGCTCCCAGGCCCTTGGATCACCGCGGGAGGAACCCTCAGGATGGGTGGAGCCTCCA
GGCTATGGGCATTGCCTGCCTGATGCCAGCACCACCTGGGCTGGGCCCTGGGCTTGGCTCGAGTTCTCCTG
CTGGTGAGGCTCCGGATCTCAGGAGCAGCCCTGAGTCTGCTTCCCAGGCTGCCCCTGCCAGGCCTGCAGCC
TCCCCAGCCAGGGCTGCTCTCTGCTGTCCCCATTCAGTGCCCTGGCCCCTGCATTCATGCCCCCCACACCC
CCTCAGGCCCTGTGCCTGGACTTTGGGGCTGGCAGCTGAAGCCTTGAGATCCTGGGCCAGCTGCCGGCACA
CAGCTAGGCAGACTCTCCCACCAGGTGCCCCTGCCCAGGCCTCCTAATCGGGGGCAGACAGGCAGGGAGGG
TGTGGCTGGGCTGGGCTGGGCGGGGCGGCCTGGGGCAGGGGTGTGGCCCCTAAATGTCCCCAACCTCAGAG
GGACCTAGAGTCCTGAGCCTCCAGTAGCTTCTCTGGGCCTGGCAGAGGTAAGGGGGAGGCAACCCTGGAGT
GTCTGGAGGCCCATGGCTGGCTGAACCCTGGATGCCTTTTCTTCCGCGTCCCCATGAATGAAAGCTGTCTG
GGCCTTCATTCTGCAGACAGGGACAAACAGCTCCATGCTGTTTGTCCTCCCAGTGCAGCCGTGCTGGGAGG
GTCTGGGGGAGCTTCCTACAAGGAGAGACTCCTGCTGCTTTGGAAAACTGAGAAAAAATAGGGGTCTAACC
CTCTCCTCCCATTTTACAAGTGGGGAAATGAGGCGTGAAAGGAGAGGCGTCTGGGTTACTCCGTGGGTCTG
GGGTCCAGGGAAGGGCCTGTATGGGGGAGGGAGCTGGGAGGGGACGGTGTCTGGCTCTACCCCTGTGGGGT
GGGGAGGTGGGGCTCCCCTGTATCACAGGACATCCCCCCTGAGAGGTCCCTCATATGTCTGGGTCCTGTGG
GTGGGGGACTAACTGCGCAATGTAGTTAGGTGCTCAATAAACGGAGTTGCCGCTGAAAAAAAAAAAAAAAA
AAAAAA

LDPPGDSTGSRQLRLLGWIRVPAGAVGGPRGLHCSPDGLLFLTAGAAPCVHVLDLEGRPICLLPCRTPGSG
AFVPEDVAVTASGLVVVSDPIHGAVHALQHTARDPGGHWVTVGTFLSPRGLAVDALNRLLVTDYLPGAVHS
FSLGPAWEPLAPASMLGLEGPCWVGPGPDGGLAVSEEFGDVRLFGSARQPLGSLGGWTGHTFGCPAGICSN
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

GCGGCCGCCGGGCGCCGGAGCCGCGTCCCCCTGCGCCGCCCCGGGCCTGCCGGCCGGAGGAGCCACCATAC
ACCCCGGGTCGCCGAGCGCCTGGCCGCCCCGCGCCCGCGCCGCGCTCCGCCTGTGGCTGGGCTGCGTCTGC
TTCGCGCTGGTGCAGGCGGACAGTCCCTCAGCCCCAGTGAACGTCACCGTCAGGCACCTCAAGGCCAACTC
TGCAGTGGTGAGCTGGGATGTTCTGGAGGATGAGGTTGTCATCGGATTTGCCATCTCCCAGCAGAAGAAGG
ATGTGCGGATGCTGCGCTTCATCCAGGAGGTGAACACCACCACCCGCTCATGTGCCCTCTGGGACCTGGAG
GAGGATACGGAGTACATAGTCCACGTGCAGGCCATCTCCATTCAGGGCCAGAGCCCAGCCAGCGAGCCTGT
GCTCTTCAAGACCCCGCGTGAGGCTGAGAAGATGGCCTCCAAGAACAAAGATGAGGTAACCATGAAAGAGA
TGGGGAGGAACCAACAGCTGCGGACAGGCGAGGTGCTGATCATCGTCGTGGTCCTGTTCATGTGGGCAGGT
GTCATTGCCCTCTTCTGCCGCCAGTATGACATCATCAAGGACAATGAACCCAATAACAACAAGGAAAAAAC
CAAGAGTGCATCAGAAACCAGCACACCAGAGCACCAGGGCGGGGGGCTTCTCCGCAGCAAGATATGAAAAC
CTTTTCAGTGCTTGCCCTCAGCAGCTAAGAAGACAGACAGTAGAGAATGTGAGAGGATCTCATGGTTCTGA
TGATGATTATCCAACAAACATCTGGCCCTCTCTACATCTCTTCCCTCCATCTCCTTGTACCCTCTGGCTTA
CTGTCCTCTCTCTGGCGCACTTTCCTGAAGCCTCTTATTAACTCCCATCTCCAGAAGCACCTCAACAATGT
CAGTGGCTGAGGCTGCACTCAGAGGGATGACTGCTGGGGGTAGACTGGGTGCCAGGGGCCATGGGCCCAGG
ACCCAGTCTTGGCCATTCAGTTGAGTGAGGCTGAAGGCTGGGTTTGAAAAGGCAAAAAGACAAGACATCCA
GGCAGGGCTTCTCTTTTCTTTCCACAAGGGACAAGAGCTTGGCTTATTTAGGCTACAGCCCTGCTGCTGCT
CCCTTCCTTCTTCTGTCCTGTTTCCAGCCTTGCAAGAAGGCTATTCACAATTAGGCCTGGCTGCTTCTCTC
ATTTTTTTCCTCCCAGTTTCCAAACAAGCCCTCAGTGAACATCATTGAAGCGTGACTGCCTGTCTGCAGGG
AGAAGGATTCCATTTTTCTTCTCAGCTGGTCCCCAGGCCCACGGGCACAGGGAGAGGGACAACTGCAGCAG
TGGGGAGGAGGCACAGCTAGCTGCACAGTTCTCTCTTCTCCTTGTCCTAGTCAGATGAAGGAGGCTGCACT
ACAAACCCAAATTCTGCAAAAAAATAAAAATAAGCCACAAAACTAAAAGGCCTGGCCCCATTCTGGAAAAG
GCAAAGCTGCATGAGACACAGCCTTCTGCCTCCTCGCCTCTCCTGGACTGGCTTCCTCTTTGAGAAAATGC
ACAAAGCCCTGGGAGATGACAAGCACAAGGACTGACTCAAGCTGTGTCTTTCAGACCAAGGAACATCAGAG
AAGCTGTGGGGCTGCCTGCCAGGCAGGATCATGGCTGCCATCAAGCCTTTTCTGGATCCAGCCATCAAGGA
CATGTTTGTGGTGTGATGCACACTTTTGCAAGCGTGTAAGATGTTACCTGGTTTGTCTCTTTTGGAAAACA
AAAATCAGAAGGCTGCATTCTAGAGGGCAGAGAAATTCCCCCGAAGACTGAGCTGGTTGCCTGCATCCTCT
ATCTTCTTTGACCCTTATGACTGAAAGATCATCAGTTTGGAAGGTACTGGTCCAATTTATTTAGGAAGTAT
CTCTTGGAGTTTCAGAAATGCTAGCTTGGACAACTGAAAAGTCACATCACAGCTGGCATTCTGGGGGCTAC
CAAAACACCCCTTCTGGAGTAGAAGCTGCTGGAAGGCAGGCCTGAGCCATTCACCACGGACAGGAAGAGCA
GCTCTGGCTATCACCACTGGCCTCTGGGGTCTTCATATCTTGCCATCTCATCCAGGGTTCCATGAAAGTTA
CCCAGGGTCCTCATGTCCTTCCTTAGAGCCTGAGTGGTGTGAGGTGACAGGTCTCTCTCTCCACTGCCCCT
TTCTGGTTTAAAAAAATGGTGCTTGATGAGGGAAGGTAGACTCTTCCCTAGGACTGACGAGTTACGGCTGC
CAGATGCCTGCATGGGAAGAGGTGGACATCTGCATCTTCCATTGGTGGTCAAGGATGGGTGTGGGAGAACC
ACACCTAGTGCAAGCCTGGTACTCAGTAAATATTTGTTGAAATGAATGATAAGAGCATTGGTCCCCAAGCC
AGAGAGCCAGAAGCCATCACCCAATGACCGCCCCTTCCTTCCGGTCTACAAGAGCTCTCAAGGCTGGGTCT
GCCACCACTCTGCTTTGCCCAAGTGTGACAGCACTGGGGAGGAGAGACAGGATAAAGGGCAGATGTCAGCA
ATACTAAGGGCTTCCTCATGGGAGGGCATGAGGCTCCACTCATTGTCTTGTGACTTCCATCCCTGCTGAAT
GGGGCTGCAAGGCCAAGGCTCCTTAGGGGAGAGGTCCTTACCTCTGATCCACTTAGAGCAATAACCACTTT
TTAAATGTAAAATAAAAAGACAAATGAAAAGGCAAAAAAAAAAAAAAAAA

IHPGSPSAWPPRARAALRLWLGCVCFALVQADSPSAPVNVTVRHLKANSAVVSWDVLEDEVVIGFAISQQK
KDVRMLRFIQEVNTTTRSCALWDLEEDTEYIVHVQAISIQGQSPASEPVLFKTPREAEKMASKNKDEVTMK
EMGRNQQLRTGEVLIIVVVLFMWAGVIALFCRQYDIIKDNEPNNNKEKTKSASETSTPEHQGGGLLRSKI

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

```
GCGAGGCAGGAGCCGGCGGCTGGGCTCCGCAGCGCAGCCAGCGCAGCGCGGCGCCCCGGGCCCGGCCCCAT
GCCCGCAGCCCCGCCGGACCGTCCTTGAGCGCGGGCGCCTAGCCCGCGCCCCCTGCCCGCCGGCACCATTG
CCCCGACGGCGCGGCCGGGCGGCCCGGCGCTCCCCAGGCTCCGCGCGGGCCGAAAGACGCTGCTAGCGGCC
GCCGCGGGTGTGGTGATGCTGCTGGTGCTGGTGGTGCTCATCCCCGTGCTGGTGAGCTCGGGCGGCCCGGA
AGGCCACTATGAGATGCTGGGCACCTGCCGCATGGTGTGCGACCCCTACCCCGCGCGGGGCCCCGGCGCCG
GCGCGCGGACCGACGGCGGCGACGCCCTGAGCGAGCAGAGCGGCGCGCCCCCGCCTTCCACGCTGGTGCAG
GGCCCCCAGGGGAAGCCGGGCCGCACCGGCAAGCCCGGCCCTCCGGGGCCTCCCGGGGACCCAGGTCCTCC
CGGCCCTGTGGGGCCGCCGGGGGAGAAGGGTGAGCCAGGCAAGCCGGGCCCTCCGGGGCTGCCGGGCGCGG
GGGGCAGCGGCGCCATCAGCACTGCCACCTACACCACGGTGCCGCGCGTGGCCTTCTACGCCGGCCTCAAG
AACCCCCACGAGGGTTACGAGGTACTCAAGTTTGACGACGTGGTCACCAACCTAGGCAACAACTACGACGC
GGCCAGCGGCAAGTTTACGTGCAACATTCCCGGCACCTACTTTTTCACCTACCATGTCCTCATGCGCGGCG
GCGACGGCACCAGTATGTGGGCAGACCTCTGCAAGAATGGCCAGGTGCGGGCCAGTGCTATTGCCCAGGAC
GCGGACCAGAACTACGACTACGCCAGCAACAGCGTGATCCTGCACCTGGACGCCGGCGACGAGGTCTTCAT
CAAGCTGGATGGAGGCAAAGCACACGGCGGCAACAGCAACAAATACAGCACGTTCTCTGGCTTCATCATCT
ACTCCGACTGAGCTCCCCACGTCTCCCTCCACCCACGTCCCTCACCCGCCGGGGTCCCCTCCGGGCGGGGC
AGACGATGACTCGCCCCTCGCCCACCCGCTCGCTGCCCGGCCCTCCCCGGCTATGACGCCCCCGGCCCGTG
CTCAACACCGCCTGGGCCACAGCTAGGCCCTCCCACCGGCTCGCTGCAGAGCCGGGCCCAGCGCGCCCTGT
CCCCGTGCCAGGGAACCGGGGTTGACCGCCCCCGCCCAGCCCGCGCTATATATTTGTACAATAGGACTGTT
TACTGCCCACCTCCGCCTGCCAGCCCACCCCAGCCTGGGGAGAGGTCGCGGCGGCGGGTTTGCTTCCTGCG
CTCTGAGATGAGCTGCCCTCGGCTCCCTCCGGGGTGGCGCGCCCGGGGGAGGGGGGAGTTGGGGGCTGGAT
AGCTTCCCAGCACCCTCAGAGCCCCCGCCCGGCTGTGCCCCGTCTGACCAAAGTTATAATAAAAACATTTT
CACCCCGCAAAAAAAAAAAAAAAAAA
```

```
IAPTARPGGPALPRLRAGRKTLLAAAAGVVMLLVLVVLIPVLVSSGGPEGHYEMLGTCRMVCDPYPARGPG
AGARTDGGDALSEQSGAPPPSTLVQGPQGKPGRTGKPGPPGPPGDPGPPGPVGPPGEKGEPGKPGPPGLPG
AGGSGAISTATYTTVPRVAFYAGLKNPHEGYEVLKFDDVVTNLGNNYDAASGKFTCNIPGTYFFTYHVLMR
GGDGTSMWADLCKNGQVRASAIAQDADQNYDYASNSVILHLDAGDEVFIKLDGGKAHGGNSNKYSTFSGFI
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

