COMPUTATIONAL ANALYSIS OF UNKNOWN DNA SEQUENCES USING INTEGRATED BIOINFORMATICS TOOLS

CT010212.1 MUS MUSCULUS FULL OPEN READING FRAME CDNA CLONE RZPDO836H0450D FOR GENE APOE, APOLIPOPROTEIN E; COMPLETE CDS, INCL. STOPCODON

DNA ANALYSIS

Mus musculus full open reading frame cDNA clone RZPDo836H0450D for gene A stopcodon

Sequence ID: CT010212.1 Length: 936 Number of Matches: 1

Range 1: 1 to 936 GenBank Graphics

ALIGNMENT

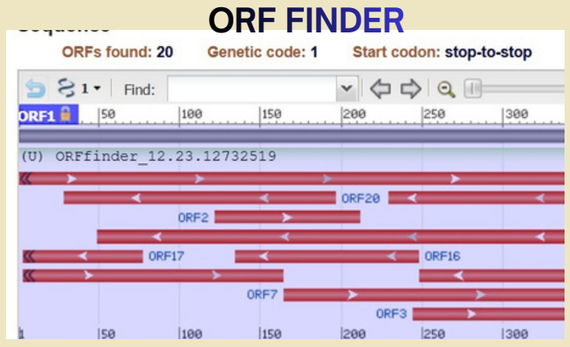
Score	Expect	Identities	Gaps	Strand
1729 bits(936)	0.0	936/936(100%)	0/936(0%)	Plus/Plus

```
Query  1    ATGAAGGCTCTGTGGGCCGTGCTGTTGGTCACATTGCTGACAGGATGCCTAGCCGAGGGA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    ATGAAGGCTCTGTGGGCCGTGCTGTTGGTCACATTGCTGACAGGATGCCTAGCCGAGGGA  60

Query  61   GAGCCGGAGGTGACAGATCAGCTCGAGTGGCAAAGCAACCAACCCTGGGAGCAGGCCCTG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   GAGCCGGAGGTGACAGATCAGCTCGAGTGGCAAAGCAACCAACCCTGGGAGCAGGCCCTG  120

Query  121  AACCGCTTCTGGGATTACCTGCGCTGGGTGCAGACGCTGTCTGACCAGGTCCAGGAAGAG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  AACCGCTTCTGGGATTACCTGCGCTGGGTGCAGACGCTGTCTGACCAGGTCCAGGAAGAG  180

Query  181  CTGCAGAGCTCCCAAGTCACACAAGAACTGACGGCACTGATGGAGGACACTATGACGGAA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  CTGCAGAGCTCCCAAGTCACACAAGAACTGACGGCACTGATGGAGGACACTATGACGGAA  240
```

DISCRIPTION

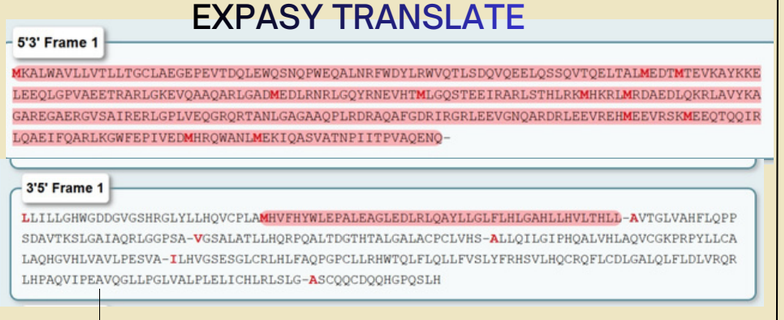
Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Mus musculus	1729	1729	100%	0.0	100.00%	936	CT010212.1
Mus musculus	1714	1714	100%	0.0	99.79%	1111	AK134921.1
Mus musculus	1714	1714	100%	0.0	99.79%	1358	NM_001305843.1
Mus musculus	1714	1714	100%	0.0	99.79%	1101	BC083351.1
Mus musculus	1714	1714	100%	0.0	99.79%	1255	NM_001305819.1
Mus musculus	1714	1714	100%	0.0	99.79%	2101	XM_030242015.2