```
AGTGTTATGATGCAGTTCCACAACACACAGCCACATTCACCCACAGACCGAGGTACAGAACGAGAGACAAC
CTCTGCCCCCCCCAGCAGCTGGCCAGCTTTGCAGCCCCAGTCTTGAGCCCCCAACTACCCTCCCCCCACCCA
CCCCCATCCCCTTCCCAATTGAAGGAGCGGAAAGAGAAGAGAGAAGAGTGAGCAGAGAGATTGAGAGATTG
AGAGAGAGAGAGAGAGATAGACGGAGATCTCTGGAGCAGACCTCAAGGTGACTTCTATTTCTATCTGGTTC
TCGTCTGGGGGGGCCCTGGCCGGGCAGCCCCCCAACACTTCTCCTGCCCTGAAACACGGCTCTAGCCAACC
TGCTCCGCTGCTTCACCTGCGACCGTCTCTGCGGGGGCTGCACGGCGCCAGCCCCTCCAGCCCACCAGGGC
ATTGTCCTCCAGCCCGTCATGCCCAGCTGTGACCCCGGTCCGGGCCCTGCCTGCCTCCCCACCAAGACTTT
```

CCGCAGCTATCTGCCCCGCTGTCACCGCACTTACAGCTGTGTCCACTGCCGTGCACACCTGGCCAAACACG
ATGAGCTTATTTCCAAGTCCTTCCAAGGGAGCCATGGCCGAGCCTACCTGTTTAACTCCGTGGTCAACGTG
GGTTGCGGGCCAGCTGAACAGCGCCTCTTGCTCACGGGGCTCCACTCGGTAGCTGACATTTTCTGTGAGAG
CTGCAAAACCACACTGGGCTGGAAATATGAGCAAGCTTTTGAGACGAGCCAGAAGTACAAGGAAGGGAAAT
ACATCATTGAAATGTCACACATGGTGAAGGACAACGGCTGGGACTGAGGGGCTCAGGCAGGGTGTGCCCTT
CCTCCGCATGCCCCTCCCTCCCCACGGCCCTGCCAAGCAGTCTATACCAGCATGAGTACTGCCCCACCCCT
GGGGGAAACCTGGCTCCAACCAACCCCTCCCCTGCCTCCACCATATCCACTACCAGGCACCCTTTAGAACA
GGGGTCTGGGGGTACCCCAGGGGTGTTAAGGCTCAGGAGTGGGCAGCAGTCAGGGAGAGACAGAACTGGGG
GAAAGGGATGGTTGTGGGTCTTTCTGTTCCCAAGATCCTGAACATGGAAGCGATGGCAGGGCATAGACTCA
GGCAGAGAGGATTGTGGGAGGAATCCGTTTTTGCTCCACCTCTTTTTGAGTGAACAGAGGACAAACCTTGG
GTCACAGGGCAAGTAGATCATGGACCACAGAACAGCAGATGAGAAAAGACTTGGGTTGGAGTGAAATTCTG
GTCTCAGACACCAGGAGACCAGAGTCTCTGAGGATGAAGTTTCCTACCCCTATTTGTAGGGAAAAGGACTT
GAGTGCAGGGAAAACTCAAATCCCAGGCCCTGGGAAATAGTAAAATAATCAAAGGGTTTTCCATTTCACTC
CACTTGTTAGTTTATCTTGGCACTGAAGAGGCACTTTCGAGTATCTAACTTTTGCCATTGGGTGGGGTGGG
GACAGCTGCTCGCGGAACAGCCCCTAGTCGGCTGCTTCCAGAGTAAGCAGTCTTTATGGGCTTTCTCTGAG
GCCCAGTCACTGCTCCTGGGACCCAGTCCCCTGGAGGGGAGGTGGAAAATCAGTGCTACGGGGCCAGTCTT
TCCCGTGGCTGCCACCAGCGAATGAAACTTTTGTATGATACATAAAGTGCTTGAGTCTATTTTTAATAAAA
AGGGAAAAAGCAACTTGAAAAAAAAAA

TALANLLRCFTCDRLCGGCTAPAPPAHQGIVLQPVMPSCDPGPGPACLPTKTFRSYLPRCHRTYSCVHCRA
HLAKHDELISKSFQGSHGRAYLFNSVVNVGCGPAEQRLLLTGLHSVADIFCESCKTTLGWKYEQAFETSQK
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

GCTGATGACATGAGGGCTCCGTCTCCCGAGTGATGGCAGCGCGCGCTGCTTCGCCGCCTCCGCCGCTCAGC
CCCGGACTCCTTACGTCAGGGTAGCGGGGTCCCCCCTCCGCGCGGGAGCCAGCGAACAGCGAGAGAGCACA
GCAGAGCGCGCCGCGGAGCCGGGGCGCCCTCACTGCGCTAGGAGCCCCACTAAACCCAGCGGAGCGGAGCC
TGGCGGGAGGCAGCGCCGCGGAGCCAGCGCCGGACGCCGCAAGCAGACTCCCCGGCCAGCGCAAGCACTCC
CTGGCCGGCGCCGGACCCTCGGGGCGCCGGATTCGTCTTGTGTGCCCCACGTCATGCGCGCGGGCGTCCGC
GGCTCTCCCGACCAGCCCCAGCGCGGTGGGCCCAGGCTCAGAGGTCCCGGCGCCTCTCGCTGAAGTAGTTG
GGTAGCCGGGGCTGGGGGTCGCCACGTCGGGGGCGCGGCCAGGACCCGCGGAGCCGGTCCCCGAGCGCGGG
GAGCGGGGCCGCCCGCGCCGCCCCACCATTACCTCCCCGGGCGGCAAGGAGGAGCTGGTGGCGGTCGCCTC
CCGGCTGTGGCAGCGGCGGCGGCGTGCCTGCCTGGCGGCCGTCGGCGTACTCTTGGCCATGGCGCTCGGGC
TGCTCATCGCCGTGCCGCTGCTGCTGCAGGCGGCGCCCCGAGGCGCCGCGCACTATGAGATGATGGGCACC
TGCCGCATGATCTGCGACCCTTACACTGCCGCGCCCGGCGGGGAGCCCCCGGGTGCAAAGGCGCAGCCACC
CGGACCCAGCACCGCCGCCCTGGAAGTCATGCAGGACCTCAGCGCCAACCCTCCTCCTCCTTTCATCCAGG
GACCCAAGGGCGACCCGGGGCGACCGGGCAAGCCAGGGCCGCGGGGGCCCCCTGGAGAGCCGGGCCCGCCT
GGACCCAGGGGCCCTCCGGGAGAGAAGGGCGACTCGGGGCGGCCCGGGCTGCCAGGGCTGCAACTGACGGC
GGGCACGGCCAGCGGCGTCGGGGTGGTGGGCGGCGGGGCCGGGGTAGGTGGCGATTCCGAGGGTGAAGTGA
CCAGTGCGCTGAGCGCCACCTTCAGCGGCCCCAAGATCGCCTTCTATGTGGGTCTCAAGAGCCCCCACGAA
GGCTATGAGGTGCTGAAGTTCGATGACGTGGTCACCAACCTCGGCAATCACTATGACCCCACCACGGGCAA

GTTCAGCTGCCAGGTACGCGGCATCTACTTCTTCACCTACCACATCCTCATGCGCGGCGGCGACGGCACCA
GCATGTGGGCGGACCTCTGCAAGAACGGGCAGGTCCGGGCCAGCGCCATTGCACAGGACGCCGACCAGAAC
TACGACTACGCCAGTAACAGCGTGGTGCTGCACTTGGATTCAGGGGACGAAGTGTATGTGAAGCTGGATGG
CGGGAAGGCTCACGGAGGCAATAATAACAAGTACAGCACGTTCTCGGGCTTTCTTCTGTACCCGGATTAGG
GGCGCGGGGGGTGCGAGGCGGGGTGGCTGCAGGCCGCCCGGTCTCCGCCCGGGCGCGGCTCCTTGGCAAAG
GCCACTCTCGATTCATAACACTTCCTGACATCTCCTTTGGAAAAGACAAATCCCTGCGTCCTCCCTGCCCC
GCTCCTGGCCTCAGTGCGTCTGCGACCCACCACGCTCAGGGCTGTGCTCCTGGTCTCCATCCCCATCCCGG
CAAGGGAGGAAGGGACGCCCGAGCCCTTGAGGCGGCGGCACAGACTTTGCAAACCTGATTAGACTGGACAG
GCAGGGCCGGGAGCCTGCCCTCCTCAGACAGCCTCCTCCCAGTGCCTAGAAGCGGAGGGCTCCGGGCCCTG
GCCAGGGAGGTAGGCCAGAGGGAGCGCGGGCTTCCTGGGGCGTCCTTCTTTGTGACCCGAAATACTTGTGC
AGATTTCCCTGTCCATCAGCCAAAACCCCACCCACAGCAGAATTCCAGCAAACAGAAAATTCACCTCTCCA
CACCGCATTCCCTCCTGACTCAGACTCACCGCGATGCATTAAATTATGTTTTTAGACTATG

ITSPGGKEELVAVASRLWQRRRRACLAAVGVLLAMALGLLIAVPLLLQAAPRGAAHYEMMGTCRMICDPYT
AAPGGEPPGAKAQPPGPSTAALEVMQDLSANPPPPFIQGPKGDPGRPGKPGPRGPPGEPGPPGPRGPPGEK
GDSGRPGLPGLQLTAGTASGVGVVGGGAGVGGDSEGEVTSALSATFSGPKIAFYVGLKSPHEGYEVLKFDD
VVTNLGNHYDPTTGKFSCQVRGIYFFTYHILMRGGDGTSMWADLCKNGQVRASAIAQDADQNYDYASNSVV
LHLDSGDEVYVKLDGGKAHGGNNNKYSTFSGFLLYPD

AUG initiation = 29.5 kDa
ACG initiation = 33.0 kDa

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

#12

NM_000314

CUG initiation conserved in mammals

PTEN phosphatase and tensin homolog

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

CCTCCCCTCGCCCGGCGCGGTCCCGTCCGCCTCTCGCTCGCCTCCCGCCTCCCCTCGGTCTTCCGAGGCGC
CCGGGCTCCCGGCGCGGCGGCGGAGGGGGCGGGCAGGCCGGCGGGCGGTGATGTGGCGGGACTCTTTATGC
GCTGCGGCAGGATACGCGCTCGGCGCTGGGACGCGACTGCGCTCAGTTCTCTCCTCTCGGAAGCTGCAGCC
ATGATGGAAGTTTGAGAGTTGAGCCGCTGTGAGGCGAGGCCGGGCTCAGGCGAGGGAGATGAGAGACGGCG
GCGGCCGCGGCCCGGAGCCCCTCTCAGCGCCTGTGAGCAGCCGCGGGGGCAGCGCCCTCGGGGAGCCGGCC
GGCCTGCGGCGGCGGCAGCGGCGGCGTTTCTCGCCTCCTCTTCGTCTTTTCTAACCGTGCAGCCTCTTCCT
CGGCTTCTCCTGAAAGGGAAGGTGGAAGCCGTGGGCTCGGGCGGGAGCCGGCTGAGGCGCGGCGGCGGCGG
CGGCACCTCCCGCTCCTGGAGCGGGGGGGAGAAGCGGCGGCGGCGGCGGCCGCGGCGGCTGCAGCTCCAGG
GAGGGGGTCTGAGTCGCCTGTCACCATTTCCAGGGCTGGGAACGCCGGAGAGTTGGTCTCTCCCCTTCTAC
TGCCTCCAACACGGCGGCGGCGGCGGCGGCACATCCAGGGACCCGGGCCGGTTTTAAACCTCCCGTCCGCC
GCCGCCGCACCCCCCGTGGCCCGGGCTCCGGAGGCCGCCGGCGGAGGCAGCCGTTCGGAGGATTATTCGTC
TTCTCCCCATTCCGCTGCCGCCGCTGCCAGGCCTCTGGCTGCTGAGGAGAAGCAGGCCCAGTCGCTGCAAC
CATCCAGCAGCCGCCGCAGCAGCCATTACCCGGCTGCGGTCCAGAGCCAAGCGGCGGCAGAGCGAGGGGCA
TCAGCTACCGCCAAGTCCAGAGCCATTTCCATCCTGCAGAAGAAGCCCCGCCACCAGCAGCTTCTGCCATC
TCTCTCCTCCTTTTTCTTCAGCCACAGGCTCCCAGACATGACAGCCATCATCAAAGAGATCGTTAGCAGAA
ACAAAAGGAGATATCAAGAGGATGGATTCGACTTAGACTTGACCTATATTTATCCAAACATTATTGCTATG
GGATTTCCTGCAGAAAGACTTGAAGGCGTATACAGGAACAATATTGATGATGTAGTAAGGTTTTTGGATTC
AAAGCATAAAAACCATTACAAGATATACAATCTTTGTGCTGAAAGACATTATGACACCGCCAAATTTAATT
GCAGAGTTGCACAATATCCTTTTGAAGACCATAACCCACCACAGCTAGAACTTATCAAACCCTTTTGTGAA
GATCTTGACCAATGGCTAAGTGAAGATGACAATCATGTTGCAGCAATTCACTGTAAAGCTGGAAAGGGACG
AACTGGTGTAATGATATGTGCATATTTATTACATCGGGGCAAATTTTTAAAGGCACAAGAGGCCCTAGATT
TCTATGGGGAAGTAAGGACCAGAGACAAAAAGGGAGTAACTATTCCCAGTCAGAGGCGCTATGTGTATTAT
TATAGCTACCTGTTAAAGAATCATCTGGATTATAGACCAGTGGCACTGTTGTTTCACAAGATGATGTTTGA

AACTATTCCAATGTTCAGTGGCGGAACTTGCAATCCTCAGTTTGTGGTCTGCCAGCTAAAGGTGAAGATAT
ATTCCTCCAATTCAGGACCCACACGACGGGAAGACAAGTTCATGTACTTTGAGTTCCCTCAGCCGTTACCT
GTGTGTGGTGATATCAAAGTAGAGTTCTTCCACAAACAGAACAAGATGCTAAAAAAGGACAAAATGTTTCA
CTTTTGGGTAAATACATTCTTCATACCAGGACCAGAGGAAACCTCAGAAAAAGTAGAAAATGGAAGTCTAT
GTGATCAAGAAATCGATAGCATTTGCAGTATAGAGCGTGCAGATAATGACAAGGAATATCTAGTACTTACT
TTAACAAAAAATGATCTTGACAAAGCAAATAAAGACAAAGCCAACCGATACTTTTCTCCAAATTTTAAGGT
GAAGCTGTACTTCACAAAAACAGTAGAGGAGCCGTCAAATCCAGAGGCTAGCAGTTCAACTTCTGTAACAC
CAGATGTTAGTGACAATGAACCTGATCATTATAGATATTCTGACACCACTGACTCTGATCCAGAGAATGAA
CCTTTTGATGAAGATCAGCATACACAAATTACAAAAGTCTGAATTTTTTTTTATCAAGAGGGATAAAACAC
CATGAAAATAAACTTGAATAAACTGAAAATGGACCTTTTTTTTTTTAATGGCAATAGGACATTGTGTCAGA
TTACCAGTTATAGGAACAATTCTCTTTTCCTGACCAATCTTGTTTTACCCTATACATCCACAGGGTTTTGA
CACTTGTTGTCCAGTTGAAAAAAGGTTGTGTAGCTGTGTCATGTATATACCTTTTTGTGTCAAAAGGACAT
TTAAAATTCAATTAGGATTAATAAAGATGGCACTTTCCCGTTTTATTCCAGTTTTATAAAAAGTGGAGACA
GACTGATGTGTATACGTAGGAATTTTTTCCTTTTGTGTTCTGTCACCAACTGAAGTGGCTAAAGAGCTTTG
TGATATACTGGTTCACATCCTACCCCTTTGCACTTGTGGCAACAGATAAGTTTGCAGTTGGCTAAGAGAGG
TTTCCGAAGGGTTTTGCTACATTCTAATGCATGTATTCGGGTTAGGGGAATGGAGGGAATGCTCAGAAAGG
AAATAATTTTATGCTGGACTCTGGACCATATACCATCTCCAGCTATTTACACACACCTTTCTTTAGCATGC
TACAGTTATTAATCTGGACATTCGAGGAATTGGCCGCTGTCACTGCTTGTTGTTTGCGCATTTTTTTTTAA
AGCATATTGGTGCTAGAAAAGGCAGCTAAAGGAAGTGAATCTGTATTGGGGTACAGGAATGAACCTTCTGC
AACATCTTAAGATCCACAAATGAAGGGATATAAAAATAATGTCATAGGTAAGAAACACAGCAACAATGACT
TAACCATATAAATGTGGAGGCTATCAACAAAGAATGGGCTTGAAACATTATAAAAATTGACAATGATTTAT
TAAATATGTTTTCTCAATTGTAACGACTTCTCCATCTCCTGTGTAATCAAGGCCAGTGCTAAAATTCAGAT
GCTGTTAGTACCTACATCAGTCAACAACTTACACTTATTTTACTAGTTTTCAATCATAATACCTGCTGTGG
ATGCTTCATGTGCTGCCTGCAAGCTTCTTTTTTCTCATTAAATATAAAATATTTTGTAATGCTGCACAGAA
ATTTTCAATTTGAGATTCTACAGTAAGCGTTTTTTTTCTTTGAAGATTTATGATGCACTTATTCAATAGCT
GTCAGCCGTTCCACCCTTTTGACCTTACACATTCTATTACAATGAATTTTGCAGTTTTGCACATTTTTTAA
ATGTCATTAACTGTTAGGGAATTTTACTTGAATACTGAATACATATAATGTTTATATTAAAAAGGACATTT
GTGTTAAAAAGGAAATTAGAGTTGCAGTAAACTTTCAATGCTGCACACAAAAAAAAGACATTTGATTTTTC
AGTAGAAATTGTCCTACATGTGCTTTATTGATTTGCTATTGAAAGAATAGGGTTTTTTTTTTTTTTTTTT
TTTTTTTTTTAAATGTGCAGTGTTGAATCATTTCTTCATAGTGCTCCCCCGAGTTGGGACTAGGGCTTCAA
TTTCACTTCTTAAAAAAAATCATCATATATTTGATATGCCCAGACTGCATACGATTTTAAGCGGAGTACAA
CTACTATTGTAAAGCTAATGTGAAGATATTATTAAAAAGGTTTTTTTTTCCAGAAATTTGGTGTCTTCAAA
TTATACCTTCACCTTGACATTTGAATATCCAGCCATTTTGTTTCTTAATGGTATAAAATTCCATTTTCAAT
AACTTATTGGTGCTGAAATTGTTCACTAGCTGTGGTCTGACCTAGTTAATTTACAAATACAGATTGAATAG
GACCTACTAGAGCAGCATTTATAGAGTTTGATGGCAAATAGATTAGGCAGAACTTCATCTAAAATATTCTT
AGTAAATAATGTTGACACGTTTTCCATACCTTGTCAGTTTCATTCAACAATTTTTAAATTTTTAACAAAGC
TCTTAGGATTTACACATTTATATTTAAACATTGATATATAGAGTATTGATTGATTGCTCATAAGTTAAATT
GGTAAAGTTAGAGACAACTATTCTAACACCTCACCATTGAAATTTATATGCCACCTTGTCTTTCATAAAAG
CTGAAAATTGTTACCTAAAATGAAAATCAACTTCATGTTTTGAAGATAGTTATAAATATTGTTCTTTGTTA
CAATTTCGGGCACCGCATATTAAAACGTAACTTTATTGTTCCAATATGTAACATGGAGGGCCAGGTCATAA
ATAATGACATTATAATGGGCTTTTGCACTGTTATTATTTTTCCTTTGGAATGTGAAGGTCTGAATGAGGGT
TTTGATTTTGAATGTTTCAATGTTTTTGAGAAGCCTTGCTTACATTTTATGGTGTAGTCATTGGAAATGGA
AAAATGGCATTATATATATTATATATATAAATATATATTATACATACTCTCCTTACTTTATTTCAGTTACC
ATCCCCATAGAATTTGACAAGAATTGCTATGACTGAAAGGTTTTCGAGTCCTAATTAAAACTTTATTTATG
GCAGTATTCATAATTAGCCTGAAATGCATTCTGTAGGTAATCTCTGAGTTTCTGGAATATTTTCTTAGACT
TTTTGGATGTGCAGCAGCTTACATGTCTGAAGTTACTTGAAGGCATCACTTTTAAGAAAGCTTACAGTTGG
GCCCTGTACCATCCCAAGTCCTTTGTAGCTCCTCTTGAACATGTTTGCCATACTTTTAAAAGGGTAGTTGA
ATAAATAGCATCACCATTCTTTGCTGTGGCACAGGTTATAAACTTAAGTGGAGTTTACCGGCAGCATCAAA
TGTTTCAGCTTTAAAAAATAAAAGTAGGGTACAAGTTTAATGTTTAGTTCTAGAAATTTTGTGCAATATGT
TCATAACGATGGCTGTGGTTGCCACAAAGTGCCTCGTTTACCTTTAAATACTGTTAATGTGTCATGCATGC
AGATGGAAGGGGTGGAACTGTGCACTAAAGTGGGGGCTTTAACTGTAGTATTTGGCAGAGTTGCCTTCTAC
CTGCCAGTTCAAAAGTTCAACCTGTTTTCATATAGAATATATATACTAAAAAATTTCAGTCTGTTAAACAG
CCTTACTCTGATTCAGCCTCTTCAGATACTCTTGTGCTGTGCAGCAGTGGCTCTGTGTGTAAATGCTATGC
ACTGAGGATACACAAAAATACCAATATGATGTGTACAGGATAATGCCTCATCCCAATCAGATGTCCATTTG
TTATTGTGTTTGTTAACAACCCTTTATCTCTTAGTGTTATAAACTCCACTTAAAACTGATTAAAGTCTCAT
TCTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAA

LERGGEAAAAAAAAAAAPGRGSESPVTISRAGNAGELVSPLLLPPTRRRRRRHIQGPGPVLNLPSAAAAPP
VARAPEAAGGGSRSEDYSSSPHSAAAAARPLAAEEKQAQSLQPSSSRRSSHYPAAVQSQAAAERGASATAK
SRAISILQKKPRHQQLLPSLSSFFFSHRLPDMTAIIKEIVSRNKRRYQEDGFDLDLTYIYPNIIAMGFPAE
RLEGVYRNNIDDVVRFLDSKHKNHYKIYNLCAERHYDTAKFNCRVAQYPFEDHNPPQLELIKPFCEDLDQW
LSEDDNHVAAIHCKAGKGRTGVMICAYLLHRGKFLKAQEALDFYGEVRTRDKKGVTIPSQRRYVYYYSYLL
KNHLDYRPVALLFHKMMFETIPMFSGGTCNPQFVVCQLKVKIYSSNSGPTRREDKFMYFEFPQPLPVCGDI
KVEFFHKQNKMLKKDKMFHFWVNTFFIPGPEETSEKVENGSLCDQEIDSICSIERADNDKEYLVLTLTKND
LDKANKDKANRYFSPNFKVKLYFTKTVEEPSNPEASSSTSVTPDVSDNEPDHYRYSDTTDSDPENEPFDED
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

TTCGCCGGAGCGCGACCCGGGGACTCCCAGGCCTGTGGGCGGGCCCTGCCCAGGACTGGGCGGTGCCATAA
CCCCTAGTTTAAAAACTCGCGGGTACCGGACCCAAGATCGGGGACCCGGCGGCGGCTCCGCGGGGGAAACA
GCGAGGCTGGCGCAGCGCCGAGGCCGCGGCCCTGGGGGCCCGCAATCCACGCCACGGAATCCCCGAGTGAG
CAGGGGTGAGCGCAGCCACTGCCCAACGCAAACCGTGAAGAAGCTTCTGGAAGAGCAGAGGCGCCGCCAGC
AGCAGCAGCCCGACGCTGGCGGGGTGCAGGGACAATTTCTCCCTCCCCCAGAGCAGCCCCTGACCCCATCT
GTGAATGAGGCTGTGACTGGCCACCCTCCCTTCCCAGCACACTCGGAGACTGTGGGTTCTGGACCTAGCAG
CCTGGGCTTTCCAGACTGGGACCCCAACACGCATGCTGCCTACACTGACAGCCCCTACTCTTGCCCTGCTT
CTGCTGCCGAAAATTTCCTGCCTCCTGACTTCTACCCACCCTCGGACCCAGGGCAGCCGTGCCCATTTCCC
CAGGGCATGGAGGCTGGACCCTGGAGAGTTTCTGCACCCCCTTCAGGACCCCCACAGTTCCCCGCTGTGGT
CCCTGGACCATCGCTGGAGGTGGCCCGAGCTCACATGCTGGCTTTGGGGCCACAGCAGCTGCTGGCCCAGG
ATGAGGAGGGGGACACGCTCCTTCACCTGTTTGCGGCTCGGGGGCTGCGCTGGGCGGCATATGCTGCGGCT
GAGGTGCTCCAGGTGTACCGGCGTCTTGACATTCGTGAGCATAAGGGCAAGACCCCTCTCCTGGTGGCGGC
TGCTGCCAACCAGCCCCTGATTGTGGAGGATCTGTTGAACCTGGGAGCAGAGCCCAATGCCGCTGACCATC
AGGGACGTTCGGTCTTGCACGTGGCCGCTACCTACGGGCTCCCAGGAGTTCTCTTGGCTGTGCTTAACTCT
GGGGTCCAGGTTGACCTGGAAGCCAGAGACTTCGAGGGCCTCACCCCGCTCCACACGGCCATCCTGGCCCT
TAACGTTGCTATGCGCCCTTCCGACCTCTGTCCCCGGGTGCTGAGCACACAGGCCCGAGACAGGCTGGATT
GTGTCCACATGTTGCTGCAAATGGGTGCTAATCACACCAGCCAGGAGATCAAGAGCAACAAGACAGTTCTG
CACTTGGCCGTGCAGGCTGCCAACCCCACTCTGGTTCAGCTGCTGCTGGAGCTGCCCCGGGGAGACCTGCG
GACCTTTGTCAACATGAAGGCCCACGGGAACACAGCCCTCCACATGGCGGCTGCCCTGCCCCCTGGGCCGG
CCCAGGAGGCCATCGTGCGGCACCTGTTGGCAGCTGGGGCGGACCCCACACTGCGCAACCTGGAGAATGAG
CAGCCCGTTCACCTGCTGCGGCCCGGGCCGGGCCCTGAGGGGCTCCGGCAGCTGTTGAAGAGGAGCCGTGT
GGCGCCGCCAGGCCTGTCCTCTTAGGACTCAAACCCAGACCCTGGACTGATTTTCCAGTCCCCACCGTCCT
GCGGGACAGCCAGCGTATGCTAATGTTGCAAACCCATGATAATGTATGTGGAATATCCTGCCATTGGGGTT
TTACATTAAAACCCCAGAATGGCTGCAGAGGGGTGAACAGGCCCCAATATTTGGGGTGCTGTGATACCCCT
CTTCTACCCACAAGGAGCCCTCTTGATGATTTCTGTGAAATCGAGGCCCCTTGATTGTTTCTGTGAAACAC
CCTGCACCCCTAGTCCTTTCCCCACTGAGATCTTTCGGGTTCTCTCCCCTAACTCAGCT

LAQRRGRGPGGPQSTPRNPRVSRGERSHCPTQTVKKLLEEQRRRQQQQPDAGGVQGQFLPPPEQPLTPSVN
EAVTGHPPFPAHSETVGSGPSSLGFPDWDPNTHAAYTDSPYSCPASAAENFLPPDFYPPSDPGQPCPFPQG
MEAGPWRVSAPPSGPPQFPAVVPGPSLEVARAHMLALGPQQLLAQDEEGDTLLHLFAARGLRWAAYAAAEV
LQVYRRLDIREHKGKTPLLVAAAANQPLIVEDLLNLGAEPNAADHQGRSVLHVAATYGLPGVLLAVLNSGV
QVDLEARDFEGLTPLHTAILALNVAMRPSDLCPRVLSTQARDRLDCVHMLLQMGANHTSQEIKSNKTVLHL

AVQAANPTLVQLLLELPRGDLRTFVNMKAHGNTALHMAAALPPGPAQEAIVRHLLAAGADPTLRNLENEQP
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

AGCATTGTGGGAGCTCCTCGGTCGGTGCCGGTCGGTGGCTGCCTATTGCGGCCTGCGGTGATCAACGAGGC
CCGGGGAGCGCGTCCCCAGTCTGCGCGCCGGTCCTGCGGCAGCTGGCCCAAGACCCGGAGCCGAAAGGAAG
TGTTGGAGCCTGAGGTCGCTCCGCGCCGCTAGGAGGACGCTGTGCCTGGCCTGGGACCTCCGCTCCCGCCC
ACCGCCCTGGAGCCGCTGAGGGACGTCCACGTGGGCCTGTCCCCGCCGAGCCGCGGCCCTGTCCGCCTGGC
GCTGCTCTCGGGCCACTACCTCTACTACCACTACGGCTGCGACGGCCTGGACGACCGCGGCTGGGGCTGCG
GCTACCGCACTCTGCAGACGCTGTGCTCGTGGCCAGAGGGCCAGCCCGCGGGCGTACCTGGACTGGCCGCC
GTACAGGCGGCCCTGGAGGACATGGGCGACAAGCCCCCCGGCTTCCGGGGCTCCCGGGACTGGATCGGCTG
CGTGGAGGCCAGCCTCTGCCTCGCTCACTTCGGAGGGCCCCAGGGACGCCTCTGCCACGTACCCCGGGGAG
TGGGGCTGCACGGGGAGCTGGAGAGGCTTTACTCGCACTTCGCAGGGGGTGGGGGCCCAGTCATGGTTGGG
GGGGACGCAGATGCCAGGTCCAAGGCCTTGCTGGGAGTCTGCGTAGGGTCAGGCACGGAAGCCTATGTCCT
GGTATTGGACCCTCACTACTGGGGCACTCCAAAAAGCCCCAGTGAACTACAGGCTGCTGGGTGGGTGGGCT
GGCAAGAGGTGAGTGCAGCCTTTGACCCCAACTCCTTCTACAACCTGTGCTTGACCAGCCTTAGCTCCCAA
CAGCAGCAGCGCACCTTGGACTGAGGACGAAGTTACAGAACTGAGATTCTCGGGTCCCAGACACGCACCTA
TGTACCTCCCACTGGTGTCCCTGCAAAGCCTGGCGCTTTTGACATCAATAATAAAAGTGGCAGGGCTGAGC
AAAAAAAAAAAAAAAAAAAAA

LEPLRDVHVGLSPPSRGPVRLALLSGHYLYYHYGCDGLDDRGWGCGYRTLQTLCSWPEGQPAGVPGLAAVQ
AALEDMGDKPPGFRGSRDWIGCVEASLCLAHFGGPQGRLCHVPRGVGLHGELERLYSHFAGGGGPVMVGGD
ADARSKALLGVCVGSGTEAYVLVLDPHYWGTPKSPSELQAAGWVGWQEVSAAFDPNSFYNLCLTSLSSQQQ
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

GTGACTCCTTGATCTTCTTCCTGCGTGGAGAGCCTTCGCGGGTGAGGCTTAACGCGCAGGAGGTCTCACGA
GAGTGGAAGCAACTCTCGCGAATTTTAAAATTTATCTTTTTGCCTAGCGACTGACAACAGGCTGGTTGCTT
GGCGTGGAATCCTAAAGTGGCCTGGCTTTGAGACTGGAGTGAGACCCCAGCCCTAGGCTGGGGTTCTTTCC
ATTATAGAGGAGACGGATTCAGAAGGGCTACAGACCAAGGTTGTTGAAAACCAAACATATGATGAGCGTCT
AGAGATTAACGACTCCGAAGAGGTTGCAAGTATTTATACTCCAACCCCAAGACACCAAGGACTTCCTCGTT
CTGCCCATCTTCCTAACAAGGCTATGGCTGATAACAGCAGTGATGAGTATGAAGAGGAAAATAGCAAGGTC
CTAAGGGAGGGCATGCCACAGGCTCCAGGCCACAGAGGCAAAGACATGGACCCTGTTCCACCTGCCCCTGC
AAGTCTCAAGTGCCACCAGACCCCTTCTCATGTACTTGAGAGGGTGGGATGGTACAGAGAGAGCCAGAAGC
ACAGAAAGGAGAAGAAGGAGACCTCACAGTTGACACCTCAACGGGGCTTTAGTGAAAATGAGGATGACGAT
GATGATGATGATGATTCATCTGAAACTGATTCTGATTCTGATGATGATGATGAAGAGCATGGAGCCCCTCT
GGAAGGGGCCTATGACCCTGCAGACTATGAGCATTTGCCAGTTTCTGCTGAAATTAAGGAACTCTTCCAGT
ACATCAGTAGGTACACACCTCAGTTGATTGACCTGGACCACAAACTGAAGCCTTTCATTCCTGATTTTATC
CCAGCTGTCGGGGATATTGATGCATTCTTAAAGGTCCCACGTCCTGATGGAAAGCCTGACAACCTTGGCCT
ATTGGTATTGGATGAACCTTCTACAAAGCAGTCAGACCCTACGGTGCTCTCACTCTGGTTAACAGAGAATT
CTAAGCAGCACAACATCACACAACATATGAAAGTAAAAAGCCTAGAAGATGCAGAAAAGAATCCCAAAGCC
ATTGACACGTGGATTGAGAGCATCTCTGAATTACACCGTTCTAAGCCCCCTGCGACTGTGCACTACACCAG
GCCCATGCCCGACATTGACACGCTGATGCAGGAATGGTCCCCGGAGTTTGAAGAGCTTTTGGGCAAGGTAA
GCCTGCCCACGGCAGAGATTGATTGCAGCCTGGCAGAGTACATTGACATGATCTGTGCCATTCTAGACATC
CCTGTCTACAAGAGTCGGATCCAGTCCCTCCATCTGCTCTTTTCCCTCTACTCAGAATTCAAGAACTCACA
GCATTTTAAAGCTCTCGCTGAAGGCAAGAAAGCATTCACTCCTTCATCCAATTCCACCTCCCAAGCTGGAG
ACATGGAGACATTAACCTTCAGCTGAGACACTTCCCAAGCTGCTGTTTCAAGGCTGAGCTGGCCCCTCTGC
CCCAGCTGAGATGGACAGATCGTTGTCAGCTACTTGATGTCCTTGCCCATGCCACAGCTTGGCTCAGGGGC
AGTGCATGTCCTGCTGCCCTCTCTGCCAGAGGGCACAGAACATGTTTGTTTAATGAACCTGCCTGCCTCAG
ATTGCTGTCCCCGGGGAGTTAATGCATCTACACCACTGTGGGGATTTGAGTTATAAGAATTGGAATTTCTG
AGATCCCATGGAGGTTAGATTGGGAGGAAAGCTTAAAAGATGTCCTTTTTGTGAGAGGGATGGAATTGTTT
TCTTTCATTCGTAAAGTTAGTGAGTAAAGATTTTATAAATCAAAAAAAAAAAAAAAAAA

IEETDSEGLQTKVVENQTYDERLEINDSEEVASIYTPTPRHQGLPRSAHLPNKAMADNSSDEYEEENSKVL
REGMPQAPGHRGKDMDPVPPAPASLKCHQTPSHVLERVGWYRESQKHRKEKKETSQLTPQRGFSENEDDDD
DDDDSSETDSDSDDDDEEHGAPLEGAYDPADYEHLPVSAEIKELFQYISRYTPQLIDLDHKLKPFIPDFIP
AVGDIDAFLKVPRPDGKPDNLGLLVLDEPSTKQSDPTVLSLWLTENSKQHNITQHMKVKSLEDAEKNPKAI
DTWIESISELHRSKPPATVHYTRPMPDIDTLMQEWSPEFEELLGKVSLPTAEIDCSLAEYIDMICAILDIP
VYKSRIQSLHLLFSLYSEFKNSQHFKALAEGKKAFTPSSNSTSQAGDMETLTFS

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

GCCGCGGCGGTGGCGGAGACTGTGGCTTTAAGAGCGTGCCGGGAGCCCGAGCCCCAGCCGGGCCGCGCTTC
GCCGCTGCGCACCCCAGCGGAGCCAAGCCCCACGCTGGCCGGACAGGGCCGCCTGTCGCCGGGCTGCTGAG
AACTAGCCCTAGACCTCTGCGTGAGGGTTCTTCTGCCGAAGACATCACCAGTGTGTGGAGCCTGCCACACC
CACCCGCTGCCAAACCACGGCCTTTACCTGTGTCTTCCGGTGTTTCCCGTGCGACCCATCCTGTGGGAGTG
CCTCGTGGGCTGCCCCAGAGTTCACCCCACACTCAGCAGCACCAATGGTGAAGATGACAAGATCGAAGACT
TTCCAGGCATATCTGCCCTCCTGCCACCGGACCTACAGCTGCATTCACTGCAGAGCTCACTTGGCCAATCA
TGATGAACTAATTTCCAAGTCATTCCAAGGAAGTCAAGGACGAGCATACCTCTTTAACTCAGTAGTTAATG
TGGGCTGTGGGCCTGCAGAAGAGCGAGTGTTGCTAACAGGACTGCATGCAGTCGCAGACATTTACTGTGAA
AACTGCAAAACCACTCTGGGCTGGAAATACGAACATGCTTTTGAAAGCAGCCAGAAATATAAAGAAGGCAA

ATACATCATTGAACTAGCACACATGATCAAGGACAATGGCTGGGACTGATTGGACAGCATCTACCCAACCC
AGTGTCCACGTGAACGCCATTCAACCGAACATTCTTCCCAAGCGTGAGAGAGTGACTGACACTTGGTTCCA
TCCATTTAGGGGCCTTGCCATCCGGGGCATCCTCCCACCCTGACGCCATCTTTCTGGTGACCGGCCTCTAA
ATCGCTGTCTCTCTGTCTCTTTGCTTTGTATCTGTTTGTGAGTTGATCCTGGCTTCTCTCTCTGTTCTAGT
TTTGGCTGAAAACAAAACAACAAAAGGAACAGATCCTTGACCGCATGGCGGCAGCCCACCTTGGTAAGGGC
CCCAGGGCCCATGCGAGAGCTGCCTGATGGCCTCTTGTCAGGAGAGCAGTGGCACGGGGGCGTGAGGAAGA
GGGAAAGGGGAAACTCTAAGGGTCCTGGCGCGGGGAAGGGGTGGAAGGGTGGAGGTAGGAACAAAATTGCG
CCGCTCCTGGAGACCTGATAACTTAGGCTTGAAATAATTGACTTGTCTAAAAGGACAAAGAGAAAAAAAAA
ATACCTCATGACTGCATTCTCTCTGACTAGAAGCTTCTGTTCCTGACACCAAATGTGCCAGGTTAGCAAAT
GAGCACAAGATGTGGCCCTGATTCTAGTTGGTGGGGCAAGGGCCTGGTTCTCCTGGGCTGAGTGGGGGAGT
GTCCTGGCAGCAGCGAGTGACCTGGGCAGTGGCCAGGTGGGTGCGATGACTCTGATGCCTCACTCAGTCTC
TGGGCAATCATCATCTTTGCCTCTAGCCACCGTAGATAAGGTGTGAAGGGACTGCTGTTTGCAATGGGCTT
ACCATCCAAATATCCCAAAGGCTTTGACCAGCAACCAAGTAAAATCAGTAATTGAGGAGAGCAGGGCACAA
AGGGGCTGCAGTTTGGGAGCTCCTGAAGAAATGGCTCAGATATTGAGTCAGAGAAATAAAAAGTAGGATCA
GTTAGCAATTCTAACTGCCCTTCCTTCTGACCCCTCATAAGAGGAGTGTGGTGAGGGAGGGGACTGGGTAG
GGGTCATCCCAGGAGGAGGGGTTTACATTGGAACCAGTTCAGGTTCGGTGCATCTTTCCTCTTCGGTTTTA
CAGTGGCTTCCGTGGGATCGTCAATTTCTTGTTCTTAGAGTTTCGGGTGTTTTTCTCCAGTCTTGTTACTG
TAGACTGTAGAAAGCACGGGCCCCAGGCTCTGAGCTTAGTAATAACCTGGCTGGTAGATTCCTCATGCCCC
TAATTGTCCCACTTAGGCCTGAATGTCTTGCATGGAGAGAAATCTCCTGTCAGTGTGGTCCAGCAGCAGGG
AGGAGTTCTGCCCAAATTCCGATATCACCCCTTCCCCCATCCAAGCATCCTTCGATTAGGGAAGTGGAGAG
CACATCCCTGTAAGGCCCATAAGAGAAAGAGGAGTTTGTTACATTTAATCAACACTGTGAAGTCTGTTCTA
CAGCAATTCAGCCATTACACAGTATATGACTGAAACTCATTTAACTGGGTTAATTTCATTTCTTAGACTGA
ATATATTATTGTTAAGATACGTGTGCGTGTTAGGTAATTCTCAGCATCTCCTCCAAGTAGGCCGACCTTCT
CGGAAAATTCACCCTAAAAGTCTCACAAAAGAATGAGTTCATGGGGAGATTCTGTAAAGTGATGAACTGAG
ATGAAAGCAGCCAACAGCCCAGGAGCTTTTCAGAATAGCGTCTGCAGCAGAACCAGTTTCCATTCAGAGCG
CGTCCTTGGTGGAAATGCTTTTTTGTGTGTCTCCACGCGCTGATGGTGGAATGGGAGCCCCAAGACGTGTG
GGCTTAGAAATCAACTTTTGTTCCCCAAGGCTTCTTGTCCAGATCTTTCCAGTGCTTTCATAGCCCTGGGA
GATCAAGTTGTTCTCCCCACTTTACTGCAAGGTAGACTGAAGTTCAGAAGAAATACTGAATTTCTGCTCCC
AGAAGAATAGTTTCTCTGGCTCACAGGCCCAAGTTCTCAATGAAATCGTTTTTTAACTTTCACATTCCTAA
GCTGGCTTCCCGGCACAGAAGCCATGGATTTCCCCTCTCTCCCTTCCCCCTCCTCAAGGAAATAGTCTTCC
TTTATGGATTTTCATTGGACTCTTTCCTCAGCGATTGTCCTGGCTGTTTATTGATAGTCCTTCCCATAAGA
AAATGGGGTTAAACATGGGGTAGGTATTTTGTCTTTCAAACTACAAATGGAATGTGGTGACATAAACTAGA
CATGGGGTGCCCTCAAGTTTCCAAGGGGACCAATGTGCCACTGTTCTTCCTTGGGGATGAGGCCTTTGACT
GTTGGATGGATCAGAGCAGGCTCCAGTCAGACCCTGGTTCTGAATGTTTTTTTTTTCGGTGACTATCCAGT
GAGCCTTCAGTGGGTGCAAGGCGCCATACTTGCTGTGAGAGAGCTGAGTAGAGTGTTGGTTTTTCCATAAC
TACAGGGGGAAAAAAAGTCATTAGGCTTTCCCTTTGTGTCAGTGAAACCAAAAGTGCTTCTTACAACGTTC
GCTCTGTTCATGGGTTGTCTATCTAACATTGAGCAGCATTGGAGAGGCCACAGCTGAGCTATGGAGATGCT
AAATTAACTCATGGCCTCAGTCAGTTCATTCTTTAATTTCCTCACCAAATTATTGACTTAGAGCATAACCA
AAGACCTCATTCATTCACCCCAGGTGGGTTGGGGTAATTGGAGTTTGTTGGTGAAGTTTGGGGGCGGGGTG
TTGGGAGTAGAGACAGGGTAAGGGGACGTGAGAAAGGAAAAGGCATGAAGTTCTATACCTCAGCCAGCAGC
TGCCTTCGTTTGGAACTGAAGTCCAGCCAGCAGACTCTCTAGCTCCATCTCCCCTGTGCCACCCTAGGTCA
TATGACCTTGGCCACCTTGGAGTAGACCCAGACCCCTCGGGACCCGGGACATTAGTCTCAGGCTGCTGATG
GATTGATTTGACATGAACCAAACACAGCCAAACTCGATACCCACAAGCTGTCAGCTGAACCTGACTGAGTG
TTCTTCCTGAGTTCACGAGGATAGGCTAGAGTGCATTTTTACTGGTGGATCAGTGTGTGCGAAAGAGATGA
CCCTTTATAAAGAGATTTTCAAGTGGATATATATAAAAGAAACAGTTGCTTGTAAAATATACTTTTGTAAA
TAATATTTAATTTTTTAAATAATATATTTGGTGCTGTTTTCTCAGATCCCCTGAGAGCACTTTTTATTTTC
CTTTTAAATTCTATGGTTTCCTTTGCATTTCTTGAAGTATATTTTAAGGGAAACAGTGATCACCAATACAT
GTTTTCAGTTTTTTTTTTTTTTAAGGTCTCTATCACTTTAATCTGGATCAAGGCTTTGAAGCAATGCCTCT
CTGCATTTTTTCCCCAGTGGAACAGACTCTGCAGTACATTAATCAGGTTGAGAATTGAAATATTTTCTTGC
ATCAGTATTGGCTAGAAAAGAAAATAAATAAAACCAAGTTAATTTAGTAGTAACAACTTACAGTGATTCTT
CCTGTTGGAAGAATTTCCAACAAATCAGAATCACGTTTTTAGTTGTGCGTGTGCGCGCACACGTGTGTAAA
AAGCACTTTCGATTGTGCCTCCTGTTTTCTCGAGTGGGGACACTTTAACTACAGTTTACACCTCGGGCGCA
TAAAGTTTTTCTTCTCTTTCTCTCTGGTTGTTTCTGTTTCTGAGTGGACCAACAGCAGAACCCACGAGGAT
TTGTTTTGAGTATGGAGCTGTTGCGGGTTTGCTCCTTTTTCTTGCTTTGCGTGCTCAGTTTTTACAGACTG
TAAAGGAGATGTGTTGTTTGTGAAGATGGAGCAGAGTCAAATCTGTGCTTCTAACTGAGATGAGAGTGTAT
TAATCACGTATCGCAGGGCTCCAGCTGTTTTAGAAGCCACATCATGTTAAACATTAACTGGTTTGGATTAA
AAGAACATTAATATTATAATACACATATCTTAGTGGTAAACAGCTTTTTTTTTTTAAGGTCAGATTGCCTC

AGGTTTAGAAAGAGGCTGAGAAATCAAATCTTGAACACAATCAACTTACATATTTTAAAGGAATCTGCCTC
AAATGAGAAAATATGCTAGTTATCTAGATAGAGGAAAGAGATATTTACTTTTTTAAAAATTAAAATAGTTA
TGAAATCTGGCAGAAAAGGTAAAGCCTAGAAGAAACTATGAAAGCTATTCTCATGTTACCAAATTCTATCT
GCGCATATGTTTTTGTATAACATTTCGGTGACAGTGGGAGTCGGTTCCCTTTCCCAACCTGCAGAGACTAT
CTTCCAATACAGAATCTGTCTATTTATGCTTGTGTTTACAAACTGTATTTGTTGGGTTTGGGTTTTTGTTT
TCTTTGGTGGCATTTTTCAGGTCACTTTGCTTCTATAACAAAGGTAATTGTTTTCAAATAATTTGTCTTCA
CCTTTTCCTGTATTTGTACATAGTGATTCAGTATTAGAGAAAAGTGCATTGTTTCTGTCATATTTCCAATC
TGTGTTGGTGCTCATTTGAGAAAATAAAAGTTTTCAAATATTAACTCTTAAAAAAAAAA

TAFTCVFRCFPCDPSCGSASWAAPEFTPHSAAPMVKMTRSKTFQAYLPSCHRTYSCIHCRAHLANHDELIS
KSFQGSQGRAYLFNSVVNVGCGPAEERVLLTGLHAVADIYCENCKTTLGWKYEHAFESSQKYKEGKYIIEL
AHMIKDNGWD

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

GGATCCCGATCTGATTCCTGATCCTTGATTCCTTGATCCTTGGTCCCGCCATGGGAGCCTGAGCGCCCCCT
ATTCCCCCCTGGCCCCCAGCCCCCGGGGCCTTGAGGGGGAAGAGGCAGCGGTCTGGGACGGAGCAGGGGGT
GACCAGACTCAAGAACCCCCCCCTCAACATCCCCCATCGCGCGCGCTGCCTGTCCAGGAGCGCCGAGTTCG
GAGCGACCCGGAGCGCTGCGGATACAAAGGCGACGGGCCGAGCGGGGCGCCCGCGGAGCCCACCCGGCAGT
TCGCAGCGGCGGGAGCGTCGCTGGATTTTCTCTGAGACAAGCCCACCCGTCCAGCAAAATAGAGTCCCTCA
GGGTGACAGTTGACTTCCTGAAGGTGCCTCTTGGCCTAAAGAAGCCGGTGCTGAAGGAGGTGGCTGTGGGG
CCCCCCAAGAGGCCCCAGCCTGCGGCCCTGGAGCGCTACAAGGCGCGGCGTTCAGACGCCATGGACACCGA
GTCCCAGTACTCGGGCTATTCCTACAAGTCGGGCCACTCCCGCAGCTCCCGCAAGCACAGGGACCGCCGGG
ACCGACACCGCTCTAAGAGTCGAGATGGGGGCCGAGGGGACAAGTCGGTGACAATCCAGGCTCCCGGGGAG
CCCCTGCTGGACAATGAGTCCACACGAGGGGATGAGCGGGATGACAACTGGGGGGAAACGACGACAGTAGT
AACGGGCACCTCAGAGCACAGCATCTCCCATGATGACCTCACACGCATCGCCAAGGACATGGAGGACAGTG
TCCCTCTGGACTGCTCCCGTCACCTGGGTGTGGCAGCGGGGGCCACCCTGGCACTGCTGTCTTTCCTCACG
CCTCTGGCCTTCCTGCTGCTGCCCCCACTGCTGTGGCGGGAGGAGCTGGAGCCTTGCGGGACGGCCTGCGA
GGGCCTCTTCATCTCTGTCGCCTTCAAGCTGCTCATCCTGCTACTGGGCAGCTGGGCTCTGTTCTTCCGCC
GGCCCAAGGCCTCGCTGCCCCGCGTCTTTGTGCTGCGTGCCCTGCTTATGGTGCTGGTTTTCCTGCTCGTG
GTCTCCTACTGGCTCTTCTATGGTGTGCGCATCCTGGATGCTCGGGAGCGCAGCTACCAGGGCGTGGTGCA
GTTCGCCGTGTCGCTGGTGGACGCCCTTCTTTTCGTGCACTACCTGGCCGTGGTCCTGCTGGAGCTGCGCC
AGCTCCAGCCTCAGTTCACGCTCAAGGTCGTGCGCTCCACCGACGGCGCCAGCCGCTTCTACAACGTTGGC
CATCTCAGCATCCAGCGCGTGGCAGTGTGGATCCTGGAGAAGTATTACCATGACTTCCCTGTCTACAACCC
TGCCCTCCTCAACCTGCCCAAGTCCGTCCTGGCCAAGAAAGTGTCTGGCTTCAAGGTGTATTCCCTCGGAG
AGGAAAACAGCACCAACAACTCCACTGGCCAGTCTCGGGCTGTGATTGCAGCGGCAGCTCGGAGGCGGGAC
AACAGTCACAATGAGTACTACTATGAGGAGGCTGAGCATGAGCGAAGGGTGCGCAAGAGGAGGGCCAGGCT
TGTAGTGGCGGTGGAGGAGGCCTTCACTCACATTAAGCGGCTGCAGGAAGAGGAGCAGAAAAACCCCAGGG
AGGTGATGGACCCCCGGGAGGCAGCCCAAGCCATCTTTGCATCCATGGCCCGTGCCATGCAGAAGTACCTT
CGGACCACCAAGCAGCAGCCCTACCACACCATGGAGAGCATCCTGCAGCACCTTGAATTCTGCATCACGCA

TGACATGACGCCCAAGGCCTTCTTGGAGCGATACTTGGCGGCTGGACCTACCATCCAGTACCACAAGGAAC
GCTGGCTGGCCAAACAGTGGACATTGGTGAGCGAGGAGCCGGTGACCAACGGCCTCAAGGATGGCATCGTT
TTCCTCTTAAAACGCCAGGACTTCAGCCTGGTGGTCAGCACCAAGAAGGTCCCATTCTTCAAACTCTCCGA
GGAATTTGTGGATCCCAAGTCACACAAGTTTGTCATGAGGCTGCAGTCTGAGACCTCAGTGTGACTGTGCA
ACAGCAGGGGGAGTGGGAAACTCTGGGGGGTCCTGAGGGGGTGGGAGGGGGCTTGGTTCTCAGGCCCAGCC
ACATTCCTGCCACCCTTCTTCTTCTTGCTCTTTTTTTTTTACTTGAATTAACGCACCCCCACCTTCTCTCC
TCGCTTCTTCCTTATTTTACCCCATGTGAACCTGGAGAGACCATCCTGCTGTCAACAGTACCTGGGAAGGA
CTCCCACCTCACCAACAACTTTTGTATTACTCTAGGCCCTGCAGGAATCAGTGCCTCTCTCCCTCTTCTTT
CCCTAGTCTTTTCCCAGATTACAGTCTCTCCTGAAAGGGCACAGGGCCCTGCTGATTGTACTTTCCCCTCC
TGAGCCCCGACTCACAAATCCAAGTTCTTAAAACATTTCTCTTCAGTGGCCCAACAGGGTTTCTCTGGGGC
ACATGGACATGACTCCAGAGAGCCACAGTGCCAAACTCCTCCAGGGCAGCAACTGGCCCTCCTGTCCCTCA
CCCCAGCCACAACAAACCCTGGGTTCTAGGGCAGGGATACTCCTGCCACACAGCCCGAGTTAGAAATCTCC
TTGCTAGGAGCATTTGCTTCCACATATATTTAGAGCAAAGAAGGATCCCATCCTTTTCCCAGAAATCTCCA
CCTAATGTTTTTGGTTTGTATGGTCACGTGACCATAGGCAACCACGTGGAAACCCTCTGTGACCACTTTTC
CAGGGACTTAGGGGAAGGTACCTTTCTTCCAATGTGTCTTCCTAGGCAGCCCCTGAGGAGGAGGGCTGAAT
AGATCCCTGAGGTTTTGGAGAGACCCCCATCACTGACTCCTGCTCCCTAACCCTACCCTCACTTTCGTCCC
CGCTCTTCCCAGTGAAGGATGGTATGTAGACTCCTGTACAGACATAGTGGCTTGCAGACCCTGACCCAGCC
CCTGTGGTCTTAGACAAATGTTTTTATTTTTGTCACCAGCCACCCCTGTCCTGCCGCCTTCTCTCGACTCC
AGAGACCTGTTGCCTCATCTCTTTTGGGGAAGAGCCGGCAGCTCCTCCTCATCCCCTGCCTTAAGTCCAGT
TCTTTGCCTCAGGGGTCTCGTTTCCTTGGCCTTCCAGGGTCCCCACCCCTTTTCTCCCTGCCTGATTCTCT
GAGCTCTGGGCTCCGTCTGTATTGGGTTGAGGGGCAAGGATTACTGCCTTTTGTAGGTACTTCACCCCTCA
CCCCATTTTAGCTTCCATAGTCTTTGCACCAAATCCAAATTCTTGATAATTTAGATCTCATTTTGAGCAAA
ATTTGCTGGCCCTCTAATAAATATTTTCAATATAAATCTGAGCCTTTGACTCAGACATTTTTGCCAAGGAG
AGTAGAATTAGGAAGTACCCATATACATCCAGCCAGGATCCACATGGAGGACCTTTCTGATGGCTGCAATG
ACTAGGCCATTCCTCTGAGTAACTCACAGTGTCCTTTTGTAGGCCCTTCTTTTCCCTGAAAGACTGGTTGG
TACTTACCTTGCAGAGCACATCCTGGGATAAGATCCCCAGTGTCTCCCCTGGGAGGCTCCCCCTCTGTGTA
GCACCAGCCCTGGGAATGATGGAGCCTAGTGATCGGGGTTTCTCCTGCTGTCCTTTCTGCAAAAGTTCACT
TGTTTACCCACCGCATGCTAGAGAGGAGCTCATTGGCCAATGCTTACCTTGTCCCCAAAGGGGTGGGTTGT
GGAGCTCACTTAGGCAGGGCCTCTGGCTGGGGCCAGGGTTATGAGATAGGCCTGTATGAAATATGTCCTGT
TCTGGGGGTCTGTCTCTTTTCTTCTCTTCAAAAACTTTGTGTCAGAGAGTCCCTTCTGAGTCACATAAATA
CCTCACTATCCTGGAAAACAGGGCCTGGATGGTGACTGGGGTCATTGCCTTTGTGGACAGGATGGAGTGTG
GTGTGGTCTGAGGAGCAGGTTGGGGTGGGGGAGAGGGAAAGGATTTGGGATCTTAGTTGCTGCCCTAGGTT
AGGGGCTGGGGAGTGTTTATTTTAAGATCCTGCCATGTTTTTAATCACTGTGATTTTTTTTTCATTCCCCT
TTCCTAAAAAAAATTTTTTTCCTCCAACTCTCTAAGCACTAAGGGCTGTGCCTGAGAATGGTAGCATTTTG
GTCTTTTGCTTCAGAACTGTGGTATCTTTGTCTTTTTTCATTATTATTATTATTATTATTATTATTATTAC
TATTGTTTTTTAAAATGTCAGGATGAATTGTCAGACATATGGCCATGTGTTTGTCCTCTGCTTCTCCCCTG
TGGGAAGTTGTCTCCATGCTGTGAACTGCTGTGGGGTGTGCAGCTGACTCAGTCCCTCTGAGCAGTTTCCC
CACTGTGTCTGTCCCATCATGCGCTGGATCTGCTCATTCTCCTGCTGTGGGGGTATGCCCACCTCTTACCC
CCTTGACACCATAGGGCTGCTGTGGCTGGGCCTCACCAGCACTGTCTTTTGTGTGACTCATGGCATCCTCG
TTCATCCCCACCGTGCCTAGCAGGCCTTCCTTTTCACCACCTCGGAACGCTTGCCTTTCCTCCCTCCACAA
CAGGACGCTGTGCCTCAGTCCTTCACCTACCTCGCCACTCTGCCACTGTCCCCATTGGTCCTTTCTCCTAA
ACTGGTCTTTGTGCTCTCTTTGTTTTTTCTTATTTCCCTCTTGTCTCTCATTTTTTCTTCCCATTCCCCTC
CCATTTCAGCCCTTAACTTTTCTCTTTCCCATCTCCACTCAGTATTCCAATGGCAAACCCTGATGATGTAA
CACCTGCGATGAGACATCGGACTCTCCGGAACTTTCTCATCTGACACGTCTTTTTCCCAGGGTTCGTACTT
CTCCTCCATTGGTCCCAGGCTAACTCCCCTGTTCCTCTGTGGTGTCTGTCAGTCCGTCTGTCTTCTCTTTC
CTCTGCCCTTCCCACAGGGCAGTATCTGCTGATGGATTCAGTCCTGGTGTGTGATTGTTGTGATTTGTTCT
TCCGTGCGCAAAAGGAAGAGGGCTTTTTGAGTCCCTTCCAAGTGAGATTGTAAATGTAGAATTTTCCACTG
TTGGATCTAGATTTTTTTTCCTTTTTTTTGGGGGGGTGGGGTTACAGAGCTGAGACCTTGTGCATGCATGT
AGAAAATTGTAAAATGTAAATTTTTTTTAATATATAAAAAGCTTGTTTCTACAGTTTGCAGTGGATCTAAA
CATTACGGCAATTTTAGGATTTTTTTCTTAAACATAGGAACTAAAACTGTACAAATTTTTTTTATATAAAA
TAAAGACATTTGACTTTTGTGGGAAAAAAAAAAAAAAA
IESLRVTVDFLKVPLGLKKPVLKEVAVGPPKRPQPAALERYKARRSDAMDTESQYSGYSYKSGHSRSSRKH
RDRRDRHRSKSRDGGRGDKSVTIQAPGEPLLDNESTRGDERDDNWGETTTVVTGTSEHSISHDDLTRIAKD
MEDSVPLDCSRHLGVAAGATLALLSFLTPLAFLLLPPLLWREELEPCGTACEGLFISVAFKLLILLLGSWA
LFFRRPKASLPRVFVLRALLMVLVFLLVVSYWLFYGVRILDARERSYQGVVQFAVSLVDALLFVHYLAVVL
LELRQLQPQFTLKVVRSTDGASRFYNVGHLSIQRVAVWILEKYYHDFPVYNPALLNLPKSVLAKKVSGFKV
YSLGEENSTNNSTGQSRAVIAAAARRRDNSHNEYYYEEAEHERRVRKRRARLVVAVEEAFTHIKRLQEEEQ

KNPREVMDPREAAQAIFASMARAMQKYLRTTKQQPYHTMESILQHLEFCITHDMTPKAFLERYLAAGPTIQ
YHKERWLAKQWTLVSEEPVTNGLKDGIVFLLKRQDFSLVVSTKKVPFFKLSEEFVDPKSHKFVMRLQSETS
V

AUG initiation = 59.7 kDa
AUA initiation = 65.0 kDa

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

#18

NM_017457

CUG initiated extension conserved in eutherian mammals

CYTH2 cytohesin 2

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

ATACCTACACCGGCTTTTTGTACGACTGTTGGCCCTGGAGAACGATCCTTTGGTGGCGAGGGCGGGGAGGA
CGAAAGCGCCCACTGTGGATTGGACAGTGTCAAAAAGAGGGGCGGTCCCTACTGAAGGGGCGGTTGGGCGA
CGAAGGGAAGAGTCTTTTCAGCGCTGAGGACTGGCGCTGAGGAGGCGGCGGTGGCTCCCGGGGCGTTTGAG
CGGGCTCACCCGAGCCCGCGGGCCAACGCGGATCCAGGCCCGACTGGCGGGACCGCCCCGGATTCCCCGCG
GGCCTTCCTAGCCGCCATGGAGGACGGCGTCTATGAACCCCCAGACCTGACTCCGGAGGAGCGGATGGAGC
TGGAGAACATCCGGCGGCGGAAGCAGGAGCTGCTGGTGGAGATTCAGCGCCTGCGGGAGGAGCTCAGTGAA
GCCATGAGCGAGGTGGAGGGGCTGGAGGCCAATGAGGGCAGTAAGACCTTGCAACGGAACCGGAAGATGGC
AATGGGCAGGAAGAAGTTCAACATGGACCCCAAGAAGGGGATCCAGTTCTTGGTGGAGAATGAACTGCTGC
AGAACACACCCGAGGAGATCGCCCGCTTCCTGTACAAGGGCGAGGGGCTGAACAAGACAGCCATCGGGGAC
TACCTGGGGGAGAGGGAAGAACTGAACCTGGCAGTGCTCCATGCTTTTGTGGATCTGCATGAGTTCACCGA
CCTCAATCTGGTGCAGGCCCTCAGGCAGTTTCTATGGAGCTTTCGCCTACCCGGAGAGGCCCAGAAAATTG
ACCGGATGATGGAGGCCTTCGCCCAGCGATACTGCCTGTGCAACCCTGGGGTTTTCCAGTCCACAGACACG
TGCTATGTGCTGTCCTTCGCCGTCATCATGCTCAACACCAGTCTCCACAATCCCAATGTCCGGGACAAGCC
GGGCCTGGAGCGCTTTGTGGCCATGAACCGGGGCATCAACGAGGGCGGGGACCTGCCTGAGGAGCTGCTCA
GGAACCTGTACGACAGCATCCGAAATGAGCCCTTCAAGATTCCTGAGGATGACGGGAATGACCTGACCCAC
ACCTTCTTCAACCCGGACCGGGAGGGCTGGCTCCTGAAGCTGGGAGGGGGCCGGGTGAAGACGTGGAAGCG
GCGCTGGTTTATCCTCACAGACAACTGCCTCTACTACTTTGAGTACACCACGGACAAGGAGCCCCGAGGAA
TCATCCCCCTGGAGAATCTGAGCATCCGAGAGGTGGACGACCCCCGGAAACCGAACTGCTTTGAACTTTAC
ATCCCCAACAACAAGGGGCAGCTCATCAAAGCCTGCAAAACTGAGGCGGACGGCCGAGTGGTGGAGGGAAA
CCACATGGTGTACCGGATCTCGGCCCCCACGCAGGAGGAGAAGGACGAGTGGATCAAGTCCATCCAGGCGG
CTGTGAGTGTGGACCCCTTCTATGAGATGCTGGCAGCGAGAAAGAAGCGGATTTCAGTCAAGAAGAAGCAG
GAGCAGCCCTGACCCCCTGCCCCCAACTCCATTATTTATTACGGAGCTGCCCCGCCTGGGTGGCCGGACCC
CTGGGCCTTGGGGCTGTGGATCCTGGTTCCCTGTTTGGAAAATTCACCACCTCTAGCTCCTCACTGTTCTT
TGTAATTAACACGCTGTTGGTAATCTTATTAATTATTTAACCACTTGGCCTGCTGACCCCCTCATTTCTTG
GGGTTGACAGAGTCGAGGTGCTCCGTGGAGCCAGCCTGTTTCCCTGGACAGGGGCCTGGACCCGCCTGTCT
CTGGGTGCTGCCTGGGCTGTCCCGGTGGGTCTGTTCTGGTTTCACCCCGAGCCCAGCAGGAGTGGAGTAAA
GGGAGAAGGTTAATATGGTGGGAGTCTGGAGTTGGAATTGGCCGGGGACAGAGTTTAGAATGCAGGGATTC
AGGGTCAAGGTCTAGCATTCATTCTAGAAGTTACTTGACAGCCATCAGCCAGTTACGCTACAGCCATCAGC
TAGTTCCCATCGTTGCTTTCATGGGGCTTGAGGTCTTTGAGGGGCAGGAGATGAAGCTGGTGAGGTTGTGA
GCTGGGCCAGGGCTTTGAGGACAACCTGGAGCTGGAAGAACATGCGACCACCTCAGGGAGGGTCAGGGAAG
GATGAGTGGGGAGGTGGCCCATGTCCTGCAAGGGCCTTGCTGATGGGATGTCCTGAAGGGCTGGGCAGCCT
TAGATCGGGCTAAGAGGGCAGGACTGTGGGCCAGTGCCAGAGCCAGGCTTGTCTGTTCTCAAAGGATCAGC
CTCCTTTGGAGGACATTTTGTGTCAAGGATAGGGCTGAGGACCTGCGTTCTGAACGTCTTTCCTGGGATTC
TTGACAGCCAAAGAAATTCGAGGATCGCCTCTATGAAGGTGAAGTCTTGGCATGTGGGGTACCCCTGAATC
CTGGGCTAAGGTCTGTGACCTTTTGCCCTCCAGTTGGGGTATGTGTGGTGTCCCCAAAGAACAAGGGCTCG
TGGATCCCGAGCAGGGTGCGATGGCGGATTGGGAGGTCCCATGTCACTCTCCCATGCCCGCCTTTGAAGCT
GAGGCGCTCTTTAGTTAACAAACTACAAGTCCCAGCAGGGACCGGGACGCGGGTGGGAGAGGCGCCTGTGG
CCCCGAGGCGTGCCGCGAGTTGTAGTCCCTCCTGCCCGCTGTGTTCGCTTTTGAGCTCTCCGATGGGATGC
GGCGCTTCGGAATTTCGGGCTTTGATCCCTGTCCCGCCCTTGGCCACAGGCACCTGCCGGCCTGAAGGCCC

CCGCGGTGGGGGTACCCTGCGCCCCTCCGCGGGAAGGTGGACTACAGTTATCGGCAGGCTGTGCGGCGCCA
AAGCCACGGTGACCCAGACCCGAGGTTTTTCCGGGCGTCGCAGTTTCCCGAGACTCCGTGGCGGCGTTTGT
CTTCTTTTTCTTAGTCAGATCCCGTACTTTTGTGGAGGGTAGAGGAGGCTTTGACCGCCGCGGCCCCGGGG
GCTGGTGGGAAATGGAGTTCCAAATGAGAAAATAGAATTCCCCACTTCTCTTTCCCACAGGTGCCAGGGAA
ACGGAGTCATCGACCCAAGAAGGTCGTGGGAGATGAGGTCCCAGGGTAAACAGCGGGTCCCGCCACTATGT
CACCCTTTCCTGCCGCCTCCCCGGATGAACTGCATGCAGGGCGGCCGGCTCCGTGGCAGGCAGAGGCAGGA
AGAGGCGCGGAGCCCGCTGGTGGCGAAGAGGGGAGTGCGGCGGGAGGGGAGGGCAGAGGCAGGCGGCCGGT
CGCGTGGGGCCTGGGCCGCCCCAGGAGGGCCTCTGGCTGGATTCTTAGCAGATGGAAGCCGTGCAAGGGCA
GGAGGCAGGGGCCTGACGTGTTTGGATTGAGGTTGCAGGAGGGGCCCCTGGCTGCTTCAGGGAGAATAATT
TGGAGGCGAGCGGGGAGGTGAGAGGAAGCTGTGCGGTGCAGGATGGCGTGGCTGTGGAGGTGGAGAACGTT
GGGCCACGCTAGGAGTCTTGAAAGAGGAGCCAAAAGGACTTGCTGAGAGCCGCCGCAGCTGGTGAATGACG
GTGCCATGACTGAGGTGGTCGAAAGGGAGTATTTAAGGGGAAATCAGGCATTCCGTTTTGACCAAATTAAG
CTGGAGATGCCTGTGGAACATTCCAGCCAGGCTGCATCACGTCACACTGACCTTCAGCATTGCCCACAGTC
CATCTTCCCCCAGGACCTGAGGATTTTGCGTCCGGCTCCCTCCTCGCCCAGGACCCCCAAGCTCCCAGCAC
GCTTCTGATTTTTTTTGTAGGTTTTTTTTTTTGTTTTTTGTTTTGTTTTGTTTTGTTTTTGAGAGGGAGTC
TCACTTTGTGCCCTAGACTGGAGTGCAATGGCGCCATCTCGGCTTACTGCAACCTCCACCTCCCAGGTGCA
AGCGATTCCCCTGCCTCAGCTTCCCGAGTAGCTGGGATTACAGATGTGAGCCACCGTACCCAGCTAATTTT
TGTATTTTTAGTAGAGACGGGGTTTCACCATATTGGCCAAGCTGGTCTCGAACTCCTGACCTCAGGCGATC
CATCTGTCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCAGGCCTCAGCCCACTTCT
TTTGGGTGGCAATGGTTTGGATATCGTTTGTCCTCACTAAAATTCATGTTGAGATTCGAGCCCCAGTGTGG
CAGGTGTTGGGATGTGGGGCCTCATAAGGAGCTGTGTGGTCTTGGAGGTGGAGTCCTCATGGATAGATTAA
TGCCTGCCTTAAGGGGTGAGTGAGTGCTCACCCTTGGGAATTTGGGAATTTGGTTTTCTCTCCCTTGCTTC
CTTCAGCACCATGTGATCTCTGTGCACACAGCTGCATCCTTCTGCCCTCCCCCAGAAGCAGCAGCAGCCTC
AGGGATGAGGGAGGTCCTCAGTGAGTGCACCTGCCCAGTCTTGAACCTTCCAGCCATCAGAATCTGAGCCA
AATCAACCTCTTCCTTTATAAAGACCCAGCCTCAGGTCTTCTGTCAGAGCAACACAAAATGGACTCAGCAC
GGATCAAATTGTGTCTCCCCCACCCCCACAAAAAACTTTATATTAAAATCCTAACCCCCAGCA

LRTGAEEAAVAPGAFERAHPSPRANADPGPTGGTAPDSPRAFLAAMEDGVYEPPDLTPEERMELENIRRRK
QELLVEIQRLREELSEAMSEVEGLEANEGSKTLQRNRKMAMGRKKFNMDPKKGIQFLVENELLQNTPEEIA
RFLYKGEGLNKTAIGDYLGEREELNLAVLHAFVDLHEFTDLNLVQALRQFLWSFRLPGEAQKIDRMMEAFA
QRYCLCNPGVFQSTDTCYVLSFAVIMLNTSLHNPNVRDKPGLERFVAMNRGINEGGDLPEELLRNLYDSIR
NEPFKIPEDDGNDLTHTFFNPDREGWLLKLGGGRVKTWKRRWFILTDNCLYYFEYTTDKEPRGIIPLENLS
IREVDDPRKPNCFELYIPNNKGQLIKACKTEADGRVVEGNHMVYRISAPTQEEKDEWIKSIQAAVSVDPFY
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

TCAGAGTAATGCCAAACTCTCTCTGAGTGGGATGAGCAGAGCAGATGCTGCAATGAGATGCCAAAGCGGCT
CCCCTTCCTCTGTGCCTTGGGTGCCTATAAATTGCTCCGGCGCGCGTTTGTCAGCCTCCTCTTCTCCTGGC
AGGTGGTACCCAGGCAGAATTCTGCCTTCAGTCTCTCTCTCGCTCCGCTCCCGGCCGTGAGGCGCTCGCCG
CTGCTCGCTCGCTCCTCCGCCCCAGCTCTGAGCCTCGCCGTGCCGACCGTGCCCGCCGCCGCCGCCGCTGG
GCGCACCCGGGGGACGCCCGGGCCCACGCGGGGCTTTGGGGTGCGCGTCTATTCGAGTTGTGGTGGTGGCA
AAGGAGGAAGAGAGAAAGGAGGCAATAAAAAAAAAAGGCAGCGGACGGGCGAACTGAGCGAGCGAAAGAAG
AGGAGGAGGAGGCAGAAAAAGGCAACTTCAGACGGAAAGTTGGTGCGAACAGGCGCAGTCTGCAAAAACGG
AAAAGTCGATCGCAGGCGGCGGCGGCATAAAAGTGTGAGCTGCCCGGGCGAGCTCAGGAGGCGGCGGCTGG
CTCTGCCCTCCCGGTGGCCGCGCCGGCGCCGGCTGCAGCCACAGGTGCGAAGGAGCTCGCGGGGGGCGAGG

GCGCCCCGCGCACCCCTCCCCCGGCCCCCACCCCGGGCTCGGGACTTCGGCTCAAGTCACTGGGCGCCCGC
GCTCCCTCCCCAGCCGCAGCCTCCGCGGGGGGAGCAGGAGTCGGCAGCAGCGGCCGGCGCACGCACGCCCA
GGGGAGTTGGGGTTCGCAGGGGGTTGTTTTCGGCTCTGAAGAGGTCCCCGCCCAACCTTCAAAATTCTGTC
CAAAAGCAGACAAGAGGATCGCCCCGCGCTGAGCCGGCTGGTGGGCAGCAGGAGGCGCCTGATCGCCGCCG
GGGCGCTGGGGGTGGTGATGGTGCTGCTGCTGGTGATCCTCATCCCGGTGCTGGTGAGCTCGGCCGGCACG
TCGGCGCACTACGAGATGCTGGGCACCTGCCGCATGGTCTGCGACCCCTACGGGGGCACCAAGGCGCCCAG
CACCGCTGCCACGCCCGACCGCGGCCTCATGCAGTCCCTGCCCACCTTCATCCAGGGCCCCAAAGGCGAGG
CCGGCAGGCCCGGGAAGGCGGGTCCGCGCGGGCCCCCCGGAGAGCCCGGGCCACCCGGCCCCATGGGGCCC
CCGGGCGAGAAGGGCGAGCCGGGCCGCCAAGGCCTGCCGGGCCCGCCCGGGGCGCCCGGCCTGAACGCGGC
CGGGGCCATCAGCGCCGCCACCTACAGCACGGTGCCCAAGATCGCCTTCTACGCCGGCCTCAAGCGGCAGC
ATGAAGGCTACGAGGTGCTCAAGTTCGACGACGTGGTCACCAACCTCGGAAACCACTACGACCCCACCACC
GGCAAGTTCACCTGCTCCATCCCGGGCATCTACTTCTTCACCTACCACGTCCTGATGCGCGGAGGGGACGG
CACCAGCATGTGGGCTGATCTCTGCAAAAACAACCAGGTGCGTGCTAGTGCAATTGCCCAAGATGCTGATC
AGAATTACGACTATGCCAGTAACAGTGTGGTTCTTCATTTGGAGCCGGGAGATGAAGTCTATATCAAATTA
GATGGCGGGAAAGCCCATGGAGGAAACAACAACAAATACAGCACGTTTTCTGGATTTATTATTTATGCTGA
CTGATAATGCAGAAACTAAGCTTATTATTCTGAGTTTGAACACTGGATTCGTATGGCTAACGTCAGTGAAT
CAAGGATCCCAGGGGATGCCAATGGCAGGGCACCTCAGTTGTGTATATGTGGGGAAATCAAATGCTACCTG
ACTCACATCTGTATCACTCAGAAACATTATGTAAAAAATATCAAAGCAAGATAAGCAGATGTGTGATCCAC
TACCGCCAAAGCAAATACTCCTTATCGTTAGTGTCCATGTGAATGAAGTCCTATATAGATCACAAATTTTT
ATAGACAAATCTAAGACATTGAATTATTTCTTCTATATATATGATACTTTGGTGTACTGTGATCTTGCTGC
TTTTATCCATATGTCAGCTTTGGTTCTTGTGAGTTTACCTGCTTATTATGATACTTGGAGTCCATTCATAG
TGTGGGGAAGAATGATTTTTGCCCTGCAGGAGAAGGTCTAATTGAAATAATGCTGCTTGTCCCCAAAGAAA
TTGTTTGCCTTGTACTCTTGTTAACCTTAGAGCTAGACCTGGGAATGATTCAACTTCAAGCCTTAACCTGG
AATTTTCTGGATTTGAGGGAATTCCCAAGCCTATGATCTTTTTCACATTTTCTTTTTCTTATATGAATTTT
CTTTTCTCTCTTTTCTGGTGATAGTCTTTAAAAATAGAGATTGACAGTGTATCATAGGATTACCCTCTAAG
TGTGAAAATGTGTAATTATGTACGGTATTTAGAAAAAGCCCTTATGTGTCCATTTTGTCAACTGAATACAT
TTAGATTC

ILSKSRQEDRPALSRLVGSRRRLIAAGALGVVMVLLLVILIPVLVSSAGTSAHYEMLGTCRMVCDPYGGTK
APSTAATPDRGLMQSLPTFIQGPKGEAGRPGKAGPRGPPGEPGPPGPMGPPGEKGEPGRQGLPGPPGAPGL
NAAGAISAATYSTVPKIAFYAGLKRQHEGYEVLKFDDVVTNLGNHYDPTTGKFTCSIPGIYFFTYHVLMRG
GDGTSMWADLCKNNQVRASAIAQDADQNYDYASNSVVLHLEPGDEVYIKLDGGKAHGGNNNKYSTFSGFII
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

ATACCACCCGCGCCCGGGAGGGAGGGGAGGAAGGTTAGGGAGGCGGAGAGGGACCGCGCCCGCAGGAGAGA
GGCGCGGGCCAGGGCTCTAGCAGGGACTGGGGCCGCGGCAGGGGTAGCAAGGTGAGTCGGTGCTTGCCAAG
AGGCAGAGCGCAAAACCTACTAGGAGATCGCGCCCGGTGAGCAGCACCCGCAGCTCAGAGCCCGGGACGTC
CGGAGCGCGGGGAGCAGTCCCCTCTCCATCAGGGAGTGGTCTATCTGGGCAGTCTGGGACCCAGGCACCGC
GCCATCCCTGAGAGAGCAGCAGTCTGGAGAGCAGGCATCTCAGATCCCTAAGAAACCAGCCGTCCGAGAAG
CCGCGGATCTCAGGTGCCCAGGATCGTTAGGACTGAACGGGAGGGTACTAGAGGACCACTGGCTCTGGACC
GTCGGGAGCTGCCCCTGACGTAACCCACGAGGGGCCTCCCCTTGACGGACGGCTTGGGGAGCGGCACCGCC
GCGGCTGGAGCCCGCAGAGGCAGGGTAAGGGGAGCGGGGGGCAGCCGTCGGGGGAGTGCAGACCCAGGCCC
AAGGCGGGTCACCGCCTCCTGGCCCGCGGAGAGCCCCGGCCCCGGCAGCCATTGCGCCCAAGAGTGAGGAA
GATTTGCTGGCCCTGGCAGCGTCGCGGCTGAGCCGCCGCAAGAGGGTGGCGGGCGCGGCCGTCGGAGTGGC

CATGGTGCTGCTGCTGCTGGTGGCCATCCCGCTGCTGGTGCACAGCTCCCGCGGGCCAGCGCACTACGAGA
TGCTGGGTCGCTGCCGCATGGTGTGCGACCCGCATGGGCCCCGTGGCCCTGGTCCCGACGGCGCGCCTGCT
TCCGTGCCCCCCTTCCCGCCAGGCGCCAAGGGAGAGGTGGGCCGGCGCGGGAAAGCAGGCCTGCGGGGGCC
CCCTGGACCACCAGGTCCAAGAGGGCCCCCAGGAGAACCCGGCAGGCCAGGCCCCCCGGGCCCTCCCGGTC
CAGGTCCGGGCGGGGTGGCGCCCGCTGCCGGCTACGTGCCTCGCATTGCTTTCTACGCGGGCCTGCGGCGG
CCCCACGAGGGTTACGAGGTGCTGCGCTTCGACGACGTGGTGACCAACGTGGGCAACGCCTACGAGGCAGC
CAGCGGCAAGTTTACTTGCCCCATGCCAGGCGTCTACTTCTTCGCTTACCACGTGCTCATGCGCGGCGGCG
ACGGCACCAGCATGTGGGCCGACCTCATGAAGAACGGACAGGTCCGGGCCAGCGCCATTGCTCAGGACGCG
GACCAGAACTACGACTACGCCAGCAACAGCGTCATTCTGCACCTGGACGTGGGCGACGAGGTCTTCATCAA
GCTGGACGGCGGGAAAGTGCACGGCGGCAACACCAACAAGTACAGCACCTTCTCCGGCTTCATCATCTACC
CCGACTGAGCCGGCCCCGCCCCGTGCCCCCGCTCGCCCCTTCTCTCCCGTCCTCACCCACCTCCTGCCCGC
CCCACCCGAGGCGCCACCCCACCCTTTGAGAGCCTGGCGGTGGGGTGGACCCTTCCGTTCCCGGAGGCGGC
CTAAATGGGCGAACTCTTGGTGCTCAAGGGTATAAGTGGCCGGGAAGAGGAGGAGACCCGGCCAGAGGAGC
AGAGCGACTTCCGGAGGGATCACCCGCACCCAAGTGCGCGCTGGACCCCATAGGGGCAGAGGTCGTGGCTT
TCTCTTTTGTACAGAGATGGGGAGCAGTTTTAATAGCGGGACTCAGAGGCCCAGAAAGCCGGAGGGAAGCC
CCCGCAGCTTGCGAGGGAAATAACAGAAACAGGAGGAGCCCATTTAGGCAAGAGAAGACATTAAAACAGGG
TAGTGCAGGTTCTCCGTCACAACTTTCTCTCGCCACCCTCTCGTCCCCTCGTCTCCACTTTCAGGCTCAGG
CTCCAGCCTTGGCAGCCTTCCTGTGAACTGGAGGAACCAGTGAATTCTTTCCTGGCATTTAAAACGCATTC
TGTACAGTCCCCATTCCCCCCTATCCGGACTAGGCCCTGGGGCTACAGCTGCTGCTGCCTCTTCTAATAAA
GTGAGG

IAPKSEEDLLALAASRLSRRKRVAGAAVGVAMVLLLLVAIPLLVHSSRGPAHYEMLGRCRMVCDPHGPRGP
GPDGAPASVPPFPPGAKGEVGRRGKAGLRGPPGPPGPRGPPGEPGRPGPPGPPGPGPGGVAPAAGYVPRIA
FYAGLRRPHEGYEVLRFDDVVTNVGNAYEAASGKFTCPMPGVYFFAYHVLMRGGDGTSMWADLMKNGQVRA
SAIAQDADQNYDYASNSVILHLDVGDEVFIKLDGGKVHGGNTNKYSTFSGFIIYPD

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

ATTAGAGGCTCCAGCCCCGCCGACTTGCAGACGTGAGATCGGGCACACCTGAGCGGCGGCGGGGCGGTCGT
GGCCACATCCGGGGCGACGTGCCTGAGTCACCCCGTCCCGCCAGCGTCTGCCAGTCCAGCCAGTCCGCCCA
GTCTCTCGCGTCCGAGACTCGCCTCCAGCCTCCCACCTCCGCCCGGGCCGCGCGAGCCTCGCGGGGGCGGG
GGCGGGGCGCCAAGGGGCGGGGCTGTCTCTTAAAGGGCCCCGGGCCGCTGCCCTTAGGCCACTTCCTGGGG
GCGGAGAGGACCTCAGCGGCTGCGGCGACACCCAGGGAAGGCGGCGCGGCCGGGTCCCGAAACTCCTGGCT
GTTTCCATCAGAGCCCTCGGACACTCCCAGCCCGGGCTGAGCACGCATCGTCGCTCCCCGGCGGATACAAG
GGGGCTCCGCCATCCGCTCCCGTCAGTTCGGCCTCCATCTCCTGGGACCCGCGCCGGCAGCCAGGCCAGGC
CTCTGAGTGGCCCCAGAGCCCTGGCTGGACTCGTCCACGGCGGCAGCGATCTGCCCGGGGTCTCGGAGGCC
ATCCCTTCAGAGTCGGCCCTGTGCTCGCCACCGTCACCTGCTGGTTGGATTCCGGAAACCCACTGTCTGAA
GACCACAGAGGGGTGTCGCTGACCACCCCAAATCGGATACGTCCAGACCTCAAGCTCCCTTCCCCTCTCTG
GCTGCCCTCTGCTCTTTTCATCTCTTCTCTCAACCTTTTGGGGATTTCTGTGTCCTGACACCACCTCCCCA
TCCACCACCAAAGTAGCCGGGGTGAGCCCCAAACCTTACTGGGTGTGCTCCACCTGTGCCTCCAACCCAGC
GAATCTGACAGCTTCGACCCAATTCTGCACACACCCAGGAAGTTCTGCCTTTTCTTTTCTTTCGGTGTCTC
CTGTACTTCCCAAAATTTCTCCTCCTCCTGTGCCCTCTTCGCCCCCCTCCTTTGGGGGCCCCGTGACCCTG
AATGTGGGGGGCACACTATATTCCACCACTTTGGAGACCCTGACCCGCTTCCCAGACTCTATGCTGGGGGC
CATGTTTAGGGCCGGCACCCCCATGCCCCCCAACCTCAATTCCCAAGGAGGCGGCCACTACTTCATCGACC
GGGATGGCAAGGCCTTCCGGCACATCCTCAATTTCCTGAGGCTGGGCCGCCTGGACCTGCCCCGTGGGTAC
GGAGAGACAGCGCTGCTCAGGGCAGAGGCTGACTTCTACCAGATCCGGCCCCTCCTGGACGCGCTGCGGGA

ACTGGAGGCCTCTCAGGGGACCCCTGCACCCACAGCTGCCCTGCTCCACGCAGATGTAGATGTCAGCCCCC
GCCTGGTGCACTTCTCTGCTCGCCGGGGACCCCATCACTATGAGCTGAGCTCCGTCCAGGTGGACACCTTC
CGAGCCAACCTTTTCTGCACCGACTCTGAGTGTCTAGGTGCTTTGCGGGCCCGATTTGGTGTGGCCAGTGG
GGATAGGGCAGAGGGGAGCCCACATTTTCATCTGGAGTGGGCCCCCCGCCCCGTGGAACTCCCCGAGGTGG
AGTATGGGAGACTGGGGCTGCAGCCGCTGTGGACTGGGGGGCCAGGAGAGCGGCGGGAGGTGGTGGGCACC
CCAAGCTTCCTGGAGGAGGTGCTGCGGGTGGCTCTCGAGCACGGCTTCCGACTAGACTCTGTCTTCCCCGA
CCCCGAAGACCTGCTCAACTCCAGGTCTCTGCGCTTTGTCCGGCACTGAGGATGCTGTTCTCAGTTTGACT
GTGGGGAGGAGAGAGAATGGGGTACTAGCACCCCTGAAGCCTCTTTCCAGCTCTGCTTCAGGAGCTATGAG
AGTCGGGACTCTCCTGCACCTGACTGGAGCTCAGATGTGGGCAGGAATTCCCAAACCTGAGCCCACCAAGG
ACTCACAAGTGGTCCAGAAGGTCTCAACCTGTGCTGACCCTGGGAGGGGTAGGGAAGGTTCTCTCAGCTTG
TTCTTGCCTAAGGCTGAGCACCTCCAGTCTCTCCTTGATTTGGAGCTCAGTGTTTAAGGGCTTGGAAAAGG
GGGGAACATCTCTTTACCCAGACTAGACCTAGCAAAACCCTGGAAGGATATTGAGGTCTGGGGAAAAGGGA
GGACTTTGCATTTTCCCAATGCGGTCTCTTGGACCATGGCTTCTACTCCTGAAGCTGGGTGGCCTGGCCTG
GCCTGACCAATGAGAGGCCAGAACACTCTGGAACATCGGAAGAGGAGTTCTTTGCTATGTTCCAAGCCATC
TACTGAGGGAGGCAGAAAGGCCACAACCCACCCTAGGTTGATGTATGGGAGCTAGGACAGTCCCCATGGCA
ATGGGGCTGGAGCATCCCTCATCTGGAAGAATCCCATACTGATGGCAGGGCTGGCCAGGGGGAAGAGGGTA
GTATCTGTGGGTCCTGGCCTTTCTTCATGTGTGCGTGCATATCAGCCCGTGTGGCTGACTGATGTATAGGT
CCCTGGCATCCTGGTTCATATCTGTGTTGCTGACTACAGTGTCTGTGATGTCCGCATGTCCAGGCCTGTTT
GGGGTTGCCTAGCGACTCTTCTGGCACAGGGTGTGTCTGTGGTATACCTGTGAGGTGGTTGACAATTAGTA
GTTTAATCACAGGGTGTGTGTGTGTGTGTGTGTGTGTGTTTATGTGCACGCATGTATATGCATCACCAC
GTAGCCAGGAGGGGCCTGTTGGGGTTTGAGTCACTGGGATCTTCCTGGTGAGAGGTAAGAGAAGTCACTGG
GCTTAGCTGGGCCTCTGAGGCCTGTATGGAACTCTTGGTTGCTGAGGCAACCATGGACCTGTTGCTAGGAG
ATAGCTGGGGAAGGCCCAAGGCCGCCCAGGGCAGAGAGAGGAGACGAAGAGTTTGGGACAGTGGGGGAGGA
GATGGGAAGGGATGGGATTTCTGGGTCCCAGAGCGGGTGGGATACTCACGCACAGCTTCTTCACTGGTGGG
GGGTGGGGCACACATTATTTCTCACTGGTCATGATTTACAAGAAGAAAAATAAAACTGCTTTTGGAACCAC
AAAAAAAAAAAAAAAAAAAAAAAAAAA

ISPPPVPSSPPSFGGPVTLNVGGTLYSTTLETLTRFPDSMLGAMFRAGTPMPPNLNSQGGGHYFIDRDGKA
FRHILNFLRLGRLDLPRGYGETALLRAEADFYQIRPLLDALRELEASQGTPAPTAALLHADVDVSPRLVHF
SARRGPHHYELSSVQVDTFRANLFCTDSECLGALRARFGVASGDRAEGSPHFHLEWAPRPVELPEVEYGRL
GLQPLWTGGPGERREVVGTPSFLEEVLRVALEHGFRLDSVFPDPEDLLNSRSLRFVRH

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

CCCGGCCCTCTCCCCGGCCCCCGCCTCCCCTCCCCCGCCTCGCCCCACCGGCTTCCCACCACGGCCTCTCT
CGGCGAGGAAACTCTGGCCTCCGCTTCCTCCTCCTCCGACTCGGACACCGGCGGAGCCTCCCCGCCCCCGC
GGAAGAAACCCCGAGCCTCGGCGGCGGAGGGAGTAGGAGAGCCCGGGGCTTCGGCAGGCAGAGCAGGCCTC
TCCCCTCCGTCCTCGTCGTCGTCGTCCTCCTCCTCCTCCTCCTCCTCCGTGGTGGTAGTGGTGGGACTTCC
CCCGGCTGCTGCCCCTCCCGCCGCCGCTGCTGTCCCCCACCGAAGTAGCGGCCACAGCCTGGTCAGCGGCA
GCATCATGCAGGCCAATGGGGCAGGAGGAGGAGGAGGAGGTGGCGGCGGCGGCGGTGGAGGAGGAGGGGGC
GGCGGGGGCCAGGGACAGACCCCGGAACTCGCCTGCTTGTCGGCCCAGAACGGGGAGTCGTCCCCCTCGTC
GTCGTCGTCCGCGGGGGACCTGGCCCACGCCAATGGGCTCCTGCCTTCCGCCCCCTCCGCCGCCAGCAACA
ATAGCAACAGCCTGAATGTCAATAACGGGGTTCCCGGCGGGGCGGCCGCCGCATCCTCAGCCACCGTCGCA
GCTGCCTCCGCCACCACCGCCGCCTCCTCTTCCTTGGCCACCCCAGAACTGGGCAGCAGCCTCAAGAAGAA
GAAGCGGCTCTCCCAGTCAGATGAGGATGTCATTAGGCTAATAGGACAGCACTTGAATGGCTTAGGGCTCA

ACCAGACTGTTGATCTCCTCATGCAAGAGTCAGGATGTCGTTTAGAACATCCTTCTGCTACCAAATTCCGA
AATCATGTCATGGAAGGAGACTGGGATAAGGCAGAAAATGACCTGAATGAACTAAAGCCTTTAGTGCATTC
TCCTCATGCTATTGTGGTAAGAGGCGCACTTGAAATCTCTCAAACGTTGTTGGGAATAATTGTGAGGATGA
AGTTTTTGCTGCTGCAGCAGAAGTACCTAGAATACCTGGAGGATGGCAAGGTCCTGGAGGCACTTCAAGTT
CTACGCTGTGAATTGACGCCGCTGAAATACAATACAGAGCGCATTCATGTTCTTAGTGGGTATCTGATGTG
TAGCCATGCAGAAGACCTACGTGCAAAAGCAGAATGGGAAGGCAAAGGGACAGCTTCCCGATCTAAACTAT
TGGATAAACTTCAGACCTATTTACCACCATCAGTGATGCTTCCCCCACGGCGTTTACAGACTCTCCTGCGG
CAGGCGGTGGAACTACAAAGGGATCGGTGCCTATATCACAATACCAAACTTGATAATAATCTAGATTCTGT
GTCTCTGCTTATAGACCATGTTTGTAGTAGGAGGCAGTTCCCATGTTATACGCAGCAGATACTTACGGAGC
ATTGTAATGAAGTGTGGTTCTGTAAATTCTCTAATGATGGCACTAAACTAGCAACAGGATCAAAAGATACA
ACAGTTATCATATGGCAAGTTGATCCGGATACACACCTGCTAAAACTGCTTAAAACATTAGAAGGACATGC
TTATGGCGTTTCTTATATTGCATGGAGTCCAGATGACAACTATCTTGTTGCTTGTGGCCCAGATGACTGCT
CTGAGCTTTGGCTTTGGAATGTACAAACAGGAGAACTAAGGACAAAAATGAGCCAGTCTCATGAAGACAGT
TTGACAAGTGTGGCTTGGAATCCAGATGGGAAGCGCTTTGTGACTGGAGGTCAGCGTGGGCAGTTCTATCA
GTGTGACTTAGATGGTAATCTCCTTGACTCCTGGGAAGGGGTAAGAGTGCAATGCCTTTGGTGCTTGAGTG
ATGGAAAGACTGTTCTGGCATCAGATACACACCAGCGAATTCGGGGCTATAACTTCGAGGACCTTACAGAT
AGGAACATAGTACAAGAAGATCATCCTATTATGTCTTTTACTATTTCAAAAAATGGCCGATTAGCTTTGTT
AAATGTAGCAACTCAGGGAGTTCATTTATGGGACTTGCAAGACAGAGTTTTAGTAAGAAAGTATCAAGGTG
TTACACAAGGGTTTTATACAATTCATTCATGTTTTGGAGGCCATAATGAAGACTTCATCGCTAGTGGCAGT
GAAGATCACAAGGTTTACATCTGGCACAAACGTAGTGAACTGCCAATTGCGGAGCTGACAGGGCACACACG
TACAGTAAACTGTGTGAGCTGGAACCCACAGATTCCATCCATGATGGCCAGCGCCTCAGATGATGGCACTG
TTAGAATATGGGGACCAGCACCTTTTATAGACCACCAGAATATTGAAGAGGAATGCAGTAGCATGGATAGT
TGATGGTGAATTTGGAGCAGACGACTTCTGTTTAACTTAAAATTAGTCGTATTTTAATGGCTTGGGATTTG
GTGCAAACAAACATGATTGATAGCTGGACAGACATGCTCGTCATGAAAAAAGAACCATTTCTGAAGCCCGA
TTGGGGCCAAACATTTACACCTTGCTTCATAGTAACCAGTTGAGATGAAGCACGTCGTTAGAACGTTGTTG
GACACCATGTTGAATTATTCCCCCATCGGTTGTGAAGAACTGTGCTACATTCAGGCTTACCCATTGAACTC
AGTATATATATTTTTTTCCTTCCTGTCTTTTGTCTGGCAGGATACCATTCTTGTTGCTCTTCTGTGTAATG
AAGTTTAAATGCTTGTTTGGAAAACTTTATTTAACAGTTTAGAAGGCTTGATAGAAAGAGTGCATTAGTCT
GAAGAGTATACATTGGATAGGAAAGAATTTCCTTCTTTTGTTTCTCCAAATCTTTCCGCCTTATTTAGCTT
GAGATCTTTGCAGCTTGGTTCATGGATTCTAGCCTTGCCCGTTGCGCAGTATATACTGATCCAGATGATAA
ACCAGTGAACTATGTCAAAAGCACTCTCAATATTACATTTGACAAAAAGTTTTGTACTTTTCACATAGCTT
GTTGCCCCGTAAAAGGGTTAACAGCACAATTTTTTAAAAATAAATTAAGAAGTATTTATAGGATTAAAGTG
ACTTCATTTGTATACATTTGGAATCTAAACCAGCTTAAAAACAGTTTCCTCAATGACTTAGATACACAGTA
TAACTGATGCTCTTCTGGAATACCACATGAGACATGGTCAGAAACAGTGCTTGGAAGGACATTACACAAGA
AATTCAGAGTAATGCTTTGAAGATTTCCCCCCTTTTGTTTTATTCCTGAAGGAACATCAGTACCCGATCTT
GAAGAAATTCAAGATTCAAAAAGAATTTTAAATACACCAACATGAGACATCAGTAGTCAGTTGGTTTTCAG
TAAAGCTTGTTCCAAGTTGTTCTCAACTTAGGAAGTAATTTTGGTGTGATCTAGCAAAAGAGTAGGAATCA
GCGATACAACCACTTTGGAAGTTTATAGTATAATTGAAATTATTAGAAGAATTCAGCAGGTTACAGACATA
CTTAAACTGGGATTAAAACCTCATAGTCATTTTTCTTAATTGCCCTTAATATTTTGACATATAGGGATACA
TAAATTTAAAGAATATTTTTTTCTCAGTTTTTTCAGATATTGCCATACTGAACCTCATTCTAAACTGGTGCT
GTGGATAGTCTTTCCCTCTCCCCTCCTGTTTTAGTTTAAGGAAAGGTTTCCTTCATGGAAAAGCACTTTTA
ATTTTGAGAATTTCACATTTAAAATAAGCCAAAAGCCTGATGTTCACATTGTGCTTTTAAACTCCGTAACA
GCTGAATACAACATAACATAGTCTCCCTTGAAATTCCCTCCCTTTTTCAGTAGAGGATAAATAGGGTGACT
CCAGCTGTTGGAATGAGAATGGGGATACCCCCAAAATAAGATTCTGGTGAAGAAAAAGTCTTGTACTAAGT
TCCCCTTTAGGTGTAGCCAAGTCTATAATATAATTATGATCATGAATAGACTCCAGAATCCACAAAGAGAA
GAGGGCTTGGATGTCAAAAATTCTTCCATTGTGGGGTAGGTGGTCGGGGATGGGAGAAAGGGAAAAGTCAG
GGAAGGGGAGAAGAAAGAAAAACACATTTGGCCTTAAAATAGCATGCATTCTCCCAGAAGCCCATAGTAAA
GAAAATGACAGATCGTGGTCAGCAGACTCTTATCAAATTGGTATGTAAAAGATTTCTTCAAGCTCCATTTT
GCAGAGACCACCACTTAGAGAATCAAAAAATTCCTGTTTTGATACTTCTAAGTTAAAATATATGTTACAGT
TTATCTGGTACTTCATTTTTCTTAACTAAAATTACTTTTTACTTTAAGCTTGAATAAAAATCTTCATTGGT
AACTGTATGTAATTAAGTGGCAACTTCTCTTTACCTCAGTATAGTTAAAAACCATTCATATTTTAAGTGAT
TCATGTATACCCTGAAGCCGTTAATACACCCCAGTCTCTATGGTAACATCACTAAATGTGGGCATAGAAAC
TCCACAGCTGTTTACAAATTTGGTTACAGTGTGCTACGGTGCATTTGTAACACTCAGATTTGTGCCCATTC
AGGGGCAATTTGGCACTCAGATTCTACTCTACTTCACCTAACCACCCCTAGATCTGAGTTTTCAGAGTGCT
TCTGAAGTACAGTTTAAAAACACTTTTTAAAAAGTGGAGTAAAAGTGAGGCACATTTTACAAGAACATAAC
TCCTATTAAAACGGAGTAACAACTATGCAAAGGTTTCTATAGCAGCTAGGTGAGTTTGTTTTCCGGGTCTG
TCTTAACTGGCAGCTTCCTGTACATACTGGTACTTATTTGCTGTAAACGTCTGTTTCATACATTTGCCATG

CAGTGACTGTGCTCGAAAGGTGAAATCAATGGTAAAGAAGCCTAGGATTTTATGGTATGAGAAACTGATCG
CAGGATTTTCTGATCCATAGCTATTAATTGAAAACTTCTGATAGTGCTGTAGGCTCTATAGTAGGCTTAAG
AGTGAGTCTTATACTATGAAGCCAGCCACAAGATACTAACAAATGTATAAATTGAAAAGTACCAATACTTT
TTGCACATAGATCAAAATAGAGTGGAGAATTTGGGCTCCAAAAATCAGGTTAGTAGGCTCATTTCCATGTG
CTTAAAAATGTTTTTGTTGTTATGCATTGATTAAGGGCTAGAGCAAATATAAAATTAAATGTCTTTTTCCT
CCAAATAATTGGTTATTAATGGTAAAATGTTGGACTTGGGGATTACTGAGATTAGTATGTGGGTAGTAAAG
AATTGGAGTGGGAGTTAGTGGGGAGGGATGACAAAGCTTCAGCTTTGGTTTTTGTATTTTTTTAGTCTTCA
TAACAGCTACATTTCTGTTGCTCTTATTGGCCAGATAAATGTGTTATATTAAACCAGGTAAGTATTTTAAA
CAGAGAACATTTTATTTGCATATATCCCTACAAATTGGCCCATTTTATTGGGTGGGGGGATAAAAAAGGGA
TTCCCTAACTTTAAATAATCAGAATATCCTGCAATTGAAAGGAGAAAAAAATCCAATTTTGTCAGTATTTT
CTTAGTCTGAAAAAAACAGCAGTAGTAATCTGGTGAATGTTATTATTAGGAGTGTTCTGCCTCTACAGTGG
CATATTCCATGACATTAATAGAAAACCCTTCTTTCCCACCCCTCCCCATCTTGACATTTTCTTTTTTAGGA
TGTTGGTTCCTGCAAACAAGAACTGACCCTATCACCTCTGCCTCACAGTGGGTAAATATATCATACATCTC
TTGAAAGGGAAAATATCCATGTCAGTTTCTTTCACAGTAGAAGAATGGCTGCTTTTGAATGGGGAGGAAAG
AAGACATTTTTCTTAATATCCTTTTTGTCCCATCTCTAAAATGTATAAATATTTTGATTTGGTTTTACTTT
AATTTCTCATTCATAATTTCGCAACTGTCCTTTTGGGTTGCCCTCCAACACATTACTTCCATTTCAGAAAC
CATATCTGATTTAATATAAACTGCTAACTTAAGAAGAGGTGAATTTGCATTTACACTTAGACACTTTCTAA
AAACAGGGTGAACTTAATTGAAATTTTTCAGTGTTGTTAAGCTATTGTTATAGTAGCTGCCCTATAGTTAC
AAAAATAAAAACTATAAAAACCCATCCCCTATCTACTTAATCCAGATAGCCTGGATTTAACAAAACCAATA
GTATAAAAATGTGTGTGTGTGTTCTGTGTGTGTGTGAGAGAGAGAGTACTTGATTCATTAAAAAGTCCTGG
AAACAGCTGTTCTAGGATCTTGTCATTTTTACTGAAAGTCTCGTGCACATGTGAAGTGCCCTCTGAGTTTA
AGGTTTGGTTAATACTGGTATATTTTTATAAGATTGAAATTGTGTCCCCATCTTTAACTTAAACATTTTCA
TTATCAGGGTCAATGTGCTAGAACTGAATTGTAACATTTTTAGGCACAGATGGAAAAAAATGTTATTGGCT
CCTTGAAAATGTGTGTGTGTGGCGAGGGGAATAGATCCACAAAAGCATGTATGTACTTACAAACCAAGCTG
TAGAGATCAAGAAAAGAACTTAAGTGTTGATCTCAAGATTTCTAAATTGTCAAGATTTACATGGCATTGTG
GTGGAACTAGTTAACACTTAGAGCTTTTGGTATGTAATAACTATTTGCTATGGACTGATTAAATGTTTCAA
AAGATTGTGTTCTTCAATTTTGGTGGGTTTTGATTTTTGTTTTTTTAACTGCCTCTCAGATTATATTTACT
TAGTTTAAATTTCTTTGCTTTATTCATTAAAGTATAAAAACTTCAGGTCTCTGATATTTATTTTCACTTGT
TTACTAATTATTTACAAAACACCCTTTGTTGACTTTTATTTTATAAATGTGTAATGTATTAAACGTCTTTA
AATTTTTGTTCAACTGAAACTACATTAACTTTGATTTGCTTTACTGGGATTTTTTTTTAAAGACACTTTTT
CCATGTCAGTGCGCAGCACTTAACCAGTCGTTTGTATTCCCTTTCTTTCAATCCAATAAACAGAATGAATA
CCTAAATAA

TASLGEETLASASSSSDSDTGGASPPPRKKPRASAAEGVGEPGASAGRAGLSPPSSSSSSSSSSSSSVVVV
VGLPPAAAPPAAAAVPHRSSGHSLVSGSIMQANGAGGGGGGGGGGGGGGGGGGQGQTPELACLSAQNGES
SPSSSSSAGDLAHANGLLPSAPSAASNNSNSLNVNNGVPGGAAAASSATVAAASATTAASSSLATPELGSS
LKKKKRLSQSDEDVIRLIGQHLNGLGLNQTVDLLMQESGCRLEHPSATKFRNHVMEGDWDKAENDLNELKP
LVHSPHAIVVRGALEISQTLLGIIVRMKFLLLQQKYLEYLEDGKVLEALQVLRCELTPLKYNTERIHVLSG
YLMCSHAEDLRAKAEWEGKGTASRSKLLDKLQTYLPPSVMLPPRRLQTLLRQAVELQRDRCLYHNTKLDNN
LDSVSLLIDHVCSRRQFPCYTQQILTEHCNEVWFCKFSNDGTKLATGSKDTTVIIWQVDPDTHLLKLLKTL
EGHAYGVSYIAWSPDDNYLVACGPDDCSELWLWNVQTGELRTKMSQSHEDSLTSVAWNPDGKRFVTGGQRG
QFYQCDLDGNLLDSWEGVRVQCLWCLSDGKTVLASDTHQRIRGYNFEDLTDRNIVQEDHPIMSFTISKNGR
LALLNVATQGVHLWDLQDRVLVRKYQGVTQGFYTIHSCFGGHNEDFIASGSEDHKVYIWHKRSELPIAELT
GHTRTVNCVSWNPQIPSMMASASDDGTVRIWGPAPFIDHQNIEEECSSMDS

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

AGTCCCTATTTGGAACTGCTCTCGCGGCAGTTCAGACCTCGTGCTCGTCCCCTTCGCCTGTCTGTGTGTGG
TATCCGTAGGTCCGGGGCACTTTTTTTTTTGGTGTGTGTGTGTGTGCGTGCGCGCGTGTGTGTGTTTGGG
TGTGTGTGTTTTGGGGGTTTGCCGGGGCGCGGAGAGGAGAGCCGGGCCGAGGGGAAGGAAGGAAGCCGTAG
AGCTCCCCGGTCTCGGGCTGCGTCTCCCCTCCAGCGCTCTCCTGCCTGTGCGCTCCGACGCGGCAGCGGCG
GTGGCGCTGAGCGCTCCTCGGTGGGGACGCGGTGCTCAAGCTGGGAGCAGCGAGAACCCTTTGCGGAGGCT
TCCCGTGCCCGCTGGACTTAGGATGCCAGGGCGTTTGATTTGCACCTACTGAGCCCGGAGTTCCTCCCCGG
CCCTGGGAAGGGCTTGCAGCGGCGGCGGCGGCGGCGGCAGCGGCAGTAGCAGACGGAACAGCAGGCTCTCC
TCTCCGAGGGGGGGCGTCGAAGCGCCCGCGCCGGGCTCCCGCCCTCACCCTCCGGCGCTCCTGCGGCCCCA
AGCCACCACCTCGTAGCCCGCCTCCTTCCTCTTCTCGCCGCGAGCCCCAGGCCCGGACGCCTCTAGGATCG
GACCCGCGCGCCCCGGATGCTTGGGGCTGCCCCGGCGCGCCCCCCTACAGGCGCGACCCGCTGAGTGACCC
GGAGGGGGCCCGAGCTCGCCCCAGGAGGGGCCCCCGGCCTCGGGACCCCGCGCCGGACCTGCTGGGGAGCC
GACGATGTCCGCGCAGTCCCAGGCATTGACGCGCTCGACGCAGCCCAGGGAAGGCGGCCGGCCGCTGCCCG
CCGAGCCCTAGGCGCACAAAGCCCGCGCCCGCCGCCCGGCCCCGGCGCCCGCCGACGACTTTGCCGCCTGC
TCCGCGGCTCTTTGTCTCCACTTGGGGCGGGCGCCCGACTCTGGGATTTCGCTGCGAGAACGAGCTGGGGG
GGCCGGGGGCGAGCTCTCGGTTTCCCGGCCGCCCCCCGCTCGGGCTCGGCTCCCCCCTCCCCCGCACCTCC
CCGGCCCGGGCTCTCGGCGCTTCCACGCTCTCGGAATCACGACCCCTCCCTGCCATGTATCCGCAGGGCAG
ACATCCGGCTCCCCATCAACCCGGGCAGCCGGGATTTAAATTCACGGTGGCTGAGTCTTGTGACAGGATCA
AAGACGAATTCCAGTTCCTGCAAGCTCAGTATCACAGCCTCAAAGTGGAGTACGACAAGCTGGCAAACGAG
AAGACGGAGATGCAGCGCCATTATGTGATGTACTATGAGATGTCCTATGGCTTGAACATTGAAATGCACAA
GCAGACAGAGATTGCGAAGAGACTGAACACAATTTTAGCACAGATCATGCCTTTCCTGTCACAAGAGCACC
AGCAGCAGGTGGCGCAGGCAGTGGAGCGCGCCAAGCAGGTCACCATGACGGAGCTGAACGCCATCATCGGG
CAGCAGCAGCTCCAGGCGCAGCACCTCTCCCATGCCACACACGGCCCCCCGGTCCAGTTGCCACCCCACCC
GTCAGGTCTCCAGCCTCCAGGAATCCCCCCAGTGACAGGGAGCAGCTCCGGGCTGCTGGCACTGGGCGCCC
TGGGCAGCCAGGCCCATCTGACGGTGAAGGATGAGAAGAACCACCATGAACTCGATCACAGAGAGAGAGAA
TCCAGTGCGAATAACTCTGTGTCACCCTCGGAAAGCCTCCGGGCCAGTGAGAAGCACCGGGGCTCTGCGGA
CTACAGCATGGAAGCCAAGAAGCGGAAGGCGGAGGAGAAGGACAGCTTGAGCCGATACGACAGTGATGGAG
ACAAGAGTGATGATCTGGTGGTGGATGTTTCCAATGAGGACCCCGCAACGCCCCGGGTCAGCCCGGCACAC
TCCCCTCCTGAAAATGGGCTGGACAAGGCCCGTAGCCTGAAAAAAGATGCCCCCACCAGCCCTGCCTCGGT
GGCCTCTTCCAGTAGCACACCTTCCTCCAAGACCAAAGACCTTGGTCATAACGACAAATCCTCCACCCCTG
GGCTCAAGTCCAACACACCAACCCCAAGGAACGACGCCCCAACTCCAGGCACCAGCACGACCCCAGGGCTC
AGGTCGATGCCGGGTAAACCTCCGGGCATGGACCCGATAGGTATAATGGCCTCGGCTCTGCGCACGCCCAT
CTCCATCACCAGCTCCTATGCGGCGCCCTTCGCCATGATGAGCCACCATGAGATGAACGGCTCCCTCACCA
GTCCTGGCGCCTACGCCGGCCTCCACAACATCCCACCCCAGATGAGCGCCGCCGCCGCTGCTGCAGCCGCT
GCCTATGGCCGATCGCCAATGGTGAGCTTTGGAGCTGTTGGTTTTGACCCTCACCCCCCGATGCGGGCCAC
AGGCCTCCCCTCAAGCCTGGCCTCCATTCCTGGAGGAAAACCAGCGTACTCATTCCATGTGAGTGCTGATG
GGCAGATGCAGCCCGTGCCCTTCCCCCACGACGCCCTGGCAGGCCCCGGCATCCCGAGGCACGCCCGGCAG
ATCAACACACTCAGCCACGGGGAGGTGGTGTGTGCCGTGACCATCAGCAACCCCACGAGGCACGTCTACAC
AGGTGGCAAGGGCTGCGTGAAGATCTGGGACATCAGCCAGCCAGGCAGCAAGAGCCCCATCTCCCAGCTGG
ACTGCCTGAACAGGGACAATTACATCCGCTCCTGCAAGCTGCTCCCTGATGGGCGCACGCTCATCGTGGGC
GGCGAGGCCAGCACGCTCACCATCTGGGACCTGGCCTCGCCCACGCCCCGCATCAAGGCCGAGCTGACGTC
CTCGGCTCCCGCCTGTTATGCCCTGGCCATTAGCCCTGACGCCAAAGTCTGCTTCTCCTGCTGCAGCGATG
GGAACATTGCTGTCTGGGACCTGCACAACCAGACCCTGGTCAGGCAGTTCCAGGGCCACACAGATGGGGCC
AGCTGCATAGACATCTCCCATGATGGCACCAAACTGTGGACAGGGGGCCTGGACAACACGGTGCGCTCCTG
GGACCTGCGGGAGGGCCGACAGCTACAGCAGCATGACTTCACTTCCCAGATCTTCTCGCTGGGCTACTGCC
CCACTGGGGAGTGGCTGGCTGTGGGCATGGAGAGCAGCAACGTGGAGGTGCTGCACCACACCAAGCCTGAC
AAGTACCAGCTGCACCTGCACGAGAGCTGCGTGCTCTCCCTCAAGTTCGCCTACTGCGGCAAGTGGTTCGT
GAGCACTGGGAAAGATAACCTTCTCAACGCCTGGAGGACGCCTTATGGAGCCAGCATATTCCAGTCTAAAG
AATCCTCGTCTGTCTTGAGTTGTGACATTTCAGCGGATGACAAATACATTGTAACAGGCTCTGGTGACAAG
AAGGCCACAGTTTATGAGGTCATCTACTAAACAAGAACTCCAGCAGGGCTGTCAAACTCTGGGAGAAACCG
ACTCGGCTCTGACAGGGAGACCCCCAGGCGAGGGGCCCCGAGGATGGCGGAGGATGGGCCGCAGGCAGCCG
AGCGTTCAGGGCTGCGCTCCGGCCGGCTGAGAGGGCACGTGCCCCGTCACAGTCTGGACTCCTGGGCCTGG
ATTGATGTGTCTCACAGACTCGGAAGGGTTCTGCTCCTCCTCCTCCCCCTGAACAATGCTGGCAGTTGCTA
CAAATAGATTTATTGGAGGCTTATGGCTCCGGTTCCCCCACAGACCCGCTCATGAGTCTCTGTTTGTTCTT
CCCTTTTCTTTTGCCCTGTCCCTCACCTTGGGTCGGGGGTGCTGGAGTGGACCACAATGTTGTGCTGGGGG
ATGGGGGGGTCTCTCTTTGCCGATTGTGCAGTGCACAAGATTTGTGAAAAATGTAAATAACAGACTCCTAT

TGCGGACTGATCAGTGGGAGAGGAGGCCCCTTCCCACCGGAAACTCTGAGTGTGTATTTCGCCTGCTGTAT
TTGTAATCCACTCGTGGTGGTGGCTTTTTTTTTTTTTTTTTTTAAATAAACAGATGCTCTCACCTGGGAA
GAGGAGACAGGGAGGGGAACCAATTGAAGAAAGAGGAGAAAAGTCTTAGAGTGTGGAAAAGGCAACCAGGT
TGGCCGTAAGGTGCCTGCTGGAATGCGTGTGCCTCCACACGGGTCTGGGCATCCGGACTGATAACCAGCCG
GCCAGACTGAGGGATGGAAGGCACTGAGATGGGGGCCCGTCCAGGCGGACACCCGCAGAAATGGAGCTTTC
TGTGGTCTCTTGCACTCTGGCTGCCTCTTGCCCTCTCTGTGTCTCTCTTTCTTGGTCTCTCCCTCTCTCCT
CCTCAGCCTGGTCTTTCTCTTTGGTGCACACTTAGTTATTGTTGTGAGCAATGGAAGTTCAAAGGAACTCC
CTCTCCAGCTCTTCTGAATCTTGGGACACAGCCTAAAAAGGACAAAAAGTTAGAAGACAGCATAGCAACTC
AGCTCAGGGAGCTACCAGAGAAAAATAGCAACTGATGTGGGTGCTTTTTTTTTTTTTTTAATTTGAATAAA
AAGAATTAGAAGTGATGTCCTTTTATAAAATGCCTTCTCCCCCTTCCCGCCTACAGTCTCTTCCTCTCCCC
TTAGAGGGGGGAAAGTGTATAAACCTACAGGGTTGTGAGTCTGAAAAGAGGATCCCCCTCACCCCCACCCT
GGGCAGAGCAGTGGGGGTTGGGGGGTGGGAGAGGGGGACACAGATCCTGGCACACTGTGGATATTTCTTGC
AGATTGCAGTCTCTTGTGGCCCAAACAGGTTAGGTAGACTATCGCCTCTGGCAGGTGCCACCTTTTGGTAC
CAACATGTTCTGAGGTGTTAGGATTTGGGTTGGGTTTTTTTTGTTTGTTTTTTTTTTCCTTTTGGTCTTTT
TTTTTTTCTCCTTTTAAAGAAAAGCTAAAGGCCGCTGTGAGTCCTGGTGGCAGGCTCTCCATGGATGTAGC
ATATCGAAGATAATTTTTATACTGCATTTTTATGGATTATTTTGTAATGTGTGATTCCGTCTGCTGAGGAG
GTGGGAGGGGCTCCAGGGAAAGCCACCCACCTTCAGTGAGGTTGCTCCCCAGCTGAGCGCACCGGGCATGG
GATGTGGAGGCTGGCGACACACCCTGTGCCTCTCCAAGGCTGGGCGCGTGGGGCGTCCAGAGTCTCTCTGG
GTCTCAGATGTCCATCTGCCACCTCTTGTTAAGGCTCTAGCCAGAAGGGAGGGTGAGGGTAGAAGAAAGTT
ATTCCCGAAGAAAAAAAGAATGAAAAGTCATTGTACTGAACTGTTTTTATATTTTTAAAAGTTACTATTTA
AAGGTTAAAAAAAAAAAAAAAAAAAAAAA

LGFRCENELGGPGASSRFPGRPPLGLGSPLPRTSPARALGASTLSESRPLPAMYPQGRHPAPHQPGQPGFK
FTVAESCDRIKDEFQFLQAQYHSLKVEYDKLANEKTEMQRHYVMYYEMSYGLNIEMHKQTEIAKRLNTILA
QIMPFLSQEHQQQVAQAVERAKQVTMTELNAIIGQQQLQAQHLSHATHGPPVQLPPHPSGLQPPGIPPVTG
SSSGLLALGALGSQAHLTVKDEKNHHELDHRERESSANNSVSPSESLRASEKHRGSADYSMEAKKRKAEEK
DSLSRYDSDGDKSDDLVVDVSNEDPATPRVSPAHSPPENGLDKARSLKKDAPTSPASVASSSTPSSKTKD
LGHNDKSSTPGLKSNTPTPRNDAPTPGTSTTPGLRSMPGKPPGMDPIGIMASALRTPISITSSYAAPFAMM
SHHEMNGSLTSPGAYAGLHNIPPQMSAAAAAAAAAYGRSPMVSFGAVGFDPHPPMRATGLPSSLASIPGGK
PAYSFHVSADGQMQPVPFPHDALAGPGIPRHARQINTLSHGEVVCAVTISNPTRHVYTGGKGCVKIWDISQ
PGSKSPISQLDCLNRDNYIRSCKLLPDGRTLIVGGEASTLTIWDLASPTPRIKAELTSSAPACYALAISPD
AKVCFSCCSDGNIAVWDLHNQTLVRQFQGHTDGASCIDISHDGTKLWTGGLDNTVRSWDLREGRQLQQHDF
TSQIFSLGYCPTGEWLAVGMESSNVEVLHHTKPDKYQLHLHESCVLSLKFAYCGKWFVSTGKDNLLNAWRT
PYGASIFQSKESSSVLSCDISADDKYIVTGSGDKKATVYEVIY

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; hIKCa1; KCNN4

GTCCTTCGGTGTCTGGGTGTGGTGAGTAGAGGTGTGTGTCACAAAGTACAGACCATTGTGTGTGACAAAGC
CCATCGTGTGTCTGTGTGTGTCTTTATCCACGTGGATGGACGTCTCTTTCTTGCTCTGCCCCAAGACACAC
CCTAGCCCCTCCTTATTCTCAAAAGGGGGAGCTGGGGAGCCTCCCCCTACCCTGGGGCCTCCCCTGCCCCT
CCCCGCCCTGCCTGGCCGTCACCACTCCCCAGAGGGCACAGGGCTCTGCTGTGCCTCAGAGCAAAAGTCCC
AGAGCCAGCAGAGCAGGCTGACGACCTGCAAGCCACAGTGGCTGCCCTGTGCGTGCTGCGAGGTGGGGGAC
CCTGGGCAGGAAGCTGGCTGAGCCCCAAGACCCCGGGGGCCATGGGCGGGGATCTGGTGCTTGGCCTGGGG
GCCTTGAGACGCCGAAAGCGCTTGCTGGAGCAGGAGAAGTCTCTGGCCGGCTGGGCACTGGTGCTGGCAGG

AACTGGCATTGGACTCATGGTGCTGCATGCAGAGATGCTGTGGTTCGGGGGGTGCTCGTGGGCGCTCTACC
TGTTCCTGGTTAAATGCACGATCAGCATTTCCACCTTCTTACTCCTCTGCCTCATCGTGGCCTTTCATGCC
AAAGAGGTCCAGCTGTTCATGACCGACAACGGGCTGCGGGACTGGCGCGTGGCGCTGACCGGGCGGCAGGC
GGCGCAGATCGTGCTGGAGCTGGTGGTGTGTGGGCTGCACCCGGCGCCCGTGCGGGGCCCGCCGTGCGTGC
AGGATTTAGGGGCGCCGCTGACCTCCCCGCAGCCCTGGCCGGGATTCCTGGGCCAAGGGGAAGCGCTGCTG
TCCCTGGCCATGCTGCTGCGTCTCTACCTGGTGCCCCGCGCCGTGCTCCTGCGCAGCGGCGTCCTGCTCAA
CGCTTCCTACCGCAGCATCGGCGCTCTCAATCAAGTCCGCTTCCGCCACTGGTTCGTGGCCAAGCTTTACA
TGAACACGCACCCTGGCCGCCTGCTGCTCGGCCTCACGCTTGGCCTCTGGCTGACCACCGCCTGGGTGCTG
TCCGTGGCCGAGAGGCAGGCTGTTAATGCCACTGGGCACCTTTCAGACACACTTTGGCTGATCCCCATCAC
ATTCCTGACCATCGGCTATGGTGACGTGGTGCCGGGCACCATGTGGGGCAAGATCGTCTGCCTGTGCACTG
GAGTCATGGGTGTCTGCTGCACAGCCCTGCTGGTGGCCGTGGTGGCCCGGAAGCTGGAGTTTAACAAGGCA
GAGAAGCACGTGCACAACTTCATGATGGATATCCAGTATACCAAAGAGATGAAGGAGTCCGCTGCCCGAGT
GCTACAAGAAGCCTGGATGTTCTACAAACATACTCGCAGGAAGGAGTCTCATGCTGCCCGCAGGCATCAGC
GCAAGCTGCTGGCCGCCATCAACGCGTTCCGCCAGGTGCGGCTGAAACACCGGAAGCTCCGGGAACAAGTG
AACTCCATGGTGGACATCTCCAAGATGCACATGATCCTGTATGACCTGCAGCAGAATCTGAGCAGCTCACA
CCGGGCCCTGGAGAAACAGATTGACACGCTGGCGGGGAAGCTGGATGCCCTGACTGAGCTGCTTAGCACTG
CCCTGGGGCCGAGGCAGCTTCCAGAACCCAGCCAGCAGTCCAAGTAGCTGGACCCACGAGGAGGAACCAGG
CTACTTTCCCCAGTACTGAGGTGGTGGACATCGTCTCTGCCACTCCTGACCCAGCCCTGAACAAAGCACCT
CAAGTGCAAGGACCAAAGGGGGCCCTGGCTTGGAGTGGGTTGGCTTGCTGATGGCTGCTGGAGGGGACGCT
GGCTAAAGTGGGTAGGCCTTGGCCCACCTGAGGCCCCAGGTGGGAACATGGTCACCCCCACTCTGCATACC
CTCATCAAAAACACTCTCACTATGCTGCTATGGACGACCTCCAGCTCTCAGTTACAAGTGCAGGCGACTGG
AGGCAGGACTCCTGGGTCCCTGGGAAAGAGGGTACTAGGGGCCCGGATCCAGGATTCTGGGAGGCTTCAGT
TACCGCTGGCCGAGCTGAAGAACTGGGTATGAGGCTGGGGCGGGGCTGGAGGTGGCGCCCCCTGGTGGGAC
AACAAAGAGGACACCATTTTTCCAGAGCTGCAGAGAGCACCTGGTGGGGAGGAAGAAGTGTAACTCACCAG
CCTCTGCTCTTATCTTTGTAATAAATGTTAAAGCCAGAA

VAALCVLRGGGPWAGSWLSPKTPGAMGGDLVLGLGALRRRKRLLEQEKSLAGWALVLAGTGIGLMVLHAEM
LWFGGCSWALYLFLVKCTISISTFLLLCLIVAFHAKEVQLFMTDNGLRDWRVALTGRQAAQIVLELVVCGL
HPAPVRGPPCVQDLGAPLTSPQPWPGFLGQGEALLSLAMLLRLYLVPRAVLLRSGVLLNASYRSIGALNQV
RFRHWFVAKLYMNTHPGRLLLGLTLGLWLTTAWVLSVAERQAVNATGHLSDTLWLIPITFLTIGYGDVVPG
TMWGKIVCLCTGVMGVCCTALLVAVVARKLEFNKAEKHVHNFMMDIQYTKEMKESAARVLQEAWMFYKHTR
RKESHAARRHQRKLLAAINAFRQVRLKHRKLREQVNSMVDISKMHMILYDLQQNLSSSHRALEKQIDTLAG
KLDALTELLSTALGPRQLPEPSQQSK

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

GAGGGAGGAGGAGGAGTGGGGACCGGGCGGGGGGTGGAGGAAGAGGCCTCGCGCAGAGGAGGGAGCAATTG
AATTTCAAACACAAACAACTGCACGAGCGCGCACCCACCGCGCCGGAGCCTTGCCCCGATCCGCGCCCGCC
CCGTCCGTGCGGCGCGCGGGCGGAGACGCCGTGGCCGCGCCGGAGCTCGGGCCGGGGGCCACCATCGAGGC
GGGGGCCGCGCGAGGGCCGGAGCGGAGCGGCGCCGCCACCGCCGCACGCGCAAACTTGGGCTCGCGCTTCC
CGGCCCGGCGCGGAGCCCGGGGCGCCCGGAGCCCCGCCATGTCGCGATCCAACCGGCAGAAGGAGTACAAA
TGCGGGGACCTGGTGTTCGCCAAGATGAAGGGCTACCCACACTGGCCGGCCCGGATTGACGAGATGCCTGA
GGCTGCCGTGAAATCAACAGCCAACAAATACCAAGTCTTTTTTTTCGGGACCCACGAGACGGCATTCCTGG
GCCCCAAAGACCTCTTCCCTTACGAGGAATCCAAGGAGAAGTTTGGCAAGCCCAACAAGAGGAAAGGGTTC

AGCGAGGGGCTGTGGGAGATCGAGAACAACCCTACTGTCAAGGCTTCCGGCTATCAGTCCTCCCAGAAAAA
GAGCTGTGTGGAAGAGCCTGAACCAGAGCCCGAAGCTGCAGAGGGTGACGGTGATAAGAAGGGGAATGCAG
AGGGCAGCAGCGACGAGGAAGGGAAGCTGGTCATTGATGAGCCAGCCAAGGAGAAGAACGAGAAAGGAGCG
TTGAAGAGGAGAGCAGGGGACTTGCTGGAGGACTCTCCTAAACGTCCCAAGGAGGCAGAAAACCCTGAAGG
AGAGGAGAAGGAGGCAGCCACCTTGGAGGTTGAGAGGCCCCTTCCTATGGAGGTGGAAAAGAATAGCACCC
CCTCTGAGCCCGGCTCTGGCCGGGGGCCTCCCCAAGAGGAAGAAGAAGAGGAGGATGAAGAGGAAGAGGCT
ACCAAGGAAGATGCTGAGGCCCCAGGCATCAGAGATCATGAGAGCCTGTAGCCACCAATGTTTCAAGAGGA
GCCCCCACCCTGTTCCTGCTGCTGTCTGGGTGCTACTGGGGAAACTGGCCATGGCCTGCAAACTGGGAACC
CCTTTCCCACCCCAACCTGCTCTCCTCTTCTACTCACTTTTCCCACTCCAAGCCCAGCCCATGGAGATTGA
CCTGGATGGGGCAGGCCACCTGGCTCTCACCTCTAGGTCCCCATACTCCTATGATCTGAGTCAGAGCCATG
TCTTCTCCCTGGAATGAGTTGAGGCCACTGTGTTCCTTCCGCTTGGAGCTATTTTCCAGGCTTCTGCTGGG
GCCTGGGACAACTGCTCCCACCTCCTGACACCCTTCTCCCACTCTCCTAGGCATTCTGGACCTCTGGGTTG
GGATCAGGGGTAGGAATGGAAAGGATGGAGCATCAACAGCAGGGTGGGCTTGTGGGGCCTGGGAGGGGCAA
TCCTCAAATGCGGGGTGGGGGCAGCACAGGAGGGCGGCCTCCTTCTGAGCTCCTGTCCCCTGCTACACCTA
TTATCCCAGCTGCCTAGATTCAGGGAAAGTGGGACAGCTTGTAGGGGAGGGGCTCCTTTCCATAAATCCTT
GATGATTGACAACACCCATTTTTCCTTTTGCCGACCCCAAGAGTTTTGGGAGTTGTAGTTAATCATCAAGA
GAATTTGGGGCTTCCAAGTTGTTCGGGCCAAGGACCTGAGACCTGAAGGGTTGACTTTACCCATTTGGGTG
GGAGTGTTGAGCATCTGTCCCCCTTTAGATCTCTGAAGCCACAAATAGGATGCTTGGGAAGACTCCTAGCT
GTCCTTTTTCCTCTCCACACAGTGCTCAAGGCCAGCTTATAGTCATATATATCACCCAGACATAAAGGAAA
AGACACATTTTTTAGGAAATGTTTTTAATAAAAGAAAATTACAAAAAAAAAATTTTAAAGACCCCTAACCCT
TTGTGTGCTCTCCATTCTGCTCCTTCCCCATCGTTGCCCCCATTTCTGAGGTGCACTGGGAGGCTCCCCTT
CTATTTGGGGCTTGATGACTTTCTTTTTGTAGCTGGGGCTTTGATGTTCCTTCCAGTGTCATTTCTCATCC
ACATACCCTGACCTGGCCCCCTCAGTGTTGTCACCAGATCTGATTTGTAACCCACTGAGAGGACAGAGAGA
AATAAGTGCCCTCTCCCACCCTCTTCCTACTGGTCTCTCTATGCCTCTCTACAGTCTCGTCTCTTTTACCC
TGGCCCCTCTCCCTTGGGCTCTGATGAAAAATTGCTGACTGTAGCTTTGGAAGTTTAGCTCTGAGAACCGT
AGATGATTTCAGTTCTAGGAAAATAAAACCCGTTGATTACTATAAAAAAAAAAA

REEEEWGPGGGWRKRPRAEEGAIEFQTQTTARARTHRAGALPRSAPAPSVRRAGGDAVAAPELGPGATIEA
GAARGPERSGAATAARANLGSRFPARRGARGARSPAMSRSNRQKEYKCGDLVFAKMKGYPHWPARIDEMPE
AAVKSTANKYQVFFFGTHETAFLGPKDLFPYEESKEKFGKPNKRKGFSEGLWEIENNPTVKASGYQSSQKK
SCVEEPEPEPEAAEGDGDKKGNAEGSSDEEGKLVIDEPAKEKNEKGALKRRAGDLLEDSPKRPKEAENPEG
EEKEAATLEVERPLPMEVEKNSTPSEPGSGRGPPQEEEEEEDEEEEATKEDAEAPGIRDHESL

AUG initiation = 26.8 kDa
GUG initiation = 31.6 kDa

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

#26

NM_013313

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

YPEL1 yippee-like 1 (Drosophila)

Synonyms: FKSG3; MGC64992; YPEL1

CAGGCGGCCCGCGGGGCGGCGGCGGGAGGATGCCTCGCGCCCTGCCTAGGCGTTAACGGCCTCAGCGCGTC
CCGGGCCGCGCCGGGAACGCCTGAGAGCCGAGCCCGCGCTGACCGGGGCCCGGGCCGGATGGGCGCTGCGG
GCCGGGGCGCGGACCGCGGAGCGGCCGTTACGTTTTTCTCTCGTCCCAGCTGTGTGGACAGTGCCACACGC
CCTCCTGGACAACGACGGCTCTTGCGTGCATTTTCCAGTGTTGCTTTTGTCAGAGAGGCCAGCAGAGCCAC
GGTTCTTCCAGAACCAGCCCTGAGCTGAGTGAGGAGTGCCCAGGAGAGATGGTGAAAATGACAAAGTCCAA
AACTTTCCAAGCGTATCTGCCGAACTGTCACCGAACGTACAGCTGTATCCACTGCAGAGCACACCTGGCCA
ATCATGACGAGCTCATCTCCAAGTCCTTTCAGGGGAGCCAGGGACGCGCCTACCTCTTCAATTCCGTGGTG
AACGTGGGCTGCGGCCCTGCAGAGGAGAGGGTCCTTCTCACCGGGCTGCATGCGGTTGCCGACATCTACTG
CGAGAACTGCAAGACCACGCTCGGGTGGAAATACGAGCATGCCTTTGAGAGCAGTCAGAAATATAAGGAAG

GAAAATTCATCATTGAGCTTGCTCATATGATCAAAGACAATGGCTGGGAGTAATGTGCGAACTTTCCCTTC
TCCTTTGAATGCTGTTTTGTGAAAGAAACTGTGAATGTAATGGAAACGTAGGAGCATCTGGTGACAGCCTT
TCTTGCCCTCTGACCTCAAAGGCTAGCTGCGCATAGCTCTTGACACTCTCGGCCATCTCTGTGGGTAAGGT
GTCCCTCGGATCTGTCCTCTTCGTGTACACAGTTGTTTCTGAAAATTTTCAATGAGCTTTTTCTAACTTCT
CAAGTTCTAGAGAAAGAATTAACCAACTGATGACTTACCTGCCTAGTTAATATCTTCCTTTCACCTTTGTC
TTCAATATAGTTGGGCTCTGCTTTTTTAAGGTTCAGTTGAAAACCAAACTGGGGCCGGGTGCGGTGGCTCA
CGCCTGTAATCCCAGCACTTTGGGAGGCCAAGATGGGTGGATCACCTGAGGTCAGGAGTTCTAGATCAGCC
TGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACGAAAATTAGCCGGGCATGGTGGCGAGTGCCTGT
AATCTTAGCTACTCAGGAGGCTAAGGCAGGAGAATCACTTGAACCTGGGACACGGAGGTTGCAGTGAGCTA
AGATCATGCCATCGCACTCCAGCCTGGGGGACAAAAGTGAGACATCGTCTCAAAAAAAAAAAAAAAAGCTG
GGTATGGTGGCGCATGCCTTTAATCCCAGCTACTCGGGAGGCTGAGGCACGAGAATCACTTGAACCCAGGA
GGCGGAGGTTGCAGTGAGCCAAGATCGCGCCACTGCACTCCAGCCTGGCAATAGGGCGAGACTCCGTCTCA
ATTTAAAACAAAAGAGAACCAGACTGAGTCTCTGAAGACCACAGGGACAGGGTCTCTTTAGATAGCAAGTC
TCACCATTCCCTTTTTTAGAGAAAAGGTATTGTAGCCCACCCTCCACCCCGCTGTTTTTCTTAAATTTGCA
GAACTTCAAATTGGCTATTCCTCTTGCAAATGAACCCTTAAAGTACAGTGTTATTTAAGAATCTTCCAGAG
GCAGTCAACAGACTTATACACTAAGGGCATTTTTGGTTTTTAGCTTGTTCAAAAACAGAGGCCAGCACAGA
TGACATTTTAGATACACTCTAAATTGAGAATGGTGTCTAGTGGAACATGTTTATTTAAGCCAGTAGATTCC
TTATCTAGAAAGCAGGTGAGCTAGCCCTTAGAGAAGGCTGTCCCGGGGCCCGCAGAGGTGCCCTTACTGAG
GTGACAGCCTCACAGGGTCTGGTACCAGGGGTTGTGCCCTCAGCAGTGACAGCAGCTTAGGTGTCAGGCAG
TTGCTGAGTGGCTGGTCCATGTCTATAGAGTAACACACTGGACCGAGGAAAAGTCAGATTTCATTTTCTAC
CCTGGATGTACTTGAAGAAAAAGAATTATTTTTGCATATGAAAGAGGCCAGAACCCACAGGAAAAACTTCA
AAACTTGACATTTGCCAGAATGTTTAAAATTTGTTCAGAAAAGGTTAAAGCAACAAGTTTAGCCTTTGTGC
ATGAAGACGCCTGGCCTGCTAGACGCGTTGCCCGTCCCTGCGTGGTGCTGTCCCATGTCACTTGAACTGAT
AGAGGGGCCTGTGCAATCTCCTAAGGCCTGTGTTTCTGCCATATATTTTATTATAAATTACAATCCACTCA
TCCACCTGCCCTCCACCAGGAGTGGGCACCCCATAAGGGTTTAGGCCACTTTGCAGAGGATGGAGGTCAAA
AACCACTCCCAGATAAGTTTGGTTTTCAACATTTAGTAACTTGTCTCAGGGCAGAGGGCAGGCAGGGGGAC
CGAGGGGCAGCAGATAGGAGAGCACTGAGCCCGGATAGTTCTCAGCCTGGCAAGTGGCTCTGAAGCTGCCT
TCAGACAAGGCTAGTCTAGGGGCAAGAGTGCGAGCTGGCTGACAATAAGAACGTGGCCACCTGCCCAGCTT
CACACCTCCCCCGACTTCAGCCCTTCCTAAACCCAGACCTGCGGTCCAGGCAGGCACTGGGCTGTGCCCAC
TCGAGCTCACTGCCCACACACAGCATGCCTTTGGGTGCCATCTCTTTGCCCAAGCCTGGAAGCCTTGGCAG
GTGGGAAATGCCGCTGCCCTGGTGGGCATGGCACTGAGATGCATCCACTCAGCAGGAGTGACAGAGGCAGA
AGTTCCTTTAAAGCACATCTTCCACTTAGGAAAGGAAGGAAATCTTTGTACTGTCTTGGAAGCCTCCACAT
CCGGCTATGGCCCTGCAAGCTGCTTTATCCCTGCGCTAGTCTCCCCCGAGGGTTTAGGCTGGCCCAGCACA
TCCTGTCCTCCTGAGCTCGCGTGCAGCCACCCAGAGCGCAGGGGTCACTGCACGCTGCAGGGCTCTTGCTG
CCATGGTCTCAAGCCTGAAGAGGCTCCGCCCACAAGCTGGCCCATGAAGTTAGCAATGCCTGTGGCTTCAG
TCAATTGTCTTGAGACTGTGAAGAGGCTGAAAGACACCTTCCCGGGTGGAAGAAGGAGTTCACTGAAAACT
TATCTTAAACTGACCCTTCCCTTTGAGTGAGTCTTCATTCCTCTCCCATGTGGGAACCCAGCCTCCGATGC
CCCGGGGACTAGGGGAAACAGTTGGAGGTTCGTGCCGTCCCCAGCCTGCCACGGGTGCGAGGACAGCCAAG
TCCTGAGTGACTCAAGATGCTTCACTTACATGGAAGAAACTTCTAAAACTCTACCGAGTGGTTTTTGTATA
TACTAAAGTTCTATTTAGAGCTTTTCTGTTTTGGGCAAGTTCGCTGCTCCTTCTATTTGGGCACTTTGGTT
TTTGTACTGTCTTTTGTGACGGCATTGATTGAACATTTTTTACTAGTAGTCTTATGACTTTTGTATTTTTT
TTTTTTTTTGTAATTTATACCAACAACACTTTTATCACTTTTTTTTTGTTGGGCTTCTGCAAAATACAAGC
TCATTTTTAAACCAAATGAACAGACCATGAGCTGGCTTCAGGGGAAGTGCTATTCACAGGACCATATCCAC
CACCCTCTTAAATTCCTAAACAATATCATCTAGGACTTCTATTTAAGTTATTTAAAATAAATCTTCCTTGA
GAGCCTTGGGAGGTGATGTCAGGGTTATAAATGGCACAGTGCATTTGCTGTAGGAATGTGGTTTGGCATTG
TTTTATACACACAGTATTTTTTATACCTTAATGCTTATTCTTGATGGCATCTGTCAGATATTAGAATTGAA
AATAAGAATCTTCCCAAAATCCTTTAATTTACCTGATGCCCTCATCAGGTCGTTAAAAATTCAAATGGTTT
TAATAGCTAAAAAACTACAAATTAAGCTCTAAAACAAACAAACTACAGAAATGTAAACCTTCATTTGCCAA
AGGTCCTTGGTGGCCTGTCCCCTGCCCTGGGAGCAGATGGCCCTGAAGCCCTTCCCTCACTGTGCAGGCCA
CCGGGTGAGGCTGGACGGTCACCCATGGTGGCTTCACTGCAAGGAGCAGGACTGCCGAGCTCAAGCACGGG
GCCTTCAGCTTCCCCTGTCCTCTGGCCACACCGCCAGCCCTTGGTCCTTATCTGTGTGAGGTTTACAAATA
AAGCTTCTGATGTCAAATGTTTAAAAAAAAAAAAAAAAAA

TALACIFQCCFCQRGQQSHGSSRTSPELSEECPGEMVKMTKSKTFQAYLPNCHRTYSCIHCRAHLANHDEL
ISKSFQGSQGRAYLFNSVVNVGCGPAEERVLLTGLHAVADIYCENCKTTLGWKYEHAFESSQKYKEGKFII
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
CCGGGCGTCCCCGCGCGGGCGCCGGAGAGGAAGGAAGGCTGGCAGCCTCGTCACGTGTCCGCTGCAGTCGC
GAAACAGTTCCGGTGGTGAGGAGACCTTTCCAAATATAAGAGGAATAAAGAAGTCACCTCCCCAGCTGTCA
TCATCTTCCAGCAGATTGAGCAAGAATATTTTGAGCACTACAGGAAAGACAGTCCATCAAACCCGAGATGA
TGATCAGCCACGTGATTTTTTCAAGAAGAGGAATAGGGTGAATGAATCTCATCAGAAAAGCAGCAATATGA
ATGCTGGCCCATCTTGGAATAAAGTGCAACATTCAAAGAATTCTTCAGGAAAAAGGCAGAGTAAATCCCAA
GTACCCCACGCTTCTTCCCAGCCGAGAAGCAGCCTCACAGCTGTCACCCAGCCTACTGAAGAAAAACTTAA
AGAAAGCATTTCCCCGGAAGCAAGACGCAAAAGGAATCCACTCGGTTCCAGGTGTCAGGGGGCCTCAGGGA
ATAAACTGTTTCTTGATTTTCAGTCAATGAAAATTATTAAAGAGAATGCTGATGAGGACAGTGCAAGTGAT
CTCTCTGATTCGGAAAGAATTCCCATTCCTCCTTCTCCCCTCACACCTCCAGATCTCAATCTTCGAGCTGA
AGAAATTGATCCAGTTTACTTTGATCTTCACCCTGGTCAGGGCCATACAAAACCTGAATACTATTATCCTA
ATTTCCTTCCATCCCCTTTCAGCTCCTGGGACCTACGAGATATGGCCCTGCTTCTGAACGCAGAGAACAAA
ACGGAAGCCGTGCCCCGAGTGGGAGGACTTCTTGGGAAGTATATCGATAGACTTATTCAGCTTGAGTGGCT
GCAAGTCCAGACTGTACAGTGTGAAAAAGCAAAGGGGGGCAAAGCAAGGCCCCCCACTGCCCCTGGGACCT
CAGGGGCACTGAAAAGCCCTGGGAGAAGTAAGCTAATTGCTAGTGCTCTGTCCAAGCCACTACCTCACCAG
GAAGGGGCTTCAAAGTCAGGCCCTTCCCGAAAGAAAGCTTTTCACCATGAAGAAATCCACCCATCACATTA
TGCATTTGAGACTTCCCCTAGACCCATTGATGTGCTTGGTGGTACCAGGTTTTGTTCTCAGAGGCAAACCC
TTGAAATGAGGACAGAAGAAAAGAAAAAGAAATCAAGTAAGAGTACGAAGCTGCAGCGCTGGGATCTGTCC
GGCAGTGGAAGCAGCTCTAAGGTGGAAACCAGCGGTCACATTCGAGTTCCCAAACAGGCAGCTGTGATTCT
GGACTCAGCAGATTCCTGTAAGGCCTCCAAAACACAAGCACATGCACATCCTAGGAAAAAGGGAAAGGCAG
AGAGCTGTGGTCATGCCACTGTATCGAGTGAGAAAAAACTGAAAACAAACGGAGTAAAGCAAAACACATAT
AAACTAAAATAAATATCTAAAATGCTGAATTGCCAAGACCTGCAGGTACCTCAATGTTAGAGCGCTTCCAA
AAGTCAAAATACTGTGAATTTTAAGGAATTTTACAAATACTGACATTTAAGTAGTTGACTGGCATTTTTGT
CCACCTTTATTTCTACCCTGAGTGGGGTTATTTTCAAAGGGAAGTGTCTTTCAATAAGCCTTTCTTTGTAT
TGTCAGTCTTAGGCAAATGAGAGCCCTTTAGATAAAAATTATGTAAAATATGTGCCATATAAAGGAATAAA
ATGGCACCTCTCCAGGGAAAGTGTCAGTGAAACCTCAGCTACAGTAGCCGGTCTGTGTAGAGCAGCTAGTG
GTGTTACCTCCCCATTTTCACATGCACGTAAGTATATGAAATAGTGCAGACTGTTTCAAATGGTGTGGAAT
CCTAAATGTTTAAAATAAGGTCCTTCTTGCCCACTCCCTCGCTTACTTTTTTATAAACTCCTCAAGCAAAA
TTTCTGTTCATTTTACCCTTAGGAGAAGCTTTAGTTCTTCCTCAAGTCAGGGAGTAGTGAGTTTGTATTTT
GAGTAGTCATTTCTCACTAAGCTGGTTGCTTTCTAGAGAGACAGTGGAATCTAGTACTTTAATACATTTTC
TCTGACATGGTTTTTTTTTTCTTTTTTGAGGGGCATTTTAAACTTAGAGGTGGTGGTAAAACCTACTTTTG
AGTTCTCCGAACTGAGGTTAAAATAACTTGCAGAATTTTCCAAAGTCAATGGGCTTAGCATGATTACTGCT
GTTTGGTGGGGCTGAGAATGAAATATTTGACATTCTGGAATTGCTGGCATGTAAAGCTTCTCCAGAGAGGC
ACCCCAGGGAAATCACTCTTTACAATTTGTAAAGGAAGGGCCTGTAAAAGGATCAAAACACATGGACCTAC
ATTCAGTGTAATAGTTACAAAGTTACTGATTTGGGTTCCACACCCTGTGGTCCTTAGTCAAAAATAATGAT
CTGTTTCAGTTTGCAAGAGCAGGATTTTATTATTTTGCTTGGGGTGAGGGGCGGGAGAGTGGAATATGAGT
AAGGTTGCTGAATGAATTCTAAACTCGCTTATCTGGTCTTCAGGCTTCCCAACTCTCTCCAAGCCTTCTTA
TTTCACTGCAGTTAAATAACATCTTCTTGTTCCTATAGTTGTGCTGTGAGTTTTCTGTTCATATTTGCGCA
GTGTATTTTAATACGGCCCATGTCATTATAGTTGATTTTATCCCTTTAAACAATTACTGTATTTGTTTTTG
ACGTAGAGGTTTCAATTTTTTCACCTTGGGGGCAAATGAAAAACTTGGCATTTTTCATTTGGGAACATATA

ATAGCTTGTAAACTTTTCAGACAGCAGTAAATGTCTGAAAAAATATCAAAAACAGCATAAAGACAAGATTA
TGTAGCTCTAATTATACGTATATAATTATAAAAAACAATGTGCAAGGGTTATATTTTAAGGTCTTTTAAAA
TCTGATTTTGATCATACCAAATGACATAATATTTTTTATGGTAGCCTTTTACTTTCAAGACTTAATTTTCA
GACTTGTACAAGTTCCTTCTTACATTCTTTCCCTCTCACACCATCCTACTGGAGAAAGCATACTTTTATGC
TAAGATCTTACTTTAAGCTTTTTATGTGAACAAAAGATGTACATATAGTAAGTATTACTTCCGTAGTCCTC
AAATTTACTATAACTTTTGTACTTAGTATATGTTTTATATTTGGAAAACAGCACTACGCTTAGTTTTCCTG
TAGTTCCTGAGTGATGTCTGTGTGTTCCTTGCCTGCCCTTTTTTGTGAGCACAGATTAGTCTGTTATCCAT
GGCTGGCACTTCACTTATGATCCTTTCTCTGCTAGATTTTTATGCAGCTCTCTATGAAGTTTCATGGCCCA
TAGATATTCAAAAGCAAGATATTCTATACATATGTGTATATGTATATATACTCCTTATGTTAATACTAAAG
TGTTTATGCTGAGTTGCTGCCTTTCCCCGTCATGTATCCATGTGCATGCTCTTAGAGACCTTGAATGGTTG
AGGGTAAAGTGATTTATTAGTAATTCTACTTGCCTTGTGTATGTCTGAGCTGAAAACAATCGTGATTAAGA
AATTTAGAGGTGGCTGGGCGTGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCGGA
TCACCTGAGGTCGGGAGTTCAAGACCAGCCTGACCAACATGGAGAAACCCTGTCTCCACTAAAAATACAAA
ATTAGCCGGGTGTGGTGGTGCATGCCTGTAATCCCAGCTACTCGGAAGTTTGAGACGAGAATCTCTTGAAC
CCGGGAGGCGGAGGTTGTGGTGAGCCAAGATCGTGCCATTGCACTCCAGCCTGGGCAACAAGAGGGAAACT
CCGTTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

IRGIKKSPPQLSSSSSRLSKNILSTTGKTVHQTRDDDQPRDFFKKRNRVNESHQKSSNMNAGPSWNKVQHS
KNSSGKRQSKSQVPHASSQPRSSLTAVTQPTEEKLKESISPEARRKRNPLGSRCQGASGNKLFLDFQSMKI
IKENADEDSASDLSDSERIPIPPSPLTPPDLNLRAEEIDPVYFDLHPGQGHTKPEYYYPNFLPSPFSSWDL
RDMALLLNAENKTEAVPRVGGLLGKYIDRLIQLEWLQVQTVQCEKAKGGKARPPTAPGTSGALKSPGRSKL
IASALSKPLPHQEGASKSGPSRKKAFHHEEIHPSHYAFETSPRPIDVLGGTRFCSQRQTLEMRTEEKKKKS
SKSTKLQRWDLSGSGSSSKVETSGHIRVPKQAAVILDSADSCKASKTQAHAHPRKKGKAESCGHATVSSEK
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

ACAGCTCATTGTTGGCAGCTGCCGGGCGGTCCTGCCGAGCTGTGAGGGCAACGGAGGGGAAATAAAAGGGA
ACGGCTCCGAATCTGCCCCAGCGGCCGCTGCGAGACCTCGGCGCCGACATCGCGACAGCGAAGCGCTTTGC
ACGCCAGGAAGGTCCCCTCTATGTGCTGCTGAGCCGGTCCTGGACGCGACGAGCCCGCCCTCGGTCTTCGG
AGCAGAAATCGCAAAAACGGAAGGACTGGAAATGGCAGACCATATGATGGCCATGAACCACGGGCGCTTCC
CCGACGGCACCAATGGGCTGCACCATCACCCTGCCCACCGCATGGGCATGGGGCAGTTCCCGAGCCCCCAT
CACCACCAGCAGCAGCAGCCCCAGCACGCCTTCAACGCCCTAATGGGCGAGCACATACACTACGGCGCGGG
CAACATGAATGCCACGAGCGGCATCAGGCATGCGATGGGGCCGGGGACTGTGAACGGAGGGCACCCCCCGA
GCGCGCTGGCCCCCGCGGCCAGGTTTAACAACTCCCAGTTCATGGGTCCCCCGGTGGCCAGCCAGGGAGGC
TCCCTGCCGGCCAGCATGCAGCTGCAGAAGCTCAACAACCAGTATTTCAACCATCACCCCTACCCCCACAA
CCACTACATGCCGGATTTGCACCCTGCTGCAGGCCACCAGATGAACGGGACAAACCAGCACTTCCGAGATT
GCAACCCCAAGCACAGCGGCGGCAGCAGCACCCCCGGCGGCTCGGGCGGCAGCAGCACCCCCGGCGGCTCT
GGCAGCAGCTCGGGCGGCGGCGCGGGCAGCAGCAACAGCGGCGGCGGCAGCGGCAGCGGCAACATGCCCGC
CTCCGTGGCCCACGTCCCCGCTGCAATGCTGCCGCCCAATGTCATAGACACTGATTTCATCGACGAGGAAG
TTCTTATGTCCTTGGTGATAGAAATGGGTTTGGACCGCATCAAGGAGCTGCCCGAACTCTGGCTGGGGCAA
AACGAGTTTGATTTTATGACGGACTTCGTGTGCAAACAGCAGCCCAGCAGAGTGAGCTGTTGACTCGATCG
AAACCCCGGCGAAAGAAATCAAACCCCCAACTTCTTCGGCGTGAATTAAAAGAAACATTCCCTTAGACACA
GTATCTCACTTTTCAGATCTTGAAAGGTTTGAGAACTTGGAAACAAAGTAAACTATAAACTTGTACAAATT
GGTTTTAAAAAAAATTGCTGCCACTTTTTTTTCCTGTTTTTGTTTCGTTTTTGTAGCCTTGACATTCACCC

ACCTCCCTTATGTAGTTGAAATATCTAGCTAACTTGGTCTTTTTCGTTGTTTGTTTTTACTCCTTTCCCTC
ACTTTCTCCAGTGCTCAACTGTTAGATATTAATCTTGGCAAACTGCTTAATCTTGTGGATTTTGTAGATGG
TTTCAAATGACTGAACTGCATTCAGATTTACGAGTGAAAGGAAAAATTGCATTAGTTGGTTGCATGAACTT
CGAAGGGCAGATATTACTGCACAAACTGCCATCTCGCTTCATTTTTTTAACTATGCATTTGAGTACAGACT
AATTTTTAAAATATGCTAAACTGGAAGATTAAACAGATGTGGGCCAAACTGTTCTGGATCAGGAAAGTCAT
ACTGTTCACTTTCAAGTTGGCTGTCCCCCCCGCCGCCCCCCCCACCCCCATATGTACAGATGATAATAGGG
TGTGGAATGTCGTCAGTGGCAAACATTTCACAGATTTTTATTTTGTTTCTGTCTTCAACATTTTTGACACT
GTGCTAATAGTTATATTCAGTACATGAAAAGATACTACTGTGTTGAAAGCTTTTTAGGAAATTTTGACAGT
ATTTTTGTACAAAACATTTTTTTGAAAAAATACTTGTTAATTTATTCTATTTTAATTTGCCAATGTCAATA
AAAAGTTAAGAAA

LDATSPPSVFGAEIAKTEGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFPSPHHHQQQQPQHAFNA
LMGEHIHYGAGNMNATSGIRHAMGPGTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQKLNN
QYFNHHPYPHNHYMPDLHPAAGHQMNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSGGGAGSSNS
GGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKELPELWLGQNEFDFMTDFVCKQ
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
CAGCGGAAGGCGGCCGCCGCTACGGCGATTCGCAGGGAGTAGCAGACGAAGACGGTGGCCGCCGCACTAGC
CACCACGTGTGGAGGATAAACGGTCTACACGGCCATTCCGGCGCCGAGTCTAGGGAAAGAGTTAGCGACGA
CGGGGAAAGAAATGTGAAGAGAGCGACCGCCGCTCCAGGGTCGCTGCAGGAAGCCTAAGTGCAGACGCCG
GCTTCTCCCGCAGTGACTTGAGAAGGGTCAGTGAAAACCTCGGCCACTGCCGCAGCGTCTCTAGGGAGAGA
GTTAGGGGAGATAGTGGCCACAGTCACAGCTGCTCTTGGGAGAGAGTTAGGGGAGACAGCACCTTCTGCAG
CAGCGACGTGAATTTTAGTGAAGTTGGAGGCCACCAAACTACCGACTCCAGGGGAACAGCCAGAGAAGACC
GAGGCCTCCGCCTCAGTGGTCCTTGGGAGGGAGTCAGTGACATTCGGGACCCGCGAACTAGTGACTTCGGG
GATAGAGTCAGTGACGATCGCAGTCGCCGCTTCAGTGGCTCCTGGGAGGGAGGGAGTGTCGAAGGCGGCCA
CAGCGTTGGTAGTTCTTGGGAGGAAGTAAGTGGAGACCGCGGCTACGCAGCCAGCGACTCCTCTGGTGTGA
GCGGCAGTGAAGACGCCAGCTACCGCTTCAGTGGCTTTTGGGAGAGAGAAAGTGAAGACGAAGGTTTCCGC
TGCAGCTTCTGGGAGAGAGCAAGAGAGGACCTTGGGCCCCGTCCTAGTGACGACGGAGAAGAGGGCCGCTG
CCGCTGCAGTGGCTCGTGGGTGAGAGCAAGTGAAGACCGCCGCAGCATCAGGGGCCTGGACTCAACTCCTC
CCCAGAGTCGGAGGTGTTGCGCCATGCCCGGGGTGGCCAATTCAGGCCCCTCCACTTCCTCTAGGGAGACT
GCAAACCCCTGTTCCAGGAAGAAGGTGCATTTTGGCAGCATACATGATGCAGTACGAGCTGGAGATGTAAA
GCAGCTTTCAGAAATAGTGTGTCTTCATTGGCTGCTCTGGCATGGAGCTGATATCACACACGTAACAACGA
GAGGTTGGACAGCATCTCACATAGCTGCAATCAGGGGTCAGGATGCTTGTGTACAGGCTCTTATAATGAAT
GGAGCAAATCTGACAGCCCAGGATGACCGGGGATGCACTCCTTTACATCTTGCTGCAACTCATGGACATTC
TTTCACTTTACAAATAATGCTCCGAAGTGGAGTGGATCCCAGTGTGACTGATAAGAGAGAATGGAGACCTG
TGCATTATGCAGCTTTTCATGGGCGGCTTGGCTGCTTGCAACTTCTTGTTAAATGGGGTTGTAGCATAGAA
GATGTGGACTACAATGGAAACCTTCCAGTTCACTTAGCAGCCATGGAAGGCCACCTTCACTGTTTCAAATT
CCTAGTCAGTAGAATGAGCAGTGCGACGCAAGTTTTAAAAGCTTTCAATGATAATGGAGAAAATGTACTGG
ATTTGGCCCAGAGGTTCTTCAAGCAGAACATTTTACAGTTTATCCAGGGGGCTGAGTATGAAGGAAAAGAC
CTAGAGGATCAGGAAACTTTAGCATTTCCAGGTCATGTGGCTGCCTTTAAGGGTGATTTGGGGATGCTTAA
GAAATTAGTGGAAGATGGAGTAATCAATATTAATGAGCGTGCTGATAATGGATCAACTCCTATGCATAAAG

CTGCTGGACAAGGCCACATAGAGTGTTTGCAGTGGTTAATTAAAATGGGAGCAGACAGTAATATTACCAAC
AAAGCAGGGGAGAGACCCAGTGATGTGGCAAAGAGGTTTGCCCATTTGGCAGCAGTGAAGCTGTTAGAGGA
GCTACAGAAATATGATATAGATGACGAAAATGAAATTGATGAAAATGATGTGAAATATTTTATAAGACATG
GTGTTGAGGGAAGCACTGATGCCAAGGATGATTTATGTCTGAGTGACTTGGATAAAACAGATGCCAGAAGA
CCATCAAAGAACTGCAGGGCCAGCTGGAGTATGAACGACTACGTAGAGAAAAATTAGAATGTCAGCTTGAT
GAATATCGAGCAGAAGTTGATCAACTCAGGGAAACACTGGAAAAAATTCAAGTCCCAAACTTTGTGGCTAT
GGTTGGTGTCCTTTTTAATACATTTATTTTTCTCAAGAAGTATATACAAGAGTGGCCAAGAGTACAAGCTT
TGGGCTGGGGCTCTTTGGACTTGAATCCTGGCTATAGTCTTTTTTGATTTGGGCTACTCATTTAACCTTTT
TGTGCCTCAACTATAAACTGGAGATGATAACTTATACTTTCTAGAGCTGATGACAGGATTAAAGGAATACA
CACACACACACACACACACACACACACACACACACACACACTCTATATGTAACTATAACTTTCTAGATACATACA
CACATCTGTCTATCTATCTTCAAACAGCAGCCTAGTTCATAAGTATTATTGTTAACATTGTACCATAGAAG
GAGAAAATGTTTTCATAAGTAATCAATCAGAATTCTAAGACAGCTAGAGGATATGTATATTTTAATATTCT
AAATAACCAAAATAAAGTGCTTTGAATGGAATCCATATTTTCTTTCCATAGGGAAGTTTCTTCATCAATCA
TCATGGATGAATTAATTCTGTTTGAGGCTTGTGTCACCTCAAATCTGATCTATTAATATTATAGAATCTTC
CTGGCTTCTTTTTATTTTGCACAAACTAGTCATTGGTCTCTTTAATAACCACTTTAAAAATATTAAAATAA
TAAAACATGTTTGTTGTATAGTGTTAAAAACAAGATAGATGTTCCTTCTGACACTAAGTTCCACTCTCCAG
AGGAAGCTATTGTTAACAATTTAGTACATACCATATTAGACAAAATTCTATTTATCTGCAAACATGTATGT
GTATTTCTATGCATATGCAAATATAACCACTTCCTCAGTTTCTCTGGGTTTTGGTGTCTCACCAGTGGTTG
ATTTTCCAAGAAATATAATTAACTATAATAATT

LERDSEKVPAEGGRRYGDSQGVADEDGGRRTSHHVWRINGLHGHSGAESRERVSDDGERKCEESDRRSRVA
AGSLSADAGFSRSDLRRVSENLGHCRSVSRERVRGDSGHSHSCSWERVRGDSTFCSSDVNFSEVGGHQTTD
SRGTAREDRGLRLSGPWEGVSDIRDPRTSDFGDRVSDDRSRRFSGSWEGGSVEGGHSVGSSWEEVSGDRGY
AASDSSGVSGSEDASYRFSGFWERESEDEGFRCSFWERAREDLGPRPSDDGEEGRCRCSGSWVRASEDRRS
IRGLDSTPPQSRRCCAMPGVANSGPSTSSRETANPCSRKKVHFGSIHDAVRAGDVKQLSEIVCLHWLLWHG
ADITHVTTRGWTASHIAAIRGQDACVQALIMNGANLTAQDDRGCTPLHLAATHGHSFTLQIMLRSGVDPSV
TDKREWRPVHYAAFHGRLGCLQLLVKWGCSIEDVDYNGNLPVHLAAMEGHLHCFKFLVSRMSSATQVLKAF
NDNGENVLDLAQRFFKQNILQFIQGAEYEGKDLEDQETLAFPGHVAAFKGDLGMLKKLVEDGVININERAD
NGSTPMHKAAGQGHIECLQWLIKMGADSNITNKAGERPSDVAKRFAHLAAVKLLEELQKYDIDDENEIDEN
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

TAGGAGCTGCTCCTTCATCACTGGAGTCGCCCCTACCTCTCTGCCCCCAGCCCGAGAGCCCCAGGCGGGGA
CCCCCGGATCGGACGTCCCCAAGCCTCCGGGCACCTGGCTCAGCAGGAGGCCCCCGGCTCGGGGCAGGGCA
GGGCCGGCGGCGGCGAGCCGGAGCCCGCCCCCTGCCCGGGCCCCGCCGAGCCCTCGGAGCCCACCCATGGG
GCACCTGCCCCTTGCGCCTCCTTGCCCGGCCGCGCCCAGCCCGGCGTCCCGAGCAGCGCAGGGGAGGATCC
CCGCGCAGTGACCCGGGAGCCACCACAGACTCTGGGAGGCTCGGCGGCTGGAGCAGCAGGCAGCTCCCCGC
AGCTCCCGGCGCTTCCAGGCAGCTCTCTGAGCCGTGCCAGAGGCCCGGCCCGCCATTCCCAGCCCCGAGCC
ATGATGAAGACTTTGTCCAGCGGGAACTGCACGCTCAGTGTGCCCGCCAAAAACTCATACCGCATGGTGGT
GCTGGGTGCCTCTCGGGTGGGCAAGAGCTCCATCGTGTCTCGCTTCCTCAATGGCCGCTTTGAGGACCAGT
ACACACCCACCATCGAGGACTTCCACCGTAAGGTATACAACATCCGCGGCGACATGTACCAGCTCGACATC
CTGGATACCTCTGGCAACCACCCCTTCCCCGCCATGCGCAGGCTGTCCATCCTCACAGGGGATGTCTTCAT
CCTGGTGTTCAGCCTGGATAACCGGGAGTCCTTCGATGAGGTCAAGCGCCTTCAGAAGCAGATCCTGGAGG
TCAAGTCCTGCCTGAAGAACAAGACCAAGGAGGCGGCGGAGCTGCCCATGGTCATCTGTGGCAACAAGAAC
GACCACGGCGAGCTGTGCCGCCAGGTGCCCACCACCGAGGCCGAGCTGCTGGTGTCGGGCGACGAGAACTG

CGCCTACTTCGAGGTGTCGGCCAAGAAGAACACCAACGTGGACGAGATGTTCTACGTGCTCTTCAGCATGG
CCAAGCTGCCACACGAGATGAGCCCCGCCCTGCATCGCAAGATCTCCGTGCAGTACGGTGACGCCTTCCAC
CCCAGGCCCTTCTGCATGCGCCGCGTCAAGGAGATGGACGCCTATGGCATGGTCTCGCCCTTCGCCCGCCG
CCCCAGCGTCAACAGTGACCTCAAGTACATCAAGGCCAAGGTCCTTCGGGAAGGCCAGGCCCGTGAGAGGG
ACAAGTGCACCATCCAGTGAGCGAGGGATGCTGGGGCGGGGCTTGGCCAGTGCCTTCAGGGAGGTGGCCCC
AGATGCCCACTGTGCGCATCTCCCCACCGAGGCCCCGGCAGCAGTCTTGTTCACAGACCTTAGGCACCAGA
CTGGAGGCCCCCGGGCGCTGGCCTCCGCACATTCGTCTGCCTTCTCACAGCTTTCCTGAGTCCGCTTGTCC
ACAGCTCCTTGGTGGTTTCATCTCCTCTGTGGGAGGACACATCTCTGCAGCCTCAAGAGTTAGGCAGAGAC
TCAAGTTACACCTTCCTCTCCTGGGGTTGGAAGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCAT
ATGGAGCCTCCTGGGACAAGCCTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAGGTCTCCTCTCTC
CTGGCATAACACCCAGCTTGGTTTGGGTGGCAGCTGGGAGAACTTCTCTCCCAGCCCTGCAACTCTTACGC
TCTGGTTCAGCTGCCTCTGCACCCCCTCCCACCCCCAGCACACACACAAGTTGGCCCCCAGCTGCGCCTGA
CATTGAGCCAGTGGACTCTGTGTCTGAAGGGGGCGTGGCCACACCTCCTAGACCACGCCCACCACTTAGAC
CACGCCCACCTCCTGACCGCGTTCCTCAGCCTCCTCTCCTAGGTCCCTCCGCCCGACAGTTGTGCTTTGTT
GTGGTTGCAGCTGTTTTCGTGTCATGTATAGTAGTAGAAATGGAAATCATTGTACTGTAAAAGCCTAGTGA
CTCCCTCCTTGGCCAGGCCCTCACCCAGTTCAGATCCACGGCCTCCACCCGGGACGCCTTCCTCCTCTGCT
CCCAAACAGGGTTTCCGTGGCCTGTTTGCAGCTAGACATTGACCTCCGCCATTGAGCTCCACGGTTTACAG
ACAATTGCACAAGCGTGGGGTGGGCAGGCCAGGACTGCTTTTTTTTAATGCTCCCATTTCACAGAGGATAC
CACCGAGACTCGGAGGGGACACGATGAGCACCAGGCCCCACCTTTGTCCCCTAGCAAATTCAGGGTACAGC
TCCACCTAGAACCAGGCTGCCCTCTACTGTGCTCGTTCCTCAAGCATTTATTAAGCACCTACTGGGTGCTG
GGTTCACTGTGTCCTAGGAAACCAAGAGGGTCCCCAGTCCTGGCCTCTGCCCGCCCCTGCTGCCCCACCAC
CTTCTGCACACACAGCGGTGGGGAGGCGGGGAGGAGCAGCTGGGACCCAGAACTGAGCCTGGGAGGGATCC
GACAGAAAAGCTCAGGGCGGGTCTTCTCCTTGTGCCCGGGATTGGGCTATGCTGGGTACCACCATGTACTC
AGGCATGGTGGGTTTTGAACCCATAAACCAAAGGCCCTTGTCATCAGCTCTTAACAAGTATATTTTGTATT
TTAATCTCTCTAAACATATTGAAGTTTTAGGGCCCTAAGGAACCTTAGTGATCTTCTATTGGGTCTTTCTG
AGGTTCAGAGAGGGTAAGTAACTTCCTCCAGGTCACACAGCAAGTCTGTGGGTGGCAGAAGCAAGCTAGCG
CTGGGCATTCAGTACATACCACGATGTGCTCCCTCTCTTGATGCTTGGCCCCTGGGGCCTTCAGGGCTTTG
GGACATCTTGTCCTCAACCCTCTCCCTAGATCAGTCTGTGAGGGTCCCTGTAGATATTGTGTACACCATGC
CCATGTATATACAAGTACACACAGATGTACACACAGATGTACACATGCTCCAGCCCCAGCTCTGCATACCT
GCACCTGCACCCCAGCCTTGGCCCCTGCCTGCGTCTGTGCTCAAAGCAGCAGCTCCAACCCTGCCTCTGTC
CCCTTCCCCACCCACTGCCTGAGCCTTCTGAGCAGACCAGGTACCTTGGCTGCACCGGTGTGTGGCCCGCT
CTCACCCAGGCACAGCCCCGCCACCATGGATCTCCGTGTACACTATCAATAAAAGTGGGTTTGTTACAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LAQQEAPGSGQGRAGGGEPEPAPCPGPAEPSEPTHGAPAPCASLPGRAQPGVPSSAGEDPRAVTREPPQTL
GGSAAGAAGSSPQLPALPGSSLSRARGPARHSQPRAMMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSI
VSRFLNGRFEDQYTPTIEDFHRKVYNIRGDMYQLDILDTSGNHPFPAMRRLSILTGDVFILVFSLDNRESF
DEVKRLQKQILEVKSCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENCAYFEVSAKKNT
NVDEMFYVLFSMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSVNSDLKYIK
AKVLREGQARERDKCTIQ

AUG initiation = 30.4 kDa
CUG initiation = 40.6 kDa

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

#31

NM_002506

UUG/CUG initiation in mammals and most tetrapods

NGF nerve growth factor (beta polypeptide)

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

CGGACGCGTGGGCGCTGGGAGCCGGAGGGGAGCGCAGCGAGTTTTGGCCAGTGGTCGTGCAGTCCAAGGGG
CTGGATGGCATGCTGGACCCAAGCTCAGCTCAGCGTCCGGACCCAATAACAGTTTTACCAAGGGAGCAGCT
TTCTATCCTGGCCACACTGAGGTGCATAGCGTAATGTCCATGTTGTTCTACACTCTGATCACAGCTTTTCT

GATCGGCATACAGGCGGAACCACACTCAGAGAGCAATGTCCCTGCAGGACACACCATCCCCCAAGTCCACT
GGACTAAACTTCAGCATTCCCTTGACACTGCCCTTCGCAGAGCCCGCAGCGCCCCGGCAGCGGCGATAGCT
GCACGCGTGGCGGGGCAGACCCGCAACATTACTGTGGACCCCAGGCTGTTTAAAAAGCGGCGACTCCGTTC
ACCCCGTGTGCTGTTTAGCACCCAGCCTCCCCGTGAAGCTGCAGACACTCAGGATCTGGACTTCGAGGTCG
GTGGTGCTGCCCCCTTCAACAGGACTCACAGGAGCAAGCGGTCATCATCCCATCCCATCTTCCACAGGGGC
GAATTCTCGGTGTGTGACAGTGTCAGCGTGTGGGTTGGGGATAAGACCACCGCCACAGACATCAAGGGCAA
GGAGGTGATGGTGGTTGGGAGAAGTGAACATTAACAACAGTGTATTCAAACAGTACTTTTTTGAGACCAAG
TGGCGGGGACCCAAATCCCGTTGACAGCGGGTGCCGGGGCATTGACTCAAAGCACTGGAACTCATATTGTA
CCACGACTCACACCTTTGTCAAGGCGCTTTACCATGGGATGGCAAGCAGGGATACCCTGGCCGGGTTATCC
CGGATGAGATACGGGCCGGGGAAAAGTGGGGCTCAACAAGGGAGGATGAAGAGAAGGAGGCTAAAGGCTGG
CGGAAGGACATCCCCTTCCCCCTGGGGCCATTCAAAAACTTCTGCGGGGGGCCCCTCCCTTAACCGGGAAC
CTGGTAAAATGGATTTTTTAAGGGTGGTAA

LASGRAVQGAGWHAGPKLSSASGPNNSFTKGAAFYPGHTEVHSVMSMLFYTLITAFLIGIQAEPHSESNVP
AGHTIPQVHWTKLQHSLDTALRRARSAPAAAIAARVAGQTRNITVDPRLFKKRRLRSPRVLFSTQPPREAA
DTQDLDFEVGGAAPFNRTHRSKRSSSHPIFHRGEFSVCDSVSVWVGDKTTATDIKGKEVMVVGRSEH

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

GTGAGTGATGATGATAAGCTCCCGCTGCAGGTGTGAATAGAGACCCCGGAGGCGCGTCCTAGCCCTCATCT
GGGGAAGCGCACCTGCATACAGACGGGTGCACCGGGGAGGAGGCGATCTGCCGCGTGTTCCTGCAAGCAGA
AAAGGAGTTAACTAGTGTCACATTTGAAGACGAGCACTGAGGATGAGGAACCAACTGAAGAATATGAAAAT
GTTGGAAATGCAGCATCTAAGTGGCCAAAAGTGGAGGATCCTATCCCTGAATCTAAGGTTGGTGACACATG
TGTTTGGGATAGCAAGGTAGAGAATCAACAGAAAAAGCCTGTGGAAAACAGGATGAAGGAGGACAAAAGCA
GCATCAGGGAAGCAATCAGCAAAGCCAAGAGTACAGCAAATATAAAGACAGAACAGGAAGGTGAGGCATCT
GAGAAGAGCTTGCATCTGAGCCCACAGCATATCACACACCAGACTATGCCTATAGGACAGAGAGGCAGTGA
GCAAGGCAAACGTGTGGAGAACATTAATGGAACCTCCTACCCTAGTCTACAGCAGAAAACCAATGCTGTTA
AGAAATTACATAAATGTGATGAATGTGGGAAATCCTTCAAATATAATTCCCGCCTTGTTCAACATAAAATT
ATGCACACTGGGGAAAAGCGCTATGAATGTGATGACTGTGGAGGGACTTTCCGGAGCAGCTCGAGCCTTCG
GGTCCACAAACGGATCCACACTGGGGAGAAGCCGTACAAGTGTGAGGAATGTGGGAAAGCCTACATGTCCT
ACTCCAGCCTTATAAACCACAAAAGCACCCATTCTGGGGAGAAGAACTGTAAATGTGATGAATGTGGAAAA
TCCTTCAATTATAGCTCTGTTCTGGACCAGCATAAAAGGATCCACACTGGGGAGAAGCCCTATGAATGTGG
TGAGTGTGGGAAGGCCTTCAGGAACAGCTCTGGGCTCAGAGTCCACAAAAGGATCCACACGGGGGAGAAGC
CCTATGAATGCGACATCTGTGGGAAAACCTTCAGTAACAGCTCTGGCCTTAGGGTCCATAAAAGGATCCAC
ACAGGTGAGAAACCTTACGAATGTGATGAGTGTGGGAAGGCCTTCATTACTTGTAGAACACTTCTCAACCA
TAAAAGCATCCACTTTGGAGATAAACCCTATAAATGTGATGAGTGTGAGAAATCTTTTAATTATAGCTCTC
TTCTCATTCAGCATAAAGTCATCCACACTGGAGAGAAACCTTATGAATGTGATGAATGTGGGAAGGCTTTC
AGGAACAGCTCAGGCCTCATAGTGCATAAAAGGATCCACACAGGAGAGAAACCTTACAAGTGTGATGTCTG
TGGCAAAGCATTCAGCTATAGCTCAGGCCTCGCAGTCCATAAAAGCATTCACCCTGGGAAGAAAGCCCATG
AATGTAAGGAGTGTGGGAAATCCTTTAGTTATAACTCACTACTTCTTCAACACAGAACTATTCATACCGGA
GAGAGACCTTATGTATGTGATGTGTGTGGGAAAACGTTCAGAAACAATGCAGGCCTCAAAGTCCACAGGAG

GCTCCATACTGGGGAAAAACCATATAAGTGTGATGTGTGTGGGAAAGCCTATATCTCACGCTCTAGCCTTA
AAAATCACAAAGGAATCCACCTTGGGGAGAAGCCCTATAAATGTAGCTATTGTGAGAAATCCTTCAACTAC
AGCTCTGCCCTTGAACAGCATAAAAGGATTCATACCAGGGAAAAACCCTTTGGGTGTGATGAGTGTGGTAA
AGCTTTCAGAAATAATTCTGGCCTTAAAGTACATAAACGAATCCACACTGGGGAACGACCTTACAAATGTG
AAGAATGTGGGAAAGCATACATCTCTCTCTCGAGCCTTATAAATCATAAAAGTGTACACCCTGGGGAGAAG
CCCTTTAAGTGTGACGAGTGTGAGAAGGCCTTCATCACATACCGAACCCTTACAAACCACAAAAAAGTTCA
TCTTGGGGAGAAGCCCTACAAATGTGATGTGTGTGAGAAATCTTTTAATTACACATCGCTCCTTTCTCAGC
ACAGAAGGGTCCACACTAGAGAGAAACCCTATGAATGTGACAGGTGTGAGAAGGTCTTCAGAAACAACTCA
AGCCTTAAAGTTCATAAAAGAATCCATACTGGGGAGAGGCCCTATGAATGTGATGTGTGTGGAAAAGCCTA
CATCTCACACTCAAGCCTTATTAACCATAAGAGTACCCACCCTGGCAAGACACCCCATACATGTGATGAAT
GTGGAAAAGCTTTTTTCTCAAGCAGAACTCTTATAAGCCATAAAAGAGTCCATCTTGGGGAGAAACCCTTC
AAGTGTGTTGAGTGTGGGAAATCTTTCAGTTACAGCTCTCTCCTTTCTCAGCACAAGAGGATCCACACAGG
GGAGAAACCCTATGTGTGTGATAGGTGTGGGAAGGCCTTCAGGAACAGCTCAGGCCTCACAGTGCATAAAA
GGATCCACACAGGTGAGAAACCCTATGAATGTGATGAGTGTGGGAAGGCATACATCTCACACTCAAGTCTT
ATCAATCATAAAAGTGTCCACCAGGGGAAGCAGCCCTATAATTGTGAGTGTGGGAAATCCTTCAATTATAG
ATCAGTCCTTGACCAGCACAAAAGGATCCACACTGGAAAGAAGCCATACCGATGTAATGAGTGTGGTAAGG
CTTTTAATATCAGATCAAATCTCACCAAGCATAAAAGAACCCATACTGGAGAGGAATCTTTAAATGTGATA
TATGTGGGAAGTTATAGTGGCACATCCCAGAAGAGAACCTATGAGGGAGGGAATGCCCTGGATGGGGGCAG
GATGAGGATGCCTCTGTAGCAGGCAGAGCTTACCAAGTCTCTCCGAACTCAAATGGAAGAAATACCTTATG
AATGTAAGAATGTAGGGGGTCATGGCTTGTAATTTACACAGTGTAAATGAAACCATCCTAGAGGATTATGA
GGAATCCTTTCTATGTGATTTTCAATCATAGCAAGCAAGAAAGGCTCCAGTGTCAAGGTAGTTCAGCTCTT
ACAGGATATAAAACAGTCCATACTTGAGAGAAAAACTTAGATCTGAGTGATGGAATGTGAAGCAAATCTTC
AAAATCAGTAGACATTTCTGGACATAAAACACAGATGAGGAAAGGGCTTCAATTAGAAGTTACGTAATCAC
CATCAGAAAGTTCATGTTTGGTAAATTCTGTTACTAGAAATGTAGGAAATTCAGGTATAGCTTTGAATCCC
AATTACACATTGGTCAGTGGGAAAACTAAGGGCCTCCAACAGGCAAATTCAGGGAGGATAGGTTTCAGGGA
ATATAAATTTATTTAATATTAGTGGTCTTTAAGTATAAACTTGATGTAATTGGTTTGGGAGGGGGCAGTGA
TGATGACTTCTGAAACAAAATTTGGATTTCCTTTTAGGAAAAGTAGAAAGCATAGACTTACAAGTCTAACA
GGAGATAGGAGAGAGTCACTCATAAAAAATGCAAATTGATGAACGTACTATTGTGATACATTAGTTGAATG
GATGAAACTTTTTTAAAGTTTCAGATGAACTCCCATAATGAATGATGAATTTGTGATGAGGGATAACCTGG
AAGTGGTATTCACACATTATGCTACAATAAAAGGTTCTACCGTGGAGAGGATTTTGACACATTCAGTAACT
AATGGAACACACCGTCAACATGAATTCGCACCTTACATGACAGAAGTGATTCAGGGATTCCTATGAATAGA
AATGCTGAGAAGGAACGCATTTTATTGCAGAAGCTAAAAAGCTAAAGTACCAGTCATCTAGAGAGAAGGAA
ATTAATGTTTCTTAATAATCCTGTTAAATGTTTGATTGTTTTTGGAATGTGTTATTGTAAAGATGTCATGC
AGGACATGTATATGTTGTCTGTTGTAAAATGTTAACGAATACTTTGTTCAGGGCTCACTCTCTCTTTGTCA
TGAAAGCCAGCTCCTTGTGGCGAGGTAAAGTGGAATTCCAATAAAGAAATTCCTTAAATCAAAA

VSHLKTSTEDEEPTEEYENVGNAASKWPKVEDPIPESKVGDTCVWDSKVENQQKKPVENRMKEDKSSIREA
ISKAKSTANIKTEQEGEASEKSLHLSPQHITHQTMPIGQRGSEQGKRVENINGTSYPSLQQKTNAVKKLHK
CDECGKSFKYNSRLVQHKIMHTGEKRYECDDCGGTFRSSSSLRVHKRIHTGEKPYKCEECGKAYMSYSSLI
NHKSTHSGEKNCKCDECGKSFNYSSVLDQHKRIHTGEKPYECGECGKAFRNSSGLRVHKRIHTGEKPYECD
ICGKTFSNSSGLRVHKRIHTGEKPYECDECGKAFITCRTLLNHKSIHFGDKPYKCDECEKSFNYSSLLIQH
KVIHTGEKPYECDECGKAFRNSSGLIVHKRIHTGEKPYKCDVCGKAFSYSSGLAVHKSIHPGKKAHECKEC
GKSFSYNSLLLQHRTIHTGERPYVCDVCGKTFRNNAGLKVHRRLHTGEKPYKCDVCGKAYISRSSLKNHKG
IHLGEKPYKCSYCEKSFNYSSALEQHKRIHTREKPFGCDECGKAFRNNSGLKVHKRIHTGERPYKCEECGK
AYISLSSLINHKSVHPGEKPFKCDECEKAFITYRTLTNHKKVHLGEKPYKCDVCEKSFNYTSLLSQHRRVH
TREKPYECDRCEKVFRNNSSLKVHKRIHTGERPYECDVCGKAYISHSSLINHKSTHPGKTPHTCDECGKAF
FSSRTLISHKRVHLGEKPFKCVECGKSFSYSSLLSQHKRIHTGEKPYVCDRCGKAFRNSSGLTVHKRIHTG
EKPYECDECGKAYISHSSLINHKSVHQGKQPYNCECGKSFNYRSVLDQHKRIHTGKKPYRCNECGKAFNIR
SNLTKHKRTHTGEESLNVIYVGSYSGTSQKRTYEGGNALDGGRMRMPL

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

AGTTGAAGCTCCTCTCCCTTCCGAACAGCTCCGCGCACCGCCCCGCGACGCAGCCCGGCGCAACTACTTTC
TTCTCTCTCCTTTCTTTCTTCCTCTCCTTTTTCCCCTGCTGGGTAGTGGCTGCGGCGGGGTGGGGGAGACT
TTGAATGACCGAGCTCGCGTCCACCTTTCTCTTCATGTCGACGTCCCTGGAAACGGCCACACGGATGCCAT
GGTTACTTTTGCCACGATCTTACAGGTGAACAAGGTGATGTCCATCTTGTTTTATGTGATATTTCTCGCTT
ATCTCCGTGGCATCCAAGGTAACAACATGGATCAAAGGAGTTTGCCAGAAGACTCGCTCAATTCCCTCATT
ATTAAGCTGATCCAGGCAGATATTTTGAAAAACAAGCTCTCCAAGCAGATGGTGGACGTTAAGGAAAATTA
CCAGAGCACCCTGCCCAAAGCTGAGGCTCCCCGAGAGCCGGAGCGGGGAGGGCCCGCCAAGTCAGCATTCC
AGCCGGTGATTGCAATGGACACCGAACTGCTGCGACAACAGAGACGCTACAACTCACCGCGGGTCCTGCTG
AGCGACAGCACCCCCTTGGAGCCCCCGCCCTTGTATCTCATGGAGGATTACGTGGGCAGCCCCGTGGTGGC
GAACAGAACATCACGGCGGAAACGGTACGCGGAGCATAAGAGTCACCGAGGGGAGTACTCGGTATGTGACA
GTGAGAGTCTGTGGGTGACCGACAAGTCATCGGCCATCGACATTCGGGGACACCAGGTCACGGTGCTGGGG
GAGATCAAAACGGGCAACTCTCCCGTCAAACAATATTTTTATGAAACGCGATGTAAGGAAGCCAGGCCGGT
CAAAAACGGTTGCAGGGGTATTGATGATAAACACTGGAACTCTCAGTGCAAAACATCCCAAACCTACGTCC
GAGCACTGACTTCAGAGAACAATAAACTCGTGGGCTGGCGGTGGATACGGATAGACACGTCCTGTGTGTGT
GCCTTGTCGAGAAAAATCGGAAGAACATGAATTGGCATCTCTCCCCATATATAAATTATTACTTTAAATTA
TATGATATGCATGTAGCATATAAATGTTTATATTGTTTTTATATATTATAAGTTGACCTTTATTTATTAAA
CTTCAGCAACCCTACAGTATATAAGCTTTTTTCTCAATAAAATCAGTGTGCTTGCCTTCCCTCAGGCCTCT
CCCATCTGTTAAAACTTGTTTTGTGATCCGGCTCTCAGGAGTCACTCTGTAAAATCTGTGTACACCAGTAT
TTTGCATTCAGTATTGTCAAGGCCATGACTGTTGTTTTAGTAAACTTGTTAAAATCAAAAAAAAAAAAAA
LNDRARVHLSLHVDVPGNGHTDAMVTFATILQVNKVMSILFYVIFLAYLRGIQGNNMDQRSLPEDSLNSLI
IKLIQADILKNKLSKQMVDVKENYQSTLPKAEAPREPERGGPAKSAFQPVIAMDTELLRQQRRYNSPRVLL
SDSTPLEPPPLYLMEDYVGSPVVANRTSRRKRYAEHKSHRGEYSVCDSESLWVTDKSSAIDIRGHQVTVLG
EIKTGNSPVKQYFYETRCKEARPVKNGCRGIDDKHWNSQCKTSQTYVRALTSENNKLVGWRWIRIDTSCVC
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

TCTTCACGTCCCAGCGCGGGTGGGCGCCGGCGGCTCCTCTTAACCACAGGTTCCAGAAGTCCTCTGCAGAA
GTGCTTCCCTCTCTCATTTCCAGGACCACAATTCCCAGAGACTTCGGCTTCACGACGTTTCTCTTTTTGCC
CGATCTCTCCCGGAGCTGGCTGGGCTTCGGCCGGCCAGAGGCCCACAGCGACGACGTGATCCGTCGTGAGC
GGGTCCCAGGGTTTCCTCGGCGGCCCTTTTTTCTCTCCCCTCGGTCGTCCCCCTTGCAGGCTGTCGCGGCT
GGCCTGAGCGGTGACCTGGCGGGTCGCGCCTGCGCTCTGCCCTGTTTCCTGCCTGGCTGGTGGCGGCGGCC
ATTTTGTTCATCCTCCTCCTCCTCCTGCTCCTCCTGGTTGGAGCGCAGTGTCCGGAGCGGGCTGGGGGGAG
AGAGCCCGAGAGCAGGGTTCGGTGCCTTTTCCTCTGTCCCCAGCCGGTGCCCAGAGCCCCCCTCCCCTTCC
TCCCCACCCCCTCCCCTCCCCAACCCTGCCCTCCCCCTTGTCCCGGGATCGCTCCGTCGCACCCACCATGA
TGGAAGACGACGGGCAGCCCCGGACTCTATACGTAGGTAACCTTTCCAGAGATGTGACAGAAGTCCTTATA
CTTCAGTTGTTCAGTCAGATTGGACCCTGTAAAAGCTGTAAAATGATAACAGAGCATACAAGCAATGACCC
ATATTGCTTTGTGGAATTTTATGAACACAGAGATGCAGCTGCTGCATTAGCTGCTATGAATGGGAGAAAAA
TTTTGGGAAAGGAGGTCAAAGTAAACTGGGCAACCACACCAAGTAGCCAGAAAAAAGATACTTCCAATCAC

TTCCATGTGTTTGTTGGGGATTTGAGTCCAGAAATTACAACAGAAGATATCAAATCAGCATTTGCCCCCTT
TGGTAAAATATCGGATGCCCGGGTAGTTAAAGACATGGCAACTGGAAAATCCAAAGGCTATGGTTTTGTAT
CTTTTTATAACAAACTGGATGCAGAAAATGCGATTGTGCATATGGGCGGTCAGTGGTTGGGTGGTCGTCAA
ATCCGAACCAATTGGGCCACTCGTAAACCACCTGCACCTAAAAGTACACAAGAAAACAACACTAAGCAGTT
GAGATTTGAAGATGTAGTAAACCAGTCAAGTCCAAAAAATTGTACTGTGTACTGTGGAGGAATTGCGTCTG
GGTTAACAGATCAGCTTATGAGACAGACATTCTCACCATTTGGACAAATTATGGAAATAAGAGTTTTCCCA
GAAAAGGGCTATTCATTTGTCAGATTTTCAACCCATGAAAGTGCAGCCCATGCCATTGTTTCGGTGAACGG
TACTACGATTGAAGGACATGTGGTTAAATGCTATTGGGGTAAAGAATCTCCTGATATGACTAAAAACTTCC
AACAGGTTGACTATAGTCAATGGGGCCAATGGAGCCAAGTGTATGGAAACCCACAACAGTATGGACAGTAT
ATGGCAAATGGGTGGCAAGTACCGCCTTATGGAGTATACGGGCAACCATGGAATCAACAAGGATTTGGAGT
AGATCAATCACCTTCTGCTGCTTGGATGGGTGGATTTGGTGCTCAGCCTCCCCAAGGACAAGCTCCTCCCC
CTGTAATACCTCCTCCTAACCAAGCCGGATATGGTATGGCAAGTTACCAAACACAGTGAGCCGGGACTCTA
AAAAAAAATTGTAATTCATGATAGGCTTCGATTTCCTGTGACACTCTGAAGACATGAAAGTAGACATCGGA
AAATGAAAATATTTATTTTAAAAATTGAAATGTTTGGAACCTTTAGCACAGATTTGCTTTGGTGAAGGACA
CGTGTCTTCTAGTTCTGCCTTTTTAAGTTTTTGTTCATGATGGATATGAACATGATTTTTCTTTATGTACA
AAAACTAAAATAAAGTCAATAAAGACAATTCTGACTACAAATTTTGATATAATAGGAAAAATGGCTAATAC
ATTTTGATTCTTAGATACTATTCCATTTTTATCTTGCTGTTCAGTATTTTAACTCACTGTGTTTTTAAAAG
AGCAAAAAAGGGAGGATCGTGAAAACCTGGGAATCACATATAAGTTCATCCTGAATCCTGATACTCCCCTC
CCCTTCCCTGAGGTGGACCACATTTGAAGTCAGCAGAGAAAAAGTGTGATATTCAGAAGAAATGCGTGATT
TTGGAGTCGCTTTGGAGGAAATATTTTCTTTCTCTATGCCTAAAGAAACTGAAGCCAGACTGAAGTTTTGC
ACCCTAAAAAAGGAACAGCATTGTTTGAGTTACTTGAGCAAATGTTGGTGGTCCACGTTAAGACATATTTT
TAAAACTTCCAAAAGTGTCGATTATTAAAATTGTAGTATTTTACATTTCATTTTGGGGGGAAATCCAAGTA
TGGTGTTTGTATTGAAGTCAGACAGTCATACTTGTGCTTTTACATGAAGTTTAAATGATACATATTGTAAA
TATTCAATAACTACAGTGTTTAAAAAGCATGCTTCAACATAGAAGTAGCAGCAATGTAATTATTTGAAGTA
ACACTTAACACACTCCGCTGCATTGAATGCAGTGGATTGATCAGAATGTTAAGACTGACATTTCCAAGGTT
GGCTACTATGTAAAATTAAAATTACACAAATTGTGCAGAAAAAGCCTTAATTTTAATTTCATACAATCTTT
GATGCATTAGTATGTTCTAAAATGTCATTGGGAATTAGTTTTTTGTTTTTGTTTTTTTTTTTTTTTTTTTT
TGCTTTACATTACTTGGTATGTAAATACCTTGATTAAAACCTTGTAAACCAATTTCAAGGTTACTATAAGT
TGTATAGTACAAGTGTTTTTTAAAAATCTTGGGGTGTTTTTAAAAATTAAGATATATTTTGCCCAAGAATT
TTTTTAACAAGATTGCTAAAAACATCTTATTTAGACACTTCAATGTACCAATTTATAATTGGATATTCAGT
TTAAATAGTACACAGAGTTGTGGCTTTTATTTTCAATTAATTTTTTTCCTTGTGGGCAGTGTGCATGGTAT
AATAAGCCTGAGCAGAGGCTTAAGTTGTATGTGTGCAGAGTTTGTAAAGGAATCAATTGGAAGATGCAGAA
GACCGAGGTTTGCTTTCAAGGTATTTTTCAGGCTGTGTGGGTAAAATTTGCCTCAAATTTCTATCAAACAG
GAATGTAAAATAGATAAAATCCTATGTATTTGAATTGTCAGAGCTAGGGAGTGCAAATGTTTTGGCAATGT
ATTCAAAATGCTGGCCTGGGCACCAAAGAGAAAATAGCCTTTTACAGTTACATAGTAAGATGCGATTAGTA
CCCACAAATTACTGTTTTCTAAACATTTGAAGTTTTACGATTAGCTTTAAAATAATGATTTTATAAATTGG
TGGTCACAATAATTTTGGTATTACTTTCCTCCTTTTCCCACTTAGCAATATAGCCAAATGTATTCAACATA
AAAATTCATAGGGTCTGAAATTCATAGCTGGGCCAAATTTTTTATGGCACCTTAGTTTTACCATAATGGTC
ATCTATTACACTCTTCTGTTATAAAATATACCCTTATTTCTTTTGTTTATAGTATCTTTGAGGAATGTTTT
TGGAAAAGTTAATTTATATTTTATAGGGAGAACACTCAATAAATTATGTTAACTGTGCCCCCGAGTTAAAA
ATTTTATGAGTATATGTGAAACTTGAACAACTGAAGACTTTTTTTAATTGATAAAAATGCTTAGTATGCCT
GTTTTGGTCTGCCAGTAAATTAAGTAGCTTATTGAGATAACTAACAGCTAAATATAGCTGTAGTGTTTCCT
GACTGTATATTCTATGATTTAATAAAATTATCCAGACTAGTTATATTGCCACAGTAAAAA

VTWRVAPALCPVSCLAGGGGHFVHPPPPPAPPGWSAVSGAGWGERAREQGSVPFPLSPAGAQSPPPLPPHP
LPSPTLPSPLSRDRSVAPTMMEDDGQPRTLYVGNLSRDVTEVLILQLFSQIGPCKSCKMITEHTSNDPYCF
VEFYEHRDAAAALAAMNGRKILGKEVKVNWATTPSSQKKDTSNHFHVFVGDLSPEITTEDIKSAFAPFGKI
SDARVVKDMATGKSKGYGFVSFYNKLDAENAIVHMGGQWLGGRQIRTNWATRKPPAPKSTQENNTKQLRFE
DVVNQSSPKNCTVYCGGIASGLTDQLMRQTFSPFGQIMEIRVFPEKGYSFVRFSTHESAAHAIVSVNGTTI
EGHVVKCYWGKESPDMTKNFQQVDYSQWGQWSQVYGNPQQYGQYMANGWQVPPYGVYGQPWNQQGFGVDQS
PSAAWMGGFGAQPPQGQAPPPVIPPPNQAGYGMASYQTQ

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