ORF FINDER



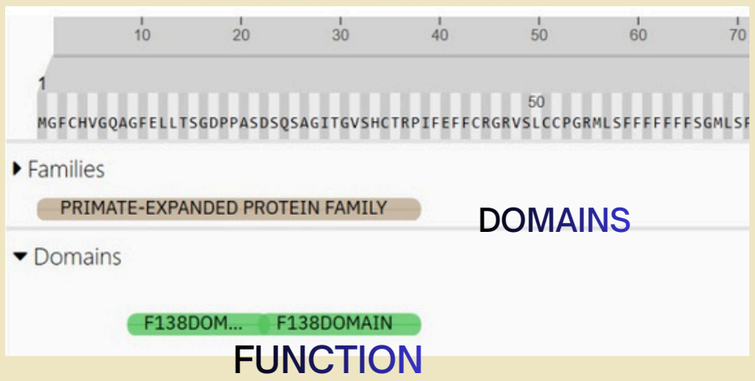
ORF1 was selected because it is the predicted ORFs with a clear start and stop codon. Its minimal overlap with other ORFs makes it a reliable candidate for protein-coding analysis.

EXPASY TRANSLATE



PROTEIN ANALYSIS

NUMBER OF AMINO ACIDS: 311
MOLECULAR WEIGHT: 35847.54
INSTABILITY INDEX (II) :46.67
ALIPHATIC INDEX: 81.93
THEORETICAL PI: 5.53

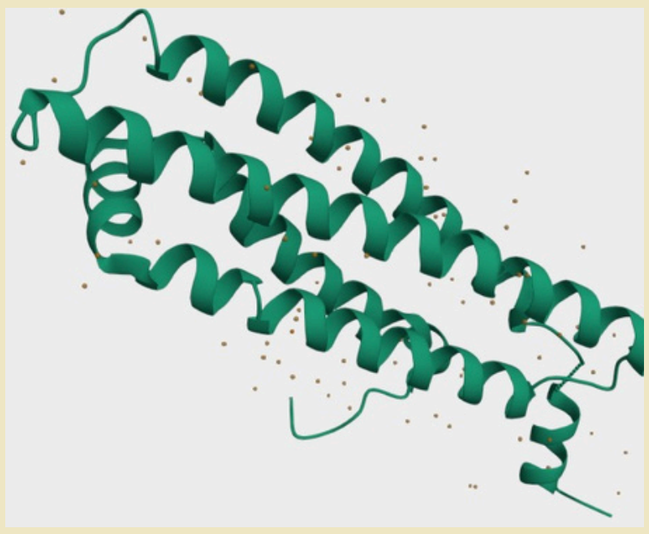


DOMAINS

FUNCTION

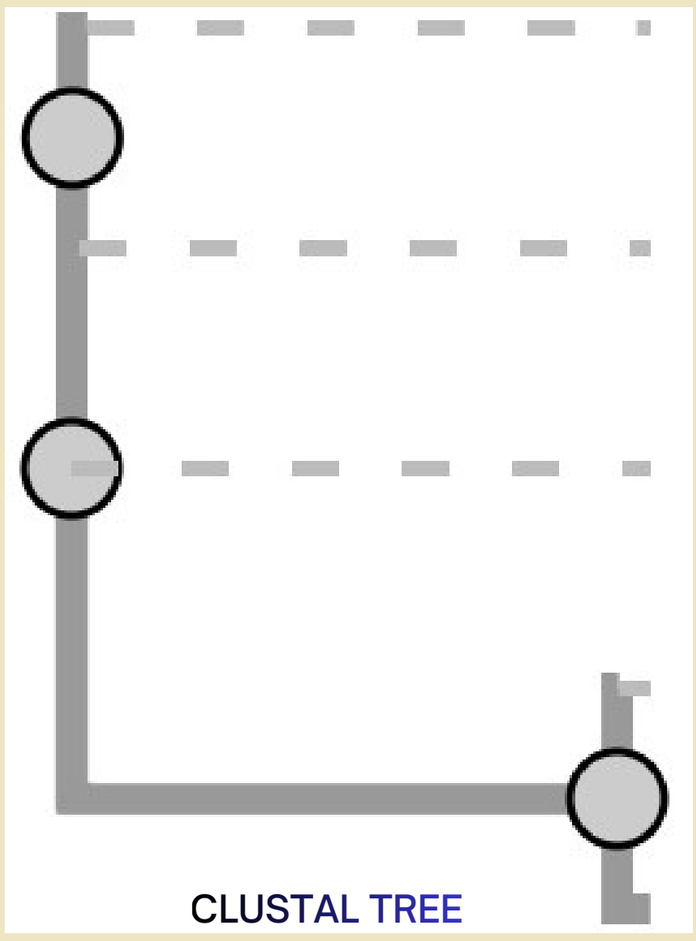
APOE is an apolipoprotein, a protein associating with lipid particles that mainly functions in lipoprotein mediated lipid transport between organs via the plasma and interstitial fluids. APOE is a core component of plasma lipoproteins and is involved in their production, conversion and clearance