```
CCATTTATTTCCCCCCTTTCCAACCTCTTCCTCCCCAGCCTCCGCACCCTACCCTTGTTTCCTGTCCCTGT
CGCGCCCAGGTGTTTACCTGGCACTCAGGTGAGTGGTGCGCTCTGGCTGTTTTCTGTCGGAGCCGCCCGCC
TCTTCCTTCAGCGCGTCCCACAAATCCCGACGGCACGGAGGGGCCCCAGGCCAAGGGCGATGGGCCCCTGA
GCCCTGACACCGCTTCGCCGCTGCTGCAGGTGCCCCTGGCCGGCAGCGCCGCCGTGGTCCCGGAGCGCGGC
GACATGCCGGAGCTGGTGGTGACCGCGCTGCTGGCGCCGTCGCGCCTGTCGCTGAAGCTGCTGCGCGCCTT
CATGTGGAGCCTGGTGTTCTCGGTGGCGCTGGTGGCCGCGGCGGTCTACGGCTGCATAGCGCTCACGCACG
TGCTGTGCCGGCCCCGGCGCGGCTGCTGCGGGCGCCGTCGGAGCGCGTCCCCCGCCTGCCTGAGCGACCCC
TCGCTGGGTGAGCACGGTTTCCTGAACCTCAAGAGCTCGGGCCTGCGTCTGCACTATGTCTCGGCTGGACG
AGGTAACGGACCCCTCATGCTGTTTCTGCACGGCTTCCCTGAGAACTGGTTCTCCTGGCGTTACCAGCTCC
GGGAGTTCCAGAGCCGCTTCCATGTTGTGGCTGTGGACTTGCGAGGCTATGGCCCCTCGGATGCACCTCGG
GATGTGGACTGCTACACAATCGACCTGCTGCTGGTGGACATCAAAGATGTCATCCTAGGCCTGGGTTACTC
GAAGTGCATCCTTGTGGCCCATGACTGGGGTGCCCTCCTTGCCTGGCATTTCTCCATCTACTACCCATCCC
TGGTCGAGCGGATGGTTGTGGTCAGTGGTGCCCCCATGTCGGTGTACCAAGACTATTCCCTGCACCACATC
AGCCAGTTCTTCCGTTCCCACTACATGTTCCTGTTCCAGCTGCCCTGGCTGCCCGAGAAGCTGCTGTCTAT
GTCTGACTTTCAGATTCTGAAGACCACCCTCACCCACCGCAAGACAGGCATCCCATGCTTGACCCCCAGCG
AGCTCGAGGCCTTCCTTTATAACTTCTCACAGCCTGGTGGCCTCACTGGGCCCCTCAACTACTACCGAAAC
CTCTTCAGGAACTTCCCCCTGGAACCCCAGGAGCTGACCACACCCACATTGCTGCTGTGGGGGGAGAAGGA
CACTTACTTGGAGCTGGGGCTGGTGGAAGCCATCGGCAGCCGCTTTGTGCCGGGCCGCTTGGAGGCCCACA
TCCTGCCAGGCATAGGGCATTGGATCCCACAGAGCAACCCCCAGGAGATGCACCAGTACATGTGGGCCTTC
TTGCAAGACCTGCTGGACTAGTGGTCCTTGCTGGCCTGCCCAGGAGGCATGGATACTCAGGAAGGAACACA
CACCCATTATCTGTGTGTGCCTGGGAGTCCATACAATGTCCATACATGGGTGAACTCTTGAATCGCTCATA
GGCATGGGACTCCTGGATCCCACACAAGCGCACCTATGGGTGCCTCGGGACACACCAACCCCTATACTCAC
ACACAGGCATGGATGCATGTGTGTGTAACAAACACTTTGACCCTGGGAACTGGGTATACCTCTCTTCCAGT
GGAGCCAGATGCCGAGACCGAGTGTCCCACCTCCTCCCTTCCTGGGGCCTCACTCTGCGCTTTGCCAAAGT
CGCTTCTCTGCCATAGCTGCACAGACCTTAAACCCTGACCTTCCTGTCCCTGCCTTCGACCTCCTGCCTGG
GTCTTCAGCTCAGTGCTACTCTGAATACAATACTGCTGACCACATTTGGCTACTTCAGCTTAAATGTATGT
TAATTGAAATGAAATACAATGAAAATGCAGTGCTCAGTCCCA

TEGPQAKGDGPLSPDTASPLLQVPLAGSAAVVPERGDMPELVVTALLAPSRLSLKLLRAFMWSLVFSVALV
AAAVYGCIALTHVLCRPRRGCCGRRRSASPACLSDPSLGEHGFLNLKSSGLRLHYVSAGRGNGPLMLFLHG
FPENWFSWRYQLREFQSRFHVVAVDLRGYGPSDAPRDVDCYTIDLLLVDIKDVILGLGYSKCILVAHDWGA
LLAWHFSIYYPSLVERMVVVSGAPMSVYQDYSLHHISQFFRSHYMFLFQLPWLPEKLLSMSDFQILKTTLT
HRKTGIPCLTPSELEAFLYNFSQPGGLTGPLNYYRNLFRNFPLEPQELTTPTLLLWGEKDTYLELGLVEAI
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

AGAGTCCTGGCTGGCTCTGCCAAGTGTAACAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAG
ACTCCCAAGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGACCTCTACAGGGAGACGGTGGGCCGGCC
CTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCAGGCCCCCAAGGA
GCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGT
TCTGCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAG
AGGATCTGGGAGTCTCCTCTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAA
GTATGAGGATTGCAAGGTGCACCAGAGAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATG
ACAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCATGACATCTGAG
CTCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGAGCCCTGCT
TGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCTGCAACCTCATCT
ACTTTGGGGAGCACCCTTTGTCCTTTGCTGCCTGTGTGAACAGTGAGGAGATCGTGCGGCTGCTCATTGAG
CATGGAGCTGACATCCGGGCCCAGGACTCCCTGGGAAACACAGTGTTACACATCCTCATCCTCCAGCCCAA
CAAAACCTTTGCCTGCCAGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCCCC
TGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTGTGATG
TTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGACTCTCTATGA
CCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGG
AGGCTCGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGG
CCGTACTTCTGCATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCG
CCCCCTCAAGCCCAGGACCAATAACCGCACGAGCCCCCGGGACAACACCCTCTTACAGCAGAAGCTACTTC
AGGAAGCCTACATGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTGGTGACTGTCATTGGGGCTATC
ATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTCGCTTCTTTGGACAGACCATCCT
TGGGGGCCCATTCCATGTCCTCATCATCACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCA
TCAGTGCCAGCGGGGAGGTGGTACCCATGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTC
GCCCGAGGATTCCAGATGCTAGGCCCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCTGATGCG
ATTCTGCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGCTTCAGCCTTCTATATCATCTTCCAGACAGAGG
ACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCTTCGAGCTGTTCCTTACC
ATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTTGC
CATCATCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCC
ATGAGCGGGATGAGCTGTGGAGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGC
TGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGGA
AGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCTTCCACACCCGGGGCTCTGAGG
ATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCAGCCCCCACCTGTCCCTTCCTATG
CCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAG
AGACCTGCGTGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTG
TTCTCACTTCGCTTCCTGGAACTTGCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCAAACACCC
AGAGGTCTCATCTCCCAGGCCCCAGGGGAGAAAGAGGAGTAGCATGAACGCCAAGGAATGTACGTTGAGAA
TCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAGCCCAGCCCAAGCACGGGGCTGGC
AGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAGGAGCACTGCATGTCAGAGCACTTT
AAAAACAGGCCAGCCTGCTTGGGCGCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCA
GGGCACCCAGGCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCT
CTGGGGGTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTCGGATCCCTGAAAA
AAAAAAAAAAAAAAAAAAAA

TGPLQGDGGPALGGADVAPRLSPVRVWPRPQAPKEPALHPMGLSLPKEKGLILCLWSKFCRWFQRRESWAQ
SRDEQNLLQQKRIWESPLLLAAKDNDVQALNKLLKYEDCKVHQRGAMGETALHIAALYDNLEAAMVLMEAA
PELVFEPMTSELYEGQTALHIAVVNQNMNLVRALLARRASVSARATGTAFRRSPCNLIYFGEHPLSFAACV
NSEEIVRLLIEHGADIRAQDSLGNTVLHILILQPNKTFACQMYNLLLSYDRHGDHLQPLDLVPNHQGLTPF
KLAGVEGNTVMFQHLMQKRKHTQWTYGPLTSTLYDLTEIDSSGDEQSLLELIITTKKREARQILDQTPVKE
LVSLKWKRYGRPYFCMLGAIYLLYIICFTMCCIYRPLKPRTNNRTSPRDNTLLQQKLLQEAYMTPKDDIRL
VGELVTVIGAIIILLVEVPDIFRMGVTRFFGQTILGGPFHVLIITYAFMVLVTMVMRLISASGEVVPMSFA
LVLGWCNVMYFARGFQMLGPFTIMIQKMIFGDLMRFCWLMAVVILGFASAFYIIFQTEDPEELGHFYDYPM
ALFSTFELFLTIIDGPANYNVDLPFMYSITYAAFAIIATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVA

TTVMLERKLPRCLWPRSGICGREYGLGDRWFLRVEDRQDLNRQRIQRYAQAFHTRGSEDLDKDSVEKLELG
CPFSPHLSLPMPSVSRSTSRSSANWERLRQGTLRRDLRGIINRGLEDGESWEYQI

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

GCGCTCCGGAGGGAGAGCTGGGGCTGGAGGTTCCTACCCCCTCGGCGGCCCGCATCTGCCCCCGCGCGCCC
GCCCTGAGCCCGCCCCGACTGGGCAGGCGGGGGAGCCCCTACTTCTCTCCCCCCGGGCGGGGGAGCCGGGG
GGCAGCGCCGGAGCCCGGGGGGAGCTCAGCCCCGCCGACCGGCCGGCCAGGGCAGGGGGCAGCTAGGACGG
CCCCGGTCCAGGTGGAGGCCGCAGAGGGCCCAGGGCAAGCAGAGGCAGCAATGGTTGGTCCTGACGGTGGC
TGAGCCCCCAGCCCCTGGAATATGCAGCCCGGGGGAGCCCCAGACAGCGGCAAGGACGAGGTGGCGGAGTG
GGGCGGGAGGCATGGTCTCCACCTACCGGGTGGCCGTGCTGGGGGCGCGAGGTGTGGGCAAGAGTGCCATC
GTGCGCCAGTTCTTGTACAACGAGTTCAGCGAGGTCTGCGTCCCCACCACCGCCCGCCGCCTTTACCTGCC
TGCTGTCGTCATGAACGGCCACGTGCACGACCTCCAGATCCTCGACTTTCCACCCATCAGCGCCTTCCCTG
TCAATACGCTCCAGGAGTGGGCAGACACCTGCTGCAGGGGACTCCGGAGTGTCCACGCCTACATCCTGGTC
TACGACATCTGCTGCTTTGACAGCTTTGAGTACGTCAAGACCATCCGCCAGCAGATCCTGGAGACGAGGGT
GATCGGAACCTCAGAGACGCCCATCATCATCGTGGGCAACAAGCGGGACCTGCAGCGCGGACGCGTGATCC
CGCGCTGGAACGTGTCGCACCTGGTACGCAAGACCTGGAAGTGCGGCTACGTGGAATGCTCGGCCAAGTAC
AACTGGCACATCCTGCTGCTCTTCAGCGAGCTGCTCAAGAGCGTCGGCTGCGCCCGTTGCAAGCACGTGCA
CGCTGCCCTGCGCTTCCAGGGCGCGCTGCGCCGCAACCGCTGCGCCATCATGTGACGCCTGCGCGCCCCTC
GGGCTGCACCGGCACTGGCCGAGCGGAGGGCGGGGCCGTACTGCGGGGCTGGGGCGGGGAGCGGGCGGGAA
ATGGAACTGTGACGGTCCCGGCCTGAGGCCCCTGCAGCCACGCACCTCCCGGTGAGAAGCAGAGCGCGAGA
GGGAGCCCTCCGTAACTGCCCAGCCCTGCCCCTTGCCCCCGTGGCTTCCTGGGACAGCCGCCTTCAGTGCT
GTATTTAGTGCAGTGCCCGGCCCGACCCGCGGGGGTGCCACAGCCTTTTGGGATGGGGGTGAGCGTGCAAT
GGAGGCTGGGGGTGGCGAGGTGCCGCCTTGGCCGGGCCCCCACGTGTCTTCTCCAGAATGTGTCTGTCTTT
GCCTGGTGTCTTCCTTTCCCGTGTCCGCCCACCCCAGCGTCTGTTGGTACTTACCTGTCTCACCTACCCTC
CAGTCCCCTCCCAGCTCCGCTCACAGGGCTCTCATTTCGTCCATCCCCTTGTCGCAGATCCTGGCAGCTTC
TTTGTGAGGCCAGGCCTTCTGACTGTCAGCACCACCGGCACAGGGCAGAGATGCGGGTGGCCCAAGGACCA
CGATCAAGGGGTCCGGGGGACCGAGGTCCCAGATCAGTGAGGGGAGAAGGTTGAGCTCTCCGGCTTCCAGG
GAGACCTCCCCGCCCAGCAGCCCCCAGAGACACAACAACCTACCTTCCAGCCTTAACTCGATGGTCCGTCC
CTGCCAGGTGCCCCTCACTCTTCCTGACCCCAAAGCCAGATCACCCCCTGGGTTAAAACTTTTTTTCTTTT
TTTTTTTTGGACAGAGTGTGGAAAGGGAGCCCCCCAAAGGATAGCTTCTTTTTCATGATGCCAGGCTCCAG
TCCTTTATTCCCTTCTGCATACTGCAATCTGATCTGTCAGACTGGGGAATGTTGGGTTCTGGGGTCTGGTC
GTGGGCAGGATGGTGCCCAGAAGGGGGTTAGGTTGTCCCAGTGAAAATTCTGTTGCCCCGTCTCAACCCCA
TCTGACTACCCCAGACTCTGCCTGCCTCAGATCTCAGACTATCCTGATTAATCTGGGGAAGAACAGAGCCA
GGGAAAGAATGGTGGGGACCCCTGTACTTGGGGGAGACACACCTGCATCTTCCTCCTGCCACAGATGGAGG
CCCTCAGGATCTGACACCCTCTTGTCCCAACACCAGTCAGCCCTATACCCTAACTCACTCCACCCCATTTT
CTCCGGCTGCCTGGCCGGGTTTCTACCTCTCGTCACCGGAGCTGATCACTGTCAGTTTTGTACCGATTTAG
AAATAACAATAATAATGAAGATTCTAGGAATGGCATGAGGGATTGATGGGGGACTTGGAGGGAGGGACAAG
TGGTGCCCTGTCCCCTGCTCCCCTGGCCAAAGAAAGCTGTCCTTGAGGCTGAGCCCTCAGCCCTGGCCTGG
TGGGGGGACAGCAAGGTCCCTTGTTATAAGAGGGGCAGAGAGGACAACTCCGCTTTGGCCAACCTAGCCAA
GGCTGCAGCATATAGACCAGGAAATCAGGTAGCCCAGACTGGTGATGGAGCAGAGTCTGGGGGAAGGGTCG
TGGGTGGGGAATTTATCACCAACATCCATTGTAGGGGGAATCTATGATTCTGCTTCCCCAGCGGATTCCCA
CTCTGTCCACCAAGTGGGGGGTAGCACAGCCTCACAGCAACCGCCCTGACCTTGGGCAGTCTAGTGTTCCT
GCATTCTAGTCCCTGCTGTGCTGCAGGACTTTGGGCAAGTGACCTGCCCTCTGTGAGCCTCCCTCTGACAC
AGAGGAGGTGGCTCCCCTTCCCCACACCTTAGAGTGGCTGGGAGGGTAACAAAGAGGGCCTGCCCCTTTAG
TCTCCTGCACCCCTGCCCCCTGGTTCACCAGAGGGAGCGGATGAAGGATGGCAGCATCTCACATGCCCCAT

CACCAACTCTGAGGCACCTGGGGTGGGGGGGCGGAGCCCAGGCCTCTGGCTGCTCCCCTGTGGGAGCCATT
GGAATGTATCCCCTGACAGGCCCCCTTCCGCCTCCACCTCAACCCAGGTCTTGGATTTCAGGTCCCTCCAC
CCCCATTCTGAGTCTCTGTCCTTCTCCTTCCACCCGCTCCCAGGGTTTCCCACCACAGGGTCTGGAAGTGT
GTGTGACGCCCATTGAGCTGTTACCCGAAGTCAGATTAAAAATCAGGGAGTGTTTTCCCTCGTTTCTGTAA
AAAAAAAAAAAAAAAAA

LVLTVAEPPAPGICSPGEPQTAARTRWRSGAGGMVSTYRVAVLGARGVGKSAIVRQFLYNEFSEVCVPTTA
RRLYLPAVVMNGHVHDLQILDFPPISAFPVNTLQEWADTCCRGLRSVHAYILVYDICCFDSFEYVKTIRQQ
ILETRVIGTSETPIIIVGNKRDLQRGRVIPRWNVSHLVRKTWKCGYVECSAKYNWHILLLFSELLKSVGCA
RCKHVHAALRFQGALRRNRCAIM

AUG initiation = 23.2 kDa
UUG initiation = 26.5 kDa

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

#38

NM_001080510

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

C17orf95 chromosome 17 open reading frame 95
(homology to AdoMet-dependent methyltransferases)
Synonyms: C17orf95

AGTTCTGCGCGTGTGAGTCTCTTTCGCCTTGCTCCGGGCTTTCTTCGCTCGCAGCGCGGCAGGGTTATCAC
CAGATCTGGGCTTTCCCCTTCTTGCCGTCAGGTGCTACGGCCACGTGGCCCGCGGCTTCCCGCTCGCGCAG
TCTGGCAGCCCGGAGCCTTCCGCGGTCCCCCACCCGCCCGGGGCCCAACGACGCCCTACTGGGCGAGCACG
ATTTCCGAGGACAGGGGGTCCGGGCCCAGCGCTTTCGATTCTCGGAGGAGCCGGGTCCGGGGGCCGACGGG
GCTGTCCTGGAGGTCCACGTCCCGCAGGTCCTGCATCTCCAGTATGGAATGTATGTTTGGCCCTGTGCTGT
GGTCCTGGCCCAGTACCTTTGGTTTCACAGAAGATCTCTGCCAGGCAAGGCCATCTTAGAGATTGGAGCTG
GAGTGAGCCTTCCAGGAATTTTGGCTGCCAAATGTGGTGCAGAAGTAATACTGTCAGACAGCTCAGAACTG
CCTCACTGTCTGGAAGTCTGTCGGCAAAGCTGCCAAATGAATAACCTGCCACATCTGCAGGTGGTAGGACT
AACATGGGGTCATATATCTTGGGATCTTCTGGCTCTACCACCACAAGATATTATCCTTGCATCTGATGTGT
TCTTTGAACCAGAAGATTTTGAAGACATTTTGGCTACAATATATTTTTTGATGCACAAGAATCCCAAGGTC
CAATTGTGGTCTACTTATCAAGTTAGGAGTGCTGGCTGGTCACTTGAAGCTTTACTCTACAAATGGGATAT
GAAATGTGTCCACATTCCTCTTGAGTCTTTTGATGCAGACAAAGAAGATATAGCAGAATCTACCCTTCCAG
GAAGACATACAGTTGAAATGCTGGTCATTTCCTTTGCAAAGGACAGTCTCTGAATTATACCTACAACCTGT
TCTGGGACAGTATCAATACTGATGAGCAACCTGGCACACAAACTATGAGCAGACCACTTCAGCTTGAGAAT
GCAGTGGGTCTGAAGATGGTCAAGTCTGTTTGCCTTAGATTTTGATGTCACCTAGACAACACTTAAACTCA
TATGAAACAAAAATTAAAATACGTATTACAAGTAAAAAAAAAAAAAAA

LAARSLPRSPTRPGPNDALLGEHDFRGQGVRAQRFRFSEEPGPGADGAVLEVHVPQVLHLQYGMYVWPCAV
VLAQYLWFHRRSLPGKAILEIGAGVSLPGILAAKCGAEVILSDSSELPHCLEVCRQSCQMNNLPHLQVVGL
TWGHISWDLLALPPQDIILASDVFFEPEDFEDILATIYFLMHKNPKVQLWSTYQVRSAGWSLEALLYKWDM
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
CCCGCTCCCCGTCCTTTGGGGCCGCGGCTGGGGAACTACAAGGCCCAGCAGGCAGCTGCAGGGGGCGGAGG
CGGAGGAGGGACCAGCGCGGGTGGGAGTGAGAGAGCGAGCCCTCGCGCCCCGCCGGCGCATAGCGCTCGGA
GCGCTCTTGCGGCCACAGGCGCGGCGTCCTCGGCGGCGGGCGGCAGCTAGCGGGAGCCGGGACGCCGGTGC
AGCCGCAGCGCGCGGAGGAACCCGGGTGTGCCGGGAGCTGGGCGGCCACGTCCGGACGGGACCGAGACCCC
TCGTAGCGCATTGCGGCGACCTCGCCTTCCCCGGCCGCGAGCGCGCCGCTGCTTGAAAAGCCGCGGAACCC
AAGGACTTTTCTCCGGTCCGAGCTCGGGGCGCCCCGCAGGGCGCACGGTACCCGTGCTGCAGTCGGGCACG
CCGCGGCGCCGGGGCCTCCGCAGGGCGATGGAGCCCGGTCTGCAAGGAAAGTGAGGCGCCGCCGCTGCGTT
CTGGAGGAGGGGGGCACAAGGTCTGGAGACCCCGGGTGGCGGACGGGAGCCCTCCCCCCGCCCCGCCTCCG
GGGCACCAGCTCCGGCTCCATTGTTCCCGCCCGGGCTGGAGGCGCCGAGCACCGAGCGCCGCCGGGAGTCG
AGCGCCGGCCGCGGAGCTCTTGCGACCCCGCCAGGACCCGAACAGAGCCCGGGGGCGGCGGGCCGGAGCCG
GGGACGCGGGCACACGCCCGCTCGCACAAGCCACGGCGGACTCTCCCGAGGCGGAACCTCCACGCCGAGCG
AGGGTCAGTTTGAAAAGGAGGATCGAGCTCACTGTGGAGTATCCATGGAGATGTGGAGCCTTGTCACCAAC
CTCTAACTGCAGAACTGGGATGTGGAGCTGGAAGTGCCTCCTCTTCTGGGCTGTGCTGGTCACAGCCACAC
TCTGCACCGCTAGGCCGTCCCCGACCTTGCCTGAACAAGCCCAGCCCTGGGGAGCCCCTGTGGAAGTGGAG
TCCTTCCTGGTCCACCCCGGTGACCTGCTGCAGCTTCGCTGTCGGCTGCGGGACGATGTGCAGAGCATCAA
CTGGCTGCGGGACGGGGTGCAGCTGGCGGAAAGCAACCGCACCCGCATCACAGGGGAGGAGGTGGAGGTGC
AGGACTCCGTGCCCGCAGACTCCGGCCTCTATGCTTGCGTAACCAGCAGCCCCTCGGGCAGTGACACCACC
TACTTCTCCGTCAATGTTTCAGATGCTCTCCCCTCCTCGGAGGATGATGATGATGATGATGACTCCTCTTC
AGAGGAGAAAGAAACAGATAACACCAAACCAAACCGTATGCCCGTAGCTCCATATTGGACATCCCCAGAAA
AGATGGAAAAGAAATTGCATGCAGTGCCGGCTGCCAAGACAGTGAAGTTCAAATGCCCTTCCAGTGGGACC
CCAAACCCCACACTGCGCTGGTTGAAAAATGGCAAAGAATTCAAACCTGACCACAGAATTGGAGGCTACAA
GGTCCGTTATGCCACCTGGAGCATCATAATGGACTCTGTGGTGCCCTCTGACAAGGGCAACTACACCTGCA
TTGTGGAGAATGAGTACGGCAGCATCAACCACACATACCAGCTGGATGTCGTGGAGCGGTCCCCTCACCGG
CCCATCCTGCAAGCAGGGTTGCCCGCCAACAAAACAGTGGCCCTGGGTAGCAACGTGGAGTTCATGTGTAA
GGTGTACAGTGACCCGCAGCCGCACATCCAGTGGCTAAAGCACATCGAGGTGAATGGGAGCAAGATTGGCC
CAGACAACCTGCCTTATGTCCAGATCTTGAAGACTGCTGGAGTTAATACCACCGACAAAGAGATGGAGGTG
CTTCACTTAAGAAATGTCTCCTTTGAGGACGCAGGGGAGTATACGTGCTTGGCGGGTAACTCTATCGGACT
CTCCCATCACTCTGCATGGTTGACCGTTCTGGAAGCCCTGGAAGAGAGGCCGGCAGTGATGACCTCGCCCC
TGTACCTGGAGATCATCATCTATTGCACAGGGGCCTTCCTCATCTCCTGCATGGTGGGGTCGGTCATCGTC
TACAAGATGAAGAGTGGTACCAAGAAGAGTGACTTCCACAGCCAGATGGCTGTGCACAAGCTGGCCAAGAG
CATCCCTCTGCGCAGACAGGTAACAGTGTCTGCTGACTCCAGTGCATCCATGAACTCTGGGGTTCTTCTGG
TTCGGCCATCACGGCTCTCCTCCAGTGGGACTCCCATGCTAGCAGGGGTCTCTGAGTATGAGCTTCCCGAA
GACCCTCGCTGGGAGCTGCCTCGGGACAGACTGGTCTTAGGCAAACCCCTGGGAGAGGGCTGCTTTGGGCA
GGTGGTGTTGGCAGAGGCTATCGGGCTGGACAAGGACAAACCCAACCGTGTGACCAAAGTGGCTGTGAAGA
TGTTGAAGTCGGACGCAACAGAGAAAGACTTGTCAGACCTGATCTCAGAAATGGAGATGATGAAGATGATC
GGGAAGCATAAGAATATCATCAACCTGCTGGGGGCCTGCACGCAGGATGGTCCCTTGTATGTCATCGTGGA
GTATGCCTCCAAGGGCAACCTGCGGGAGTACCTGCAGGCCCGGAGGCCCCCAGGGCTGGAATACTGCTACA
ACCCCAGCCACAACCCAGAGGAGCAGCTCTCCTCCAAGGACCTGGTGTCCTGCGCCTACCAGGTGGCCCGA
GGCATGGAGTATCTGGCCTCCAAGAAGTGCATACACCGAGACCTGGCAGCCAGGAATGTCCTGGTGACAGA
GGACAATGTGATGAAGATAGCAGACTTTGGCCTCGCACGGGACATTCACCACATCGACTACTATAAAAAGA
CAACCAACGGCCGACTGCCTGTGAAGTGGATGGCACCCGAGGCATTATTTGACCGGATCTACACCCACCAG
AGTGATGTGTGGTCTTTCGGGGTGCTCCTGTGGGAGATCTTCACTCTGGGCGGCTCCCCATACCCCGGTGT
GCCTGTGGAGGAACTTTTCAAGCTGCTGAAGGAGGGTCACCGCATGGACAAGCCCAGTAACTGCACCAACG

AGCTGTACATGATGATGCGGGACTGCTGGCATGCAGTGCCCTCACAGAGACCCACCTTCAAGCAGCTGGTG
GAAGACCTGGACCGCATCGTGGCCTTGACCTCCAACCAGGAGTACCTGGACCTGTCCATGCCCCTGGACCA
GTACTCCCCCAGCTTTCCCGACACCCGGAGCTCTACGTGCTCCTCAGGGGAGGATTCCGTCTTCTCTCATG
AGCCGCTGCCCGAGGAGCCCTGCCTGCCCCGACACCCAGCCCAGCTTGCCAATGGCGGACTCAAACGCCGC
TGACTGCCACCCACACGCCCTCCCCAGACTCCACCGTCAGCTGTAACCCTCACCCACAGCCCCTGCTGGGC
CCACCACCTGTCCGTCCCTGTCCCCTTTCCTGCTGGCAGGAGCCGGCTGCCTACCAGGGGCCTTCCTGTGT
GGCCTGCCTTCACCCCACTCAGCTCACCTCTCCCTCCACCTCCTCTCCACCTGCTGGTGAGAGGTGCAAAG
AGGCAGATCTTTGCTGCCAGCCACTTCATCCCCTCCCAGATGTTGGACCAACACCCCTCCCTGCCACCAGG
CACTGCCTGGAGGGCAGGGAGTGGGAGCCAATGAACAGGCATGCAAGTGAGAGCTTCCTGAGCTTTCTCCT
GTCGGTTTGGTCTGTTTTGCCTTCACCCATAAGCCCCTCGCACTCTGGTGGCAGGTGCCTTGTCCTCAGGG
CTACAGCAGTAGGGAGGTCAGTGCTTCGTGCCTCGATTGAAGGTGACCTCTGCCCCAGATAGGTGGTGCCA
GTGGCTTATTAATTCCGATACTAGTTTGCTTTGCTGACCAAATGCCTGGTACCAGAGGATGGTGAGGCGAA
GGCCAGGTTGGGGGCAGTGTTGTGGCCCTGGGGCCCAGCCCCAAACTGGGGGCTCTGTATATAGCTATGAA
GAAAACACAAAGTGTATAAATCTGAGTATATATTTACATGTCTTTTTAAAAGGGTCGTTACCAGAGATTTA
CCCATCGGGTAAGATGCTCCTGGTGGCTGGGAGGCATCAGTTGCTATATATTAAAAACAAAAAAGAAAAAA
AAGGAAAATGTTTTTAAAAAGGTCATATATTTTTTGCTACTTTTGCTGTTTTATTTTTTTAAATTATGTTC
TAAACCTATTTTCAGTTTAGGTCCCTCAATAAAAATTGCTGCTGCTTCATTTATCTATGGGCTGTATGAAA
AGGGTGGGAATGTCCACTGGAAAGAAGGGACACCCACGGGCCCTGGGGCTAGGTCTGTCCCGAGGGCACCG
CATGCTCCCGGCGCAGGTTCCTTGTAACCTCTTCTTCCTAGGTCCTGCACCCAGACCTCACGACGCACCTC
CTGCCTCTCCGCTGCTTTTGGAAAGTCAGAAAAAGAAGATGTCTGCTTCGAGGGCAGGAACCCCATCCATG
CAGTAGAGGCGCTGGGCAGAGAGTCAAGGCCCAGCAGCCATCGACCATGGATGGTTTCCTCCAAGGAAACC
GGTGGGGTTGGGCTGGGGAGGGGGCACCTACCTAGGAATAGCCACGGGGTAGAGCTACAGTGATTAAGAGG
AAAGCAAGGGCGCGGTTGCTCACGCCTGTAATCCCAGCACTTTGGGACACCGAGGTGGGCAGATCACTTCA
GGTCAGGAGTTTGAGACCAGCCTGGCCAACTTAGTGAAACCCCATCTCTACTAAAAATGCAAAAATTATCC
AGGCATGGTGGCACACGCCTGTAATCCCAGCTCCACAGGAGGCTGAGGCAGAATCCCTTGAAGCTGGGAGG
CGGAGGTTGCAGTGAGCCGAGATTGCGCCATTGCACTCCAGCCTGGGCAACAGAGAAAACAAAAAGGAAAA
CAAATGATGAAGGTCTGCAGAAACTGAAACCCAGACATGTGTCTGCCCCCTCTATGTGGGCATGGTTTTGC
CAGTGCTTCTAAGTGCAGGAGAACATGTCACCTGAGGCTAGTTTTGCATTCAGGTCCCTGGCTTCGTTTCT
TGTTGGTATGCCTCCCCAGATCGTCCTTCCTGTATCCATGTGACCAGACTGTATTTGTTGGGACTGTCGCA
GATCTTGGCTTCTTACAGTTCTTCCTGTCCAAACTCCATCCTGTCCCTCAGGAACGGGGGGAAAATTCTCC
GAATGTTTTTGGTTTTTTGGCTGCTTGGAATTTACTTCTGCCACCTGCTGGTCATCACTGTCCTCACTAAG
TGGATTCTGGCTCCCCCGTACCTCATGGCTCAAACTACCACTCCTCAGTCGCTATATTAAAGCTTATATTT
TGCTGGATTACTGCTAAATACAAAAGAAAGTTCAATATGTTTTCATTTCTGTAGGGAAAATGGGATTGCTG
CTTTAAATTTCTGAGCTAGGGATTTTTTGGCAGCTGCAGTGTTGGCGACTATTGTAAAATTCTCTTTGTTT
CTCTCTGTAAATAGCACCTGCTAACATTACAATTTGTATTTATGTTTAAAGAAGGCATCATTTGGTGAACA
GAACTAGGAAATGAATTTTTAGCTCTTAAAAGCATTTGCTTTGAGACCGCACAGGAGTGTCTTTCCTTGTA
AAACAGTGATGATAATTTCTGCCTTGGCCCTACCTTGAAGCAATGTTGTGTGAAGGGATGAAGAATCTAAA
AGTCTTCATAAGTCCTTGGGAGAGGTGCTAGAAAAATATAAGGCACTATCATAATTACAGTGATGTCCTTG
CTGTTACTACTCAAATCACCCACAAATTTCCCCAAAGACTGCGCTAGCTGTCAAATAAAAGACAGTGAAAT
TGACCTGAAAAAAAAAAAAAAAAA

TADSPEAEPPRRARVSLKRRIELTVEYPWRCGALSPTSNCRTGMWSWKCLLFWAVLVTATLCTARPSPTLP
EQAQPWGAPVEVESFLVHPGDLLQLRCRLRDDVQSINWLRDGVQLAESNRTRITGEEVEVQDSVPADSGLY
ACVTSSPSGSDTTYFSVNVSDALPSSEDDDDDDDSSSEEKETDNTKPNRMPVAPYWTSPEKMEKKLHAVPA
AKTVKFKCPSSGTPNPTLRWLKNGKEFKPDHRIGGYKVRYATWSIIMDSVVPSDKGNYTCIVENEYGSINH
TYQLDVVERSPHRPILQAGLPANKTVALGSNVEFMCKVYSDPQPHIQWLKHIEVNGSKIGPDNLPYVQILK
TAGVNTTDKEMEVLHLRNVSFEDAGEYTCLAGNSIGLSHHSAWLTVLEALEERPAVMTSPLYLEIIIYCTG
AFLISCMVGSVIVYKMKSGTKKSDFHSQMAVHKLAKSIPLRRQVTVSADSSASMNSGVLLVRPSRLSSSGT
PMLAGVSEYELPEDPRWELPRDRLVLGKPLGEGCFGQVVLAEAIGLDKDKPNRVTKVAVKMLKSDATEKDL
SDLISEMEMMKMIGKHKNIINLLGACTQDGPLYVIVEYASKGNLREYLQARRPPGLEYCYNPSHNPEEQLS
SKDLVSCAYQVARGMEYLASKKCIHRDLAARNVLVTEDNVMKIADFGLARDIHHIDYYKKTTNGRLPVKWM
APEALFDRIYTHQSDVWSFGVLLWEIFTLGGSPYPGVPVEELFKLLKEGHRMDKPSNCTNELYMMMRDCWH
AVPSQRPTFKQLVEDLDRIVALTSNQEYLDLSMPLDQYSPSFPDTRSSTCSSGEDSVFSHEPLPEEPCLPR
HPAQLANGGLKRR

AUG initiation = 91.9 kDa

ACG initiation = 96.7 kDa

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

#40

NM_153369

CUG initiation conserved in mammals. Extension initated 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
GTCTTTTGCCACGAACACCTGCGGCGTGCCGAAGTTCCTCTTCTCTGCCCCGGTTCCGGGCGGTCCGCTGG
GAGGTGGCGCCCGGAGGGACACCCGTGCCTGGGGTCGGGGCCAGTGCTCTTCTCGGAGCTCCAGAAGTTCT
GCTGAGAAGCGCGGCGGCAGCAAGACGACTTCTCCGGAGCCGCCGAGCTGGAGTTAGAGGTGGAGCTCCGT
GGGGCCGGGCCCCCGGCTGCGGGGCAGCGGCTCCTGCAGGCGGAGGCCCCGGCGGAGAATGAGCCGGAGCC
GGAGGTGGTGGTGGTCTCCTGGCAGAGCGGTGGGACCGGGAGCAAGCTGCGGTGGTTCACCACCTTGATGC
TGTGTGCCTCCTTCCTGGGGCTGGGATTGAGTGTTGCTATAGTGGGACCCACGTTTCAAGATTTGGCAACA
AACGTGAACCGAAATATCAGTAGTCTGTCTTTCATTTTTGTGGGTCGTGCCTTGGGATATTTGAGTGGCTC
TGTGATTGGTGGATTTCTTGTCGATGTCATGAATTATTTTTTACTTTTGGGAATCTCAATGTCGGCTACCA
CCGTTGGTCTTTATCTTGTTCCTTTTTGCAAGACAGCAATATTACTCACTGTCATGATGTCTATCTTCGGT
GTTTCAATTGGCATTCTGGATACAGGTGGTAACGTCCTTATCTTGGCTATTTGGGGGGACAAAGGAGCCCC
ACATATGCAGGCCTTACACTTCTCTTTTGCCTTGGGTGCCTTTTTGGCTCCACTGCTAGCTAAACTGGCTT
TGGGTCCGACAGCGTCTGCTGAAAACCACACAGAGTCTGACTTCCATCCTGCACTCAACCAATCATCTGAT
GCTGACTCAGAAGCTCTGTTTGGAGTACCTAATGATAAGAATTTACTGTGGGCTTATGCTGTTATCGGTAC
TTACATGTTCTTAGTTTCTGTCATTTTTTTTTGTCTGTTTTTAAAGAATAGCTCAAAGCAAGAAAAAGCAA
GAGCATCTGCTGAGACATTTCGAAGAGCAAAATATCACAACGCCCTTCTTTGTCTCCTTTTTCTGTTCTTC
TTTTTTTATGTTGGAGCTGAGGTAACATATGGCTCTTATGTTTTCTCATTTGCAACCACCCATGCTGGCAT
GAAAGAAAGTGAAGCTGCTGGGTTGAACTCCATCTTCTGGGGGACATTTGCAGCCTGCAGGGGCCTGGCAA
TCTTTTTTGCTACCTGTTTACAGCCTGGAACCATGATTGTGTTGAGCAACATTGGCAGCCTGACTTCATCT
TTATTTCTGGTGCTTTTTGACAAGAACCCAATTTGTCTCTGGATAGCAACTTCAGTGTATGGGGCTTCAAT
GGCAACCACATTTCCGAGTGGTGTTTCTTGGATTGAGCAGTACACGACCATCCATGGGAAATCTGCAGCAT
TTTTTGTAATTGGTGCTTCCCTGGGAGAAATGGCTATTCCTGCAGTCATTGGAATTCTTCAAGGAAAATAC
CCTGATTTGCCTGTAGTTCTGTATACCTCTTTGGGAGCATCAATAGCTACTGGTATTTTATTTCCTGTGCT
ATATAAATTAGCCACTTCACCTCTTGATCGCCAGCGAAAAGAAGACAGAAAGAGTGAGGACCAGAAAGCTC
TGCTCTCTAGCTCCGGGCTAAATGAATATGAGGAAGAGAATGAAGAGGAGGATGCAGAAAAATGGAATGAA
ATGGATTTTGAAATGATTGAAACGAATGATACAATGAGGCATTCTATAATAGAGACATCTAGAAGTAGTCT
GACGGAGCCCACAGCTGAAGTCTATAATCAATACCCATCAAATGCACTGGTGTTTGAGTCTTCTCCTTTTA
ATACTGGCAGTGCCCATGTGAAGCACTTGCCAGAAACCAGGACAAAAGGGACTAACGTTTAGAGAAGATGG
ATTACTCACTGACATCTTTGAATAACTGCCACTTCTAAGGACGCCATTCAGAGCAGAGCATTAGCAGAATT
TCACCATGCTTTGGGGATTAAAATTTTTAGGTCCATATTATAGCAAATGCTAAAAATTTTTGAAATGTTCT
GTAATTCCCAGAGTCTTCCATAAATAACCAAATGGTCTCATACACGTACAATAGGATCTTGTTATAACGCT
ATCATTTGACATGTGACTGATTAGGAAATATTTTATGGTCTCCACCCCTCAGAGCACCCAAAGAAGGCATT
GTTTTTGAGAAGGGGGGTATTCTCAGAATGAGACCATGTAGGAGTTGAATTGATTGGTTGAACAAATTATC
AGGTATCAATATTTTTCCAGGAAAGTTGAGGAGAATAGTTTCATAAAAGAGAAATGGGCCGGGTGCAGTGG
CTAAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACGGGCAGATCACGAGGTCAGGAGATCGAGACCAT
CCTGGCTAACACAGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCAGGCGTGGTGGCAGGCACC
TGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGTGGAGCTTGCAGTGAG
CCGAGATCATGCCACACCGTACTCCAGCCTGGGCGACATCTAGACTCCGTCTCAAAAAAGAAATGGAGAGG
TATTTTTATTAATTGCTTTACCTTATCTGTAGGGTTGGATCAAAGAATGTAATGACTCATTTGATTTCTAA
TTTCAATGAAGAGAACACAGAACAGGGCCCAAGGGGTCGGGCTTCAAGCCTTGACTTCTTCTAGCCCTGCA
CTCTGTCACTTAACGTAGGCCTTGTTTATCTCGCCTCCTGATCTGTGATTCTAGAAATGAAAATGAGATGA
TTTCCAAATTGGGAATCTGTAAGTTTTAAAGTTTGTGCAAAATGTCTGCTAGAAATTAGTACAAGTGACAG

CATGTTCTTCATTTCAAATGAGAAGTGAAACCAAAATGATAAATAGCTTTAAGAATGTGCTAATGATAAAT
GATTACATGTCAATTTAATGTACTTAATGTTTAATACCTTATTTGAATAATTACCTGAAGAATATATTTTT
TAGTACTGCATTTCATTGATTCTAAGTTGCACTTTTTACCCCCATACTGTTAACATATCTGAAATCAGAAT
GTGTCTTACAATCAGTGATCGTTTAACATTGTGACAAAGTTTAATGGACAGTTTTTTCCCATATGTATATA
TAAAATAATGTGTTTTACAATCAGTGGCTTAGATTCAGTGAAATACAGTAATTCATTCAATTATGATAGTA
TCTTTACAGACATTTTAAAAATAAGTTATTTTTATATGCTAATATTCTATGTTCAAGTGGAATTTGGAGAC
ACTATGTTCTAATTTTAAGAACTGTTTCAGTGTTATTAGATTAGTAGTAGGGCTGTAGATGGGGTCTTTCG
AGGTCCTTTGCCTACCTCAGTAACAGAGGCTGGGTCTGAGCCTTGCTCGAGATTAGCTCTTCACCAGGTCT
GTGTGGTAATCATATCGTAGTTACCGTAGGTCACGCTGCAAGTGTAATTTACAGTGAGGCATAGTCTTTTC
CAGTGATTGGAATACTTGTGTCACCGATCCAGGGTTTCTCTTTCCTGAGTCTGAAATACCTGGGAAACCAG
TTAGAAGGTGACTGGGTAGCCTAGGAGTCAAGGATCAGAAAGGGTACTTCATATATATGTGGCAGAATTGG
TGCAGGTGATCCTTTGAAAAATGACAAGAATTGATGAAATTTTGTTTACTTCCTCCAATTTTTTGCTGAAT
AACATCTCCTTTCTACCTGATTGGAGTAGGCTTGCTTTCTCTTACTTAATTTCTATAAATTTGCCATTGTT
TTGTCTCAAGTAAAAAAAAAAAAAA

LELEVELRGAGPPAAGQRLLQAEAPAENEPEPEVVVVSWQSGGTGSKLRWFTTLMLCASFLGLGLSVAIVG
PTFQDLATNVNRNISSLSFIFVGRALGYLSGSVIGGFLVDVMNYFLLLGISMSATTVGLYLVPFCKTAILL
TVMMSIFGVSIGILDTGGNVLILAIWGDKGAPHMQALHFSFALGAFLAPLLAKLALGPTASAENHTESDFH
PALNQSSDADSEALFGVPNDKNLLWAYAVIGTYMFLVSVIFFCLFLKNSSKQEKARASAETFRRAKYHNAL
LCLLFLFFFYVGAEVTYGSYVFSFATTHAGMKESEAAGLNSIFWGTFAACRGLAIFFATCLQPGTMIVLS
NIGSLTSSLFLVLFDKNPICLWIATSVYGASMATTFPSGVSWIEQYTTIHGKSAAFFVIGASLGEMAIPAV
IGILQGKYPDLPVVLYTSLGASIATGILFPVLYKLATSPLDRQRKEDRKSEDQKALLSSSGLNEYEEENEE
EDAEKWNEMDFEMIETNDTMRHSIIETSRSSLTEPTAEVYNQYPSNALVFESSPFNTGSAHVKHLPETRTK
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
GGGGCAGTTCTCGGTGAGGAGGAAGAGAGCAGCGGACGGCACAGCACCCGCGCGGGCCCTCCCACAACAGC
TCCAGCTGGCAGCATCACTTCCCGCCAATTTATCCAACTTCTGCCAAGGCTCTGAAATGCCAACAACGTCG
AGGCCTGCACTTGATGTCAAGGGTGGCACCTCACCTGCGAAGGAGGATGCCAACCAAGAGATGAGCTCCGT
GGCCTACTCCAACCTTGCGGTGAAAGATCGCAAAGCAGTGGCCATTCTGCACTACCCTGGGGTAGCCTCAA
ATGGAACCAAGGCCAGTGGGGCTCCCACTAGTTCCTCGGGATCTCCAATAGGCTCTCCTACAACCACCCCT
CCCACTAAACCCCCATCCTTCAACCTGCACCCCGCCCCTCACTTGCTGGCTAGTATGCACCTGCAGAAACT
TAATAGCCAGTATCAGGGGATGGCTGCTGCCACTCCAGGCCAACCCGGGGAGGCAGGACCCCTGCAAAACT
GGGACTTTGGGGCCCAGGCGGGAGGGGCAGAATCACTCTCTCCTTCTGCTGGTGCCCAGAGCCCTGCTATC
ATCGATTCGGACCCAGTGGATGAGGAAGTGCTGATGTCGCTGGTGGTGGAACTGGGGTTGGACCGAGCCAA
TGAGCTTCCGGAGCTGTGGCTGGGGCAGAATGAGTTTGACTTCACTGCGGACTTTCCATCTAGCTGCTAATT
GCCAAGTGTCCCTAAAGATGGAGGAATAAAGCCACCAATTCTGTTGTAAATAAAAATAAAGTTACTTACAA
AGAGACGGGCCAAAAAAAAAA

LAASLPANLSNFCQGSEMPTTSRPALDVKGGTSPAKEDANQEMSSVAYSNLAVKDRKAVAILHYPGVASNG
TKASGAPTSSSGSPIGSPTTTPPTKPPSFNLHPAPHLLASMHLQKLNSQYQGMAAATPGQPGEAGPLQNWD
FGAQAGGAESLSPSAGAQSPAIIDSDPVDEEVLMSLVVELGLDRANELPELWLGQNEFDFTADFPSSC

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
CGGTCCAGTCATCGTTCTGTGGTCCTTTCTTTTATGATTCACAAGGAATGACCCTCTTCATCGCCTCTCCT
AATTCAGTCCTCACAACAGTCCTTTTACAAATGGGACAACAGGTTAGAGGAAGTCAGGCAGATTTCCAGCA
TCATAGAGAGTAAAGGACCAGGGAAGGATCAGGATTCAAGGACTGCACCCAGGCTCTGCTTCCAGCTTGCT
GTGTGACTTTGGGTAATTTTGTTCCCTTAGGGAACTGAGCTTTCTCATTTGTAAATGCAAACAGGCTGTTG
GGAGGATCAAATGAGATCCAGGGGTGAAAACAGCTTAGTTTACTTTCAGGAATTTACCCACGCGGTATATA
AAGGCAAAATATTATTATAGTCAGGTGATTGTAGATTGAGGAACCCATTTCCTCATTCTGCAAATTGCAAA
CCTGAGGGCCCAAAGAGGGACAGGGGCTTGCCCCAGGTCTCAGCAGGCTGTGAGCAAGAGCTAAAGCCTAA
TCCTCCTGCCTTTGGGCCTGGAGCCCTTCCTTGTACCCCAGGGGTCAGTGTCTTTGTTGGATACAGGCTTA
GATTGACTGACTGTACCCTGAGAACCTAGGGGAGTCCCTGTTCCCAATTCTTCTCCTACCCCCACCTTGGC
CTGATGGAGGAAGACCCTGCTGTGTTGAGATGAGCACCAGAGCCAAGAAGCTGAGGAGGATCTGGAGAATT
CTGGAGGAAGAGGAGAGTGTTGCTGGAGCTGTACAGACCCTGCTTCTCAGGTCCCAGGAAGGTGGCGTCAG
CATCTGCAGCCGCGTCGACGTTGTCGGAGCCTCCGCGGAGGACCCAGGAGAGCCGGACTAGGACCAGGGCC
CTGGGCCTCCCCACACTCCCCATGGAGAAGCTGGCGGCCTCTACAGAGCCCCAAGGGCCTCGGCCGGTCCT
GGGCCGTGAGAGTGTCCAGGTGCCCGATGACCAAGACTTTCGCAGCTTCCGGTCAGAGTGTGAGGCTGAGG
TGGGCTGGAACCTGACCTATAGCAGGGCTGGGGTGTCTGTCTGGGTGCAGGCTGTGGAGATGGATCGGACG
CTGCACAAGATCAAGTGCCGGATGGAGTGCTGTGATGTGCCAGCCGAGACACTCTACGACGTCCTACACGA
CATTGAGTACCGCAAGAAATGGGACAGCAACGTCATTGAGACTTTTGACATCGCCCGCTTGACAGTCAACG
CTGACGTGGGCTATTACTCCTGGAGGTGTCCCAAGCCCCTGAAGAACCGTGATGTCATCACCCTCCGCTCC
TGGCTCCCCATGGGCGCTGATTACATCATTATGAACTACTCAGTCAAACATCCCAAATACCCACCTCGGAA
AGACTTGGTCCGAGCTGTGTCCATCCAGACGGGCTACCTCATCCAGAGCACAGGGCCCAAGAGCTGCGTCA
TCACCTACCTGGCCCAGGTGGACCCCAAAGGCTCCTTACCCAAGTGGGTGGTGAATAAATCTTCTCAGTTC
CTGGCTCCCAAGGCCATGAAGAAGATGTACAAGGCGTGCCTCAAGTACCCCGAGTGGAAACAGAAGCACCT
GCCTCACTTCAAGCCGTGGCTGCACCCGGAGCAGAGCCCGTTGCCGAGCCTGGCGCTGTCGGAGCTGTCGG
TGCAGCATGCGGACTCACTGGAGAACATCGACGAGAGCGCGGTGGCCGAGAGCAGAGAGGAGCGGATGGGC
GGCGCGGGCGGCGAGGGCAGCGACGACGACACCTCGCTCACCTGAGCGCCGCACCGCTTCAGGGACGGAGA
CAGGACCGGGCGAGCCCTGGGGCGGCGGCCGCTCCTGCACTTTCTCCCCTCCCCCACCCGGCACCTGGTGG
CACCGGGCCAGGCCCAGGCGGGTGCTGCAGCCTGGCTGGACAGAGCCCCAATAAACGATCCCACAGCCTCA

VASASAAASTLSEPPRRTQESRTRTRALGLPTLPMEKLAASTEPQGPRPVLGRESVQVPDDQDFRSFRSEC
EAEVGWNLTYSRAGVSVWVQAVEMDRTLHKIKCRMECCDVPAETLYDVLHDIEYRKKWDSNVIETFDIARL
TVNADVGYYSWRCPKPLKNRDVITLRSWLPMGADYIIMNYSVKHPKYPPRKDLVRAVSIQTGYLIQSTGPK
SCVITYLAQVDPKGSLPKWVVNKSSQFLAPKAMKKMYKACLKYPEWKQKHLPHFKPWLHPEQSPLPSLALS
ELSVQHADSLENIDESAVAESREERMGGAGGEGSDDDTSLT

AUG initiation = 33.0 kDa
GUG initiation = 36.6 kDa

B)

#1

NM_002097

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

GTF3A general transcription factor IIIA

Synonyms: AP2; TFIIIA; GTF3A

```
TGCGCGATCTCCCGGAGCATGCGCAGCAGCGGCGCCGACGCGGGGCGGTGCCTGGTGACCGCGCGCGCTCC
CGGAAGTGTGCCGGCGTCGCGCGAAGGTTCAGCAGGGAGCCGTGGGCCGGGCGCGCCGGTTCCCGGCACGT
GTCTCGGCACGTGGCAGCGCGCCTGGCCCTGGGCTTGGAGGCGCCGGCGCCCTGGATCCGCCGGCCGTGGT
CGCCGAGTCGGTGTCGTCCTTGACCATCGCCGACGCGTTCATTGCAGCCGGCGAGAGCTCAGCTCCGACCC
CGCCGCGCCCCGCGCTTCCCAGGAGGTTCATCTGCTCCTTCCCTGACTGCAGCGCCAATTACAGCAAAGCC
TGGAAGCTTGACGCGCACCTGTGCAAGCACACGGGGGAGAGACCATTTGTTTGTGACTATGAAGGGTGTGG
CAAGGCCTTCATCAGGGACTACCATCTGAGCCGCCACATTCTGACTCACACAGGAGAAAAGCCGTTTGTTT
GTGCAGCCAATGGCTGTGATCAAAAATTCAACACAAAATCAAACTTGAAGAAACATTTTGAACGCAAACAT
GAAAATCAACAAAAACAATATATATGCAGTTTTGAAGACTGTAAGAAGACCTTTAAGAAACATCAGCAGCT
GAAAATCCATCAGTGCCAGCATACCAATGAACCTCTATTCAAGTGTACCCAGGAAGGATGTGGGAAACACT
TTGCATCACCCAGCAAGCTGAAACGACATGCCAAGGCCCACGAGGGCTATGTATGTCAAAAAGGATGTTCC
TTTGTGGCAAAAACATGGACGGAACTTCTGAAACATGTGAGAGAAACCCATAAAGAGGAAATACTATGTGA
AGTATGCCGGAAAACATTTAAACGCAAAGATTACCTTAAGCAACACATGAAAACTCATGCCCCAGAAAGGG
ATGTATGTCGCTGTCCAAGAGAAGGCTGTGGAAGAACCTATACAACTGTGTTTAATCTCCAAAGCCATATC
CTCTCCTTCCATGAGGAAAGCCGCCCTTTTGTGTGTGAACATGCTGGCTGTGGCAAAACATTTGCAATGAA
ACAAAGTCTCACTAGGCATGCTGTTGTACATGATCCTGACAAGAAGAAAATGAAGCTCAAAGTCAAAAAAT
CTCGTGAAAAACGGAGTTTGGCCTCTCATCTCAGTGGATATATCCCTCCCAAAAGGAAACAAGGGCAAGGC
TTATCTTTGTGTCAAAACGGAGAGTCACCCAACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTACT
TACCCTTGGCTAAGAACTGCACTGCTTTGTTTAAAGGACTGCAGACCAAGGAGCGAGCTTTCTCTCAGAGC
ATGCTTTTCTTTATTAAAATTACTGATGCAGAACATTTGATTCCTTATCATTTC

LDPPAVVAESVSSLTIADAFIAAGESSAPTPPRPALPRRFICSFPDCSANYSKAWKLDAHLCKHTGERPFV
CDYEGCGKAFIRDYHLSRHILTHTGEKPFVCAANGCDQKFNTKSNLKKHFERKHENQQKQYICSFEDCKKT
FKKHQQLKIHQCQHTNEPLFKCTQEGCGKHFASPSKLKRHAKAHEGYVCQKGCSFVAKTWTELLKHVRETH
KEEILCEVCRKTFKRKDYLKQHMKTHAPERDVCRCPREGCGRTYTTVFNLQSHILSFHEESRPFVCEHAGC
GKTFAMKQSLTRHAVVHDPDKKKMKLKVKKSREKRSLASHLSGYIPPKRKQGQGLSLCQNGESPNCVEDKM
LSTVAVLTLG
```

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

#2

NM_001418

Region conserved from insects to human.

EIF4G2 eukaryotic translation initiation factor 4 gamma

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

```
GCAGACGGCAACCGGGCCGCTGATTGGGCGGCGAAGGAGCCATTCGGGGAGACTCTGGTGGGTTCGGCTGC
CCCAAGAGTGATAAGTTCGGCTTCAGACACGCCTTAGCGCCAGCAGTGAGTCGGAGCTCTATGGAGGTGGC
AGCGGGTACCGAGTGGCGGCTGCAGCAGCGACTCCTCTGAGCTGAGTTTGAGGCCGTCCCCGACTCCTTCC
TCCCCCTTCCCTCCCCCTTTTTTTTGTTTTCCGTTCCCCTTTCCCCTCCCTTCCCTATCCCCGACGACCGG
ATCCTGAGGAGGCAGCTGCGGTGGCAGCTGCTGAGTTCTCGGTGAAGGTATTTCATTTCTCCTGTCCCCTC
```

CCCTCCCCACCCCATCTATTAATATTATTCTTTTGAAGATTCTTCGTTGTCAAGCCGCCAAAGTGGAGAGT
GCGATTGCAGAAGGGGGTGCTTCTCGTTTCAGTGCTTCTTCGGGCGGAGGAGGAAGTAGGGGTGCACCTCA
GCACTATCCCAAGACTGCTGGCAACAGCGAGTTCCTGGGGAAAACCCCAGGGCAAAACGCTCAGAAATGGA
TTCCTGCACGAAGCACTAGACGAGATGACAACTCCGCAGCAAACAACTCCGCAAACGAAAAAGAACGACAT
GATGCAATCTTCAGGAAAGTAAGAGGCATACTAAATAAGCTTACTCCTGAAAAGTTTGACAAGCTATGCCT
TGAGCTCCTCAATGTGGGTGTAGAGTCTAAACTCATCCTTAAAGGGGTCATACTGCTGATTGTGGACAAAG
CCCTAGAAGAGCCAAAGTATAGCTCACTGTATGCTCAGCTATGTCTGCGATTGGCAGAAGATGCACCAAAC
TTTGATGGCCCAGCAGCAGAGGGTCAACCAGGACAGAAGCAAAGCACCACATTCAGACGCCTCCTAATTTC
CAAATTACAAGATGAATTTGAAAACCGAACTAGAAATGTTGATGTCTATGATAAGCGTGAAAATCCCCTCC
TCCCCGAGGAGGAGGAACAGAGAGCCATTGCTAAGATCAAGATGTTGGGAAACATCAAATTCATTGGAGAG
CTTGGCAAGCTTGATCTTATTCACGAATCTATCCTTCATAAGTGCATCAAAACACTTTTGGAAAAGAAGAA
GAGAGTCCAACTCAAAGATATGGGAGAGGATTTGGAGTGCCTCTGTCAGATAATGAGGACAGTGGGACCTA
GATTAGACCATGAACGAGCCAAGTCCTTAATGGATCAGTACTTTGCCCGAATGTGCTCCTTGATGTTAAGT
AAGGAATTGCCAGCAAGGATTCGTTTCCTGCTGCAGGATACCGTAGAGTTGCGAGAACACCATTGGGTTCC
TCGCAAGGCTTTTCTTGACAATGGACCAAAGACGATCAATCAAATTCGTCAAGATGCAGTAAAAGATCTAG
GGGTGTTTATTCCTGCTCCTATGGCTCAAGGGATGAGAAGTGACTTCTTTCTGGAGGGACCGTTCATGCCA
CCCAGGATGAAAATGGATAGGGACCCACTTGGAGGACTTGCTGATATGTTTGGACAAATGCCAGGTAGCGG
AATTGGTACTGGTCCAGGAGTTATCCAGGATAGATTTTCACCCACCATGGGACGTCATCGTTCAAATCAAC
TCTTCAATGGCCATGGGGGACACATCATGCCTCCCACACAATCGCAGTTTGGAGAGATGGGAGGCAAGTTT
ATGAAAAGCCAGATTAGCCTGAGGCCTGCTCAGTCGTTCCTAATGAATAAAAATCAAGTGCCAAAGCTTCA
GCCCCAGATAACTATGATTCCTCCTAGTGCACAACCACCACGCACTCAAACACCACCTCTGGGACAGACAC
CTCAGCTTGGTCTCAAAACTAATCCACCGCTTATCCAGGAAAAGCCTGCCAAGACCAGCAAAAAGCCACCA
CCGTCAAAGGAAGAACTCCTTAAACTAACTGAAACTGTTGTGACTGAATATCTAAATAGTGGAAATGCAAA
TGAGGCTGTCAATGGTGTAAGAGAAATGAGGGCTCCTAAACACTTTCTTCCTGAGATGTTAAGCAAAGTAA
TCATCCTGTCACTAGATAGAAGCGATGAAGATAAAGAAAAAGCAAGTTCTTTGATCAGTTTACTCAAACAG
GAAGGGATAGCCACAAGTGACAACTTCATGCAGGCTTTCCTGAATGTATTGGACCAGTGTCCCAAACTGGA
GGTTGACATCCCTTTGGTGAAATCCTATTTAGCACAGTTTGCAGCTCGTGCCATCATTTCAGAGCTGGTGA
GCATTTCAGAACTAGCTCAACCACTAGAAAGTGGCACCCATTTTCCTCTCTTCCTACTTTGTCTTCAGCAG
TTAGCTAAATTACAAGATCGAGAATGGTTAACAGAACTTTTTCAACAAAGCAAGGTCAATATGCAGAAAAT
GCTCCCAGAAATTGATCAGAATAAGGACCGCATGTTGGAGATTTTGGAAGGAAAGGGACTGAGTTTCTTAT
TCCCACTCCTCAAATTGGAGAAGGAACTGTTGAAGCAAATAAAGTTGGATCCATCCCCTCAAACCATATAT
AAATGGATTAAAGATAACATCTCTCCCAAACTTCATGTAGATAAAGGATTTGTGAACATCTTAATGACTAG
CTTCTTACAGTACATTTCTAGTGAAGTAAACCCCCCCAGCGATGAAACAGATTCATCCTCTGCTCCTTCCA
AAGAACAGTTAGAGCAGGAAAAACAACTACTACTATCTTTCAAGCCAGTAATGCAGAAATTTCTTCATGAT
CACGTTGATCTACAAGTCAGTGCCCTGTATGCTCTCCAGGTGCACTGCTATAACAGCAACTTCCCAAAAGG
CATGTTACTTCGCTTTTTTGTGCACTTCTATGACATGGAAATTATTGAAGAAGAAGCTTTCTTGGCTTGGA
AAGAAGATATAACCCAAGAGTTTCCGGGAAAAGGCAAGGCTTTGTTCCAGGTGAATCAGTGGCTAACCTGG
TTAGAAACTGCTGAAGAAGAAGAATCAGAGGAAGAAGCTGACTAAAGAACCAGCCAAAGCCTTAAATTGTG
CAAAACATACTGTTGCTATGATGTAACTGCATTTGACCTAACCACTGCGAAAATTCATTCCGCTGTAATGT
TTTCACAATATTTAAAGCAGAAGCACGTCAGTTAGGATTTCCTTCTGCATAAGGTTTTTTTGTAGTGTAAT
GTCTTAATCATAGTCTACCATCAAATATTTTAGGAGTATCTTTAATGTTTAGATAGTATATTAGCAGCATG
CAATAATTACATCATAAGTTCTCAAGCAGAGGCAGTCTATTGCAAGGACCTTCTTTGCTGCCAGTTATCAT
AGGCTGTTTTAAGTTAGAAAACTGAATAGCAACACTGAATACTGTAGAAATGCACTTTGCTCAGTAATACT
TGAGTTGTTGCAATATTTGATTATCCATTTGGTTGTTACAGAAAAATTCTTAACTGTAATTGATGGTTGTT
GCCGTAATAGTATATTGCCTGTATTTCTACCTCTAGTAATGGGCTTTATGTGCTAGATTTTAATATCCTTG
AGCCTGGGCAAGTGCACAAGTCTTTTTAAAAGAAACATGGTTTACTTGCACAAAACTGATCAGTTTTGAGA
GATCGTTAATGCCCTTGAAGTGGTTTTTGTGGGTGTGAAACAAATGGTGAGAATTTGAATTGGTCCCTCCT
ATTATAGTATTGAAATTAAGTCTACTTAATTTATCAAGTCATGTTCATGCCCTGATTTTATATACTTGTAT
CTATCAATAAACATTGTGATACTTGATGTAGTGA

VESAIAEGGASRFSASSGGGGSRGAPQHYPKTAGNSEFLGKTPGQNAQKWIPARSTRRDDNSAANNSANEK
ERHDAIFRKVRGILNKLTPEKFDKLCLELLNVGVESKLILKGVILLIVDKALEEPKYSSLYAQLCLRLAED
APNFDGPAAEGQPGQKQSTTFRRLLISKLQDEFENRTRNVDVYDKRENPLLPEEEEQRAIAKIKMLGNIKF
IGELGKLDLIHESILHKCIKTLLEKKKRVQLKDMGEDLECLCQIMRTVGPRLDHERAKSLMDQYFARMCSL
MLSKELPARIRFLLQDTVELREHHWVPRKAFLDNGPKTINQIRQDAVKDLGVFIPAPMAQGMRSDFFLEGP
FMPPRMKMDRDPLGGLADMFGQMPGSGIGTGPGVIQDRFSPTMGRHRSNQLFNGHGGHIMPPTQSQFGEMG

GKFMKSQISLRPAQSFLMNKNQVPKLQPQITMIPPSAQPPRTQTPPLGQTPQLGLKTNPPLIQEKPAKTSK
KPPPSKEELLKLTETVVTEYLNSGNANEAVNGVREMRAPKHFLPEMLSKVIILSLDRSDEDKEKASSLISL
LKQEGIATSDNFMQAFLNVLDQCPKLEVDIPLVKSYLAQFAARAIISELVSISELAQPLESGTHFPLFLLC
LQQLAKLQDREWLTELFQQSKVNMQKMLPEIDQNKDRMLEILEGKGLSFLFPLLKLEKELLKQIKLDPSPQ
TIYKWIKDNISPKLHVDKGFVNILMTSFLQYISSEVNPPSDETDSSSAPSKEQLEQEKQLLLSFKPVMQKF
LHDHVDLQVSALYALQVHCYNSNFPKGMLLRFFVHFYDMEIIEEEAFLAWKEDITQEFPGKGKALFQVNQW
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

CCTCGCCCGCCTGCCGCCTTTTTGTGCGCGTGTGAGTGTGGGCCCCAGCGTGCCCTCCCGGGGGTGGGTTC
CGGGCGGAAGGCGGAGGCCCGGCGCGCAGCCCGCCGCCCGCCTGCCCGCGGACCGGGGAGCCGGGGTGCTT
GGAGCGGGGGACGCCAGGCGTGGGCTGGCGGCGGGACCAGGAGGAGGAGGAGGAGGAGGAGGAGCGCGG
GCTGGCGCTTGCCCGGGCGCAGTCGGCGGGGACCGAGTCGTACTTCCTGTGCGAAAGGCGGCCCGACCCTA
ACCGCCACCCCCTCCCCCTGTCTCCCTCTCTGAACCCGCCCATTGGGGGTAGGACACTCAGCCGTCACCGC
TCGCTCTGCTGGCCGCTACCTGCAGCAAGATAGGGCCGCCATCGCCGGGCGACGACGAGGAGGAGGCGGCC
GCCGCAGCCGGGGCCCCCGCCGCCGCCGGAGCGACAGGTGATTTGGCTTCTGCACAGTTAGGAGGAGCACC
AAACCGATGGGAGGTTTTGTCAGCCACACCTACAACTATAAAAGAGATGAAGCTGGTAATCTAGTCCAGATTC
CAAGTGCTGCTACTTCAAGTGGGCAGTATGTTCTTCCCCTTCAGAATTTGCAGAATCAACAAATATTTTCC
GTTGCACCAGGATCAGATTCATCAAATGGTACAGTGTCCAGTGTTCAATATCAAGTGATACCACAGATCCA
GTCAGCAGATGGTCAGCAGGTTCAAATTGGTTTCACAGGCTCTTCAGATAATGGGGGTATAAATCAAGAAA
GCAGTCAAATTCAGATCATTCCTGGCTCTAATCAAACCTTACTTGCCTCTGGAACACCTTCTGCTAACATC
CAGAATCTCATACCACAGACTGGTCAAGTCCAGGTTCAGGGAGTTGCAATTGGTGGTTCATCTTTTCCTGG
TCAAACCCAAGTAGTTGCTAATGTGCCTCTTGGTCTGCCAGGAAATATTACGTTTGTACCAATCAATAGTG
TCGATCTAGATTCTTTGGGACTCTCGGGCAGTTCTCAGACAATGACTGCAGGCATTAATGCCGACGGACAT
TTGATAAACACAGGACAAGCTATGGATAGTTCAGACAATTCAGAAAGGACTGGTGAGCGGGTTTCTCCTGA
TATTAATGAAACTAATACTGATACAGATTTATTTGTGCCAACATCCTCTTCATCACAGTTGCCTGTTACGA
TAGATAGTACAGGTATATTACAACAAAACACAAATAGCTTGACTACATCTAGTGGGCAGGTTCATTCTTCA
GATCTTCAGGGAAATTATATCCAGTCGCCTGTTTCTGAAGAGACACAGGCACAGAATATTCAGGTTTCTAC
AGCACAGCCTGTTGTACAGCATCTACAACTTCAAGAGTCTCAGCAGCCAACCAGTCAAGCCCAAATTGTGC
AAGGTATTACACCACAGACAATCCATGGTGTGCAAGCCAGTGGTCAAAATATATCACAACAGGCTTTGCAA
AATCTTCAGTTGCAGCTGAATCCTGGAACCTTTTTAATTCAGGCACAGACAGTGACCCCTTCTGGACAGGT
AACTTGGCAAACGTTTCAAGTACAAGGGGTCCAGAACTTGCAGAATTTGCAAATACAGAATACTGCTGCCC
AACAAATAACTTTGACGCCTGTTCAAACCCTCACACTTGGTCAAGTTGCGGCAGGTGGAGCCTTCACTTCA
ACTCCAGTTAGTCTAAGCACTGGTCAGTTGCCAAATCTACAAACAGTTACAGTGAACTCTATAGATTCTGC
TGGTATACAGCTACATCCAGGAGAGAATGCTGACAGTCCTGCAGATATTAGGATCAAGGAAGAAGAACCTG
ATCCTGAAGAGTGGCAGCTCAGTGGTGATTCTACCTTGAATACCAATGACCTAACACACTTAAGAGTACAG
GTGGTAGATGAAGAAGGGGACCAACAACATCAAGAAGGAAAAAGACTTCGGAGGGTAGCTTGCACCTGTCC
CAACTGTAAAGAAGGTGGTGGAAGAGGTACCAATCTTGGGAAAAAGAAGCAACACATTTGTCATATACCAG
GATGTGGTAAAGTCTATGGGAAGACCTCACATCTGAGAGCTCATCTGCGTTGGCATTCTGGAGAACGCCCT
TTTGTTTGTAACTGGATGTACTGTGGTAAAAGATTTACTCGAAGTGATGAATTACAGAGGCACAGAAGAAC
ACATACAGGTGAGAAGAAATTTGTTTGTCCAGAATGTTCAAAACGCTTTATGAGAAGTGACCACCTTGCCA
AACATATTAAAACACACCAGAATAAAAAAGGTATTCACTCTAGCAGTACAGTGCTGGCATCTGTGGAAGCT
GCGCGAGATGATACTTTGATTACTGCAGGAGGAACAACGCTTATCCTTGCAAATATTCAACAAGGTTCTGT
TTCAGGGATAGGAACTGTTAATACTTCCGCCACCAGCAATCAAGATATCCTTACCAACACTGAAATACCTT
TACAGCTTGTCACAGTTTCTGGAAATGAGACAATGGAGTAAATATTACACAAATACTTATTCATTGTGGTT
ATTTTTATACAGTAGTGAGAAGAATATTGTTCCTAAGTTCTTAGATATCTTTTTATTGATGTGCAAAAATT

TTTGGATTGACAGTAACTTGGTTATACATGACACTGAAATGCCTTACTTTGTATGATATTCCATAGTATAT
TAAAAATGGTAAAATTGCATGGGTTTTGTAGGTACTTTTGGAATCTAGAAGAAATGAAATTTTACCAAGTT
ATATAAAGAGAAAATTGAATTTAACAATGCGAATGGTAGTCTAACCAAATGCATCAATCCTGTGTGGTTTA
GTGTAAAAATGAGAACATGTTGGTATTTATCTATTGTAAGATAAAAAAGCTGGTGGGTGAAAGAAATCATG
TTATGATAAAAAATTTTGTAATTTTCTTGATGACTGGAATTTTTATTATGCATAACTGACAAATCAAGTTT
CCAAGCAAATGTTACATAGTGTAGGCTTTACTTAGCTTATCAATTTGTCATTTTGAAGCTAATTATTTTAA
TTAGGTTAACTATGTACAATATTTTAAGCATTACTCTTGTAAGATTTTGAAAACTACATTTTAACATGGAA
CTCTAGGGATAGTCACCTTTTAAATCCTGTTGAAAAGCCATGTTTAAGATTTAATTTGCCAAAATAATGTC
TTGTTAATATTCTTTCAATAACGAAGTTGGGCAATATAACCAATGTTTAAAAAAGTTTAAAATGTATAAGT
TGAGGCATTTGGGTGGTAAGAGAATGTTATAGTGAATTATCCCTTTTCTTGACTATTGGAGGACCAAAAAA
ATAAGGTGTATTGCGTCTTAGCAGTGATTTTTATCCAATCTTGTTTCCAAAAACCATGTCTCCCAGGGCCT
TAAAAGCCATCATGTAAATTACCAGTAAAGTGTAACATATGCAAACATAACAAAATCACTTCCATAGTGAC
GATACTCCAACCATATGGATATTAGTCATAGAAGAACTAGAGGTTTTATGATATTTTTTTAAGTCTTTTTT
TTTTTTGTCTAGGTAGTCAGTCTGCACTTAAATATCAATCATTTTCCTTTTTTGCTTCTTCCCTTAAAATT
TATATGTATCCAGTACATTTAATTGAGAAGCGTATGTTTTTTATTATGCTGTATTTTCTTTTTATTTTTTA
ATTATTGTTTATATTTTCAATTCAAAAATGTACAAAATAAAGTTACATTGCTGGTCTTGTAAGAGCTATAC
AGTTTTCCTAAATGTATACCTGTAACTGCAGCAGTTCACCTATTTCAAAAATTTGGAATTCTGTTCATTTG
TTATTCTTAAGACCACCTCAAATTTAAAGGCTACCTTATTGTACGTTTAAAGTGTATTATAACAGTGTGGT
AGTTAATAAAACACTATTTTTTTTTCTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAA

IGPPSPGDDEEEAAAAAGAPAAAGATGDLASAQLGGAPNRWEVLSATPTTIKDEAGNLVQIPSAATSSGQY
VLPLQNLQNQQIFSVAPGSDSSNGTVSSVQYQVIPQIQSADGQQVQIGFTGSSDNGGINQESSQIQIIPGS
NQTLLASGTPSANIQNLIPQTGQVQVQGVAIGGSSFPGQTQVVANVPLGLPGNITFVPINSVDLDSLGLSG
SSQTMTAGINADGHLINTGQAMDSSDNSERTGERVSPDINETNTDTDLFVPTSSSSQLPVTIDSTGILQQN
TNSLTTSSGQVHSSDLQGNYIQSPVSEETQAQNIQVSTAQPVVQHLQLQESQQPTSQAQIVQGITPQTIHG
VQASGQNISQQALQNLQLQLNPGTFLIQAQTVTPSGQVTWQTFQVQGVQNLQNLQIQNTAAQQITLTPVQT
LTLGQVAAGGAFTSTPVSLSTGQLPNLQTVTVNSIDSAGIQLHPGENADSPADIRIKEEEPDPEEWQLSGD
STLNTNDLTHLRVQVVDEEGDQQHQEGKRLRRVACTCPNCKEGGGRGTNLGKKKQHICHIPGCGKVYGKTS
HLRAHLRWHSGERPFVCNWMYCGKRFTRSDELQRHRRTHTGEKKFVCPECSKRFMRSDHLAKHIKTHQNKK
GIHSSSTVLASVEAARDDTLITAGGTTLILANIQQGSVSGIGTVNTSATSNQDILTNTEIPLQLVTVSGNE
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
GTTGGCCAAGACGCCGAATATCAAAATCTTCAGCGGCAGCTCCCACCAGGACTTATCCCAGAAAATTGCTG
ACCGCCTGGGCCTGGAGCTAGGCAAGGTGGTGACTAAGAAATTCAGCAACCAGGAGACCTGCGTGGAAATT
GATGAGAGTGTGCGTGGAGAGGATGTCTACATCGTTCAGAGTGGTTGTGGCGAAATCAACGACAGTCTAAT
GGAGCTTTTGATCATGATTAATGCCTGCAAGATTGCTTCAGCTAGCCGAGTTACTGCAGTCATCCCATGCT
TCCCTTATGCCCGACAGGATAAGAAGGATAAGAGCCGGTCCCCAATCTCTGCCAAGCTTGTTGCAAATATG
CTCTCTATAGCAGGTGCGGATCATATCATCACCATGGACCTACATGCTTCTCAAATTCAGGGCTTTTTTGA
TATCCCAGTAGACAACTTGTATGCAGAGCCAACTGTCCTGAAGTGGATAAGGGAGAATATCCCTGAGTGGA
AGAACTGCATTATTGTCTCGCCAGATGCTGGTGGAGCTAAAAGAGTGACCTCCATTGCAGACCAGTTGAAT

GTGGACTTTGCTTTGATTCATAAAGAACGGAAGAAGGCCAATGAAGTGGACTGCATAGTGCTAGTGGGAGA
TGTGAATGATCGTGTGGCTATCCTTGTAGATGACATGGCAGACACTTGTGTTACAATCTGCCTCGCAGCTG
ACAAACTTCTCTCAGCTGGAGCAACCAGAGTTTATGCTATCTTGACTCATGGAATCTTTTCTGGCCCAGCC
ATTTCTCGCATCAACACTGCATGCTTTGAAGCAGTGGTAGTCACCAATACCATACCTCAAGATGAGAAGAT
GAAGCATTGCTCCAAAATACGAGTAATTGACATCTCCATGATCCTTGCAGAAGCCATAAGGAGAACTCATA
ATGGGGAATCTGTTTCCTACCTGTTCAGCCATGTTCCTTTATAACAGAATAACTTCTAGGTTATGCTATTT
TAAAATAAAATAAGATTAATAAAAAA

TPNIKIFSGSSHQDLSQKIADRLGLELGKVVTKKFSNQETCVEIDESVRGEDVYIVQSGCGEINDSLMELL
IMINACKIASASRVTAVIPCFPYARQDKKDKSRSPISAKLVANMLSIAGADHIITMDLHASQIQGFFDIPV
DNLYAEPTVLKWIRENIPEWKNCIIVSPDAGGAKRVTSIADQLNVDFALIHKERKKANEVDCIVLVGDVND
RVAILVDDMADTCVTICLAADKLLSAGATRVYAILTHGIFSGPAISRINTACFEAVVVTNTIPQDEKMKHC
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

GTGGCCCCGCCGCCACTCCCTCCGGCTCCCTCCCTCCCGCCGCGGCGCGCATCTCATTCCAGCCCTCATTC
CGCGCATTCCAGCGTCCTCCTCGCACACTCGAGGCCAGGGGGCGGGAGGGCCGCAGCTCCGGCGCCGCCGC
GTCCCGCCAGAACGATCGCCGCGGCCGGAAGAGTTGGCGCTCGGGGCGGACTCCTTGGAACTGGCTTAGCG
CACCCATCCCACCTTCCCGCACCCTGGGACCGGTCCAACGAGCGCTCCTCCAAGCGGAGCCTTGGAGGGCA
CGGCCGGCACCATTACCTCCAACGAGTGGAGCTCTCCCACCTCCCCTGAGGGGAGCACCGCCTCTGGGGGC
AGTCAGGCACTGGACAAGCCCATCGACAATGACGCAGAGGGCGTGTGGAGCCCGGATATTGAGCAGAGTTT
CCAGGAGGCCCTCGCCATCTACCCGCCCTGTGGCAGGCGCAAAATCATCCTGTCGGACGAGGGCAAGATGT
ATGGTCGGAACGAGCTGATTGCCCGCTACATCAAGCTCCGGACAGGGAAGACCCGCACCAGGAAGCAGGTC
TCCAGCCACATCCAGGTGCTGGCTCGTCGCAAAGCTCGCGAGATCCAGGCCAAGCTAAAGGACCAGGCAGC
TAAGGACAAGGCCCTGCAGAGCATGGCTGCCATGTCGTCTGCACAGATCATCTCCGCCACGGCCTTCCACA
GTAGCATGGCCCTCGCCCGGGGCCCCGGCCGCCCAGCAGTCTCAGGGTTTTGGCAAGGAGCTTTGCCAGGC
CAAGCCGGAACGTCCCATGATGTGAAGCCTTTCTCTCAGCAAACCTATGCTGTCCAGCCTCCGCTGCCTCT
GCCAGGGTTTGAGTCTCCTGCAGGGCCCGCCCCATCGCCCTCTGCGCCCCCGGCACCCCCATGGCAGGGCC
GCAGCGTGGCCAGCTCCAAGCTCTGGATGTTGGAGTTCTCTGCCTTCCTGGAGCAGCAGCAGGACCCGGAC
ACGTACAACAAGCACCTGTTCGTGCACATTGGCCAGTCCAGCCCAAGCTACAGCGACCCCTACCTCGAAGC
CGTGGACATCCGCCAAATCTATGACAAATTCCCGGAGAAAAAGGGTGGACTCAAGGATCTCTTCGAACGGG
GACCCTCCAATGCCTTTTTTCTTGTGAAGTTCTGGGCAGACCTCAACACCAACATCGAGGATGAAGGCAGC
TCCTTCTATGGGGTCTCCAGCCAGTATGAGAGCCCCGAGAACATGATCATCACCTGCTCCACGAAGGTCTG
CTCTTTCGGCAAGCAGGTGGTGGAGAAAGTTGAGACAGAGTATGCTCGCTATGAGAATGGACACTACTCTT
ACCGCATCCACCGGTCCCCGCTCTGTGAGTACATGATCAACTTCATCCACAAGCTCAAGCACCTCCCTGAG
AAGTACATGATGAACAGCGTGCTGGAGAACTTCACCATCCTGCAGGTGGTCACCAACAGAGACACACAGGA
GACCTTGCTGTGCATTGCCTATGTCTTTGAGGTGTCAGCCAGTGAGCACGGGGCTCAGCACCACATCTACA
GGCTGGTGAAAGAATGAGAGACTCGGGGAGCAGGGAGGGGGGAAGAGACGTGTGTGCAGGAAACGGGGACG
TGGGGAGGGGACCTGCAGGGGCAGCCCCCTGAAGTGCCAAGAGAGCTGAGAGGAGCAGTTGTGACTCTACC
CAGGAACAAACTGTGCCTGAACCTGAGGTGCCCAACCCCAAATAAACCCAAGATGCTGTGTATTTTCAGAG
GAAAA

LEGTAGTITSNEWSSPTSPEGSTASGGSQALDKPIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDE
GKMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLARRKAREIQAKLKDQAAKDKALQSMAAMSSAQIISAT
AFHSSMALARGPGRPAVSGFWQGALPGQAGTSHDVKPFSQQTYAVQPPLPLPGFESPAGPAPSPSAPPAPP

WQGRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDKFPEKKGGLKDL
FERGPSNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPENMIITCSTKVCSFGKQVVEKVETEYARYENG
HYSYRIHRSPLCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTNRDTQETLLCIAYVFEVSASEHGAQH
HIYRLVKE

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

#6

NM_003214

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

TEAD3 TEA domain family member 3

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

TCCTCAACACAAACTTTCCGTCCCGCTCGCTCCCTCCTCCGCGCTCGGCGCCTCCCGCTCCAGCCCGGCTC
ATTCCGCACATTCCGGCCAGCCCCCTCCCCACGACCCCCCTTCCCCGGCCCCCCTTGCGGCTCCCTCGGGC
CCGGCGGAGCGGCCCGGCCGGAGCGCCCCCGCGAGCTCGGACCAGGCTCAGCCGCCCAGTGGGCTCAGGCC
CAGAGCCCAGAGCAACCAGCACAATAGCGTCCAACAGCTGGAACGCCAGCAGCAGCCCCGGGGAGGCCCGG
GAGGATGGGCCCGAGGGCCTGGACAAGGGGCTGGACAACGATGCGGAGGGCGTGTGGAGCCCGGACATCGA
GCAGAGCTTCCAGGAGGCCCTGGCCATCTACCCGCCCTGCGGCCGGCGGAAGATCATCCTGTCAGACGAGG
GCAAGATGTACGGCCGAAATGAGTTGATTGCACGCTATATTAAACTGAGGACGGGGAAGACTCGGACGAGA
AAACAGGTGTCCAGCCACATACAGGTTCTAGCTCGGAAGAAGGTGCGGGAGTACCAGGTTGGCATCAAGGC
CATGAACCTGGACCAGGTCTCCAAGGACAAAGCCCTTCAGAGCATGGCGTCCATGTCCTCTGCCCAGATCG
TCTCTGCCAGTGTCCTGCAGAACAAGTTCAGCCCACCTTCCCCTCTGCCCCAGGCCGTCTTCTCCACTTCC
TCGCGGTTCTGGAGCAGCCCCCCTCTCCTGGGACAGCAGCCTGGACCCTCTCAGGACATCAAGCCCTTTGC
ACAGCCAGCCTACCCCATCCAGCCGCCCCTGCCGCCGACGCTCAGCAGTTATGAGCCCCTGGCCCCGCTCC
CCTCAGCTGCTGCCTCTGTGCCTGTGTGGCAGGACCGTACCATTGCCTCCTCCCGGCTGCGGCTCCTGGAG
TATTCAGCCTTCATGGAGGTGCAGCGAGACCCTGACACGTACAGCAAACACCTGTTTGTGCACATCGGCCA
GACGAACCCCGCCTTCTCAGACCCACCCCTGGAGGCAGTAGATGTGCGCCAGATCTATGACAAATTCCCCG
AGAAAAAGGGAGGATTGAAGGAGCTCTATGAGAAGGGGCCCCCTAATGCCTTCTTCCTTGTCAAGTTCTGG
GCCGACCTCAACAGCACCATCCAGGAGGGCCCGGGAGCCTTCTATGGGGTCAGCTCTCAGTACAGCTCTGC
TGATAGCATGACCATCAGCGTCTCCACCAAGGTGTGCTCCTTTGGCAAACAGGTGGTAGAGAAGGTGGAGA
CTGAGTATGCCAGGCTGGAGAACGGGCGCTTTGTGTACCGTATCCACCGCTCGCCCATGTGCGAGTACATG
ATCAACTTCATCCACAAGCTGAAGCACCTGCCCGAGAAGTACATGATGAACAGCGTGCTGGAGAACTTCAC
CATCCTGCAGGTGGTCACGAGCCGGGACTCCCAGGAGACCCTGCTTGTCATTGCTTTTGTCTTCGAAGTCT
CCACCAGTGAGCACGGGGCCCAGCACCATGTCTACAAGCTCGTCAAAGACTAGGGTGCCCTCTGCGCCTCC
TTAAGGATGCAGGGTGAGCATCTCCTCTCCACACCTGCCTGGCACCCCTGGGGGGGTCCAGGATTGAGGAT
TCATCTACCTGCCAGGCCTCAGGCCCAGGACCCAGGAGGCCTCCCCACCTACCCCAGCACACACACTCCCT
GCCACTGTTCTGCGCTTTAATTGTGGGAGAAGAGAGGAGAGGAGGGCTCAGCGGTGGGGCAGCCTGTCCGG
GGCGCTGACCCACCATCACCCTGCTCTGCCCAGCCTCGCGTGACCTCAGAGAGGTGGGGATAGGGGACACC
TTCAGCCTCCAGCATGTGTGGCCACTGTACCCCCACCCACCCTTGGGGGAGCATGATGGGCAGGTGAGGGC
AGGATGGAGACCAAGGGAGTCAGTGAGCAGAGGCCCTGGGAGTGTCCGGTTGGGGTTGGACTGAGGACAGA
GGGGCCCACACTTCCTTGCCCCTTTGTGTCCCAGGCCTGGTGCCCAGACTCCTTGCATGGCTTGTGTGGTC
CTCAGACTCCGCACAGCGAGCGTAGGTCTCTGGGTTTCAGATGAAGTGCCCAGGCTCCAGGAAGTTGAGGG
ACCCACAGGAGAGGTGGGCAGAGCTGGAGTTCTCATCCAGGGCTGCTTGTCCCCAGAGCCCAGGTTTATAC
TACCTCCCTGGGGCGGGGGCTGGCCGCAGGGTAGGGGAGAGGCTCTGCAGTGTGGAGTGGAGCCTCATCGA
GGGGCGCTGGGTTAGGGGAGCACCTGTTTCAGACTGGGCATGAAGAAGGGAGCACAGCAGCTACTAGACCC
CATTAGCACCTCATTAGCCCACAAGCCAGCCAGGGGCCCCAGGAAGATGGGGCACCCCCCAGCACCCTCCA
GATTGAGAGCAAGGTAGAGGAAGGAGTCCCAGCCTCTGGGCAGACCAGAGGCCCAGAGGGAGAGAGTAGCA
GAAGGCTTTTGATTTTTCTCTTGCCTGAGGCTTGAATCTGACAAACCCTTGGTGGGCACTGCTCCCTTAGG
TTCTTCCCCACCTCAATCTACCTGCCTAGAGTAGCAGCTCCCAGACCCAGTTCTGGGACTGAAGGTTAACC
CTTCACCTGCTGTCCCTTCTTAACACCCAGGCCCCCAGAGCCAGCTGGGCCTGTCCAGCAGCCACCTGTGG
GTATTTATGAGTTTCATATGAAGTACTGTGCCCCTTCCCTTCCTCATCCCGACCCTGCCCGAGCTTCCTGA
AGGTCCTCACTGTTTGCATATCGCTCAGGCCACCTCCAAACCCCACCTAGGTTTTATAATGTATATTATAT
ATTTTTTTGTGTATTTTTAAAATCCAGCTGTGATGGGTTATATCATAAATGCAGCTTGGGGTTGGAGCAGG

GGCCCTCAAAGGCCCAGCTCCTGCTCAAAAAAAAAAAAAAAAAAAAATTAAAGTTATTTGTTTGTGGGTCA
GTCATGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

IASNSWNASSSPGEAREDGPEGLDKGLDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRNE
LIARYIKLRTGKTRTRKQVSSHIQVLARKKVREYQVGIKAMNLDQVSKDKALQSMASMSSAQIVSASVLQN
KFSPPSPLPQAVFSTSSRFWSSPPLLGQQPGPSQDIKPFAQPAYPIQPPLPPTLSSYEPLAPLPSAAASVP
VWQDRTIASSRLRLLEYSAFMEVQRDPDTYSKHLFVHIGQTNPAFSDPPLEAVDVRQIYDKFPEKKGGLKE
LYEKGPPNAFFLVKFWADLNSTIQEGPGAFYGVSSQYSSADSMTISVSTKVCSFGKQVVEKVETEYARLEN
GRFVYRIHRSPMCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTSRDSQETLLVIAFVFEVSTSEHGAQ
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

CCCCGCTTCTGCCTGCGCTGTGAACCCCCCCCCAGCCGCCGGCACGGCCCCGCCCCCGCTGCCCCGGTGGT
GGCCCACGGCCCCCCGGCTGCCCGTGGTCAAACTGGAGTCGCTGAAGCGCTGGAACGAAGAGCGGGGCCTC
TGGTGCGAGAAGGGGGTGCAGGTGCTGCTGACGACGGTGGGCGCCTTCGCCGCCTTCGGCCTCATGACCAT
CGCCATCAGCACTGACTACTGGCTCTACACGCGCGCCCTCATCTGCAACACCACCAACCTCACGGCCGGCG
GCGACGACGGGACCCCCCACCGCGGGGGCGGCGGCGCCTCGGAGAAGAAGGACCCCGGCGGCCTCACGCAC
TCGGGCCTCTGGAGGATCTGCTGCCTGGAAGGGTTGAAAAGAGGCGTCTGCGTGAAGATCAATCATTTCCC
GGAGGACACGGACTACGACCACGACAGCGCGGAGTATCTACTCCGAGTTGTCCGGGCCTCCAGCATCTTCC
CCATCCTTAGCGCCATCCTGCTGCTGCTCGGGGGTGTGTGCGTGGCGGCCTCCCGCGTCTACAAGTCCAAG
AGGAACATCATTCTGGGCGCAGGGATCCTGTTCGTGGCAGCAGGCCTGAGCAACATCATCGGCGTGATCGT
GTACATCTCCGCCAACGCGGGCGAGCCGGGCCCGAAGCGGGACGAGGAGAAGAAAAACCACTACTCGTACG
GCTGGTCCTTCTACTTCGGCGGGCTGTCGTTCATCCTGGCCGAGGTGATAGGCGTGCTGGCCGTCAACATC
TACATCGAGCGCAGCCGCGAGGCGCACTGCCAGTCTCGCTCGGACCTGCTCAAGGCCGGCGGGGGCGCGGG
CGGCAGTGGCGGGAGCGGCCCCTCGGCCATCCTCCGTCTGCCCAGTTACCGCTTCCGCTACCGCCGCCGCT
CCCGCTCTAGCTCCCGCTCCAGCGAGCCGTCGCCGTCGCGGGACGCGTCTCCCGGCGGCCCCGGGGGCCCG
GGCTTTGCCTCCACGGACATCTCCATGTACACGCTCAGCCGCGACCCCTCCAAGGGCAGCGTGGCCGCGGG
GCTGGCGGGGGCCGGCGGCGGCGGCGGCGGCGCCGTGGGGGCGTTCGGCGGCGCGGCCGGGGGCGCCGGGG
GCGGCGGCGGAGGCGGCGGCGGGGCGGGTGCCGAGCGGGACCGCGGGGGGGCGTCCGGCTTCCTCACGCTG
CACAACGCCTTCCCCAAGGAGGCGGGCGGCGGCGTCACGGTCACGGTCACCGGGCCGCCCGCCCCGCCCGC
GCCCGCGCCACCCGCGCCCTCTGCGCCCGCCCCCGGGACCCTGGCCAAGGAGGCCGCCGCCTCCAACACCA
ACACGCTCAACAGGAAAACCACGCCTGTGTAGGGGCGCGGCGGGGGAGCCGAGGGGCGTGTCCGGGGC

LESLKRWNEERGLWCEKGVQVLLTTVGAFAAFGLMTIAISTDYWLYTRALICNTTNLTAGGDDGTPHRGGG
GASEKKDPGGLTHSGLWRICCLEGLKRGVCVKINHFPEDTDYDHDSAEYLLRVVRASSIFPILSAILLLLG
GVCVAASRVYKSKRNIILGAGILFVAAGLSNIIGVIVYISANAGEPGPKRDEEKKNHYSYGWSFYFGGLSF
ILAEVIGVLAVNIYIERSREAHCQSRSDLLKAGGGAGGSGGSGPSAILRLPSYRFRYRRRSRSSSRSSEPS
PSRDASPGGPGGPGFASTDISMYTLSRDPSKGSVAAGLAGAGGGGGGAVGAFGGAAGGAGGGGGGGGAGA
ERDRGGASGFLTLHNAFPKEAGGGVTVTVTGPPAPPAPAPPAPSAPAPGTLAKEAAASNTNTLNRKTTPV

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

#8

NM_016178

CUG initiation well conserved in mammals.

OAZ3 ornithine decarboxylase antizyme 3

Synonyms: AZ3; OAZ-t; TISP15; OAZ3

GTTGCCTAAACCTCTGCCACCCACCTGTGAACTTCACTTTGCCACAGGGAGGCACTGAACTGAAAAAACTGC
CTTGTAAGAGGTGTCGCCCCTCTGTCTACTCCCTTTCTTATATCAAGAGGGGAAAAACACGTAACTACCTC
TACCCGATCTGGTCACCATACGCCTATTACCTTTACTGTTACAAGTACCGGATCACTCTCCGGGAGAAGAT
GCTGCCTCGTTGTTATAAAAGCATCACTTATAAGGAAGAGGAGGACTTGACACTCCAGCCCCGTTCCTGCC
TCCAGTGCTCCTGAGTCCCTAGTAGGCCTCCAGGAGGGCAAAAGCACCGAGCAGGGTAACCACGACCAGCT
TAAAGAACTGTATTCGGCTGGGAACTTGACGGTGCTGGCTACTGACCCCCTGCTCCACCAGGACCCAGTAC
AGTTAGACTTTCACTTCCGCCTTACCTCCCAGACCTCTGCCCATTGGCACGGCCTTCTCTGTGACCGTCGA
CTCTTCCTGGATATCCCATATCAGGCCTTGGATCAAGGCAACCGGGAAAGTTTGACTGCAACCCTGGAGTA
CGTGGAAGAGAAGACAAATGTGGACTCTGTGTTTGTGAACTTCCAGAATGATCGGAACGACAGAGGTGCCC
TGCTGCGGGCCTTCAGCTACATGGGCTTTGAGGTGGTCAGACCAGATCACCCTGCCCTCCCTCCCTTGGAC
AATGTCATCTTTATGGTGTATCCCCTTGAAAGGGATGTTGGCCACCTGCCCAGTGAGCCTCCTTGAACATG
CTTATTCCAACGCTTTGAGGGGCTGGAAGCCTTGACACATGGAATCAGGGGCCCGGGATGTGATTCAGGAC
ACTTTCCATCCTAGGAATAAAGGGTAGTGCAATCAAAAAAAAAAAAAAA

LPCKRCRPSVYSLSYIKRGKTRNYLYPIWSPYAYYLYCYKYRITLREKMLPRCYKSITYKEEEDLTLQPRS
CLQCSESLVGLQEGKSTEQGNHDQLKELYSAGNLTVLATDPLLHQDPVQLDFHFRLTSQTSAHWHGLLCDR
RLFLDIPYQALDQGNRESLTATLEYVEEKTNVDSVFVNFQNDRNDRGALLRAFSYMGFEVVRPDHPALPPL
DNVIFMVYPLERDVGHLPSEPP

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

#9

NM_021961

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

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

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

TTCCGAACATTCTTAGCATCGCTCGCGCCGCGCCGCGCCGCCTGAGCCGAGCCGAGCCTCTGCTGCCGCCG
CCGCGGCCCCGCCGCCCGCCGCGGGCGCCCACCAAGCACTTTGCAGACTCGCTTCCACCCTGCGGGCCATT
CCGCGCGGCGGGGCCCGGGCCCGGGGCGGCCGCGTCCAGGCACAGGCCATGCAGTGACGCCCCCCCACCCC
TCCACCTTTGCCCGGAGCGCGGGCAGCAGCCCAGCGCGCCAGCCGGCCCCGGGGCAGGAGCGGTGCTAGGC
AGGGGTGGGGTGGCCGGGCCCAGGGACCGGGAGCCGGGGAGGGAGCCGGGCACCGAGCAGAGGGCGGGGGA
AGCGGCGCCGAAGTTTGCCTCGGACTCGCCGGGCGCTGCGGTGGCTCCCTGGGCCGAGGACTGTTGCTGCC
GCTGCCGCCGCCGCTTCATTGCACATTCAAGTGGAAAATTTTCAGGAGTCAGCAGAAACATTGTGTCCAAA
AAAGACTGAGTCGCAGTTACCACCAAACCCAGGAGGAGACTCTCCCTGGAAAACTTCCCTTCCCTTTCGGT
TTATTTTCTTGAAAAGGCTCCAGGCTTCGGCTTGGAAAATCCCACCGCCAAAATTGAGCCCAGCAGCTGGA
GCGGCAGTGAGAGCCCTGCCGAAAACATGGAAAGGATGAGTGACTCTGCAGATAAGCCAATTGACAATGAT
GCAGAAGGGGTCTGGAGCCCCGACATCGAGCAAAGCTTTCAGGAGGCCCTGGCTATCTATCCACCATGTGG
GAGGAGGAAAATCATCTTATCAGACGAAGGCAAAATGTATGGTAGGAATGAATTGATAGCCAGATACATCA
AACTCAGGACAGGCAAGACGAGGACCAGAAAACAGGTGTCTAGTCACATTCAGGTTCTTGCCAGAAGGAAA
TCTCGTGATTTTCATTCCAAGCTAAAGGATCAGACTGCAAAGGATAAGGCCCTGCAGCACATGGCGGCCAT
GTCCTCAGCCCAGATCGTCTCGGCCACTGCCATTCATAACAAGCTGGGGCTGCCTGGGATTCCACGCCCGA
CCTTCCCAGGGGCGCCGGGGTTCTGGCCGGGAATGATTCAAACAGGGCAGCCAGGATCCTCACAAGACGTC
AAGCCTTTTGTGCAGCAGGCCTACCCCATCCAGCCAGCGGTCACAGCCCCCATTCCAGGGTTTGAGCCTGC
ATCGGCCCCAGCTCCCTCAGTCCCTGCCTGGCAAGGTCGCTCCATTGGCACAACCAAGCTTCGCCTGGTGG
AATTTTCAGCTTTTCTCGAGCAGCAGCGAGACCCAGACTCGTACAACAAACACCTCTTCGTGCACATTGGG

CATGCCAACCATTCTTACAGTGACCCATTGCTTGAATCAGTGGACATTCGTCAGATTTATGACAAATTTCC
TGAAAAGAAAGGTGGCTTAAAGGAACTGTTTGGAAAGGGCCCTCAAAATGCCTTCTTCCTCGTAAAATTCT
GGGCTGATTTAAACTGCAATATTCAAGATGATGCTGGGGCTTTTTATGGTGTAACCAGTCAGTACGAGAGT
TCTGAAAATATGACAGTCACCTGTTCCACCAAAGTTTGCTCCTTTGGGAAGCAAGTAGTAGAAAAAGTAGA
GACGGAGTATGCAAGGTTTGAGAATGGCCGATTTGTATACCGAATAAACCGCTCCCCAATGTGTGAATATA
TGATCAACTTCATCCACAAGCTCAAACACTTACCAGAGAAATATATGATGAACAGTGTTTTGGAAAACTTC
ACAATTTTATTGGTGGTAACAAACAGGGATACACAAGAAACTCTACTCTGCATGGCCTGTGTGTTTGAAGT
TTCAAATAGTGAACACGGAGCACAACATCATATTTACAGGCTTGTAAAGGACTGAACATGGTTATTTATAT
ATATAGATATCTGTATATACACACACACATATGTGCACACACACACTCTCTCTCCATTATCGAACGACTGA
CTGTAAACCTCACCACACAGGGTGGTGCCCTGGCCCCGAGGTCACCCCGACTTTTCTAAATCTTGTTTGAG
TGAAGTCATTTTTTCATGTGTTCATACTATCATTGTAGCTGTGAAGTTCTGGTACAGTTGTAAAAAGAGAA
ATTGAGTTGTTTCTCTATGTTCTTCAGATGTGCAGCCCACAATTCCTCGGGAAAGGTGAACCTGAACAACC
CAAGTCTCTCTCTGCAGAGCCCTGTTTCTAATTGTGGTAGAAAATATTGAGACAGAGCATTTGCCATGGGA
CATTTACAGCCTTTATACAAATGTATTTAGTTCTCTTTTTTCCAACATAAAATTCTTGTTTTAAGATACAA
GTAAAATTAATCTTTAAATATAAATGTAAATTAGTACACAAAACTAAGAATCTTTAGACTTATCTTTGTAA
CTAATTAGGGTGGAAGTTATGAAAGAATGTAATTCACTAAATTATTTTTTAAATGAAACCTTTTTTTTTCT
TTTTGAAACCAAATGTTAAACTATAGCCTTAAGAAATGCTTGGTAGAAGTGTCCTAATGAGACAAATTTGT
ACTTTTATCCTCAAGGTTAACACTAATCTCCTAATCCATTAAACTCTTGAACAGGTATTACAAAGGAAGAA
AACTTCACCCCTTATCCTTAACATATATAGTATATTTAAAAAATATAAAATTGTATTGTACTAATGTGATG
ATGGATTATTTAATGAAAAAGAAAAAATGGCTCTTTTTGCAATAAGTAGATACATACTGAAAAAATCTAAA
CTTACAATGTTTATAGTCTTGTGTGTGCAGTTATATTTTATATGGACGACCAAATTTTTTATTAAGATGAG
TAAATATTTGAACCACTGAATTTTAATAACAAAATTTTAAAATTGGCATGAATACGGAATACTGCACTGTG
AGATGCAAAGTATACAGAATCTGTGGCTGGGAGAAAATTTCATCAAATAGACAAGTAAAAGGCTCATCAGT
TTTAGCATCTCTGCTCCCCAGAAAATTGTAAGCATCCTCACCAGCCTGTGGATACATTCTTTATTTCTAGT
GACCCAATATGCATATTAACCTGCTATAACTAGGGCTATATGTGTAGGTATGTGTATACATATACACAAAT
GCACATATAGAGTTAACACATTTAGTGAACACTTGTTTAGTGTCACTCAGTTTGCTAGGTGCTGATATGTA
CGTATATCTCAATGTGTCTGTAGACTTAGATACATCCTCTTGAAGCACATCCATTTCTTTAGCGTCTCTCA
GTAAGTTACAGTACTTGTTTGACTTAGGTTTAAGAGGCCCAGCTACCTATCTCTGACCTTTTCAAATAGGC
TCATTTGGGAGATTCTTTTGCCAGGAGAGATTCAACTTTCCAATCTAAGTATTCCAGAGCATTGCCCAGGC
AGAGTTGGTTTGATGTGGCCAGATGTTTTGAGTTATTTCCCTTAAGTGTTTCACTGGGGAGAGAACAGGGA
GTGCTCCTCCAGCTTCCCAAAGAAATATGTTTTTGTAAGTGGTAGGAACATGTGCACACAATAGAACATGA
AATAAGTTTTTTAACTTGTAAAACATGTCAAGATTTTTCCACCAAGCTAGAAAATAAAAAACTTAGTTCTA
CCACATCCAATTAACTTACACACCCCCTTCCCTGTCTCAACACCTGCTTTGACCCTGCTTTTCTATTATTA
CATCAGTCAGCATCTTGTGGTCCCTAACATGAGGATGTGGCTGGCTCGTGGGAAACAGCAAAACACTAAGC
CTGACCTCTCCCAAATTGGGAAGACCAGAGGAGAAAGTGCAAAACTGTCCCCATTTGGAATGCCCATTCCT
TCTAGAAACCAGTTGGACAGTGCTCCTCTGCCCTTCATAAACAGACTACTGTTGGGTCCCTGATTCCAGGC
TGGCCTGTGAAGGATTGCCCCAGGTGTCCCCTTTCACGGTTGTCACATTTACAGTGACTTCTGTTGAACAC
CCCTCTTAGGGATGTTTCTTTTGCTCTTATTTCCTGCATCTTTCCTTAAGGGAAGCCCCATCCTCTCCCAG
GACCAGGAGTTTATGACCAGGCGAGCACAAATGGCTAAAAGCCAAGCTGTCCTAGAACTTCAGTGGGAGAG
CTGTCTGGTTCATATTCTACCCAGGAATGGTACTTTTCAGTGCAGCCAGGAGGGCTCTTGGGATTTCCTTT
CCAAAGCACAAAAATACTGGGACCCAAGAAGAACAGCTAGAGGACAACTCTGTTGGCACAGAGACGGGGAC
AGCCCAGTCTGCTGACCTCACAGGGTCAGCTGGGCCCCCCTGGTGCTTCACCACCTGCATCCTCTTGCTCA
GAATGCCTTTGCAGTTGAGTTTTCTGGGTTTCTATGATTGACCTTGAGGTTTACTCCTTGCTCTTACAACA
TTTCTAAGGATTTTTAAAAGTTTACTTCTTGTCTTGTTCTTCTAAAGCTTTCTCCAGGACAGATATTTTCC
CTGTCTTAACCACTGGTCCAGTCATCCCAGTGGGCTTCTCTTTGTCTCTCCCAGATTAGACCTTTGGGTGA
GATTGGCATCACAACATCTAATCTGAGTCTGTCTTTTGTCCTTCATTCTGTATGGCAGTCTCCCTTTGTTA
TAAAAGCTTTCTAAAGCATACTAAAGAAGCCTTCCCAGAGCCCCGTCTTGCTTCTCTTCCAGGTGCTCTAT
CCCCTCGAGACCCTCTGGTGCCAGGCTTGCTTCACGGCCATCTTGTGTTGTCACTGCAGAGTTTGGAGGCC
AGTTTTCCACAGCCTAAACAGGGAGGAGCTGCAGAATGGGGCTCTGGTCTCTGGGCATTCATTTCCCTCAT
AGAGGCTGAGAATAAAACAAGGACTTATTCACACATGTTCTAGAACCCCAGAATGGCCCAAGTTACCTGAG
ACCAGGGTTTCTCAACCTTGACACCATTGACATTTTGGACTGGGTAATTCTTTGTTCTGCAGAGCTGTCCT
TTGCACTGTAGGAGATTTACTAATATCCCTGGCCTCTACCCAGTAGTACCACTAGCACCTATTCCCCACCC
AGCGTGTCTCCAGATATTGTCAAATATCCCATCGGGTGCAAAATGATCCCTGGTCAAGATCTGTTGCCCAA
GATGTTACAGGTCACAATGACCACATTTGAAATTGTTTTCCCTTTCATTTTACCCTGTGAAAGCATCTCTC
CTAGAGCCTTGCAAGAGGCAGGTGACATTGTGTCCATATTTCTTCCTGTTTCAGAACTTCTGTTTCACAAC
AATTTCTCTCTCGCTACAAGTATTCTTTCACTCAGCACTGGGGAAGTTGGGAACAGCTGGTCACCATCATC
CCTTTAATCAACTCACACCTGTTTAAAGAGTGTTTCTGATTTGACCTTCATCCCTTAGTTTACTGGCGTTA

AAAAAAGTCTCAGCAATTTTCATTATTTCTCGTGGGTCTCATTATCAAACCTTTACTTATTTCGGCATATT
TCCTCTGGGCTTCTTCTAGTTTCTGCCTTACAAGCAATGCTGTTCTGTAAATTTATTGAAACCTCTGGAAC
ATTTCACCTTTAGAGATGGAGGATGGAAGGATTGGTACCAGAAGAGGGCTAAGATACGTTTTCTGTCTTGA
GCTGAAAGCACAGTCTACTCTCCTTCGTTTTGTCGATGAGAAAGTTGAGGCCAGAGGGGAGGTGACATGTT
TAGAGTCACCCAGCTGGTTAGTGACAGAAAAAGCGTGAGAGTTGTCTAGGATTCCTGCCACTTTGGTCCCT
GGCCTCTCCTGGGGGAGGCTGCTGTTCTTAGGTGCTCTAAGCTTAATCCCTCAGAATGTGTGGACAGGTCA
GCTTAGAAGAGATGGGGAGATTCAGGATCCCCCTGTGCCAGAGCACAGCCTCACCGGATGCTGCTTCCCAC
ACTGAAGTGTCCTGTCCGACCATTGCTATCTGAGGCATCCACAAGCAGGTAGGAAAGCTGGCGAGCCATTT
TACTTCCTGAGGACAATTCCCCAGCCACAGGCTCTGAGTCAAATTTCTATTTGGTAAGCATCCTAGCAGCA
AAGTCCTGCACTCAGACCAGCCAAAAAACAGCCCCCATTCCAAGTACTTGGTGTCAAAAGTCCCCGAACGA
CTTTTAAACCCAAGTCTTCTTAAGGTTTCAGTACTGTGGTGGCTTTAGCAGTTGTTTTTGTGCAACTATAA
ATTATTTAAATCATCTGAGATGACAGTCAATTTTACAAACCAGGTACATATTAATTTGTATAATTTTGTAT
ATGCTCTGGTACACTACCTGAACTAACGAAGGGTAGAACTAATTCTGTTTGTCAGTGTTCACACCTGTAAC
ATTAGGAGGATATGTCTGCATTGCTTATTTCTTTATGTTGGTGTTTCTGTGGCAAAGCCCTGCACATGGCA
TTTCTGAAAAGCCTTAAATCTTTAAGATGTTGCATGTAGGGTATGCAGTGCAAAAGGCTGCCTCAGAACTG
TGAGCCCTTTTGTAAGCTGGAAGCATTTCTCTTACTACTGTTACTTTTGTAGGAAGTTTTCAATTCAGAGC
TGCCAAAGTGTTCCCGTAAGCAGTGCCTTAGTAATACCTTAGTCATGCCGCCAGCCTTTTCTTACACCAAT
TCCTAATGTTCATTTACGAATTGGCCCAATATTGGAAACAAAACAAGCAAAAATTGTCTTCATTTTTGTTT
TGTAAGCCCATTTTTTCTCCAGTTCTATAGGAAACTGACTGCTTGGTGTAAAATCCGAAACTGGACACAAG
TCAGTTCTTTCACCACACTCAAATGTATATACCAAAACAAAAGGTTGCAACTTCATAGTTTACTATGAAAA
GCAAATTGTACTTTTTAATGTTGCCTTTTAAATTCATGACCAAATACTTAGCTATTTGTGAATCTTCTGCA
CTCTAGCATGAAAGTGCCTTTGGTTTGAGATTCCAGCTTAGAAAAGTGCTGCCATAATAACGATAATTTGT
AGAGAGACCAAAAATATTTTGAGATCACCGTAATGCCTTTGGTTTACCGGGATGAGTAACCAACCACAGGC
CTCTGTTCACAAGAGCACGACGTGGTCCCCGCCTGCTGCTAGTCTGTCTGCCACTGGGGGCCTCCCAACAT
CCATAGCACACTTCAGCGGAAGGACCCCAGAAACTGTTGTGTTTGTGTGTGCTGATGACCTAGTGTGTCAT
TTCACCTCGTCACCCAGCCCTGCGTCCGGATGAGGGGACTTCTGCACAAATGACAGAATCTCGGCTGGTGG
ACAGATACTACAGCTTTCTCCTCCTCCTTGTGTTCGTGTTCAGTCTCTGTGGAGACTTTCTTTTCCATTCA
AATGACAGTGCGCACTTATCTGGTTTACACAATGATACCATTTTGAAAGTTGGAAGCCTCAAACTGAGACG
ACAGTGCAGAACAAAACAAAAGTGAGTTAGGGTCGTTAAAATTGAAGTGTTCTTCTTAGGGCAAACATGTT
GACTCCGAGTATTGTGTATGAATGTGCTACGAGAAACTTCCAAAGAGCACCATTCACAATTTGGCATTTTC
AAAGAATGTTCCAGCCCTCAAAGGGGCAACTCTTTAAAGTCCTTGTTGGCTTTTATCCAAACCTTGTAGAA
ATTGGGAAAGCTGATAGAGGTAAGGAAGACGAGTGAAAAGGACAAGAAGGCCAAACACCAGCCAAAAAGAA
ACTAGGAAAAAAAGATTTTCTTTGCTAATATAGATGTAAAAATAACATCAGACATCTTTGAAAATTAGCCT
CTAAACTCTTAATACATACGTTCTGTGTGTCTCTACCTGGCGTCTTTAAGAATATCCTCTCTGGGCTCTGA
AATTTTAGGAGTGATTCTTATCCACTCCAAGTTGTAAGTATTTGTAGAAATTTGTGCAAACAAACAAAAAC
TATCAAATGAAAAGAAAATGTACTCAACCTAACTTATAGTTAGCAGCTGGAATTCTCAACTCTTCCCTGCC
AGCACTATACCACAGTGTGGAAGAAATTAGTCAAATGCTTGTTTTCCTGCTTCTCTTTTCAACTGTTACTG
TGCTTTGTTTGAAAGTAGTTTTCTCTCTCAAAGCCGTTGCTTATATCGTTAAGAATGAAGGTTTGTGTTTA
AAATTTATTGCATTGCAAAGGGTAGTTTCACTGAAGTCATGCACCATTAAATAAGATGAAATATTTGTATT
TATTGTCCTACTTCCTAAGCCGTAACTTCTTTTCCTCTGTGAATTTGCATTGAGTCACTCATGCTACACTA
CATCGCTTTAGTATTTGAGATGGCATTTATGTTTCCTCTCGTTTATCATGAAATGGGGTCAGATTCCATCA
GATTCCACCTCTGTCAGGTGGACTCTTGTCTGCCTTCCATGATGAGATTTTTTTTCTCCTTCCCCTTTCTT
TAAGAGAGGCTGACAGATCTAGGTGTCAATCAATTGGAAACCAGTCTCTGATTTTTTTTCATTAGTTATTT
TCTATCATTAGTTTCACTGTGTAAATTAGATATCAACTGCACTTCTTTAAAAAAAAATACATCTCCCTATT
ACCTCCTTGAAAGATTTACTTCTGTAGGCCTTTTTCAATAGGCTCATGACTGCAGACAAGGAAAAAAAAAG
TAAAAACAAAAACAGTATGTGCCTGAAAATGACAAAAAAAAAATTTGTAACATTTAAAAAAGAAACCTGAA
TAGCCTTTAATTCTTTAATAATACACTTAAATTTTATGTAAATCGGTTTTCGCCACGTGTGTTTGTTCACA
TTCTAAATGACTTAATGGGATTCTCACGGTCTGTGTCTTTGTGTCACGTGTATAAAATGGGCTTGTGATGT
AAGCGTTTCATCTGGTCAGTGGTTCCTTTGATATTGTACTGCTGCTGGGAGTGGGCTGTGGAACCTGCCTT
CGGGTAACTGGGTTCCTCTTGGGTAGATTGGAGAGATGGGGGTGGGCGTGGGCAAATTCTCACACATGTTT
TCTTAACCTATTTGCAGAAACTTTCAAAAGGCATTTGATTAAACCTCTTGGCAGTACAGTATTCTTGTATT
TGTTAACGTCTGTGTTTAGGTACTGGTACCTTTTTGTTTTAAAATGTTCTAAGTGTTGGCTTTAAAGTGAA
TTTATCTTTAGTATGATAGTTATATGAAAATTATAGGATTTGTGTGCAGAGAATTTTTTTATAAAGTGCTT
TGTAAAAAAAAAAAAAATGTATTCTAGCTTTTGCGGTACATATGTGTGATAACTTTAATACCCATGACAGTT
AAGTGCAATTATTTCATCACTCTAAAAATGCTATTTTTGTGTCAGTTCCTGCAGGTGTTTTCATGTCTTTG
CAAAGTGACACATTTTGATGCCTTCTTGATAAAGTGGTAGACATTTTGTAGCTTTCTAGAAACTTTGTATT

CATACGGTATCAATGAAAAATAAAGAAAATGAAAGTGTGGGTCACCTTTTTTATCTGCAAAAAAAAAAAAAA
AAAAA

LENPTAKIEPSSWSGSESPAENMERMSDSADKPIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEG
KMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLARRKSRDFHSKLKDQTAKDKALQHMAAMSSAQIVSATA
IHNKLGLPGIPRPTFPGAPGFWPGMIQTGQPGSSQDVKPFVQQAYPIQPAVTAPIPGFEPASAPAPSVPAW
QGRSIGTTKLRLVEFSAFLEQQRDPDSYNKHLFVHIGHANHSYSDPLLESVDIRQIYDKFPEKKGGLKELF
GKGPQNAFFLVKFWADLNCNIQDDAGAFYGVTSQYESSENMTVTCSTKVCSFGKQVVEKVETEYARFENGR
FVYRINRSPMCEYMINFIHKLKHLPEKYMMNSVLENFTILLVVTNRDTQETLLCMACVFEVSNSEHGAQHH
IYRLVKD

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

#10

NM_001098504

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

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

Synonyms: P72; RH70; DKFZp761H2016; DDX17

GTTAAGTTGGAGCCGACTCAGCGGCGGCCGCCATTTTGTGCAGTCGCTGGGAAGGAAGGAGACGCCTAAAC
CGCGGCACTGCCCGGTTTGAGCGTAGCCAAACCTGCCCACCGGCTTTGTAGCCCCGATTCTCTGTGTTTTG
CTCCCGTCTCCGACGAGAGAGGCGGCGACGGTGGCGTCTGCGACGGGAGACAGCGCGTCGGAGCGAGAGAG
CGCTGCGCCTGCCGCCGCCCCAACAGCGGAGGCGCCGCCGCCATCGGTCGTCACCAGACCGGAGCCGCAGG
CCCTCCCGAGCCCGGCCATCCGTGCCCCGCTCCCAGATCTCTATCCTTTTGGGACCATGCGCGGAGGAGGC
TTTGGGGACCGGGACCGGGATCGTGACCGTGGAGGATTTGGAGCAAGAGGTGGTGGTGGCCTTCCCCCGAA
GAAATTTGGTAATCCTGGGGAGCGTTTGCGTAAAAAAAAGTGGGATTTGAGTGAGCTCCCCAAGTTTGAGA
AAAATTTTTATGTGGAACATCCGGAAGTAGCAAGGCTGACACCATATGAGGTTGATGAGCTACGCCGAAAG
AAGGAGATTACAGTGAGGGGGGGAGATGTTTGTCCTAAACCCGTGTTTGCCTTCCATCATGCTAACTTCCC
ACAATATGTAATGGATGTGTTGATGGATCAGCACTTTACAGAACCAACTCCAATTCAGTGCCAGGGATTTC
CGTTGGCTCTTAGTGGCCGGGATATGGTGGGCATTGCTCAGACTGGCTCTGGGAAGACGTTGGCGTATCTC
CTGCCTGCAATTGTTCATATTAACCACCAGCCATACTTGGAAAGGGGAGATGGCCCAATCTGTCTAGTTCT
GGCTCCTACCAGAGAGCTTGCCCAGCAAGTACAGCAGGTGGCCGATGACTATGGCAAATGTTCTAGATTGA
AGAGTACTTGTATTTATGGAGGTGCTCCTAAAGGTCCCCAGATTCGAGACTTGGAAAGAGGTGTTGAGATC
TGCATAGCCACTCCTGGACGTCTGATAGATTTCCTGGAGTCAGGAAAGACAAATCTTCGCCGATGTACTTA
CCTTGTATTGGACGAAGCTGACAGAATGCTTGATATGGGGTTTGAACCCCAGATCCGTAAAATTGTTGACC
AAATCAGGCCTGATAGGCAGACACTGATGTGGAGTGCAACCTGGCCAAAAGAAGTAAGACAGCTTGCAGAG
GATTTCCTTCGTGATTACACCCAGATCAACGTAGGCAATCTGGAGTTGAGTGCCAACCACAACATCCTCCA
GATAGTGGATGTCTGCATGGAAAGTGAAAAAGACCACAAGTTGATCCAACTAATGGAAGAAATAATGGCTG
AAAAGGAAAACAAAACAATAATATTTGTGGAGACAAAGAGACGCTGTGATGATCTGACTCGAAGGATGCGC
AGAGATGGTTGGCCAGCTATGTGTATCCATGGAGACAAGAGTCAACCAGAAAGAGATTGGGTACTTAATGA
GTTCCGTTCTGGAAAGGCACCCATCCTTATTGCTACAGATGTAGCCTCCCGTGGGCTAGATGTGGAAGATG
TCAAGTTTGTGATCAACTATGACTATCCAAACAGCTCAGAGGATTATGTGCACCGTATTGGCCGAACAGCC
CGTAGCACCAACAAGGGTACCGCCTATACCTTCTTCACCCCAGGGAACCTAAAACAGGCCAGAGAGCTTAT
CAAAGTGCTGGAAGAGGCCAATCAGGCTATCAATCCAAAACTGATGCAGCTTGTGGACCACAGAGGAGGCG
GCGGAGGCGGGGGTAAGGGTGGTCGTTCTCGTTACCGGACCACTTCTTCAGCCAACAATCCCAATCTGATG
TATCAGGATGAGTGTGACCGAAGGCTTCGAGGAGTCAAGGATGGTGGCCGGAGAGACTCTGCAAGCTATCG
GGATCGTAGTGAAACCGATAGAGCTGGTTATGCTAATGGCAGTGGCTATGGAAGTCCAAATTCTGCCTTTG
GAGCACAAGCAGGCCAATACACCTATGGTCAAGGCACCTATGGGGCAGCTGCTTATGGCACCAGTAGCTAT
ACAGCTCAAGAATATGGTGCTGGCACTTATGGAGCTAGTAGCACCACCTCAACTGGGAGAAGTTCACAGAG
CTCTAGCCAGCAGTTTAGTGGGATAGGCCGGTCTGGGCAGCAGCCACAGCCACTGATGTCACAACAGTTTG
CACAGCCTCCGGGAGCTACCAATATGATAGGTTACATGGGGCAGACTGCCTACCAATACCCTCCTCCTCCT
CCCCCTCCTCCTCCTTCACGTAAATGAAACCACTCAAGTGGTAGTGACTCCAGCAGACTTAATTACATTTT

AAGGAACACTGTCTTTCCTTTTTTTTTCCTCTTCGCCTTTTCTTTTTTTTTCCTTTTTTCTTTTTTTTTTT
TTAATTTTTCCCCCCAACCATCGTGATTTGTCTTTTCATGCAGATTAGTTAGAATTCACTGCCAGGTTTCT
TCTGCCCACCAAAATGATCCAGTCTGGAATAACATTTTGTAAAAAAAAAAAAAATATATATATATATATAT
AGCTGACTGGAAGAGATTAATTTCTTCCCCCAACTTCTTGCATGTTGAAGATATTTGAGCTATTTTTCATC
TAAAAGAGTAAGGTATTAGGCCCTTTTGTGGGAGCCCCATGTTTTGTTTTTCTGAGTTGGTGGGGAGGGAG
GGAGGGGGAGGGCTGAATTGTTTTGCAGAGGAAGATGGCATCTGTGCTTTAAATTTCTCATTACTGGGTTA
GAAAACAAAGAGGGATTGCCCTGCACATTTTCTTTTGTGCTTTTAAATGTTTCTTAAGTTGGAACAGGTTT
CCTCGGGCCTGTTTTGACTGATTGCTGGAGTGCATTTGATAGTTAAAAATTACTAATTGGTTTTATTTCCC
TTCACACTCTGCCTCCCCACTTCTCCCCCCGTTACTGAAAAATAACCATTTTAGTGTCAGGCTAGAAATTG
AATTGCTGAGTTTTGTGTATCCTTTAAATTAAAAACCACAAGTGTTTATTGTAGTGGTTAAACTGTAGCAT
CTCAGCATCTGGGTGGAAGCTGCCTATATTTCTTCCCAGTTTAACTGGGGACCATCTGTGAAATTAATTTT
CCATCCAGACAGCTGCTGTGAGCAAATGAACATAAATGCTCGCTGGAAATTTACTAACCAGTTTTTATATT
GACCTGCAGTGTAAAAAGCACATTTAATTATAAACAATATATTCAAAATGGGCAAATTTTATTTTCAAATG
CAGTGTAGAGCTAGATTAAAAGCAACTCTTTGCCACCTACTCTGCCCTTTTGGCAAAGTTACCTTGAACAA
AGAATCTTAAGGGTTTATTAAGAACTCTTTATTTTCTTCATACCCTGTTCTCTGCAGTGCTTTCTAACAGC
TTCTGGGTGCAGATTTTCTTCGGCATCCTTTTGCACTCAGCTTATTACAGGTAGGTAGTGCTTAAGAAAAG
TCATGGAGGACTAAAGCCTAAGTCCTTTTCACTTTTCCTCCATCTGAAGGTAGGTGAGTTCATCCTCTTCA
TGGTAATGCTGTTTTACCAAGACTTTATAGCAGATGGACCCAGAAAGAATTTTCTGCTATTGTGTTCACTA
CAACAGGATAGGGACATCAGACAGCCCCAGAAACCCCTTCCAGATCTGATATGGGACTATTAATTTTTATG
CTGTTAATTGGTATTCATTCACAATGCAGTTGAAGGGGGAAGGCTCCACTGCATTCTTTGGCTAAGGCCTG
AATGCTTGCTCATCTGTAAGATCTATACTCGAGGTTTTGTTTTCCTTTTAAAATTCTTTAGGGAGAGAGGG
ATGGTTTCTGAGGGGTTCTGAAAGTATGATTCAATGTGCAACATACAGGTAGGTCTTCAGCATAAGCTGAA
ATATATGCATGTAAAAACTTTGACATCTTTTTTTTTAATTTTCCACTTTCTTCTTAACTTTACTTCTCTTT
TTGTCCCCCCCCCATCTTACAGAAGTTGAGGCCAAGGGAGAATGGTAGGCACAGAAGAAACATGGCAAACT
GCTCTGTGCTTTCAAACCAAAGTGTTCCCCCCAACCCCAAATTTGTCTAAGCACTGGCCAGTCTGTTGTGG
GCATTGTTTTCTACAACCAAATTCTGGGTTTTTTTCTTCTTTCTTTAAACATAGAGGTACCACCACAAGGG
ATGCCCTACTCTCTCGCAGCTCTTGAAAGCATCTGTTTGAGGGAAAGGTCTCTGGGCAAGCAAGTGGTTAT
TTGGATTGCTTGCTTCCCTTTTTCCACCTGGGACATTGTAATCATAAAATAACAGTAAATTCCAAACCTCA
AAAACTATTATGGCCTGAGCACAGCTGAAATCTAGCAGAGTTTAACTCTTCTGCCTCCATGTCTGTCACTT
ATAATTCAGGTTCTGCTGTTGGCTTCAGAACATGAGCAGAAGAATCGTTTTATGCTAGTTATTGCATTCAT
GGTTGAAACTCAACTTAGGGAAAGGGTTCCAATGTATTAAGCAATGGGCTGCTTCTCCCCAATCCTCCCTA
ACAATTCGTTGTGTGGACTTCTCATCTAAAAGGTTAGTGGCTTTTGCTTGGGATCAGTGCTCTCTATTGAT
GTTCTTGCTGGTCTCCAGACACATTCCTGTTGCATTAAGACTTGAAAGACTTGTAGATGTGTGATGTTCAG
GCACAGGATGCTGAAAGCTATGTTACTATTCTTAGTTTGTAAATTGTCCTTTTGATACCATCATCTTGTTT
TCTTTTTGTAGGTATAAATAAAAACACTGTTGACAATAAAAAAAAAAAAAAAAA

TVASATGDSASERESAAPAAAPTAEAPPPSVVTRPEPQALPSPAIRAPLPDLYPFGTMRGGGFGDRDRDRD
RGGFGARGGGGLPPKKFGNPGERLRKKKWDLSELPKFEKNFYVEHPEVARLTPYEVDELRRKKEITVRGGD
VCPKPVFAFHHANFPQYVMDVLMDQHFTEPTPIQCQGFPLALSGRDMVGIAQTGSGKTLAYLLPAIVHINH
QPYLERGDGPICLVLAPTRELAQQVQQVADDYGKCSRLKSTCIYGGAPKGPQIRDLERGVEICIATPGRLI
DFLESGKTNLRRCTYLVLDEADRMLDMGFEPQIRKIVDQIRPDRQTLMWSATWPKEVRQLAEDFLRDYTQI
NVGNLELSANHNILQIVDVCMESEKDHKLIQLMEEIMAEKENKTIIFVETKRRCDDLTRRMRRDGWPAMCI
HGDKSQPERDWVLNEFRSGKAPILIATDVASRGLDVEDVKFVINYDYPNSSEDYVHRIGRTARSTNKGTAY
TFFTPGNLKQARELIKVLEEANQAINPKLMQLVDHRGGGGGGKGGRSRYRTTSSANNPNLMYQDECDRRL
RGVKDGGRRDSASYRDRSETDRAGYANGSGYGSPNSAFGAQAGQYTYGQGTYGAAAYGTSSYTAQEYGAGT
YGASSTTSTGRSSQSSSQQFSGIGRSGQQPQPLMSQQFAQPPGATNMIGYMGQTAYQYPPPPPPPPSRK

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

#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

GGCTTGGGGCAGCCGGGTAGCTCGGAGGTCGTGGCGCTGGGGGCTAGCACCAGCGCTCTGTCGGGAGGCGC
AGCGGTTAGGTGGACCGGTCAGCGGACTCACCGGCCAGGGCGCTCGGTGCTGGAATTTGATATTCATTGAT
CCGGGTTTTATCCCTCTTCTTTTTTCTTAAACATTTTTTTTTAAAACTGTATTGTTTCTCGTTTTAATTTA
TTTTTGCTTGCCATTCCCCACTTGAATCGGGCCGACGGCTTGGGGAGATTGCTCTACTTCCCCAAATCACT
GTGGATTTTGGAAACCAGCAGAAAGAGGAAAGAGGTAGCAAGAGCTCCAGAGAGAAGTCGAGGAAGAGAGA
GACGGGGTCAGAGAGAGCGCGCGGGCGTGCGAGCAGCGAAAGCGACAGGGGCAAAGTGAGTGACCTGCTTT
TGGGGGTGACCGCCGGAGCGCGGCGTGAGCCCTCCCCCTTGGGATCCCGCAGCTGACCAGTCGCGCTGACG
GACAGACAGACAGACACCGCCCCCAGCCCCAGCTACCACCTCCTCCCCGGCCGGCGGCGGACAGTGGACGC
GGCGGCGAGCCGCGGGCAGGGGCCGGAGCCCGCGCCCGGAGGCGGGGTGGAGGGGGTCGGGGCTCGCGGCG
TCGCACTGAAACTTTTCGTCCAACTTCTGGGCTGTTCTCGCTTCGGAGGAGCCGTGGTCCGCGCGGGGGAA
GCCGAGCCGAGCGGAGCCGCGAGAAGTGCTAGCTCGGGCCGGGAGGAGCCGCAGCCGGAGGAGGGGGAGGA
GGAAGAAGAGAAGGAAGAGGAGAGGGGGCCGCAGTGGCGACTCGGCGCTCGGAAGCCGGGCTCATGGACGG
GTGAGGCGGCGGTGTGCGCAGACAGTGCTCCAGCCGCGCGCGCTCCCCAGGCCCTGGCCCGGGCCTCGGGC
CGGGGAGGAAGAGTAGCTCGCCGAGGCGCCGAGGAGAGCGGGCCGCCCCACAGCCCGAGCCGGAGAGGGAG
CGCGAGCCGCGCCGGCCCCGGTCGGGCCTCCGAAACCATGAACTTTCTGCTGTCTTGGGTGCATTGGAGCC
TTGCCTTGCTGCTCTACCTCCACCATGCCAAGTGGTCCCAGGCTGCACCCATGGCAGAAGGAGGAGGGCAG
AATCATCACGAAGTGGTGAAGTTCATGGATGTCTATCAGCGCAGCTACTGCCATCCAATCGAGACCCTGGT
GGACATCTTCCAGGAGTACCCTGATGAGATCGAGTACATCTTCAAGCCATCCTGTGTGCCCCTGATGCGAT
GCGGGGGCTGCTGCAATGACGAGGGCCTGGAGTGTGTGCCCACTGAGGAGTCCAACATCACCATGCAGATT
ATGCGGATCAAACCTCACCAAGGCCAGCACATAGGAGAGATGAGCTTCCTACAGCACAACAAATGTGAATG
CAGACCAAAGAAAGATAGAGCAAGACAAGAAAAAAAATCAGTTCGAGGAAAGGGAAAGGGGCAAAAACGAA
AGCGCAAGAAATCCCGGTATAAGTCCTGGAGCGTGTACGTTGGTGCCCGCTGCTGTCTAATGCCCTGGAGC
CTCCCTGGCCCCCATCCCTGTGGGCCTTGCTCAGAGCGGAGAAAGCATTTGTTTGTACAAGATCCGCAGAC
GTGTAAATGTTCCTGCAAAAACACAGACTCGCGTTGCAAGGCGAGGCAGCTTGAGTTAAACGAACGTACTT
GCAGATGTGACAAGCCGAGGCGGTGAGCCGGGCAGGAGGAAGGAGCCTCCCTCAGGGTTTCGGGAACCAGA
TCTCTCACCAGGAAAGACTGATACAGAACGATCGATACAGAAACCACGCTGCCGCCACCACACCATCACCA
TCGACAGAACAGTCCTTAATCCAGAAACCTGAAATGAAGGAAGAGGAGACTCTGCGCAGAGCACTTTGGGT
CCGGAGGGCGAGACTCCGGCGGAAGCATTCCCGGGCGGGTGACCCAGCACGGTCCCTCTTGGAATTGGATT
CGCCATTTTATTTTTCTTGCTGCTAAATCACCGAGCCCGGAAGATTAGAGAGTTTTATTTCTGGGATTCCT
GTAGACACACCCACCCACATACATACATTTATATATATATATATTATATATATATAAAAATAAATATCTCT
ATTTTATATATATAAAATATATATATTCTTTTTTTAAATTAACAGTGCTAATGTTATTGGTGTCTTCACTG
GATGTATTTGACTGCTGTGGACTTGAGTTGGGAGGGGAATGTTCCCACTCAGATCCTGACAGGGAAGAGGA
GGAGATGAGAGACTCTGGCATGATCTTTTTTTTGTCCCACTTGGTGGGGCCAGGGTCCTCTCCCCTGCCCA
GGAATGTGCAAGGCCAGGGCATGGGGGCAAATATGACCCAGTTTTGGGAACACCGACAAACCCAGCCCTGG
CGCTGAGCCTCTCTACCCCAGGTCAGACGGACAGAAAGACAGATCACAGGTACAGGGATGAGGACACCGGC
TCTGACCAGGAGTTTGGGGAGCTTCAGGACATTGCTGTGCTTTGGGGATTCCCTCCACATGCTGCACGCGC
ATCTCGCCCCCAGGGGCACTGCCTGGAAGATTCAGGAGCCTGGGCGGCCTTCGCTTACTCTCACCTGCTTC
TGAGTTGCCCAGGAGACCACTGGCAGATGTCCCGGCGAAGAGAAGAGACACATTGTTGGAAGAAGCAGCCC
ATGACAGCTCCCCTTCCTGGGACTCGCCCTCATCCTCTTCCTGCTCCCCTTCCTGGGGTGCAGCCTAAAAG
GACCTATGTCCTCACACCATTGAAACCACTAGTTCTGTCCCCCCAGGAGACCTGGTTGTGTGTGTGTGAGT
GGTTGACCTTCCTCCATCCCCTGGTCCTTCCCTTCCCTTCCCGAGGCACAGAGAGACAGGGCAGGATCCAC
GTGCCCATTGTGGAGGCAGAGAAAAGAGAAAGTGTTTTATATACGGTACTTATTTAATATCCCTTTTTAAT
TAGAAATTAAAACAGTTAATTTAATTAAAGAGTAGGGTTTTTTTTCAGTATTCTTGGTTAATATTTAATTT
CAACTATTTATGAGATGTATCTTTTGCTCTCTCTTGCTCTCTTATTTGTACCGGTTTTTGTATATAAAATT
CATGTTTCCAATCTCTCTCTCCCTGATCGGTGACAGTCACTAGCTTATCTTGAACAGATATTTAATTTTGC
TAACACTCAGCTCTGCCCTCCCCGATCCCCTGGCTCCCCAGCACACATTCCTTTGAAATAAGGTTTCAATA
TACATCTACATACTATATATATATTTGGCAACTTGTATTTGTGTGTATATATATATATATATGTTTATGTA
TATATGTGATTCTGATAAAATAGACATTGCTATTCTGTTTTTTATATGTAAAAACAAAACAAGAAAAAATA
GAGAATTCTACATACTAAATCTCTCTCCTTTTTTAATTTTAATATTTGTTATCATTTATTTATTGGTGCTA
CTGTTTATCCGTAATAATTGTGGGGAAAAGATATTAACATCACGTCTTTGTCTCTAGTGCAGTTTTTCGAG
ATATTCCGTAGTACATATTTATTTTTAAACAACGACAAAGAAATACAGATATATCTTAAAAAAAAAAAAGC
ATTTTGTATTAAAGAATTTAATTCTGATCTCAAAAAAAAAAAAA

LTDRQTDTAPSPSYHLLPGRRRTVDAAASRGQGPEPAPGGGVEGVGARGVALKLFVQLLGCSRFGGAVVRA
GEAEPSGAARSASSGREEPQPEEGEEEEEKEEERGPQWRLGARKPGSWTGEAAVCADSAPAARAPQALARA
SGRGGRVARRGAEESGPPHSPSRRGSASRAGPGRASETMNFLLSWVHWSLALLLYLHHAKWSQAAPMAEGG
GQNHHEVVKFMDVYQRSYCHPIETLVDIFQEYPDEIEYIFKPSCVPLMRCGGCCNDEGLECVPTEESNITM
QIMRIKPHQGQHIGEMSFLQHNKCECRPKKDRARQEKKSVRGKGKGQKRKRKKSRYKSWSVYVGARCCLMP
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

TTCTTAACCCTTTCCAGCTTTCCCACCCTCTTTGGCTTTAGCCATGGCCTTCTGATCTGTGTTTCTCAGGG
GACCTGCAGGCCCCAGATATAGCCCCATGCTGTCCTCCTACCCCAGAGCACACTGTTCAGGCTACTTCCAC
TGGTACTGAAATCCAGTATTTCACTTACTCTTTTTCTTTCCAATATCCTCATGACATTCAATATTTCACTT
ACTCTAGGTCCTCCCTGCCTAAGGCCCAAGTCAACTTTCTGTCCAGTGGGATTTGTAATCCAATACCTCCT
AGCCCTAGCAGAATCCCATGTGGATAATCAGAAATGTGACTGGAAAAAGGACAGAGCTCTATGGCTGTGGG
TCCCAGTCCCCACTGCTGGCAGTAAGTCCCCAGCAGTGAGCTGTGTAAGCACCTTACATTCTGCGCTTGGT
TGAAAACAGCAAGGCAAGCATCCACTTGAGAAATGTCAACCCCTAGGAAATCCCAGCCTCAAGTCTTTCTC
ATCCCTTGGGAAGTGCAAATTGGATAGAGAAGAAACCAATTAAAAACAAAACAAACAAATCATACTTAGAT
ATTCTGGCTTTTCTCACCAGGGCTGGATTAAAGCATGTACTTCAAAATAATAACAACTTAAGTCAATAAAT
AAATGTAAGGAAGTCCAAATGTTCACCTGAAGACAACTGTGGTCATTTTTTGGCAATCCCAGGTTCTCTTT
TCTACCTGTTTGCTCAATCGTGGTCTCCCTCTCCCTCTCTTGTTGGGGCCCATGCCCCTGCTTTACTGTTG
CCAGAGGCTTGTACTTGTTTGCCTTTTAGGTAGGAGCAGTTACTTCCACTCCCCTCACCTGCCATAAAGCA
TCTTTATAAACAAAGCAAGTAGAAGAAACACATCCTGGTATCCACCACATTCGGCTTTTGTTGATTCTGTT
CACTTGGGAGCACCTGCTGCTAGGGAATAAGAAGGTTGAGGCTGAAGAGTGAGGACTCTTCAGCTCCCCTC
TGGCAGGACCCGGGAGAGGAAAGAGCCCTCAGCTGGTCCATCCTCCCCACTCCTGGTCAGCCTTCTGTTCT
GAGATCAAAGTGGTGGGGTCACATTCTCGAGAACTGTGCTCAGCCCCCTCATCTCACACCCTTTCCCTCTC
CCTGTGTGCCTGCCCCCCTCTTACATAACCATGCTGGTGATTGGCACCGTCATAAATCAATACTTTGCTCA
CTTTCACATCAAGTAACACTATCCAGGGAGGTGGTTTCAACAAAGGAGGAAGTATAAGGAGATCTAGGTTC
AAATTAATGTTGCCCCTAGTGGTAAAGGACAGAGACCCTCAGACTGATGAAATGCACTCAGAATTACTTAG
ACAAAGCGGATATTTGCCACTCTCTTCCCCTTTTCCTGTGTTTTTGTAGTGAAGAGACCTGAAAGAAAAAA
GTAGGGAGAACATAATGAGAACAAATACGGTAATCTCTTCATTTGCTAGTTCAAGTGCTGGACTTGGGACT
TAGGAGGGGCAATGGAGCCGCTTAGTGCCTACATCTGACTTGGACTGAAATATAGGTGAGAGACAAGATTG
TCTCATATCCGGGGAAATCATAACCTATGACTAGGACGGGAAGAGGAAGCACTGCCTTTACTTCAGTGGGA
ATCTCGGCCTCAGCCTGCAAGCCAAGTGTTCACAGTGAGAAAAGCAAGAGAATAAGCTAATACTCCTGTCC
TGAACAAGGCAGCGGCTCCTTGGTAAAGCTACTCCTTGATCGATCCTTTGCACCGGATTGTTCAAAGTGGA
CCCCAGGGGAGAAGTCGGAGCAAAGAACTTACCACCAAGCAGTCCAAGAGGCCCAGAAGCAAACCTGGAGG
TGAGACCCAAAGAAAGCTGGAACCATGCTGACTTTGTACACTGTGAGGACACAGAGTCTGTTCCTGGAAAG
CCCAGTGTCAACGCAGATGAGGAAGTCGGAGGTCCCCAAATCTGCCGTGTATGTGGGGACAAGGCCACTGG
CTATCACTTCAATGTCATGACATGTGAAGGATGCAAGGGCTTTTTCAGGAGGGCCATGAAACGCAACGCCC
GGCTGAGGTGCCCCTTCCGGAAGGGCGCCTGCGAGATCACCCGGAAGACCCGGCGACAGTGCCAGGCCTGC
CGCCTGCGCAAGTGCCTGGAGAGCGGCATGAAGAAGGAGATGATCATGTCCGACGAGGCCGTGGAGGAGAG
GCGGGCCTTGATCAAGCGGAAGAAAAGTGAACGGACAGGGACTCAGCCACTGGGAGTGCAGGGGCTGACAG
AGGAGCAGCGGATGATGATCAGGGAGCTGATGGACGCTCAGATGAAAACCTTTGACACTACCTTCTCCCAT
TTCAAGAATTTCCGGCTGCCAGGGGTGCTTAGCAGTGGCTGCGAGTTGCCAGAGTCTCTGCAGGCCCCATC
GAGGGAAGAAGCTGCCAAGTGGAGCCAGGTCCGGAAAGATCTGTGCTCTTTGAAGGTCTCTCTGCAGCTGC
GGGGGGAGGATGGCAGTGTCTGGAACTACAAACCCCCAGCCGACAGTGGCGGGAAAGAGATCTTCTCCCTG
CTGCCCCACATGGCTGACATGTCAACCTACATGTTCAAAGGCATCATCAGCTTTGCCAAAGTCATCTCCTA

CTTCAGGGACTTGCCCATCGAGGACCAGATCTCCCTGCTGAAGGGGGCCGCTTTCGAGCTGTGTCAACTGA
GATTCAACACAGTGTTCAACGCGGAGACTGGAACCTGGGAGTGTGGCCGGCTGTCCTACTGCTTGGAAGAC
ACTGCAGGTGGCTTCCAGCAACTTCTACTGGAGCCCATGCTGAAATTCCACTACATGCTGAAGAAGCTGCA
GCTGCATGAGGAGGAGTATGTGCTGATGCAGGCCATCTCCCTCTTCTCCCCAGACCGCCCAGGTGTGCTGC
AGCACCGCGTGGTGGACCAGCTGCAGGAGCAATTCGCCATTACTCTGAAGTCCTACATTGAATGCAATCGG
CCCCAGCCTGCTCATAGGTTCTTGTTCCTGAAGATCATGGCTATGCTCACCGAGCTCCGCAGCATCAATGC
TCAGCACACCCAGCGGCTGCTGCGCATCCAGGACATACACCCCTTTGCTACGCCCCTCATGCAGGAGTTGT
TCGGCATCACAGGTAGCTGAGCGGCTGCCCTTGGGTGACACCTCCGAGAGGCAGCCAGACCCAGAGCCCTC
TGAGCCGCCACTCCCGGGCCAAGACAGATGGACACTGCCAAGAGCCGACAATGCCCTGCTGGCCTGTCTCC
CTAGGGAATTCCTGCTATGACAGCTGGCTAGCATTCCTCAGGAAGGACATGGGTGCCCCCCACCCCCAGTT
CAGTCTGTAGGGAGTGAAGCCACAGACTCTTACGTGGAGAGTGCACTGACCTGTAGGTCAGGACCATCAGA
GAGGCAAGGTTGCCCTTTCCTTTTAAAAGGCCCTGTGGTCTGGGGAGAAATCCCTCAGATCCCACTAAAGT
GTCAAGGTGTGGAAGGGACCAAGCGACCAAGGATGGGCCATCTGGGGTCTATGCCCACATACCCACGTTTG
TTCGCTTCCTGAGTCTTTTCATTGCTACCTCTAATAGTCCTGTCTCCCACTTCCCACTCGTTCCCCTCCTC
TTCCGAGCTGCTTTGTGGGCTCCAGGCCTGTACTCATCGGCAGGCGCATGAGTATCTGTGGGAGTCCTCTA
GAGAGATGAGAAGCCAGGAGGCCTGCACCAAATGTCAGAAGCTTGGCATGACCTCATTCCGGCCACATCAT
TCTGTGTCTCTGCATCCATTTGAACACATTATTAAGCACCGATAATAGGTAGCCTGCTGTGGGGTATACAG
CATTGACTCAGATATAGATCCTGAGCTCACAGAGTTTATAGTTAAAAAAACAAACAGAAACACAAACAATT
TGGATCAAAAGGAGAAATGATAAGTGACAAAAGCAGCACAAGGAATTTCCCTGTGTGGATGCTGAGCTGTG
ATGGCGGGCACTGGGTACCCAAGTGAAGGTTCCCGAGGACATGAGTCTGTAGGAGCAAGGGCACAAACTGC
AGCTGTGAGTGCGTGTGTGTGATTTGGTGTAGGTAGGTCTGTTTGCCACTTGATGGGGCCTGGGTTTGTTC
CTGGGGCTGGAATGCTGGGTATGCTCTGTGACAAGGCTACGCTGACAATCAGTTAAACACACCGGAGAAGA
ACCATTTACATGCACCTTATATTTCTGTGTACACATCTATTCTCAAAGCTAAAGGGTATGAAAGTGCCTGC
CTTGTTTATAGCCACTTGTGAGTAAAAATTTTTTTGCATTTTCACAAATTATACTTTATATAAGGCATTCC
ACACCTAAGAACTAGTTTTGGGAAATGTAGCCCTGGGTTTAATGTCAAATCAAGGCAAAAGGAATTAAATA
ATGTACTTTTGGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

LEVRPKESWNHADFVHCEDTESVPGKPSVNADEEVGGPQICRVCGDKATGYHFNVMTCEGCKGFFRRAMKR
NARLRCPFRKGACEITRKTRRQCQACRLRKCLESGMKKEMIMSDEAVEERRALIKRKKSERTGTQPLGVQG
LTEEQRMMIRELMDAQMKTFDTTFSHFKNFRLPGVLSSGCELPESLQAPSREEAAKWSQVRKDLCSLKVSL
QLRGEDGSVWNYKPPADSGGKEIFSLLPHMADMSTYMFKGIISFAKVISYFRDLPIEDQISLLKGAAFELC
QLRFNTVFNAETGTWECGRLSYCLEDTAGGFQQLLLEPMLKFHYMLKKLQLHEEEYVLMQAISLFSPDRPG
VLQHRVVDQLQEQFAITLKSYIECNRPQPAHRFLFLKIMAMLTELRSINAQHTQRLLRIQDIHPFATPLMQ
ELFGITGS

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

#11

NM_001172131

Region conserved in mammals.

HCK hemopoietic cell kinase

Synonyms: JTK9; HCK

GTCCCAGCTCGGGAGCACATCAGAGGCTTAGAGGCGAGTGGGAAGGGACTCAGACAGTGCAGGACGAGAAA
CGCCCGCGGCACCAAAGCCCCTCAGAGCGTCGCCCCCGCCTCTAGTTCTAGAAAGTCAGTTTCCCGGCACT
GGCACCCCGGAACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTCAAGCTGCGAGGATCCGGGCTGCCCGCGA
GACGAGGAGCGGGCGCCCAGGATGGGGTGCATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTC
AAAAACTGAAACCAGCGCCAGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGG
GGCCTAATAGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC
CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATGGTGGTCCTAGA
GGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATG
TCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGC
CAACTGCTGGCTCCCGGCAACATGCTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTA

CTCTTTGTCCGTGCGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGG
ACAACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCACTACAAG
AAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAGCCCCAGAAGCCTTGGGA
GAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAGAAGAAACTTGGAGCTGGGCAGTTTGGGG
AAGTCTGGATGGCCACCTACAACAAGCACACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCG
GTGGAGGCCTTCCTGGCAGAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGC
GGTGGTCACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTGA
AAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATTGCAGAAGGCATG
GCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCCAACATCTTGGTCTCTGCATCCCT
GGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGG
CCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTC
TGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCC
TGAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAGCTCTACA
ACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTG
GATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCAGGGCCA
GGGGGTGCCCAGGTGGTGGCTGCAAGGTGGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTC
CCAGACACCCACCCTCGCTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTC
TTTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGAA
TGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGAATATTTAAATAA
AAGATATAAATGCCAAAGTCTTTACCAAAAAAAAAAAAAAAAAAAA

LGGRSSCEDPGCPRDEERAPRMGCMKSKFLQVGGNTFSKTETSASPHCPVYVPDPTSTIKPGPNSHNSNTP
GIREAGSEDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWWKARSLATRKEGYIPSNYVARVDSLETE
EWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYSLSVRDYDPRQGDTVKHYKIRTLDNGGFYISPR
STFSTLQELVDHYKKGNDGLCQKLSVPCMSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKH
TKVAVKTMKPGSMSVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPL
PKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGLARVIEDNEYTAREGAKFPIKWTAP
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

CCAGGCAGCTGGGGTAAGGAGTTCAAGGCAGCGCCCACACCCGGGGGCTCTCCGCAACCCGACCGCCTGTC
CGCTCCCCCACTTCCCGCCCTCCCTCCCACCTACTCATTCACCCACCCACCCACCCAGAGCCGGGACGGCA
GCCCAGGCGCCCGGGCCCCGCCGTCTCCTCGCCGCGATCCTGGACTTCCTCTTGCTGCAGGACCCGGCTTC
CACGTGTGTCCCGGAGCCGGCGTCTCAGCACACGCTCCGCTCCGGGCCTGGGTGCCTACAGCAGCCAGAGC
AGCAGGGAGTCCGGGACCCGGGCGGCATCTGGGCCAAGTTAGGCGCCGCGAGGCCAGCGCTGAACGTCTC
CAGGGCCGGAGGAGCCGCGGGGCGTCCGGGTCTGAGCCGCAGCAAATGGGCTCCGACGTGCGGGACCTGAA
CGCGCTGCTGCCCGCCGTCCCCTCCCTGGGTGGCGGCGGCGGCTGTGCCCTGCCTGTGAGCGGCGCGGCGC
AGTGGGCGCCGGTGCTGGACTTTGCGCCCCCGGGCGCTTCGGCTTACGGGTCGTTGGGCGGCCCCGCGCCG
CCACCGGCTCCGCCGCCACCCCCGCCGCCGCCGCCTCACTCCTTCATCAAACAGGAGCCGAGCTGGGGCGG
CGCGGAGCCGCACGAGGAGCAGTGCCTGAGCGCCTTCACTGTCCACTTTTCCGGCCAGTTCACTGGCACAG
CCGGAGCCTGTCGCTACGGGCCCTTCGGTCCTCCTCCGCCCAGCCAGGCGTCATCCGGCCAGGCCAGGATG
TTTCCTAACGCGCCCTACCTGCCCAGCTGCCTCGAGAGCCAGCCCGCTATTCGCAATCAGGGTTACAGCAC

GGTCACCTTCGACGGGACGCCCAGCTACGGTCACACGCCCTCGCACCATGCGGCGCAGTTCCCCAACCACT
CATTCAAGCATGAGGATCCCATGGGCCAGCAGGGCTCGCTGGGTGAGCAGCAGTACTCGGTGCCGCCCCCG
GTCTATGGCTGCCACACCCCCACCGACAGCTGCACCGGCAGCCAGGCTTTGCTGCTGAGGACGCCCTACAG
CAGTGACAATTTATACCAAATGACATCCCAGCTTGAATGCATGACCTGGAATCAGATGAACTTAGGAGCCA
CCTTAAAGGGCCACAGCACAGGGTACGAGAGCGATAACCACACAACGCCCATCCTCTGCGGAGCCCAATAC
AGAATACACACGCACGGTGTCTTCAGAGGCATTCAGGATGTGCGACGTGTGCCTGGAGTAGCCCCGACTCT
TGTACGGTCGGCATCTGAGACCAGTGAGAAACGCCCCTTCATGTGTGCTTACCCAGGCTGCAATAAGAGAT
ATTTTAAGCTGTCCCACTTACAGATGCACAGCAGGAAGCACACTGGTGAGAAACCATACCAGTGTGACTTC
AAGGACTGTGAACGAAGGTTTTCTCGTTCAGACCAGCTCAAAAGACACCAAAGGAGACATACAGGTGTGAA
ACCATTCCAGTGTAAAACTTGTCAGCGAAAGTTCTCCCGGTCCGACCACCTGAAGACCCACACCAGGACTC
ATACAGGTAAAACAAGTGAAAAGCCCTTCAGCTGTCGGTGGCCAAGTTGTCAGAAAAAGTTTGCCCGGTCA
GATGAATTAGTCCGCCATCACAACATGCATCAGAGAAACATGACCAAACTCCAGCTGGCGCTTTGAGGGGT
CTCCCTCGGGGACCGTTCAGTGTCCCAGGCAGCACAGTGTGTGAACTGCTTTCAAGTCTGACTCTCCACTC
CTCCTCACTAAAAAGGAAACTTCAGTTGATCTTCTTCATCCAACTTCCAAGACAAGATACCGGTGCTTCTG
GAAACTACCAGGTGTGCCTGGAAGAGTTGGTCTCTGCCCTGCCTACTTTTAGTTGACTCACAGGCCCTGGA
GAAGCAGCTAACAATGTCTGGTTAGTTAAAAGCCCATTGCCATTTGGTGTGGATTTTCTACTGTAAGAAGA
GCCATAGCTGATCATGTCCCCCTGACCCTTCCCTTCTTTTTTTATGCTCGTTTTCGCTGGGGATGGAATTA
TTGTACCATTTTCTATCATGGAATATTTATAGGCCAGGGCATGTGTATGTGTCTGCTAATGTAAACTTTGT
CATGGTTTCCATTTACTAACAGCAACAGCAAGAAATAAATCAGAGAGCAAGGCATCGGGGGTGAATCTTGT
CTAACATTCCCGAGGTCAGCCAGGCTGCTAACCTGGAAAGCAGGATGTAGTTCTGCCAGGCAACTTTTAAA
GCTCATGCATTTCAAGCAGCTGAAGAAAAAATCAGAACTAACCAGTACCTCTGTATAGAAATCTAAAAGAA
TTTTACCATTCAGTTAATTCAATGTGAACACTGGCACACTGCTCTTAAGAAACTATGAAGATCTGAGATTT
TTTTGTGTATGTTTTTGACTCTTTTGAGTGGTAATCATATGTGTCTTTATAGATGTACATACCTCCTTGCA
CAAATGGAGGGGAATTCATTTTCATCACTGGGAGTGTCCTTAGTGTATAAAAACCATGCTGGTATATGGCT
TCAAGTTGTAAAAATGAAAGTGACTTTAAAAGAAAATAGGGGATGGTCCAGGATCTCCACTGATAAGACTG
TTTTTAAGTAACTTAAGGACCTTTGGGTCTACAAGTATATGTGAAAAAAATGAGACTTACTGGGTGAGGAA
ATCCATTGTTTAAAGATGGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGTGTTGTGTTTTGTT
TTTTAAGGGAGGGAATTTATTATTTACCGTTGCTTGAAATTACTGTGTAAATATATGTCTGATAATGATTT
GCTCTTTGACAACTAAAATTAGGACTGTATAAGTACTAGATGCATCACTGGGTGTTGATCTTACAAGATAT
TGATGATAACACTTAAAATTGTAACCTGCATTTTTCACTTTGCTCTCAATTAAAGTCTATTCAAAAG

LDFLLLQDPASTCVPEPASQHTLRSGPGCLQQPEQQGVRDPGGIWAKLGAAEASAERLQGRRSRGASGSEP
QQMGSDVRDLNALLPAVPSLGGGGGCALPVSGAAQWAPVLDFAPPGASAYGSLGGPAPPPAPPPPPPPPPH
SFIKQEPSWGGAEPHEEQCLSAFTVHFSGQFTGTAGACRYGPFGPPPPSQASSGQARMFPNAPYLPSCLES
QPAIRNQGYSTVTFDGTPSYGHTPSHHAAQFPNHSFKHEDPMGQQGSLGEQQYSVPPPVYGCHTPTDSCTG
SQALLLRTPYSSDNLYQMTSQLECMTWNQMNLGATLKGHSTGYESDNHTTPILCGAQYRIHTHGVFRGIQD
VRRVPGVAPTLVRSASETSEKRPFMCAYPGCNKRYFKLSHLQMHSRKHTGEKPYQCDFKDCERRFSRSDQL
KRHQRRHTGVKPFQCKTCQRKFSRSDHLKTHTRTHTGKTSEKPFSCRWPSCQKKFARSDELVRHHNMHQRN
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
TTCCCGGCTGCGCGCCCTTCGGCCAGGCCGGGAGCCGCGCCAGTCGGAGCCCCCGGCCCAGCGTGGTCCGC

CTCCCTCTGGGCGTCCACCTGCCCGGAGTACTGCCAGCGGGCATGACCGACCCACCAGGGGCGCCGCCGCC
GGCGCTCGCAGGCCGCGGATGAAGAAGAAAACCCGGCGCCGCTCGACCCGGAGCGAGGAGTTGACCCGGAG
CGAGGAGTTGACCCTGAGTGAGGAAGCGACCTGGAGTGAAGAGGCGACCCAGAGTGAGGAGGCGACCCAGG
GCGAAGAGATGAATCGGAGCCAGGAGGTGACCCGGGACGAGGAGTCGACCCGGAGCGAGGAGGTGACCAGG
GAGGAAATGGCGGCAGCTGGGCTCACCGTGACTGTCACCCACAGCAATGAGAAGCACGACCTTCATGTTAC
CTCCCAGCAGGGCAGCAGTGAACCAGTTGTCCAAGACCTGGCCCAGGTTGTTGAAGAGGTCATAGGGGTTC
CACAGTCTTTTCAGAAACTCATATTTAAGGGAAAATCTCTGAAGGAAATGGAAACACCGTTGTCAGCACTT
GGAATACAAGATGGTTGCCGGGTCATGTTAATTGGGAAAAAGAACAGTCCACAGGAAGAGGTTGAACTAAA
GAAGTTGAAACATTTGGAGAAGTCTGTGGAGAAGATAGCTGACCAGCTGGAAGAGTTGAATAAAGAGCTTA
CTGGAATCCAGCAGGGTTTTCTGCCCAAGGATTTGCAAGCTGAAGCTCTCTGCAAACTTGATAGGAGAGTA
AAAGCCACAATAGAGCAGTTTATGAAGATCTTGGAGGAGATTGACACACTGATCCTGCCAGAAAATTTCAA
AGACAGTAGATTGAAAAGGAAAGGCTTGGTAAAAAAGGTTCAGGCATTCCTAGCCGAGTGTGACACAGTGG
AGCAGAACATCTGCCAGGAGACTGAGCGGCTGCAGTCTACAAACTTTGCCCTGGCCGAGTGAGGTGTAGCA
GAAAAAGGCTGTGCTGCCCTGAAGAATGGCGCCACCAGCTCTGCCGTCTCTGGAGCGGAATTTACCTGATT
TCTTCAGGGCTGCTGGGGGCAACTGGCCATTTGCCAATTTTCCTACTCTCACACTGGTTCTCAATGAAAAA
TAGTGTCTTTGTGATTTTGAGTAAAGCTCCTATCTGTTTTCTCCTTCTGTCTCTGTGGTTGTACTGTCCAG
CAATCCACCTTTTCTGGAGAGGGCCACCTCTGCCCAAATTTTCCCAGCTGTTTGGACCTCTGGGTGCTTTC
TTTGGGCTGGTGAGAGCTCTAATTTGCCTTGGGCCAGTTTCAGGTTTATAGGCCCCCTCAGTCTTCAGATA
CATGAGGGCTTCTTTGCTCTTGTGATCGTGTAGTCCCATAGCTGTAAAACCAGAATCACCAGGAGGTTGCA
CCTAGTCAGGAATATTGGGAATGGCCTAGAACAAGGTGTTTGGCACATAAGTAGACCACTTATCCCTCATT
GTGACCTAATTCCAGAGCATCTGGCTGGGTTGTTGGGTTCTAGACTTTGTCCTCACCTCCCAGTGACCCTG
ACTAGCCACAGGCCATGAGATACCAGGGGGCCGTTCCTTGGATGGAGCCTGTGGTTGATGCAAGGCTTCCT
TGTCCCCAAGCAAGTCTTCAGAAGGTTAGAACCCAGTGTTGACTGAGTCTGTGCTTGAAACCAGGCCAGAG
CCATGGATTAGGAAGGGCAAAGAGAAGGCACCAGAATGAGTAAAGCAGGCAGGTGGTGAAGCCAACCATAA
ACTTCTCAGGAGTGACATGTGCTTCCTTCAAAGGCATTTTTGTTAACCATATCCTTCTGAGTTCTATGTTT
CCTTCACAGCTGTTCTATCCATTTTGTGGACTGTCCCCCACCCCCACCCCATCATTGTTTTTAAAAAATTA
AGGCCTGGCGCAGCAGCTCATGCCTATAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACTTGAGG
CCAGGAGTTTGAGACCAGCCCAGGCAACATAGCAAAACCCCATTCTGCTTTAAAAAAAAAAAAAAAAAAAA
TTAGCTTGGCGTAGTGGCATGTGCCTATAATCCCAGCTACTGGGGAGGCTGAGGCACAAGAATCATTTGAA
CCTGGGAGGTAGAGGTTGCTGTGAGCCGAGATTACGCCCCTGCACTCCAGCCTGGGTCACAGAGTGAGACT
CCATCTCAGAAAAAAAAAAAATTGAGTCAGGTGCAGTAGCTCCTTCCTGTAGTCCCAGCTACTTGGGAGGC
TGAGGCTAGAGGATCACTTGAGCCCAGGAGTTTGAGTCTAGTCTGGGCAACATAGCAAGACCCCATCTCTA
AAATTTAAGTAAGTAAAAGTAGATAAATAAAAAGAAAAAAAAACTGTTTATGTGCTCATCATAAAGTAGAA
GAGTGGTTTGCTTTTTTTTTTTTTTTTGGATTAATGAGGAAATCATTCTGTGGCTCTAGTCATAATTTATG
CTTAATAACATTGATAGTAGCCCTTTGCGCTATAACTCTACCTAAAGACTCACATCATTTGGCAGAGAGAG
AGTCGTTGAAGTCCCAGGAATTCAGGACTGGGCAGGTTAAGACCTCAGACAAGGTAGTAGAGGTAGACTTG
TGGACAAGGCTCGGGTCCCAGCCCACCGCACCCCAACTTTAATCAGAGTGGTTCACTATTGATCTATTTTT
GTGTGATAGCTGTGTGGCGTGGGCCACAACATTTAATGAGAAGTTACTGTGCACCAAACTGCCGAACACCA
TTCTAAACTATTCATATATATTAGTCATTTAATTCTTACATAACTTGAGAGGTAGACAGATATCCTTATTT
TAGAGATGAGGAAACCAAGAGAACTTAGGTCATTAGCGCAAGGTTGTAGAGTAAGCGGCAAAGCCAAGACA
CAAAGCTGGGTGGTTTGGTTTCAGAGCCAGTGCTTTTCCCCTCTACTGTACTGCCTCTCAACCAACACAGG
GTTGCACAGGCCCATTCTCTGATTTTTTTCCTCTTGTCCTCTGCCTCTCCCTCTAGCTCCCACTTCCTCTC
TGCTCTAGTTCATTTTCTTTAGAGCAGCCCGAGTGATCATGAAGTGCAAATCTTGCCATGTCAGTCCCCTG
CTTAGAACCCTCCAATGGCTCACTTTCTCTTTAGGCAAAAGTCTTTACCCCATGCCTTCTCCCATCTCATC
TCAACCCCCTCATTTGTTGGCTGTCTGCTGTCAGCCACTCTTCTTTCAGGTCCTCAGATGCACTGCACCCT
CTCCTGCCTGGGGGTCTTTGCTCCTGCTACTACCTCTGCTTGAACAGCTCCTCACCTTCCTTCCTCCAACC
CTACCCTTGTATAGGTGACTTTTGTTCATCCTTCAGAATTCAACTCACATGTCTCTTGCATGGAGAACCCT
CACCTACTGTGTTGAGACCCTGTCCAGCCCCCAGGTGGGATCCTCTCTCGACTTCCCATACATTTCTTTCA
CAGCATTTACATAGTCCATGATAGTTTACTTGTGGGATTATTTGGTTAATCTTTGCCTTTAACACCAGGGT
TCCTTGGGTGAAGGAGCTTCTTTATCTTGGTAACAGCATTATTTCAAGCATAACTTGTAATATAGTTATAT
TACATATATAACATATATATATATAACATAACATATATAACATATATAACAAGCATAACTTGTTATATAGT
CTTGTATATAGTAAGACCTCAATAAATATTTGGAGAACAAAAAAAAAAAAA

LAQRGGARRPRGDRERLGSRLRALRPGREPRQSEPPAQRGPPPSGRPPARSTASGHDRPTRGAAAGARRPR
MKKKTRRRSTRSEELTRSEELTLSEEATWSEEATQSEEATQGEEMNRSQEVTRDEESTRSEEVTREEMAAA
GLTVTVTHSNEKHDLHVTSQQGSSEPVVQDLAQVVEEVIGVPQSFQKLIFKGKSLKEMETPLSALGIQDGC

RVMLIGKKNSPQEEVELKKLKHLEKSVEKIADQLEELNKELTGIQQGFLPKDLQAEALCKLDRRVKATIEQ
FMKILEEIDTLILPENFKDSRLKRKGLVKKVQAFLAECDTVEQNICQETERLQSTNFALAE

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

#16

NM_001099456

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

NPW neuropeptide W

Synonyms: L8; L8C; PPL8; PPNPW; NPW

GTACCCACTCCCTGGCACTTCCACTCCTAGAGGGAGGAGGCTGAGCAGGCAGAGAATGGGACGTGTCCCCT
CAGAGGAGCCTCGAGCCCAGTTCCAGCCAGCGGCCCACTCAGTGAGGTGCTCAAGTACCCACGTCCCCCGC
CAGCTGCCAGGGTTCCCTCTCCTCCCTCCGTCCCTCCCCCCATCTGGGGAGCCCAGCGGTACTGAGGGGGC
GGAACGAGGCGGGGCCACCGAGCGGTTATAGCTGGGCCTGCAGGGGACCCACGGCTCGCCTCCAGCCTCCT
GCGCTCCGGTACCTGGGCGTCCCAACTCCACTGCGCGCCCAAACCCAGCCGAGCCGGTTCGTGGCCCGCCC
CGCCGGGCGGCCGTCGACGCGAGCGCCCTGGCGTGGCGCCCAGGGGAGCGGGGGGCTCCCGCGAGCCGGCC
GCGGCTGGCACTGCTGCTGCTTCTGCTCCTGCTGCCGCTGCCCTCCGGCGCGTGGTACAAGCACGTGGCGA
GTCCCCGCTACCACACGGTGGGCCGCGCCGCTGGCCTGCTCATGGGGCTGCGTCGCTCACCCTATCTGTGG
CGCCGCGCGCTGCGCGCGGCCGCCGGGCCCCTGGCCAGGGACACCCTCTCCCCCGAACCCGCAGCCCGCGA
GGCTCCTCTCCTGCTGCCCTCGTGGGTTCAGGAGCTGTGGGAGACGCGACGCAGGAGCTCCCAGGCAGGGA
TCCCCGTCCGTGCGCCCCGGAGCCCGCGCGCCCCAGAGCCTGCGCTGGAACCGGAGTCCCTGGACTTCAGC
GGAGCTGGCCAGAGACTTCGGAGAGACGTCTCCCGCCCAGCGGTGGACCCCGCAGCAAACCGCCTTGGCCT
GCCCTGCCTGGCCCCCGGACCGTTCTGACAGCGTCCCCCGCCCGCCCGTGGCGCCTCCGCGCCTGACCCAG
GAGGAGTGGCCGCGCGCTTCCAGGAGCCGCTCATAGACCCCGCCTGCCGTCCGGTCAATAAAATCCGCCTG
ACTCCTGCGCCCCCGCATGCGAAAAAAAAAAAAAAAAAA

LAWRPGERGAPASRPRLALLLLLLLLPLPSGAWYKHVASPRYHTVGRAAGLLMGLRRSPYLWRRALRAAAG
PLARDTLSPEPAAREAPLLLPSWVQELWETRRRSSQAGIPVRAPRSPRAPEPALEPESLDFSGAGQRLRRD
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

ACCCCCGAGCTGTGCTGCTCGCGGCCGCCACCGCCGGGCCCCGGCCGTCCCTGGCTCCCCTCCTGCCTCGA
GAAGGGCAGGGCTTCTCAGAGGCTTGGCGGGAAAAAGAACGGAGGGAGGGATCGCGCTGAGTATAAAAGCC
GGTTTTCGGGGCTTTATCTAACTCGCTGTAGTAATTCCAGCGAGAGGCAGAGGGAGCGAGCGGGCGGCCGG
CTAGGGTGGAAGAGCCGGGCGAGCAGAGCTGCGCTGCGGGCGTCCTGGGAAGGGAGATCCGGAGCGAATAG
GGGGCTTCGCCTCTGGCCCAGCCCTCCCGCTGATCCCCCAGCCAGCGGTCCGCAACCCTTGCCGCATCCAC
GAAACTTTGCCCATAGCAGCGGGCGGGCACTTTGCACTGGAACTTACAACACCCGAGCAAGGACGCGACTC
TCCCGACGCGGGGAGGCTATTCTGCCCATTTGGGGACACTTCCCCGCCGCTGCCAGGACCCGCTTCTCTGA
AAGGCTCTCCTTGCAGCTGCTTAGACGCTGGATTTTTTTCGGGTAGTGGAAAACCAGCAGCCTCCCGCGAC
GATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCT
ACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAG

GATATCTGGAAGAAATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTC
GCCCTCCTACGTTGCGGTCACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCA
CGGCCGACCAGCTGGAGATGGTGACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGAC
CCGGACGACGAGACCTTCATCAAAAACATCATCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGC
CAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCTGCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCC
GCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGATCTGAGCGCCGCCGCCTCAGAGTGCATC
GACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCAAGTCCTGCGCCTCGCAAGACTC
CAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCGCAGGGCAGCCCCGAGC
CCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGAAGATGAGGAA
GAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTTCTGC
TGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACA
ACTACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGA
GTCCTGAGACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAA
GAGGCGAACACACAACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTG
ACCAGATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATAC
ATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAACGACGAGAACA
GTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCGTAAGGAAAAGTAAGGAAAACGATTCCTTCTA
ACAGAAATGTCCTGAGCAATCACCTATGAACTTGTTTCAAATGCATGATCAAATGCAACCTCACAACCTTG
GCTGAGTCTTGAGACTGAAAGATTTAGCCATAATGTAAACTGCCTCAAATTGGACTTTGGGCATAAAAGAA
CTTTTTTATGCTTACCATCTTTTTTTTTTCTTTAACAGATTTGTATTTAAGAATTGTTTTTAAAAAATTTT
AAGATTTACACAATGTTTCTCTGTAAATATTGCCATTAAATGTAAATAACTTTAATAAAACGTTTATAGCA
GTTACACAGAATTTCAATCCTAGTATATAGTACCTAGTATTATAGGTACTATAAACCCTAATTTTTTTTAT
TTAAGTACATTTTGCTTTTTAAAGTTGATTTTTTTCTATTGTTTTTAGAAAAAATAAAATAACTGGCAAAT
ATATCATTGAGCCAAAAAAAAAAAAAAAAAAAA

LDFFRVVENQQPPATMPLNVSFTNRNYDLDYDSVQPYFYCDEEENFYQQQQQSELQPPAPSEDIWKKFELL
PTPPLSPSRRSGLCSPSYVAVTPFSLRGDNDGGGGSFSTADQLEMVTELLGGDMVNQSFICDPDDETFIKN
IIIQDCMWSGFSAAAKLVSEKLASYQAARKDSGSPNPARGHSVCSTSSLYLQDLSAAASECIDPSVVFPYP
LNDSSSPKSCASQDSSAFSPSSDSLLSSTESSPQGSPEPLVLHEETPPTTSSDSEEEQEDEEEIDVVSVEK
RQAPGKRSESGSPSAGGHSKPPHSPLVLKRCHVSTHQHNYAAPPSTRKDYPAAKRVKLDSVRVLRQISNNR
KCTSPRSSDTEENVKRRTHNVLERQRRNELKRSFFALRDQIPELENNEKAPKVVILKKATAYILSVQAEEQ
KLISEEDLLRKRREQLKHKLEQLRNSCA

C)

#1

NM_002006

Does not appear conserved beyond humans.

FGF2 fibroblast growth factor 2 (basic)

Synonyms: BFGF; FGFB; HBGF-2; FGF2

CGGCCCCAGAAAACCCGAGCGAGTAGGGGGCGGCGCGCAGGAGGGAGGAGAACTGGGGGCGCGGGAGGCTG
GTGGGTGTGGGGGGTGGAGATGTAGAAGATGTGACGCCGCGGCCCGGCGGGTGCCAGATTAGCGGACGCGG
TGCCCGCGGTTGCAACGGGATCCCGGGCGCTGCAGCTTGGGAGGCGGCTCTCCCCAGGCGGCGTCCGCGGA
GACACCCATCCGTGAACCCCAGGTCCCGGGCCGCCGGCTCGCCGCGCACCAGGGGCCGGCGGACAGAAGAG
CGGCCGAGCGGCTCGAGGCTGGGGGACCGCGGGCGCGGCCGCGCGCTGCCGGGCGGGAGGCTGGGGGGCCG
GGGCCGGGGCCGTGCCCCGGAGCGGGTCGGAGGCCGGGGCCGGGGCCGGGGGACGGCGGCTCCCCGCGCGG
CTCCAGCGGCTCGGGGATCCCGGCCGGGCCCCGCAGGGACCATGGCAGCCGGGAGCATCACCACGCTGCCC
GCCTTGCCCGAGGATGGCGGCAGCGGCGCCTTCCCGCCCGGCCACTTCAAGGACCCCAAGCGGCTGTACTG
CAAAAACGGGGGCTTCTTCCTGCGCATCCACCCCGACGGCCGAGTTGACGGGGTCCGGGAGAAGAGCGACC
CTCACATCAAGCTACAACTTCAAGCAGAAGAGAGAGGAGTTGTGTCTATCAAAGGAGTGTGTGCTAACCGT

TACCTGGCTATGAAGGAAGATGGAAGATTACTGGCTTCTAAATGTGTTACGGATGAGTGTTTCTTTTTTGA
ACGATTGGAATCTAATAACTACAATACTTACCGGTCAAGGAAATACACCAGTTGGTATGTGGCACTGAAAC
GAACTGGGCAGTATAAACTTGGATCCAAAACAGGACCTGGGCAGAAAGCTATACTTTTTCTTCCAATGTCT
GCTAAGAGCTGATTTTAATGGCCACATCTAATCTCATTTCACATGAAAGAAGAAGTATATTTTAGAAATTT
GTTAATGAGAGTAAAAGAAAATAAATGTGTATAGCTCAGTTTGGATAATTGGTCAAACAATTTTTTATCCA
GTAGTAAAATATGTAACCATTGTCCCAGTAAAGAAAAATAACAAAAGTTGTAAAATGTATATTCTCCCTTT
TATATTGCATCTGCTGTTACCCAGTGAAGCTTACCTAGAGCAATGATCTTTTTCACGCATTTGCTTTATTC
GAAAAGAGGCTTTTAAAATGTGCATGTTTAGAAACAAAATTTCTTCATGGAAATCATATACATTAGAAAAT
CACAGTCAGATGTTTAATCAATCCAAAATGTCCACTATTTCTTATGTCATTCGTTAGTCTACATGTTTCTA
AACATATAAATGTGAATTTAATCAATTCCTTTCATAGTTTTATAATTCTCTGGCAGTTCCTTATGATAGAG
TTTATAAAACAGTCCTGTGTAAACTGCTGGAAGTTCTTCCACAGTCAGGTCAATTTTGTCAAACCCTTCTC
TGTACCCATACAGCAGCAGCCTAGCAACTCTGCTGGTGATGGGAGTTGTATTTTCAGTCTTCGCCAGGTCA
TTGAGATCCATCCACTCACATCTTAAGCATTCTTCCTGGCAAAAATTTATGGTGAATGAATATGGCTTTAG
GCGGCAGATGATATACATATCTGACTTCCCAAAAGCTCCAGGATTTGTGTGCTGTTGCCGAATACTCAGGA
CGGACCTGAATTCTGATTTTATACCAGTCTCTTCAAAAACTTCTCGAACCGCTGTGTCTCCTACGTAAAAA
AAGAGATGTACAAATCAATAATAATTACACTTTTAGAAACTGTATCATCAAAGATTTTCAGTTAAAGTAGC
ATTATGTAAAGGCTCAAAACATTACCCTAACAAAGTAAAGTTTTCAATACAAATTCTTTGCCTTGTGGATA
TCAAGAAATCCCAAAATATTTTCTTACCACTGTAAATTCAAGAAGCTTTTGAAATGCTGAATATTTCTTTG
GCTGCTACTTGGAGGCTTATCTACCTGTACATTTTTGGGGTCAGCTCTTTTTAACTTCTTGCTGCTCTTTT
TCCCAAAAGGTAAAAATATAGATTGAAAAGTTAAAACATTTTGCATGGCTGCAGTTCCTTTGTTTCTTGAG
ATAAGATTCCAAAGAACTTAGATTCATTTCTTCAACACCGAAATGCTGGAGGTGTTTGATCAGTTTTCAAG
AAACTTGGAATATAAATAATTTTATAATTCAACAAAGGTTTTCACATTTTATAAGGTTGATTTTTCAATTA
AATGCAAATTTGTGTGGCAGGATTTTTATTGCCATTAACATATTTTTGTGGCTGCTTTTTCTACACATCCA
GATGGTCCCTCTAACTGGGCTTTCTCTAATTTTGTGATGTTCTGTCATTGTCTCCCAAAGTATTTAGGAGA
AGCCCTTTAAAAAGCTGCCTTCCTCTACCACTTTGCTGGAAAGCTTCACAATTGTCACAGACAAAGATTTT
TGTTCCAATACTCGTTTTGCCTCTATTTTTCTTGTTTGTCAAATAGTAAATGATATTTGCCCTTGCAGTAA
TTCTACTGGTGAAAAACATGCAAAGAAGAGGAAGTCACAGAAACATGTCTCAATTCCCATGTGCTGTGACT
GTAGACTGTCTTACCATAGACTGTCTTACCCATCCCCTGGATATGCTCTTGTTTTTTCCCTCTAATAGCTA
TGGAAAGATGCATAGAAAGAGTATAATGTTTTAAAACATAAGGCATTCGTCTGCCATTTTTCAATTACATG
CTGACTTCCCTTACAATTGAGATTTGCCCATAGGTTAAACATGGTTAGAAACAACTGAAAGCATAAAAGAA
AAATCTAGGCCGGGTGCAGTGGCTCATGCCTATATTCCCTGCACTTTGGGAGGCCAAAGCAGGAGGATCGC
TTGAGCCCAGGAGTTCAAGACCAACCTGGTGAAACCCCGTCTCTACAAAAAAACACAAAAAATAGCCAGGC
ATGGTGGCGTGTACATGTGGTCTCAGATACTTGGGAGGCTGAGGTGGGAGGGTTGATCACTTGAGGCTGAG
AGGTCAAGGTTGCAGTGAGCCATAATCGTGCCACTGCAGTCCAGCCTAGGCAACAGAGTGAGACTTTGTCT
CAAAAAAAGAGAAATTTTCCTTAATAAGAAAAGTAATTTTTACTCTGATGTGCAATACATTTGTTATTAAA
TTTATTATTTAAGATGGTAGCACTAGTCTTAAATTGTATAAAATATCCCCTAACATGTTTAAATGTCCATT
TTTATTCATTATGCTTTGAAAAATAATTATGGGGAAATACATGTTTGTTATTAAATTTATTATTAAAGATA
GTAGCACTAGTCTTAAATTTGATATAACATCTCCTAACTTGTTTAAATGTCCATTTTTATTCTTTATGTTT
GAAAATAAATTATGGGGATCCTATTTAGCTCTTAGTACCACTAATCAAAAGTTCGGCATGTAGCTCATGAT
CTATGCTGTTTCTATGTCGTGGAAGCACCGGATGGGGGTAGTGAGCAAATCTGCCCTGCTCAGCAGTCACC
ATAGCAGCTGACTGAAAATCAGCACTGCCTGAGTAGTTTTGATCAGTTTAACTTGAATCACTAACTGACTG
AAAATTGAATGGGCAAATAAGTGCTTTTGTCTCCAGAGTATGCGGGAGACCCTTCCACCTCAAGATGGATA
TTTCTTCCCCAAGGATTTCAAGATGAATTGAAATTTTTAATCAAGATAGTGTGCTTTATTCTGTTGTATTT
TTTATTATTTTAATATACTGTAAGCCAAACTGAAATAACATTTGCTGTTTTATAGGTTTGAAGAACATAGG
AAAAACTAAGAGGTTTTGTTTTTATTTTTGCTGATGAAGAGATATGTTTAAATATGTTGTATTGTTTTGTT
TAGTTACAGGACAATAATGAAATGGAGTTTATATTTGTTATTTCTATTTTGTTATATTTAATAATAGAATT
AGATTGAAATAAAATATAATGGGAAATAATCTGCAGAATGTGGGTTTTCCTGGTGTTTCCCTCTGACTCTA
GTGCACTGATGATCTCTGATAAGGCTCAGCTGCTTTATAGTTCTCTGGCTAATGCAGCAGATACTCTTCCT
GCCAGTGGTAATACGATTTTTTAAGAAGGCAGTTTGTCAATTTTAATCTTGTGGATACCTTTATACTCTTA
GGGTATTATTTTATACAAAAGCCTTGAGGATTGCATTCTATTTTCTATATGACCCTCTTGATATTTAAAAA
ACACTATGGATAACAATTCTTCATTTACCTAGTATTATGAAAGAATGAAGGAGTTCAAACAAATGTGTTTC
CCAGTTAACTAGGGTTTACTGTTTGAGCCAATATAAATGTTTAACTGTTTGTGATGGCAGTATTCCTAAAG
TACATTGCATGTTTTCCTAAATACAGAGTTTAAATAATTTCAGTAATTCTTAGATGATTCAGCTTCATCAT
TAAGAATATCTTTTGTTTTATGTTGAGTTAGAAATGCCTTCATATAGACATAGTCTTTCAGACCTCTACTG
TCAGTTTTCATTTCTAGCTGCTTTCAGGGTTTTATGAATTTTCAGGCAAAGCTTTAATTTATACTAAGCTT
AGGAAGTATGGCTAATGCCAACGGCAGTTTTTTTCTTCTTAATTCCACATGACTGAGGCATATATGATCTC
TGGGTAGGTGAGTTGTTGTGACAACCACAAGCACTTTTTTTTTTTTTAAAGAAAAAAAGGTAGTGAATTTT

TAATCATCTGGACTTTAAGAAGGATTCTGGAGTATACTTAGGCCTGAAATTATATATATTTGGCTTGGAAA
TGTGTTTTTCTTCAATTACATCTACAAGTAAGTACAGCTGAAATTCAGAGGACCCATAAGAGTTCACATGA
AAAAAATCAATTTATTTGAAAAGGCAAGATGCAGGAGAGAGGAAGCCTTGCAAACCTGCAGACTGCTTTTT
GCCCAATATAGATTGGGTAAGGCTGCAAAACATAAGCTTAATTAGCTCACATGCTCTGCTCTCACGTGGCA
CCAGTGGATAGTGTGAGAGAATTAGGCTGTAGAACAAATGGCCTTCTCTTTCAGCATTCACACCACTACAA
AATCATCTTTTATATCAACAGAAGAATAAGCATAAACTAAGCAAAAGGTCAATAAGTACCTGAAACCAAGA
TTGGCTAGAGATATATCTTAATGCAATCCATTTTCTGATGGATTGTTACGAGTTGGCTATATAATGTATGT
ATGGTATTTTGATTTGTGTAAAAGTTTTAAAAATCAAGCTTTAAGTACATGGACATTTTTAAATAAAATAT
TTAAAGACAATTTAGAAAATTGCCTTAATATCATTGTTGGCTAAATAGAATAGGGGACATGCATATTAAGG
AAAAGGTCATGGAGAAATAATATTGGTATCAAACAAATACATTGATTTGTCATGATACACATTGAATTTGA
TCCAATAGTTTAAGGAATAGGTAGGAAAATTTGGTTTCTATTTTTCGATTTCCTGTAAATCAGTGACATAA
ATAATTCTTAGCTTATTTTATATTTCCTTGTCTTAAATACTGAGCTCAGTAAGTTGTGTTAGGGGATTATT
TCTCAGTTGAGACTTTCTTATATGACATTTTACTATGTTTTGACTTCCTGACTATTAAAAATAAATAGTAG
ATACAATTTTCATAAAGTGAAGAATTATATAATCACTGCTTTATAACTGACTTTATTATATTTATTTCAAA
GTTCATTTAAAGGCTACTATTCATCCTCTGTGATGGAATGGTCAGGAATTTGTTTTCTCATAGTTTAATTC
CAACAACAATATTAGTCGTATCCAAAATAACCTTTAATGCTAAACTTTACTGATGTATATCCAAAGCTTCT
CATTTTCAGACAGATTAATCCAGAAGCAGTCATAAACAGAAGAATAGGTGGTATGTTCCTAATGATATTAT
TTCTACTAATGGAATAAACTGTAATATTAGAAATTATGCTGCTAATTATATCAGCTCTGAGGTAATTTCTG
AAATGTTCAGACTCAGTCGGAACAAATTGGAAAATTTAAATTTTTATTCTTAGCTATAAAGCAAGAAAGTA
AACACATTAATTTCCTCAACATTTTTAAGCCAATTAAAAATATAAAAGATACACACCAATATCTTCTTCAG
GCTCTGACAGGCCTCCTGGAAACTTCCACATATTTTTCAACTGCAGTATAAAGTCAGAAAATAAAGTTAAC
ATAACTTTCACTAACACACACATATGTAGATTTCACAAAATCCACCTATAATTGGTCAAAGTGGTTGAGAA
TATATTTTTTAGTAATTGCATGCAAAATTTTTCTAGCTTCCATCCTTTCTCCCTCGTTTCTTCTTTTTTTG
GGGGAGCTGGTAACTGATGAAATCTTTTCCCACCTTTTCTCTTCAGGAAATATAAGTGGTTTTGTTTGGTT
AACGTGATACATTCTGTATGAATGAAACATTGGAGGGAAACATCTACTGAATTTCTGTAATTTAAAATATT
TTGCTGCTAGTTAACTATGAACAGATAGAAGAATCTTACAGATGCTGCTATAAATAAGTAGAAAATATAAA
TTTCATCACTAAAATATGCTATTTTAAAATCTATTTCCTATATTGTATTTCTAATCAGATGTATTACTCTT
ATTATTTCTATTGTATGTGTTAATGATTTTATGTAAAAATGTAATTGCTTTTCATGAGTAGTATGAATAAA
ATTGATTAGTTTGTGTTTTCTTGTCTCCC

LVGVGGGDVEDVTPRPGGCQISGRGARGCNGIPGAAAWEAALPRRRPRRHPSVNPRSRAAGSPRTRGRRTE
ERPSGSRLGDRGRGRALPGGRLGGRGRGRAPERVGGRGRGRGTAAPRAAPAARGSRPGPAGTMAAGSITTL
PALPEDGGSGAFPPGHFKDPKRLYCKNGGFFLRIHPDGRVDGVREKSDPHIKLQLQAEERGVVSIKGVCAN
RYLAMKEDGRLLASKCVTDECFFFERLESNNYNTYRSRKYTSWYVALKRTGQYKLGSKTGPGQKAILFLPM
SAKS

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

#2

NM_014293

No conservation found.

NPTXR neuronal pentraxin receptor

Synonyms: NPR; NPTXR

CGGCCGCGGCGACAGCTCCAGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCCGCGCCTGCCCCGCTC
GGCCCAGCGCGCCCGGGCTCCGCGCCCCGACCCCGCCGCCGCGCCTGCCGGGGGCCTCGGGCGCCCCCGCC
GCCCGCCTCACGCTGAAGTTCCTGGCCGTGCTGCTGGCCGCGGGCATGCTGGCGTTCCTCGGTGCCGTCAT
CTGCATCATCGCCAGCGTGCCCCTGGCGGCCAGCCCGGCGCGGGCGCTGCCCGGCGGCGCCGACAATGCTT
CGGTCGCCTCGGGCGCCGCCGCGTCCCCGGGCCCGCAGCGGAGCCTGAGCGCGCTGCACGGCGCGGGCGGT
TCAGCCGGGCCCCCCGCGCTGCCCGGGGCACCCGCGGCCAGCGCGCACCCGCTGCCGCCCGGGCCCCTGTT
CAGCCGCTTCCTGTGCACGCCGCTGGCTGCTGCCTGCCCGTCGGGGGCCCAGCAGGGGGACGCGGCGGGCG
CTGCGCCGGGCGAGCGCGAAGAGCTGCTGCTGCTGCAGAGCACGGCCGAGCAGCTGCGCCAGACGGCGCTG
CAGCAGGAGGCGCGCATCCGCGCCGACCAGGACACCATCCGTGAGCTCACCGGCAAGCTGGGCCGCTGCGA
GAGCGGCCTGCCGCGCGGCCTCCAGGGCGCCGGGCCCCGCCGCGACACCATGGCCGACGGGCCCTGGGACT

CGCCTGCGCTCATTCTGGAGCTGGAGGACGCCGTGCGCGCCCTGCGGGACCGCATCGACCGCCTGGAGCAG
GAGCTTCCAGCCCGTGTGAACCTCTCAGCTGCCCCAGCCCCAGTCTCTGCTGTGCCCACCGGCCTACACTC
CAAGATGGACCAGCTGGAGGGGCAGCTGCTGGCCCAGGTGCTGGCACTGGAGAAGGAGCGTGTGGCCCTCA
GCCACAGCAGCCGCCGGCAGAGGCAGGAAGTGGAAAAGGAGTTGGACGTCCTGCAGGGTCGTGTGGCTGAG
CTGGAGCACGGGTCCTCAGCCTACAGTCCTCCAGATGCCTTCAAGATCAGCATCCCCATCCGTAACAACTA
CATGTACGCCCGCGTGCGGAAGGCTCTGCCCGAGCTCTACGCATTCACCGCCTGCATGTGGCTGCGGTCCA
GGTCCAGCGGCACCGGCCAGGGCACCCCCTTCTCCTACTCAGTGCCCGGGCAGGCCAACGAGATTGTACTG
CTAGAGGCGGGCCATGAGCCCATGGAGCTGCTGATCAACGACAAGGTGGCCCAGCTGCCCCTGAGCCTGAA
GGACAATGGCTGGCACCACATCTGCATCGCCTGGACCACAAGGGATGGCCTATGGTCTGCCTACCAGGACG
GGGAGCTGCAGGGCTCCGGTGAGAACCTGGCTGCCTGGCACCCCATCAAGCCTCATGGGATCCTTATCTTG
GGCCAGGAGCAGGATACCCTGGGTGGCCGGTTTGATGCCACCCAGGCCTTTGTCGGTGACATTGCCCAGTT
TAACCTGTGGGACCACGCCCTGACACCAGCCCAGGTCCTGGGCATTGCCAACTGCACTGCGCCACTGCTGG
GCAACGTCCTTCCCTGGGAAGACAAGTTGGTGGAGGCCTTTGGGGGTGCAACAAAGGCTGCCTTCGATGTC
TGCAAGGGGAGGGCCAAGGCATGAGGGGCCACCTCATCCAGGGCCCCTCCCTTGCCTGCCACTTTGGGGAC
TTGAGGGGGGTCATATTCCCTCCTCAGCCTGCCCACGCACTGGCCTTCCCTCCTGCCCCACTCCTGGCTGT
GCCTCCCATTTCCCCTCACCTGTACCCACACCTCCAGAATGCCCTGCCCTGCGAGTGTGTCCCCTGTCCCC
ACCTGAGTGGGGAGGAGCGTCTCAAGTGAACAGTGGGAGCCTGCCCACCTGGCACTGCACTGGAGTTGTCT
CTTACCCCACCCTCCCTGCCCATCAACTGTATCTGATTTCACTAATTTTGACAGCACCCCCAGTAGGGTAG
GATTGTGTATGAGGGGGACCCCACTATCTCAGTGGTGGGGGTGGCCGCCCGCCCCCTTGTCCCCCATGCAA
CAGGCCCAGTGGCTTCCCCTTCAGGGCCACAACAGGCTGTAGAAGGGGATGACGAGGACATCAGAGGTTAG
ACTTACCCTCCTCCCTCTTTCCACCAGCTGCCAGTCAAGGGCAGTGGGATCTCGATGGAGCCTCCCCCCCC
CCCCACCCATGCCTCCCTCTTCCTCCTCTTTCCTCCTCTCTTTGTGTGTAGCGGTTTGAATGTTGGTTCCA
TGCCTGGCCCAGCCCCACCTCAGTCTCCAGGACATTCCTTTCCCAGCTCCAGCCTGGAGGGAAGGGGACAA
AGACCCCAGGAGGCCAAAGGGCTGCAGTCACCCCTTGTGCTCACCCATAGTGATGGCCACTGGTATAGTCA
TCGCTCTCCCTCCATGCCAAGGACAGGACTTGGACCGCTTCAGCCTGGGCTGGGAGCAGCCCTAAGGTAGA
GGCCTCATGGCCCAGGAGACCCCACCTCTGGCAGAGCCACATTACCTACCCTGTGCATGGTCCTGGGGCAG
CAAGGAAGAAGCTCAGAGGGTGGGGAGAAGCATGAAGCAGTGAGCAGAGCACTGGGTGAGAGGGAGAAGAC
CTTGGTTCCTAGCCAGCCCTGCTAATGTGCTGTGTGGCCTTCTGTAAGTCCCTGCCCTCTCTGGGCCTGGC
CTTCCTCATTCGTGAGCTGAGGCCCTCGCTTTGGTCATTTGCTCTCCAGATTGGGTGTGAGCTTCTCTGTG
ATTCCAGGTGGATATGTGGGGAAAGCTCTGGTGACCCTGGGCTTCGCAGGGGTAGATCCCAGGACTCGGCA
GTGGATGGGATGCAGCCAGTCATGGGTTAGGGTCAGCAGAGACTCAGAGTCCAGGGCAAGGTTCAAGGCAG
ACTAACCTCATGCATGGATTGTAAAAAACCAGCTCCCTTTGGATCAACCCAGCCTGGCACCCTTGCCTGTC
TGAGAGTGTCTCAAAGGGCTGATGGCTTCCTGGTCCCCTTGAGTCATCACCAGCTTCCCCAAGAGAGTGTC
AGAATCTTAAGAGCTGAGAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGA
CAGGCAGATCACTTGAGGTCAGGAGTTCGAAGTCAGCCTGGCCAACGTGGTGAAACCCCATCTTCACTAAA
AATACAAAACTTAGCTGGTTAGGTGGTGCATGCCTGTAGTCCCAGCTACTCGGGAGGCCGAGGCAGAAGAA
TCTCTTGAACTGAGGAGGTGGAGGTTGCAGTGAGCCGAGATCACGCCATTGCACTCCAGCCTGGGCAACAG
AGCAAGACTCCATCTCAAAAAAATAATAATAATCTTAAAGATGAGAAAAGCCACCCCATCTGGCACCACAG
CTGCATCTTGCTTGTGAGAAATGGGGAAGAGTTCAGGGAGGACACGTGACCTGCACAGGATCACAGAGCAT
GGGGCAGAGCCAGGACTAGAGCTCAGGGCATCTGACTCCCTCTTCAGTGTTCTTCCCCCTCCATGTTGCCT
GCCCCTGAAGACCTTTGAGTTCAGTCTACACCTAAGCAGGTAGACATCCGCGAGGTCAGATGCTTTCCAAC
ATGACACCTGAACATCTTCCTTTATGCAACACCCAAACATCTTGGCATCCCCACCCCAGGAAGTGCGGGGA
GGAGGTTATGATCCCTGGGCGCTTCGGCAGAATGGAGAGCTGAGGTGTCCCTCCCCTGCTAGTCACCTACC
AGGTGTCTGAGCAGCTGCATGCTCCCTGGCTCAAGTGGGCACTGTACCTTTTGCCTGCCTTTTTGTTCCCT
ATCTCCACTCCCTGAGGCCACTTAGCCTGAGACATGATGCAAGAGCTGCAGGCCGGGGGGCTCAGTGCCAT
GGAAGCTACTCCAAGTTGCATTGCCTCCCGCGCCCAGATCCTGCTTTCCATTTCGAGAACATAAATAGATT
GCCCAGCCCCTCCAGTACAATCCCACTGGAAGAAAAGGCAATGGCGGGCTTCAGCCAGACCTGCTGAGACC
TAGGTTGCCACGGTAACAGCCAAAGACATCAACCCAAGTGCTGGGTCAAGTGTCTCATCATACTGGCACTG
TTGCTGGGGTGACGGCAGAATTCAGAACTTCAATTTCAGTGACGCCAAGCTTGATGTGTTTCTGTTATTGT
TTTGAAGAAGGTAGCTCTTGTGGAGGACTTGGGAGAAGGATGGGGTCTTAGGAAGGAGGTGACAGCACTTG
CATGGTCACTTGAGCCCACACACACGCTCAACCCCAAGTCCTTTATGCTTTGTCACAGTGAAGATGAGACC
TCTGACGTCCAAGCCTTGTTCCTGTGCTGCATCACCCACTCAGCCTTCCAAAGGGAACAGGAACAAATTTC
CCCAGCACCACTGTTTGGGTCCCGCTTTTCCTATCTTCTGCTGCCCCTGAGCACATCCAAGCAGACAGGGA
AAGAGGAGTCAGACATGGCCCAGTCACATCCTGAGCTGCTCCTGGCTGATAACCACGATGGAGCCCGTGTT
TGTCCTGCCATCTGGCACTGCACTGAGTGTGGCACAGGCACCGTCCTGTTGATCTCACAACACAGTTCTAA
GTTAGGACGTTCTTGGCTCCGTTAGACAGGTGAGGAAACTGGGGCACAGAGAGGTGATGTCATCTGCCTGG
TGTCAATCAGCTAGCAAGTGATGGAGCCCAGATTTCAAACCAAAGGGGGTTACGTCCAGGGGCTGAGTTCC

CACTCACCTGTGTAGAGTGCCATCTGGGCACCATTGCTCCAGACGTGTTCCGACCCCTTTCCCAGCCCACA
GGGCTTGAAGTGAAGGAACAGAGGCAGGGGGTGGGCCAGCCCCAGGGCCAGGGTCCCCTTGGTGAAGCCGT
GCCAGGGGGCTCAGCTGCTTCAGGGAATGTGTCCCTCCCACCATGGGCCAGAGCTTCAGCCCTTCTTTAGC
TCAGCTAGAGTTCACAGGAGAGCCAAAAAAGAAAAGGAAGCTGAGCATCTCCCGAGTCCTGGGCAGGGAAG
GGGAGGGAAATTGCTGCTTCTCCAACTCTTGCTTGGGGCCAAGCCCTGCACCAGTTGCTTCCCAGCTGTTA
TCTGCCAGATCTTCCCATCTTGTGGCATGTGGTGCCCCCACCAACATCCCAAGGGGACCAATCCCCTTGCC
ACCACTTTGCATCACCTGGGACCACAGATTTGGACAGGAAGGGCTCTGAGAAGAGGCCAAAGCCCTCATTT
TACAGATGAGGAAGCTGAAGCCCGGGGAGGGGAGCGACCCTCAAGGCCACCCAGCTGGACACGGGAGACTT
GAGCCCAGCCTTCTGACTGCATTCAGCCCTCTCTAGGACGCAGCAGCCTCTCCCCAGCACTGAGTCCCCCC
TCCTTTGTGTGTCCCAGCACCCTTGGCCTGAGTAAACTTGGAAAGGGGCTCCCTCCCAGAGAAGGGACTAC
TCTCTTCACCCCTTTATTCCAGCTGCCTGCCACCCCAGACCCCCACCTCCCACCCTGACCCCCGACCCCTG
GGTGGGGAAGGGGCTCACATGGGCCCAGGCTGAGTGTGAGTGAGCATGTCAAGTTGTCTGACACTGTGACA
TTAGTGCACCCTACTGACAACCCCTCCCCAGCCTTGCCCCTTTCTCCTCTCCCTGTTTTGTACATAAATTG
ACATGAGCTGCAACATGTGTGCGTGTGTGTGCGTGTGTGTGTGTGTGTATGTGTGTGTGATCTGTGTCATG
GTTTTGTTACCTTTTTGTTTTTGTAAACTTGAATGTTCAAAATAAACATGCTGTTTACTCTGA

LKFLAVLLAAGMLAFLGAVICIIASVPLAASPARALPGGADNASVASGAAASPGPQRSLSALHGAGGSAGP
PALPGAPAASAHPLPPGPLFSRFLCTPLAAACPSGAQQGDAAGAAPGEREELLLLQSTAEQLRQTALQQEA
RIRADQDTIRELTGKLGRCESGLPRGLQGAGPRRDTMADGPWDSPALILELEDAVRALRDRIDRLEQELPA
RVNLSAAPAPVSAVPTGLHSKMDQLEGQLLAQVLALEKERVALSHSSRRQRQEVEKELDVLQGRVAELEHG
SSAYSPPDAFKISIPIRNNYMYARVRKALPELYAFTACMWLRSRSSGTGQGTPFSYSVPGQANEIVLLEAG
HEPMELLINDKVAQLPLSLKDNGWHHICIAWTTRDGLWSAYQDGELQGSGENLAAWHPIKPHGILILGQEQ
DTLGGRFDATQAFVGDIAQFNLWDHALTPAQVLGIANCTAPLLGNVLPWEDKLVEAFGGATKAAFDVCKGR
AKA

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

#3

NM_021182

No conservation found.

HMHB1 histocompatibility (minor) HB-1

Synonyms: HB-1; HLA-HB1; HMHB1

CCACATCCCAGGAGGCCGAGGCGGCTTGCCCCGCATCTCAGAAGCCGGGCAGGCCCTGAGCCTTCTGACCT
CACATCCTCTGCCACACCACAGTGGAGAAACCAGAAGTGGAGGAGCAGCCAGAATGCAGAGAAGAAAAAAG
AGGTTCTCTGCATGTTTGGAAGTCGGAATTGGTTGAAGTTGAAGATGATGTGTATCTGAGGCACAGCTCTT
CCCTGACTTATAGGCTTTGACACTGCTGTTGAGGTTTGACTCGAAGCCCAGAGTTTTGGTGTGGATGAGCA
GGGACAAATTGCTGAGCATGAAGAAGAGTAAAATTAAGCAAGTGGAACATATGCCCTTTGCCTCTGCTCTG
CACAGTGAAATGAAAAGTCAACCTTTGAAAAAAAAAAAAAAA

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

#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

GGCGGAGCGTGGTACTACGACCAGCGCGGGCCGGAGGGGGCGGGGGGATGCGCCGCGGCGGCGGCGGCGCG
GGAGCTGGGGTTGGTGTTTGGCGGCGCCAGAGCAGCGGATCCCGGTCTCGCCGCAGCAGCAGCGCGGGTGT
CGTGCACCGCCTGAAGACGCCGTACCTTTCTACCCCCCACCTTTTTTTTTTTTTTTTTTAAATAACCGGAA
CCAATGAACGCAGCCGGGATCAGAGCTCCGGAGGCCGCCGGTGCCGATGGGACCAGGCTGGCGCCCGGCGG
GAGCCCGTGTCTGAGGCGGCGGGGGCGGCCGGAGGAGTCGCCGGCGGCGGTGGTGGCGCCTCGCGGAGCCG
GCGAGCTGCAGGCGGCCGGGGCGCCGCTGCGCTTTCACCCGGCTTCTCCTCGGCGCCTTCATCCCGCCTCG
ACTCCTGGCCCAGCGTGGGGCTGGCTGCTGCGGCGGCGGCGCTGGGCTGCGTTGCTGGTGCTCGGGCTGCT
GGTAGCCGGAGCGGCGGACGGATGCGAGCTTGTGCCCCGGCACCTCCGCGGGCGGCGGGCGACTGGCTCTG
CCGCAACTGCCGCCTCCTCTCCCGCCGCGGCGGCCGGCGATAGCCCGGCGCTCATGACAGATCCCTGCATG
TCACTGAGTCCACCATGCTTTACAGAAGAAGACAGATTTAGTCTGGAAGCTCTTCAAACAATACATAAACA
AATGGATGATGACAAAGATGGTGGAATTGAAGTAGAGGAAAGTGATGAATTCATCAGAGAAGATATGAAAT
ATAAAGATGCTACTAATAAACACAGCCATCTGCACAGAGAAGATAAACATATAACGATTGAGGATTTATGG
AAACGATGGAAAACATCAGAAGTTCATAATTGGACCCTTGAAGACACTCTTCAGTGGTTGATAGAGTTTGT
TGAACTACCCCAATATGAGAAGAATTTTAGAGACAACAATGTCAAAGGAACGACACTTCCCAGGATAGCAG
TGCACGAACCTTCATTTATGATCTCCCAGTTGAAAATCAGTGACCGGAGTCACAGACAAAAACTTCAGCTC
AAGGCATTGGATGTGGTTTTGTTTGGACCTCTAACACGCCCACCTCATAACTGGATGAAAGATTTTATCCT
CACAGTTTCTATAGTAATTGGTGTTGGAGGCTGCTGGTTTGCTTATACGCAGAATAAGACATCAAAAGAAC
ATGTTGCAAAAATGATGAAAGATTTAGAGAGCTTACAAACTGCAGAGCAAAGTCTAATGGACTTACAAGAG
AGGCTTGAAAAGGCACAGGAAGAAAACAGAAATGTTGCTGTAGAAAAGCAAAATTTAGAGCGCAAAATGAT
GGATGAAATCAATTATGCAAAGGAGGAGGCTTGTCGGCTGAGAGAGCTAAGGGAGGGAGCTGAATGTGAAT
TGAGTAGACGTCAGTATGCAGAACAGGAATTGGAACAGGTTCGCATGGCTCTGAAAAAGGCCGAAAAAGAA
TTTGAACTGAGAAGCAGTTGGTCTGTTCCAGATGCACTTCAGAAATGGCTTCAGTTAACACATGAAGTAGA
AGTGCAATACTACAATATTAAAAGACAAAACGCTGAAATGCAGCTAGCTATTGCTAAAGATGAGGCAGAAA
AAATTAAAAAGAAGAGAAGCACAGTCTTTGGGACTCTGCACGTTGCACACAGCTCCTCCCTAGATGAGGTA
GACCACAAAATTCTGGAAGCAAAGAAAGCTCTCTCTGAGTTGACAACTTGTTTACGAGAACGACTTTTTCG
CTGGCAACAAATTGAGAAGATCTGTGGCTTTCAGATAGCCCATAACTCAGGACTCCCCAGCCTGACCTCTT
CCCTTTATTCTGATCACAGCTGGGTGGTGATGCCCAGAGTCTCCATTCCACCCTATCCAATTGCTGGAGGA
GTTGATGACTTAGATGAAGACACACCCCCAATAGTGTCACAATTTCCCGGGACCATGGCTAAACCTCCTGG
ATCATTAGCCAGAAGCAGCAGCCTGTGCCGTTCACGCCGCAGCATTGTGCCGTCCTCGCCTCAGCCTCAGC
GAGCTCAGCTTGCTCCACACGCCCCCCACCCGTCACACCCTCGGCACCCTCACCACCCGCAACACACACCA
CACTCCTTGCCTTCCCCTGATCCAGATATCCTCTCAGTGTCAAGTTGCCCTGCGCTTTATCGAAATGAAGA
GGAGGAAGAGGCCATTTACTTCTCTGCTGAAAAGCAATGGGAAGTGCCAGACACAGCTTCAGAATGTGACT
CCTTAAATTCTTCCATTGGAAGGAAACAGTCTCCTCCTTTAAGCCTCGAGATATACCAAACATTATCTCCG
CGAAAGATATCAAGAGATGAGGTGTCCCTAGAGGATTCCTCCCGAGGGGATTCGCCTGTAACTGTGGATGT
GTCTTGGGGTTCTCCCGACTGTGTAGGTCTGACAGAAACTAAGAGTATGATCTTCAGTCCTGCAAGCAAAG
TGTACAATGGCATTTTGGAGAAATCCTGTAGCATGAACCAGCTTTCCAGTGGCATCCCGGTGCCTAAACCT
CGCCACACATCATGTTCCTCAGCTGGCAACGACAGTAAACCAGTTCAGGAAGCCCCAAGTGTTGCCAGAAT
AAGCAGCATCCCACATGACCTTTGTCATAATGGAGAGAAAAGCAAAAAGCCATCAAAAATCAAAAGCCTTT
TTAAGAAGAAATCTAAGTGAACTGGCTGACTTGATGGAATCATGTTCAAGTGGCATCTGTAAACTATTATC
CCCCACCCTCCACTCCCCACCTTTTTTTTGGTTTAATTTTAGGAATGTAACTCCATTGGGGCTTTCCAGGC
CGGATGCCATAGTGGAACATCCAGAAGGGCAACTGTCTACTGTCTGCTTATTTAAGTGACTATATATAATC
AATTCATCAAGCCAGTTATTACTGAAAAATCATTGAAATGAGACAGTTTACAGTCATTTCTGCCTATTTAT
TTCTGCTTTGTTCTCAGTGATGTATATGCAACATTTTGTTGAAAGCCACGATGGACTTACAAGCTTTAATG
GACTCGTAAGCCAGCATGGGCTTGCAAAAATTTCTTGTTTACCAGAGCATCTTCTTATCTTTCCACAGAGC
TATTTACATCCTGGACTATATAACTTAAAAGAAGTAAAACGTAATTGCACTACTGTTTTCCAGACTGGAAA
AAAAAAAAATCTCTGCAAGTGAAACTGTATAGAGTTTATAAAATGACTATGGATAGGGGACTGTTTTCACT
TTTAGATCAAAATGGGTTTTTAAGTAGAACCTAGGGTTTCTAATTGACTTGATTTCTGGAAATGAAAACCC
GCGCTTTTATTATGGGAAGCTTCTTGAACTGCATTTACTATTGTGAAGTTTCAAGTCCCGCTGTAAAGATC
ATGTTGTTTTGTTTTCCCCAGGGCTTTCACTGTGATTTACTGCATTGCAGGCTGTATGATAAAACACACAT
AATTTAAAGAGAGAAGGCTCTTGATTCCTTATGCAAGTGGAAGAGTTGAAACTTGATTGAAGGACTTAAAA
CATTCACAACCTTAAGCCGAGGTGGGGGGATATGGGGATTCAGGCAATTGTTTACACACTTTGAATAACTG
CAAAGGATTTACGGTTTGTGAAAAATGTGTACTGTGGAAAAGATAATAAATTGAAGACATTATTGTGTGGG
ATTGTGCTGATTTTTGTTGATAACACAAAAAACACTATGTTTTCTGGAGAGCTGTGTAAGCTGTCTTGTTG
CTTAGTTGCAATATAAGAAATAGTGATGTTTTGGACGTAAGTTGTCAACAAATTTCTATTTTATATTGTTA
TATTTTTATGTAGTTTGAAATGTAAAAATGTTCTAATATCAAGATTAACAAATATAAATTTATGGTGCATT
TAGAAAAAAAAAAAAAAA

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

#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

AGGAGCCGCCGCCAGTGGAGGGCCGGGCGCTGCGGCCGCGGCCGGGGCGGGCGCAGGGCCGAGCGGACGGG
GGGGCGCGGGCCCCCCGGGAGGCCGCGGCCACTCCCCCCCGGGCCGGCGCGGCGGGGGAGGCGGAGGATGG
AAACACCCTTCTACGGCGATGAGGCGCTGAGCGGCCTGGGCGGCGGCGCCAGTGGCAGCGGCGGCAGCTTC
GCGTCCCCGGGCCGCTTGTTCCCCGGGGCGCCCCCGACGGCCGCGGCCGGCAGCATGATGAAGAAGGACGC
GCTGACGCTGAGCCTGAGTGAGCAGGTGGCGGCAGCGCTCAAGCCTGCGGCCGCGCCGCCTCCTACCCCCC
TGCGCGCCGACGGCGCCCCCAGCGCGGCACCCCCCGACGGCCTGCTCGCCTCTCCCGACCTGGGGCTGCTG
AAGCTGGCCTCCCCCGAGCTCGAGCGCCTCATCATCCAGTCCAACGGGCTGGTCACCACCACGCCGACGAG
CTCACAGTTCCTCTACCCCAAGGTGGCGGCCAGCGAGGAGCAGGAGTTCGCCGAGGGCTTCGTCAAGGCCC
TGGAGGATTTACACAAGCAGAACCAGCTCGGCGCGGGCGCGGCCGCTGCCGCCGCCGCCGCCGCCGGG
GGGCCCTCGGGCACGGCCACGGGCTCCGCGCCCCCCGGCGAGCTGGCCCCGGCGGCGGCCGCGCCCGAAGC
GCCTGTCTACGCGAACCTGAGCAGCTACGCGGGCGGCGCCGGGGGCGCGGGGGGCGCCGCGACGGTCGCCT
TCGCTGCCGAACCTGTGCCCTTCCCGCCGCCGCCACCCCCAGGCGCGTTGGGGCCGCCGCGCCTGGCTGCG
CTCAAGGACGAGCCACAGACGGTGCCCGACGTGCCGAGCTTCGGCGAGAGCCCGCCGTTGTCGCCCATCGA
CATGGACACGCAGGAGCGCATCAAGGCGGAGCGCAAGCGGCTGCGCAACCGCATCGCCGCCTCCAAGTGCC
GCAAGCGCAAGCTGGAGCGCATCTCGCGCCTGGAAGAGAAAGTGAAGACCCTCAAGAGTCAGAACACGGAG
CTGGCGTCCACGGCGAGCCTGCTGCGCGAGCAGGTGGCGCAGCTCAAGCAGAAAGTCCTCAGCCACGTCAA
CAGCGGCTGCCAGCTGCTGCCCCAGCACCAGGTGCCCGCGTACTGAGTCCGCGCGCGGGGCGCATGCGCGG
CCACCCTCCCCAAGGGGCGGGCTCGCGGGGGGGTGTCGTGGGCGCCCCGGACTTGGAGAGGGTGCGGCCCT
GGGGACCCCCCCTCCCCGAGTGTGCCCAGGAACTCAGAGAGGGCGCGGCCCCCGGGGATTCCCCCCCCCCG
AGGGTGCCCAGGACTCGACAAGCTGGACCCCCTGCTCCCGGGGGGGCGAGCGCATGACCCCCCCGCCCTCG
CGCTGCCTCTTTCCCCCGCGCGGCCGCCCCGTGTTGCACAAACCCGCGCGTCTCGGCTGCCCCTTTGTACA
CCGCGCCGCGGAAGGGGGCTCCGAGGGGGCGCAGCCTCAAACCCTGCCTTTCCTTTACTTTTACTTTTTTT
TTTTTTTCTTTGGAAGAGAGAAGAACAGAGTGTTCGATTCTGCCCTATTTATGTTTCTACTCGGGAACAAA
CGTTGGTTGTGTGTGTGTGTTTTCTTGTGTTGGTTTTTTAAAGAAATGGGAAGAAGAAAAAAAAATTCT
CCGCCCCTTTCCTCGATCTCGCTCCCCCCTTCGGTTCTTTCGACCGGTCCCCCCTCCCTTTTTTGTTCTGT
TTTGTTTTGTTTTGCTACGAGTCCACATTCCTGTTTGTAATCCTTGGTTCGCCCGGTTTTCTGTTTTCAGT
AAAGTCTCGTTACGCCAGCTCGGCAAAAAAAAAAAAAAAAAAA

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

#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.

MRVI1 murine retrovirus integration site 1 homolog

Synonyms: IRAG; JAW1L; MRVI1

GGCGCTGGCCCAGGGTCTTCTCCAGGCTCACATGCTTGCTGGAGGCTCCAGGCGGGTGGGGAGATCTGCAG
GGTCCCTTTGTTCCCGTCCGGCCAGGACAGTGTCCAGTGGCACGAGTTGCCATGGGGACTGCCATCTGCTG
TGTGGATACCCATCGGAGGGGCCCAGGGTGAGCGACTGCCCCAAAGGCCCAGCCTGCCCTGAGCCCTGAGG
CCAGCTGGTCGGCTTTCCCTGCCATTATCCCCCGAGCCAGGCCGGAGTGGGGCTCCCAGAGGAGCTCTCGG
CTCTAAGTCCCAGCCCCTCTCAGAGGAGGCGTGCGGCCGAGGTTCTGCAGCCAGAGGCTTCTCGGCTCTGA
GCTCCGGAGCCAGATGTAACATTGACCTTAAATGGTAAAAGCTCCCCAGAGTGAAGAGAGGCTGGCCAGAG
GAGGAAAGGAGAATAACTCAGTTTTAGACAGGGTCTCGCACTGTCACCCAGGCTGGAGTGCAGTGGTGCAA
CCCCACCTCCCGGGTTCAAGCAATTCTCATGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGGTCCTCCTA
CCGCAGGAGTATCTTGCAGTCCAACTCCCACGATTGTCCTGACTGGGGATGCCACTTCACCAGAAGGAGAA
ACCGACAAAAACCTGGCCAACAGAGTTCACAGTCCCCACAAGAGGCTTTCTCACCGACACTTGAAGGTGTC
CACTGCCTCCCTGACATCTGTGGACCCCGCGGGGCACATCATTGACCTGGTGAATGACCAGCTGCCAGACA
TCAGCATCTCAGAGGAGGACAAGAAGAAAAACCTGGCGCTGCTGGAAGAAGCCAAGTTGGTGAGTGAGCGA
TTCCTGACCCGCCGTGGGAGGAAGTCCAGGAGCAGCCCCGGAGACTCCCCATCAGCTGTTTCCCCGAACCT
CAGCCCCAGCGCTTCTCCTACATCCTCTCGGAGCAACTCACTTACAGTCCCCACCCCGCCAGGTTTGGATG
TGTGCAGTGGCCCGCCATCCCCTCTGCCTGGAGCACCACCACAGAAGGGGGGATGAGGCCGACGTCTCTTCA
CCTCACCCTGGCGAGCCTAACGTCCCCAAAGGGCTAGCTGACAGGAAGCAGAATGACCAGAGGAAAGTGTC
TCAGGGCAGGCTGGCTCCTCGTCCTCCTCCAGTTGAGAAGTCCAAAGAGATTGCAATAGAACAAAAGGAAA
ACTTCGATCCCCTCCAGTACCCCGAGACCACACCCAAAGGCCTAGCTCCTGTTACAAACAGCAGTGGGAAA
ATGGCCCTGAACAGCCCTCAGCCTGGCCCCGTGGAGAGCGAGCTGGGGAAGCAGCTCTTGAAAACGGGCTG
GGAGGGCAGCCCTCTGCCGAGAAGTCCAACCCAGGATGCGGCAGGAGTGGGTCCCCCAGCCTCCCAGGGGA
GAGGCCCAGCTGGAGAGCCGATGGGGCCCGAGGCTGGCTCCAAAGCTGAGCTTCCACCCACTGTGTCCCGG
CCCCCGCTGCTGCGAGGGCTCTCCTGGGACAGTGGCCCTGAAGAACCTGGCCCCCGGCTGCAGAAAGTGCT
TGCCAAGCTGCCACTGGCAGAGGAAGAAAAGCGTTTTGCAGGCAAGGCCGGCGGCAAGCTGGCCAAGGCCC
CTGGTCTCAAAGACTTTCAGATACAAGTGCAGCCCGTGCGGATGCAGAAACTGACCAAGCTCCGAGAGGAG
CACATCCTGATGAGAAATCAGAACTTAGTGGGGCTCAAGCTTCCAGACCTTAGTGAAGCAGCTGAGCAGGA
AAAAGGGCTTCCTTCTGAACTCTCCCCAGCTATTGAGGAAGAAGAGTCAAAGAGTGGCTTAGATGTCATGC
CTAATATTTCTGATGTGCTGCTGCGCAAACTGCGGGTCCACAGGAGTCTCCCTGGAAGTGCCCCTCCACTC
ACTGAAAAGGAAGTTGAGAACGTGTTTGTGCAACTGTCCTTGGCCTTTAGAAATGACAGCTACACTCTGGA
ATCTAGAATTAACCAGGCTGAAAGGGAACGCAACCTGACAGAGGAGAACACTGAGAAAGAACTGGAAAACT
TCAAAGCTTCCATTACGTCCTCAGCTTCACTCTGGCACCACTGTGAGCACCGGGAAACCTACCAGAAGTTG
CTGGAGGACATCGCTGTCCTGCACCGCCTGGCTGCCCGCCTCTCCAGCCGAGCTGAGGTGGTAGGCGCCGT
CCGCCAGGAAAAGCGCATGTCGAAAGCAACGGAAGTGATGATGCAGTATGTGGAGAATCTAAAGAGGACGT
ATGAGAAGGACCATGCGGAGCTCATGGAGTTTAAAAAGCTTGCAAATCAGAATTCAAGCCGCAGCTGTGGC
CCCTCTGAAGATGGGGTCCCTCGCACGGCACGGTCCATGTCCCTCACGCTGGGAAAGAATATGCCTCGCCG
GAGGGTCAGCGTTGCTGTGGTTCCTAAGTTTAATGCCCTGAATCTGCCTGGCCAAACTCCCAGCTCATCAT
CCATTCCCTCCTTACCAGCCTTGTCGGAATCACCCAATGGGAAAGGCAGCCTACCTGTCACTTCAGCACTG
CCTGCACTTTTGGAAAATGGAAAGACAAATGGGGACCCAGATTGTGAAGCCTCTGCTCCTGCGCTGACCCT
GAGCTGCCTGGAGGAGCTTAGTCAGGAGACCAAGGCCAGGATGGAGGAAGAAGCCTACAGCAAGGGATTCC
AAGAAGGTCTAAAGAAGACCAAAGAACTTCAAGACCTGAAGGAGGAGGAGGAAGAACAGAAGAGTGAGAGT
CCTGAGGAACCTGAAGAGGTAGAAGAAACTGAGGAAGAGGAAAAGGGCCCAAGAAGCAGCAAACTTGAAGA
ATTGGTCCATTTCTTACAAGTCATGTATCCCAAACTGTGTCAGCACTGGCAAGTGATCTGGATGATGGCTG
CAGTGATGCTGGTCTTGACTGTTGTGCTGGGGCTCTACAATTCCTATAACTCTTGTGCAGAGCAGGCTGAT
GGGCCCCTTGGAAGATCCACTTGCTCGGCAGCCCAGAGGGACTCCTGGTGGAGCTCAGGACTCCAGCATGA
GCAGCCTACAGAGCAGTAGGAAACCTCACACCTAGCCAGTGCCCTGCTCTGAGACACTCAGACTACCACCC
TTTCCCCAAGTATAACGTCAGGCCCAAGTGTGGACACACTGCCGCCCATCCCATCAGGTCATGAGGAAGGG
TTCTTTTAACACTCGGCACTTCTGTGGGAGCTATTCATACACAGTGACTTGATGTTCTTGGAGGATCAACA
AAACTGCCCTGGGAAAGCATCCAGTGGATGAAGAAGTCACCTTCACCAAGGAACTCTATTGGAAGGGAAGG
TCTCCTGCCCCTAGCTCAGGTGGCTGGGGAGAACTAAAACACCTTCACTGGTGGTTGGGGGTAAGGAGCGG
GGCACGGGGGAGGAGGAGGTAGGGGGCAGTAAAAAACTTACTCTCTTTTTTCCTCTCTGTAATTGGTTATC
AGGAAGAATTTGCTTAATGACTAACACCCTAAGCATCAGACCTGGAATTTGGAGTTGCAAAGTGACTATCT
TCCCATTTCCCATCTCATTTTCAATAACTTCAGCCTCCCATTCTTTCCTTTGGAATGAGAGTTTCTTTTTA
CAGAAGTAGGAAAGGCTTCTCAGAAAAAAAAAAAAAAGTATAGGCTGAATTTAGCTCAGTGCTTGAAATGG
GAAGATATGAATTATTATATACGCATCTGTCCACACATACACACATACTGTTGTGTACACACACACAACAT
GCCTGTGCACAGAGCCAACAACCCTTCAAAAGTGTGCTCTGGGTGTGTACCTCTGGATAAATAAGATGCAT
GCCAAGCCAACCCACAGATTTTCACCAGTGTGGGGCAGTCACCAGGCACCTGTTCAATGAGCTGTCCACAT

GGATTGAAGATGTTTTAAAAACACAGAAAACTCATGGCTTCAATGGCAGACTTACTAGTCTCCATTTCAAA
TGCCAACTCTGAGCTGCTGTACAGCACAATCTATTCCCTATTCTCTCTTTGAAAACAGTTAACCCACCTCA
CAGGTGAATGAGGAGAGAAGATGTGCTTTCTGCTTCAGTCTCTTACTCTGTGTGTGACCACATGCAAGAGT
AAACTTGCACCTCAGTGCTTCAGTTCAAATGGGGTTTCCAACCCCAGTATAATTAGGGGTGTTTCAGAGCA
TCCCCAGTTATTTAGCACAACACTGAAGGAGCACATCCCCTCTCCATTTTGACTTCTCTCCCCACTTTTAC
AGCCACTGCCTTCATCAGTTTTGTAGAGGTTTGATTTCCATGTGGGTTTTGTTGTCATTGTTTTGCATTTT
TGTTTTGTTATTGATATTGTTTGCTTTCATTGCTAAAACTCATATACGACTTACTATGAGCCAAGCACTGT
TCTCAGTATTACATAGGTATGAATTCATTTAAGTCCTGAAGAAAAAGAAAAAAAAATACGAAGTGGATATT
ACCCTTCCCATTTTCAAATAAGGAAACTGAAGCACAAAAAGAACAAGTAACTTGACAAGGACACCCCGGTA
GTAAATCATGGGGCTGGAGCTCAACCCCAGGGTAGGCTGGCTCCAGAGCTGTGCTCTCCTTGACTCTTCTG
ATGGTCTCCTAGCTGGAAGCCTCACATTTCAGTCTCATTCCCCCAAGTGGCCCATCAGCTACTCCATCTCT
GGCTCCCCAACTAAACAGTTTCTCTCATAGTGCTGGACCTCCACTCACTAGTTTTTTTTCCAGCTGTTCTT
CTCTTTTCTTCAGGTCACTCTTCTCGACCGAGTGCAAAAATTATCCCCTCCATACCAGCTTTGATGACCTT
CCTTCCATACTCCTCACCAGACACAACATAATAGGTCACACACTCCTCTGTGCTTTCTGGCACGTTTTAAA
CATTATTATTATTGACCTTTACCTATAGTATACCATGGCCTATTTATGTATCCATCTCCCCTAGCATTTTT
CCTCAAAGACAAGAACCATGTCTTACCCATCTCTTGGGTAAGTGCCTAGCATGGTGGCTGACGCTTGGGAG
GGTGTCATTAAATGTTGCTCAAAAGAACAAGCAAACATTTAAGGTGGTGGAGAGCAGCCTGGGGACAGCTG
ACATGCTGCATGCTTCTCAGTACCAGCACCATCACAATGCAAAAAGCAACATCTTTCTTAACCTCAGCTTA
TTCTGTTTTTCAGTCTACTCTGTGAGAGAGCAGGAATGAGACCAGACTAGCAACACCATTGCCAAGCTCAA
GGACTGGGCTCAATGCAGTCACTCCTTCAGAGAGACCCCCCACCCCAAGCATGCCCCACTTTAAAATAGCA
TGTTTATTGAAGGGGGCATCCTTTACAGTAGCTAGAAAATGACTGAGGCCCAAGCCAGGGTTGATCAAGGA
TGTGCCATTAAGGTAAAGAGTTACAGAGCAGGGCAGAGGGACTCTGGGGGCAGAAGTGGATGATTTGCCCG
GCCTCTTCCAGGGGGTCTGGATACAACTGAAGGAGCTTTAGCTACATGAGGCCCTCAGAGCCAAAGACAGG
ATGCAAATAGAGTTCTAGAGAGTGGCCGTGGAAGCAGAACTCCAGGTGGGGAATGTTCAATCTCTGCCTCC
CTTAAAGCAGGGCCAGGCTCAGCTGGCCCCATTGTTCACTTGGTCACAAGTTTCCTACCTTTGTTTCTGGA
TGAGTCAAAGGCCAGGAAGGCAGTTATGGAGAGCTCCTGCACCTCCAGCTGCCCCACAGAAAAGCCTGCAA
GAGTACTTCCAGGCACAGGCCCTCTCCCACCCTATTCCATTTGTAAGCAAGGGAGGTCGAGGAAAAGGACA
TCTCCAAAAGGGAGCATAAGAATAGCCATATATACAGGGGCTGAAAAAGTGCTCATGAGTCCATCTTTTCT
GGAAAGCAAAGACCAGCCTGAAGCAGTGGGAGCTGCTGCCCAAGCGGTAGTGAACTGGAGAGAAACAGGCC
CTGGATTTTCAGTGAGGATGTGGATCTGAAGAGTCCCCCAAATGCCTCTGAAGTCTGACATCTCTGCTTAG
CCCTAGGAGTCTGGTTCCCTGCCTTCAGTTGCAGAGTGATGTGTTTGTGTCATTGTTGATGTCACCTCCTA
AAAAGACCTTCACTTTCTGGCTGCCACAAAGCCATATGTGTTGCTCCCCATATACAGCCTGACAGAGTAAA
TGGAGAGGAAGTGCTGGATTTGTGTATCACTGGCTATCAGTTCCTCATGTTGTTAAGCCTCACACAGGTGT
GCTAGCATTGAACTGTAGAGTGTCACATACCTGAGTTTGAAAATAAAAGCACATTTCCAAACCTCAAAAAA
AAAA

LTGDATSPEGETDKNLANRVHSPHKRLSHRHLKVSTASLTSVDPAGHIIDLVNDQLPDISISEEDKKKNLA
LLEEAKLVSERFLTRRGRKSRSSPGDSPSAVSPNLSPSASPTSSRSNSLTVPTPPGLDVCSGPPSPLPGAP
PQKGDEADVSSPHPGEPNVPKGLADRKQNDQRKVSQGRLAPRPPPVEKSKEIAIEQKENFDPLQYPETTPK
GLAPVTNSSGKMALNSPQPGPVESELGKQLLKTGWEGSPLPRSPTQDAAGVGPPASQGRGPAGEPMGPEAG
SKAELPPTVSRPPLLRGLSWDSGPEEPGPRLQKVLAKLPLAEEEKRFAGKAGGKLAKAPGLKDFQIQVQPV
RMQKLTKLREEHILMRNQNLVGLKLPDLSEAAEQEKGLPSELSPAIEEEESKSGLDVMPNISDVLLRKLRV
HRSLPGSAPPLTEKEVENVFVQLSLAFRNDSYTLESRINQAERERNLTEENTEKELENFKASITSSASLWH
HCEHRETYQKLLEDIAVLHRLAARLSSRAEVVGAVRQEKRMSKATEVMMQYVENLKRTYEKDHAELMEFKK
LANQNSSRSCGPSEDGVPRTARSMSLTLGKNMPRRRVSVAVVPKFNALNLPGQTPSSSSIPSLPALSESPN
GKGSLPVTSALPALLENGKTNGDPDCEASAPALTLSCLEELSQETKARMEEEAYSKGFQEGLKKTKELQDL
KEEEEEQKSESPEEPEEVEETEEEEKGPRSSKLEELVHFLQVMYPKLCQHWQVIWMMAAVMLVLTVVLGLY
NSYNSCAEQADGPLGRSTCSAAQRDSWWSSGLQHEQPTEQ

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

#7

NM_001085386

No conservation found.

NF-E4 transcription factor NF-E4

Synonyms: NF-E4

GACAATTCCTGTTTACGGAAGACTATAAAACCCTGCCCCCTACTCATTTGGTGCTGACGCCATTTTAGGCC
TTAGCCTGCCTGCACCCAGGCGCTCATTAAAACAGCAGGTTGCTCCACACTGCCTCGTGTTGTCTGTTGGC
ACACTCTCAAGAGTTTGAACGGATACAAGAATCTTTCATCTGGTGCCGAAACCCGGGAGGGGCTCCGGTCT
TCGTCCCCCGTGGACCTACCCCTCCGCCCCAGAAAGCAGGCCACAGCAGCCGGACAAAGGAAGCTCCTCAG
CCTCCAGTTGCTTCTCTGTGCATGCACATCAGTCACTGATCTCACCTACTGGGGCCCTGCAGGCCATGGGG
CCACAGCTCCACACAGAAGCCTCCTAGCAATCCACCTCCACCTGGTGCCTGCTTCAAGTGCGGCAATGAAG
GCCACTGGCCCACACAATGCCCAAACCCAGGTAAACCCACGAGGCCATGCCCCCTCTGCGGAGGACCCCAC
TGGAAGTTGGACTGTGAGCGGCCCCTGCAAGGACCACCCCCATCCCTTCCTGAGCCAATCAAACCCTCCTA
CTCGGATCTCGTCAGCCTTGCCGCTGAAGACTGATAGTGCCTTGGAACAGACACCCCAGCAACTACCATCG
CTTCATCTGAGCCAAGGGTAACCCTGATGGTGGCAGGCCCAGTCCCAAGGCGCAAGGCCACTTGTGCCAGC
AGTGTGAGTCAGCAAGATAGCAGAAGCAGGAAGAGAGCCGGCCAGAAGACACCTACTCTGACTGGGAGACA
CGTACCCCTGAAGATTGAGAAAGAGGCCATCCAGGTACCACATAGCAGTTACATCAGACTGGGACATTTCC
TGTTTACAGGAGACTATAAAACCCCTGCCCCCTACTCATTTGGTGCTGACGCCATTTTAGGCCTCAGCCTG
CCTGCACCCAGGTACTCATTAAAACAGCGTGTTGCTCCAAAAA

LPRVVCWHTLKSLNGYKNLSSGAETREGLRSSSPVDLPLRPRKQATAAGQRKLLSLQLLLCACTSVTDLTY
WGPAGHGATAPHRSLLAIHLHLVPASSAAMKATGPHNAQTQVNPRGHAPSAEDPTGSWTVSGPCKDHPHPF
LSQSNPPTRISSALPLKTDSALEQTPQQLPSLHLSQG

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

#8

NM_199072

No conservation found.

MDFIC MyoD family inhibitor domain containing

Synonyms: HIC; MDFIC

CCCAGGCCGGCTCTGGCCTCCTGACCCAGACAGCGCAGGGCGCGAGGGATCGCGCGGCCGAGCCCGGGTCG
CGCCGCTCCCAGCATCGGGGCCGCTAGCCAAGAGTTCGAGGCCTTCCCGATCCGGATGTGATGAAAAAGAG
CAACAGAGGGAGAAGTGTTTCAGGATTGTAGGAGTGGAAGAGGGGAAAGAGAGGCAGAGAGGGGGAAGGCC
CCCTCGCAGGGGAGCCGGCTGGAGTGAGCTGGCTGGAAAGAGGGGGCGGAGTGCGCGGAGTCAGAGCCGCC
ACCGCTGCCGCAGTTGCCGCCACTGCGGCGTCTGGGCTGAGCCGGAGGGAGGCGGGAGGACGCGCAGGGGC
GGCCGCCGCCGTCGTCAGGCCACCGGGGCGAAAATGCGGCCGCTGCCGGAGGCTCGCTAACTTTCCGGGGC
GGAAGAGGAGGAGGAGGAGGAAGGGGCTTGGAGCGACTACGGGGGGATGCGGAGAAGCAGTCAGTTCC
CTGCACCCAGCACCTCACAGCCCTTCCTCCGTGCGCCCTGCCGGGCGGCGAGCTAGGCGGCAGCGGCGCGG
CGCGGGCTCGGCGGAGCGGCCCATGTCCGGCGCGGGCGAAGCCCTCGCTCCCGGGCCCGTGGGGCCGCAGC
GCGTGGCCGAGGCGGGCGGCGGCCAGCTGGGCTCCACAGCCCAGGGAAAATGTGATAAAGACAATACTGAG
AAAGATATAACTCAAGCTACCAATAGCCACTTCACACATGGAGAGATGCAAGACCAGTCCATTTGGGGAAA
TCCTTCGGATGGTGAACTCATTAGAACCCAACCTCAGCGCTTGCCTCAGCTTCAGACTTCAGCCCAGGTGC
CAAGTGGTGAGGAAATAGGCAAGATAAAGAACGGCCACACAGGTCTGAGCAATGGAAATGGAATTCACCAC
GGGGCCAAACACGGATCCGCAGATAATCGCAAACTTTCAGCACCTGTTTCTCAAAAAATGCATAGAAAAAT
TCAGTCCAGCTTGTCTGTAAACAGCGATATCAGTAAGAAGAGCAAAGTAAATGCTGTCTTTTCCCAAAAGA
CAGGCTCTTCACCTGAAGATTGTTGTGTCCACTGTATCCTGGCTTGCTTGTTCTGCGAATTCCTGACCCTT
TGCAACATTGTCCTGGGACAAGCGTCATGTGGCATCTGCACCTCAGAAGCCTGCTGCTGTTGCTGTGGTGA
CGAGATGGGGGATGATTGTAACTGCCCTTGTGATATGGACTGTGGCATCATGGATGCCTGTTGTGAATCAT
CAGACTGCTTGGAAATCTGTATGGAATGCTGTGGAATTTGTTTTCCTTCATAAATATTTATCTTTTGTTTG
TGTTAAAACTGGAGAGTGTTTAAAAATTTCCTTTTGGGGGGAAGAAAAGCACATTGTAAGATTCTCATGAA
ACAACATGGAATTTGCACTGTTAACTCATTATTGTAAGTAATCTCTGAAAGCCTTTTTACTTTAACCAAAT
CTACATGGTTTAATATGTGAAATTTTAACTACTTTAACTAGTTTTATAAATTTCTTAATATGTTACAATAA
CTTAGGGACATTTTGACACCCCCCTTCCCAAATGTTAAATGCCTTCTCCTTTTTACCGATATTTCTGTTTC

TTTTAACCGTTCTCAGGAGCACTTTGCTCCAAATATATTATTTTTCAGTGTGTATTTAAACGAGGCAGTTT
ATTTTGATATGTATCTATTCATGATTGAAAGGAAGCAGTCTTGGCCAGGCACGGTGGCTTACACCTGTAAC
CCTGGCATTTTGGGAGGCCAAGGTGGGCAGATTGCCTGAGCTCAGGAGTTCGAGACCAGCCAGGGCAACAT
GGTGAAACCCCATCTCTACTAAAATACAAAAAGTTAGCTGGGCTTGGCGGTGTGCGCCTGTAGTCCCAGCT
ACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGAGAGGCGGAAGTTGCAGTGAGCCGAGATTGTGCC
ACTGAACTCCAACCTGCACTCCAGCCTGGGCAACAGAGCGAGACTCCATCTCTAAATAAATAAATAAATAA
ATAAATAAATAAATAAATAAATAAACAAACCAGTCTTTATTTTAAAAGAAACTTTAGGAAACAAACCCACA
TAATAGTTGGGAACCAGTGTTGATCTCTCTCCCTTACCTTCTCCACTTGTTCAACAGACTCTGAATGCCGA
CTGTGTGGACTCTCTTCCTCAGACTGTGGGGACAGATACAATTCCACTCCTGTCCACAGGAACATGAGATT
TAGCAGACTAAGGAGATCTGTAAAGAATGAACCATACCACAAGGCATACTGAAGTGAGGATTATAAGAGAA
ATAAACTCAAAATGCTGTTGGAATATGCAGAGAATTGCTACCAGAATATTCAGTAAGGTTTCAGGGAGAAT
GTGGCATTTGAGGACTCTCTTAGAATGAGTGATTCACCTGCTATTTAAATGAATTATTTAGATTTTTGACA
AAGATTTAGGTGGACACCCTAAACTGTGTGTGCCTTTAACCAGTTAAAAGAACAGTGCCTTCAGCATACTT
TTTTATTAGTTGTAGGAATACAGCTTTTTGAAAAAGCTATAAAGTTTAAATTAACTAAAAATATGCATTTT
CTTACACATAATTTAAATGTTATCATACTTTTTTGATGAAAACATAATGCCTTAGTAAAATAGCTCTATTT
AATAAAGAAGATTGAGTACTCTGACACATTTCATTTAAATTAGGAAATTTTTAATATTAAAATCCCAGTGT
TCTGAGTTATTGAAAGGCTTTCTTTTATTTTGAGAGCTTTAGGTCTTTTTGGGATGAGAACATTTTAGTTG
TTTAGTTTGTTTCTTAAGCAGTGCTATTTTTTGTAAACACAGATAAATGGAAACCATTCTTTTCAATGCAG
AAGAAATCTAGATATCCCCTACTGTGACCAAATTTCTGTATTACGATTTTATGTTAAATTAAACTAATATG
GCAGGTTATAATGATCCTTAAGTGTAAAGAAATCAGTCAATTACAAGAGTAATTGTATAGTTATTGAGACC
TATAGTGTGTGGCTTAGATGAAAGGGAGAGTAAATTTTCATACCATGCTCTCTCCTACTCAGTTTGATCTC
TCTAAAATTGTAGTTTGGTTTGATTTAATATAATTCTTAGTAGAAATTTTGAAAGTATGCTTTGGGATTAA
TAATTATTTTTAATTTTTCTGGCTGAATATCAAATTGATAGTAACAACAGAAGCATAATTTTAGGAAGGCT
TTCGCAAACCTAGCCTTTTAAGAGAGGTTTTTAACCTGAAGCATGAGAATATATCACCTGTGGTTTTTCCT
TTGAGATGAAACGTAGTTTCTAGTTATATCATTACTTAAAGGGCTTAAAAAGAAAAAACTTAGCAAACTTT
TGAATCTTTCTTTTATTGCTATTTACACATACATACACACATACAAAACCTTTAAATTTTGGGATCTGAAT
ATAATTCTGGTAAACAGCTGTCTTCATTTTTCTCCTCTAAAGAACTTAATTCATTTGTTACATAAAATATA
AGGAAATCTTTATACTATTTTACAGTAACCACAATCTAAATATTTACATATACCCAAAATTAACTTATGCT
CATATATTAGGATGTGAGAATATCATCTGTTTATGGACACATGAAACCTCCTAATGACCTGGAATTGTTAG
AATATTTGACTTTTTATATGCAAAGTTTTTCAACCAAGTGGTTTGTCTAATATTTAAACATGTACTGGCAC
AATTTGTGATGAAAATATTAGCACATTTGCAATAATGTTTCTCCATAACAGAGAATGTTAATGGATACCAG
AATTTTATTTTTGTATTTATGTTCATAGTACTTTTCCTCTTGTCTACTCCAGACAGTTATTCCATAAAGCA
TTTGTATAATTAAAAGGAAAACAGAAAAAGGAAAAGTAGGCAAATGTGAAAATAGTTTCAATATATCTTAT
GATTTCTTAATGTAAAATGTTTTGTTGAAGTATATGGCTATCATGACTAAGTGCTAGAATTTATAGTTACA
GGCGGTGTCCTTTTAAATGTGGAAAGGCTTTTAAAATATTTTAAAACTGGACCTGTATTATCCTGAATACA
CTATTTTGAAAATTTTTAAAAATGACTTCTTTATTTTGCTTTACCGTATGTTTATATCTAATTGACATATT
GACTAATGTTTGAAAGAATTCAACCATAAGTTAAAATCTGAAGGTTATCTTTATCATGTTTCATCCCTGTC
TGAAGATTTCCTAGTCTTCTTATGTAAATCACATGACTCATGTCCGTAAATGAACTATGAAAGATATCGAT
CAGTTTATGATCATTGACATGTGATTTCAAAACACAGTGTTCTTTTAAAAATCTATAATATGTCAAAATAC
AAGTTTTTTTTTTTTACATCGTTTTAGTAAGTTAATTTCATTTATTTACTTTGGAGCTATATTTCCACTTA
GAAAAACTAAGGTAATTTTACAATATATGCTGAGATTAAAAACCAAGGTAAAAATGATCAAACATATATGA
AATTGAGTCTTAGATTTAATGAATTTCACTCGAAAATAAATGATCAGAAGAATTTTCATCTAA

VRGVRAATAAAVAATAASGLSRREAGGRAGAAAAVVRPPGRKCGRCRRLANFPGRKRRRRRRKGLGATTGG
CGEAVSSLHPAPHSPSSVRPAGRRARRQRRGAGSAERPMSGAGEALAPGPVGPQRVAEAGGGQLGSTAQGK
CDKDNTEKDITQATNSHFTHGEMQDQSIWGNPSDGELIRTQPQRLPQLQTSAQVPSGEEIGKIKNGHTGLS
NGNGIHHGAKHGSADNRKLSAPVSQKMHRKIQSSLSVNSDISKKSKVNAVFSQKTGSSPEDCCVHCILACL
FCEFLTLCNIVLGQASCGICTSEACCCCCGDEMGDDCNCPCDMDCGIMDACCESSDCLEICMECCGICFPS

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

#9

NM_005583

Not well conserved beyond humans.

LYL1 lymphoblastic leukemia derived sequence 1

Synonyms: bHLHa18; LYL1

CAGCCTGGCCCTTATCTGCACTGGGCCAGCATCCTCCGGCCGCTGCGCCGCCAGGGGTGAGAGGGAGGAAA
CCGGGCCGCCGGGGGCGGGGAGAAGGCGGGCCGGCCCGGGAGCCGCTCACTTTCCCTGGGGGGGACCTACG
CGGAGACCTCGGCTATCCTGGCCTTCCGAGGCCCACGAGGAGGCGCGGCCCAACGCCGGGGCCTGGAGCAT
TGAGGCCGGACCCTCGCGAGACAGCAGAGCCTGGCCTGACGCTGGAAACCACACCCTGGCCCAGACTGCCA
GCCCTGACGGGACAGAGCCAGGGCACTCACCAGGCTGCAAGAACAGTGCTGGGGTGAGTACCCCCACGTCG
GGGTCCATGTGCCCGCCTCAGGCACAGGCAGAGGTGGGCCCCACCATGACTGAGAAGGCAGAGATGGTGTG
TGCCCCCAGCCCAGCGCCTGCCCCACCCCCTAAGCCTGCCTCGCCTGGGCCCCCGCAGGTGGAGGAGGTGG
GCCACCGAGGAGGCTCCTCGCCCCCCAGGCTGCCACCTGGTGTACCAGTGATCAGCCTGGGCCACAGCAGG
CCCCCAGGGGTAGCCATGCCCACCACAGAGCTGGGCACTCTGCGGCCCCCGCTGCTGCAACTCTCCACCCT
GGGAACTGCCCCGCCCACTTTGGCCCTGCACTACCACCCTCACCCCTTCCTCAACAGTGTCTACATTGGGC
CAGCAGGACCTTTTAGCATCTTCCCTAGCAGCCGGTTGAAGCGGAGACCAAGCCACTGTGAGCTGGACCTG
GCTGAGGGGCACCAGCCCCAGAAGGTGGCCCGGCGCGTGTTCACCAACAGCCGGGAGCGCTGGCGGCAGCA
GAACGTTAACGGCGCCTTCGCCGAGCTGAGGAAGCTGCTGCCGACGCACCCGCCCGACCGGAAGCTGAGCA
AGAACGAGGTGCTCCGCCTAGCCATGAAGTACATCGGCTTCCTGGTGCGGCTGCTGCGCGACCAAGCCGCA
GCTCTGGCCGCAGGCCCCACCCCTCCCGGGCCTCGCAAACGGCCGGTGCACCGGGTCCCAGACGACGGCGC
CCGCCGGGGATCCGGACGCAGGGCCGAGGCGGCAGCGCGCTCGCAGCCCGCGCCCCCGGCCGACCCCGACG
GCAGCCCCGGTGGAGCGGCCCGGCCCATCAAGATGGAGCAAACCGCTTTGAGCCCAGAGGTGCGGTGACCG
CACGCGGCAGCACCTCTGAGCCGGAGGGCACCAGGGACTCGGCCCAGGGCCGTCAAGGAAAGGGCAGTGGA
CGTGCTGCGCATGTTCGGGAGCGAACTCCCCCGAAGAAGGACCAGTGAAGACGTCAGGGGCAAGGTCTCGG
GGGTCCGGAAGGGTGATCATCGACCCCCAAGGGACCCGCAGACCCTTAAAAAAATCACCCACAACCCTCTG
GAAGTGGCCTTGCCCGGTCCCCTTCCCAGGGGCGAGGTCGGCAAAGCAACATGGCAGAGCAGTCATAGGAA
AAAAAAAAAAAAAAAAAA

LAQTASPDGTEPGHSPGCKNSAGVSTPTSGSMCPPQAQAEVGPTMTEKAEMVCAPSPAPAPPPKPASPGPP
QVEEVGHRGGSSPPRLPPGVPVISLGHSRPPGVAMPTTELGTLRPPLLQLSTLGTAPPTLALHYHPHPFLN
SVYIGPAGPFSIFPSSRLKRRPSHCELDLAEGHQPQKVARRVFTNSRERWRQQNVNGAFAELRKLLPTHPP
DRKLSKNEVLRLAMKYIGFLVRLLRDQAAALAAGPTPPGPRKRPVHRVPDDGARRGSGRRAEAAARSQPAP
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

GGCTCCCCACTCTGCCAGAGCGAGGCGGGGCAGTGAGGACTCCGCGACGCGTCCGCACCCTGCGGCCAGAG
CGGCTTTGAGCTCGGCTGCGTCCGCGCTAGGCGCTTTTTCCCAGAAGCAATCCAGGCGCGCCCGCTGGTTC
TTGAGCGCCAGGAAAAGCCCGGAGCTAACGACCGGCCGCTCGGCCACTGCACGGGGCCCCAAGCCGCAGAA
GGACGACGGGAGGGTAATGAAGCTGAGCCCAGGTCTCCTAGGAAGGAGAGAGTGCGCCGGAGCAGCGTGGG
AAAGAAGGGAAGAGTGTCGTTAAGTTTACGGCCAACGGTGGATTATCCGGGCCGCTGCGCGTCTGGGGGCT
GCGGAATGCGCGAGGAGAACAAGGGCATGCCCAGTGGGGGCGGCAGCGATGAGGGTCTGGCCAGCGCCGCG
GCGCGGGGACTAGTGGAGAAGGTGCGACAGCTCCTGGAAGCCGGCGCGGATCCCAACGGAGTCAACCGTTT
CGGGAGGCGCGCGATCCAGGTCATGATGATGGGCAGCGCCCGCGTGGCGGAGCTGCTGCTGCTCCACGGCG
CGGAGCCCAACTGCGCAGACCCTGCCACTCTCACCCGACCGGTGCATGATGCTGCCCGGGAGGGCTTCCTG
GACACGCTGGTGGTGCTGCACCGGGCCGGGGCGCGGCTGGACGTGCGCGATGCCTGGGGTCGTCTGCCCGT
GGACTTGGCCGAGGAGCGGGGCCACCGCGACGTTGCAGGGTACCTGCGCACAGCCACGGGGGACTGACGCC
AGGTTCCCCAGCCGCCCACAACGACTTTATTTTCTTACCCAATTTCCCACCCCCACCCACCTAATTCGATG
AAGGCTGCCAACGGGGAGCGGCGGAAAGCCTGTAAGCCTGCAAGCCTGTCTGAGACTCACAGGAAGGAGGA

GCCGACCGGGAATAACCTTCCATACATTTTTTTCTTTGTCTTATCTGGCCCTCGACACTCACCATGAAGCG
AAACACAGAGAAGCGGATTTCCAGGGATATTTAGGAGTGTGTGACATTCCAGGGGTCGTTTGCTTTTCAGG
GTTTTCTGAGGGAAAGTGCATATGAAATCCTTGACTGGACCTGGTGGCTACGAATCTTCCGATGGATGAAT
CTCCCACTCCAGCGCTGAGTGGGAGAAGGCAGTGATTAGCACTTGGGTGACGGCAGTCGATGCGTTCACTC
CAATGTCTGCTGAGGAGTTATGGTGAACCCACAACTTAGGCCCTAGCGGCAGAAAGGAAAACCTGAAGACT
GAGGACAAAGTGGAGGAGGGCCGAGGTGGGCTTCAGTAAGTCCCCGGCGGCGCTTTAGTTTGAGCGCATGG
CAAGTCACATGCGTAAACGACACTCTCTGGAAGCCCTGGAGACCCTCGCCCAACTCCACCAGATAGCAGAG
GGGTAAGAGAGGATGTGCAAGCGACGACAGATGCTAAAATCCCTGGATCACGACGCTGCAGAGCACCTTTG
CACAGGATGCTGGCCTTTGCTCTTACTACACTGAGGAGAGATTCCCGCGGGTTCCGCAGGCAGACTACACA
GGATGAGGTGGTGGAGTGGAGTGAGAGCAATTGTAACGGTTAACTGTAACGTTTTCTTTCACACACACACA
CACACACACACACACACACATGCTAGGATGCGGAAATCCCCTTATGACTTGCTACTTTTTGATTTTGTGATAT
TTTGTACTTTTTAGTTGTTCAGCAACTGTCTTATTTAATGGGGAGATTTTAAGTAACATAACTAGTGGCTC
TCAGTTAAAATGTGAGGAAGAACTACAGCTCTTAAATGTAGCAATGGCACTGTTGCAAACTCAGTGCAAAC
GCCTAGATTGCTTTCTTCTTAACCTATTTATTTCTTTGTTAAATTTTTCTGATTGTTTCCTTTATAGAGTG
TCTCAGGGTGCAGAGGTCAGACTAAGAAATATTCCAAATGTCTTTTAGAAGATAGATGCACTTATGCAGTA
AATTATCTTGGGATAGTTCCCAAAAGATTGCTGAAAAAGTAGATTGAGTATAAAAACTTGAAAATATATGA
TGGCTCGTGGGATGTCCTACTATCACTGAACAAACTAAAGGTGCACTGCTTTGGGATTTAATTTCCAGGGT
TGCTTGATCATTATATCATTGGAACAACTGATACTTCACTACTTTAATAAAGAATTAACAGAGATTGAACT
CCAAGAGGTGGGTAATTTGGTTTAAAAATACATGTTCATGGGTTTACCACTAACTCCTGAGAAATGTTAAA
GGTTCACAGGGGTTCCCTTCTCTCAATGTTTGTAATAATTGCTCATAAGCAATACCAGCAATTCATAAAAA
CTGCTTACTTATGCCATAGAAAATTAAACACAAAGTGTATACATGTATTATGCTTCTAAATGCTCATTCTA
CCAGATACACATTTAAAAGAGAAAAAAGGAACAGAAACAAGTCATTTGAGAGTGGAGACTTATAAGAAGGA
GTACATTTGAGTTGAATACACAAATCTTTACTTCTCTACCAATTCCTATTCCCAAAATGAACATATTACTG
GGGAAAGTTAGTTGAGAATCAGAGCATATGTTATTGGGGAAAGGATATGTTTATTGACACATAATCTGTAC
CAGGTATGCATTAAAATATATTTGTTAATTTAATATTTAAACCTGAGAGATAGGTATTGTTTCCCAGATGA
GGACAATGAGGCAAAGAAATATCAAGTAACTTGCCAAAGGTTACAAGATATTCATTCCATGGATGCACAAA
GAAGTGCATCTAGTTCCACAGCTGATTATGGTTGTCTTGCTTTTCTTCCCATTGCACCAGCTTGTCCTCCA
AAATCATGAATGATACACATGAAGATAACTTTTTTTAAAAAAAAGCAGAAATACACAATGATCTCCCTTGT
AAGCTCCTAAGGTGGCTTTTCTTTCTCTAACTTCTAGTAAATATAAACGGTTTGTTTGAAAACTATTTTAA
AATGTCAACAATATGGAGAATAACCCCCCCCAACACACCTATAAAAACCCAAATTTTTGGAACAAAGATAA
TGGAACCTCCATTTTCAAACTGAAGCACAGGGACAGAAAATATATTTCTAGTTATCACTTAAGCACTCAAT
CATTAGAGGCTACAAGAATAATATTTTTAAAGTTACAGTATTTTACAATTATTAGAAAACATTCTATATAA
AAGAAGTCAGTTGATACTTTAAAATCTCCCATTTGGTTTATAAAATCCCTTAATTTGACCTCTATATCTTA
AATTCCAAGATGTTTAAATTTGCTAGTTGCATTATACTGGGTCATGAAAAATTATCCCTTGAAATAGATAT
GAAACATGTTACTTCATTTCTGGTTTAAATAACTTGTGGAATCTTTCCTAATGACAACCTGATATTAAGGG
AAACTAAAGAAAATGTTATTGTGGATCCCACAGTACTATATTACACTGTTTTTTTTGTTTGTTTTGTTAGT
TTTTTTTATTTAAAGCAAACCTCAAACATTATTGGGTATCAATTACCACCTGGTTGTATTGAAATAGTAAC
TTATCAATGCCATGTAAAAATTAATTCCATTTTCGAAGCCACCTGGCAGACAGGTTTAGCTGTTTCATCAG
CAGCCTAATATATACTGTTAAATTTGTTAAGGATTTCACTTTGAAGGATACATGCAAAACATATAGTTACT
ATTTTCATGAGTCCTGCTTCTAGCTCCATTGTGGAATACAGAAAATTAAATATACCTGTTAAGTTCGTATC
TAAACCTAAGACATTACCAAGGTTTGTACAAATTCTACTACCTGACATTTATTCCAAGAAGATCTGGAAAG
TTAAATAAATTTATAAATTTAATAACAAAAAAAAAAAAAAAAAA

TVDYPGRCASGGCGMREENKGMPSGGGSDEGLASAAARGLVEKVRQLLEAGADPNGVNRFGRRAIQVMMMG
SARVAELLLLHGAEPNCADPATLTRPVHDAAREGFLDTLVVLHRAGARLDVRDAWGRLPVDLAEERGHRDV
AGYLRTATGD

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

#11

NM_001085386

Apparently conserved in primates only.

NFE4 transcription factor NF-E4

Synonyms: NF-E4

GACAATTCCTGTTTACGGAAGACTATAAAACCCTGCCCCCTACTCATTTGGTGCTGACGCCATTTTAGGCC
TTAGCCTGCCTGCACCCAGGCGCTCATTAAAACAGCAGGTTGCTCCACACTGCCTCGTGTTGTCTGTTGGC
ACACTCTCAAGAGTTTGAACGGATACAAGAATCTTTCATCTGGTGCCGAAACCCGGGAGGGGCTCCGGTCT
TCGTCCCCCGTGGACCTACCCCTCCGCCCCAGAAAGCAGGCCACAGCAGCCGGACAAAGGAAGCTCCTCAG
CCTCCAGTTGCTTCTCTGTGCATGCACATCAGTCACTGATCTCACCTACTGGGGCCCTGCAGGCCATGGGG
CCACAGCTCCACACAGAAGCCTCCTAGCAATCCACCTCCACCTGGTGCCTGCTTCAAGTGCGGCAATGAAG
GCCACTGGCCCACACAATGCCCAAACCCAGGTAAACCCACGAGGCCATGCCCCCTCTGCGGAGGACCCCAC
TGGAAGTTGGACTGTGAGCGGCCCCTGCAAGGACCACCCCCATCCCTTCCTGAGCCAATCAAACCCTCCTA
CTCGGATCTCGTCAGCCTTGCCGCTGAAGACTGATAGTGCCTTGGAACAGACACCCCAGCAACTACCATCG
CTTCATCTGAGCCAAGGGTAACCCTGATGGTGGCAGGCCCAGTCCCAAGGCGCAAGGCCACTTGTGCCAGC
AGTGTGAGTCAGCAAGATAGCAGAAGCAGGAAGAGAGCCGGCCAGAAGACACCTACTCTGACTGGGAGACA
CGTACCCCTGAAGATTGAGAAAGAGGCCATCCAGGTACCACATAGCAGTTACATCAGACTGGGACATTTCC
TGTTTACAGGAGACTATAAAACCCCTGCCCCCTACTCATTTGGTGCTGACGCCATTTTAGGCCTCAGCCTG
CCTGCACCCAGGTACTCATTAAAACAGCGTGTTGCTCCAAAAA

LPRVVCWHTLKSLNGYKNLSSGAETREGLRSSSPVDLPLRPRKQATAAGQRKLLSLQLLLCACTSVTDLTY
WGPAGHGATAPHRSLLAIHLHLVPASSAAMKATGPHNAQTQVNPRGHAPSAEDPTGSWTVSGPCKDHPHPF
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*, *Dasypus 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

CGGAGGTGCGCGGGCGCGGGCGAGCAGGGTCTCCGGGTGGGCGGCGGCGACGCCCCGCGCAGGCTGGAGGC
CGCCGAGGCTCGCCATGCCGGGAGAACTCTAACTCCCCCATGGAGTCGGCCGACTTCTACGAGGCGGAGCC
GCGGCCCCCGATGAGCAGCCACCTGCAGAGCCCCCCGCACGCGCCCAGCAGCGCCGCCTTCGGCTTTCCCC
GGGGCGCGGGCCCCGCGCAGCCTCCCGCCCCACCTGCCGCCCCGGAGCCGCTGGGCGGCATCTGCGAGCAC
GAGACGTCCATCGACATCAGCGCCTACATCGACCCGGCCGCCTTCAACGACGAGTTCCTGGCCGACCTGTT
CCAGCACAGCCGGCAGCAGGAGAAGGCCAAGGCGGCCGTGGGCCCCACGGGCGGCGGCGGCGGCGGCGACT
TTGACTACCCGGGCGCGCCCGCGGGCCCCGGCGGCGCCGTCATGCCCGGGGGAGCGCACGGGCCCCCGCCC
GGCTACGGCTGCGCGGCCGCCGGCTACCTGGACGGCAGGCTGGAGCCCCTGTACGAGCGCGTCGGGGCGCC
GGCGCTGCGGCCGCTGGTGATCAAGCAGGAGCCCCGCGAGGAGGATGAAGCCAAGCAGCTGGCGCTGGCCG
GCCTCTTCCCTTACCAGCCGCCGCCGCCGCCGCCGCCCTCGCACCCGCACCCGCACCCGCCGCCCGCGCAC
CTGGCCGCCCCGCACCTGCAGTTCCAGATCGCGCACTGCGGCCAGACCACCATGCACCTGCAGCCCGGTCA
CCCCACGCCGCCGCCCACGCCCGTGCCCAGCCCGCACCCCGCGCCCGCGCTCGGTGCCGCCGGCCTGCCGG
GCCCTGGCAGCGCGCTCAAGGGGCTGGGCGCCGCGCACCCCGACCTCCGCGCGAGTGGCGGCAGCGGCGCG
GGCAAGGCCAAGAAGTCGGTGGACAAGAACAGCAACGAGTACCGGGTGCGGCGCGAGCGCAACAACATCGC
GGTGCGCAAGAGCCGCGACAAGGCCAAGCAGCGCAACGTGGAGACGCAGCAGAAGGTGCTGGAGCTGACCA
GTGACAATGACCGCCTGCGCAAGCGGGTGGAACAGCTGAGCCGCGAACTGGACACGCTGCGGGGCATCTTC
CGCCAGCTGCCAGAGAGCTCCTTGGTCAAGGCCATGGGCAACTGCGCGTGAGGCGCGCGGCTGTGGGACCG
CCCTGGGCCAGCCTCCGGCGGGGACCCAGGGAGTGGTTTGGGGTCGCCGGATCTCGAGGCTTGCCCGAGCC
GTGCGAGCCAGGACTAGGAGATTCCGGTGCCTCCTGAAAGCCTGGCCTGCTCCGCGTGTCCCCTCCCTTCC

TCTGCGCCGGACTTGGTGCGTCTAAGATGAGGGGGCCAGGCGGTGGCTTCTCCCTGCGAGGAGGGGAGAAT
TCTTGGGGCTGAGCTGGGAGCCCGGCAACTCTAGTATTTAGGATAACCTTGTGCCTTGGAAATGCAAACTC
ACCGCTCCAATGCCTACTGAGTAGGGGGAGCAAATCGTGCCTTGTCATTTTATTTGGAGGTTTCCTGCCTC
CTTCCCGAGGCTACAGCAGACCCCCATGAGAGAAGGAGGGGAGCAGGCCCGTGGCAGGAGGAGGGCTCAGG
GAGCTGAGATCCCGACAAGCCCGCCAGCCCCAGCCGCTCCTCCACGCCTGTCCTTAGAAAGGGGTGGAAAC
ATAGGGACTTGGGGCTTGGAACCTAAGGTTGTTCCCCTAGTTCTACATGAAGGTGGAGGGTCTCTAGTTCC
ACGCCTCTCCCACCTCCCTCCGCACACACCCCACCCCAGCCTGCTATAGGCTGGGCTTCCCCTTGGGGCGG
AACTCACTGCGATGGGGGTCACCAGGTGACCAGTGGGAGCCCCCACCCCGAGTCACACCAGAAAGCTAGGT
CGTGGGTCAGCTCTGAGGATGTATACCCCTGGTGGGAGAGGGAGACCTAGAGATCTGGCTGTGGGGCGGGC
ATGGGGGGTGAAGGGCCACTGGGACCCTCAGCCTTGTTTGTACTGTATGCCTTCAGCATTGCCTAGGAACA
CGAAGCACGATCAGTCCATCCCAGAGGGACCGGAGTTATGACAAGCTTTCCAAATATTTTGCTTTATCAGC
CGATATCAACACTTGTATCTGGCCTCTGTGCCCCAGCAGTGCCTTGTGCAATGTGAATGTGCGCGTCTCTG
CTAAACCACCATTTTATTTGGTTTTTGTTTTGTTTTGGTTTTGCTCGGATACTTGCCAAAATGAGACTCTC
CGTCGGCAGCTGGGGGAAGGGTCTGAGACTCCCTTTCCTTTTGGTTTTGGGATTACTTTTGATCCTGGGGG
ACCAATGAGGTGAGGGGGGTTCTCCTTTGCCCTCAGCTTTCCCCAGCCCCTCCGGCCTGGGCTGCCCACAA
GGCTTGTCCCCCAGAGGCCCTGGCTCCTGGTCGGGAAGGGAGGTGGCCTCCCGCCAACGCATCACTGGGGC
TGGGAGCAGGGAAGGACGGCTTGGTTCTCTTCTTTTGGGGAGAACGTAGAGTCTCACTCTAGATGTTTTAT
GTATTATATCTATAATATAAACATATCAAAGTCAA

VRGRGRAGSPGGRRRRPAQAGGRRGSPCRENSNSPMESADFYEAEPRPPMSSHLQSPPHAPSSAAFGFPRG
AGPAQPPAPPAAPEPLGGICEHETSIDISAYIDPAAFNDEFLADLFQHSRQQEKAKAAVGPTGGGGGGDFD
YPGAPAGPGGAVMPGGAHGPPPGYGCAAAGYLDGRLEPLYERVGAPALRPLVIKQEPREEDEAKQLALAGL
FPYQPPPPPPPSHPHPHPPPAHLAAPHLQFQIAHCGQTTMHLQPGHPTPPPTPVPSPHPAPALGAAGLPGP
GSALKGLGAAHPDLRASGGSGAGKAKKSVDKNSNEYRVRRERNNIAVRKSRDKAKQRNVETQQKVLELTSD
NDRLRKRVEQLSRELDTLRGIFRQLPESSLVKAMGNCA