SECONDARY STRUCTURE



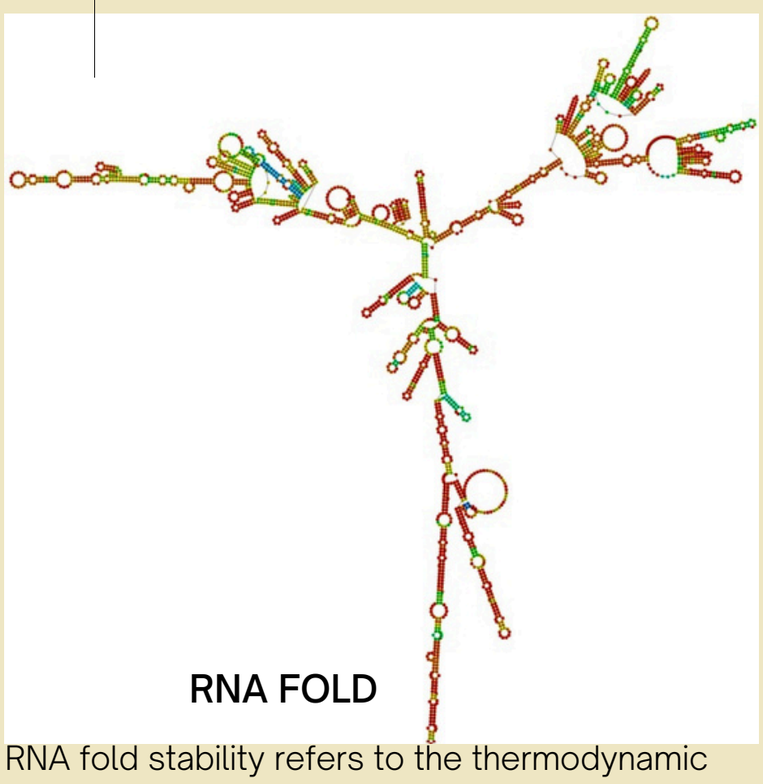
PHYLOGENETIC ANALYSIS

```
CLUSTAL O(1.2.4) multiple sequence alignment


GCTCAGACCCCGGAGGCTAAGGAGTTGTTTCGGAAGGAGCTGACTGGCCAATCACAACTGGGAAGATGAA
GGCTCTGTGGGCCCTGCTGTTGGTCCCATTGCTGACAGGATGCCTGG 47
GCAGGAAGCCCGATAGAAGACTTAGGGTGGCGGGGAGACAAACTAAGATCGTGAGACTGGCCAATCACAA
CTGGGAAGATGAAGGCTCTGTGGGCCCTGCTGTTGGTCCCATTGCTGACAGGATGCCTGG     60
GCTCAGACCCCGGAGGCTAAGGAGTTGCTTCGGAAGGAGCTGACTGGCCAATCACAACTGGGAAGATGAA
GGCTCTGTGGGCCCTGCTGTTGGTCCCATTGCTGACAGGATGCCTGG 47
ATGAAGGCTCTGTGGGCCGTGCTGTTGGTCACATTGCTGACAGGATGCCTAGCCGAGGGAGAGCCGGAGG
---------------------------------------------   0
GGCTCAGACCCTGGAGGCTAAGGACTTGTTTCGGAAGGAGCTGACTGGCCAATCACAATTGCGAAGATGA
AGGCTCTGTGGGCCGTGCTGTTGGTCACATTGCTGACAGGATGCCTAG         48
```



CLUSTAL TREE

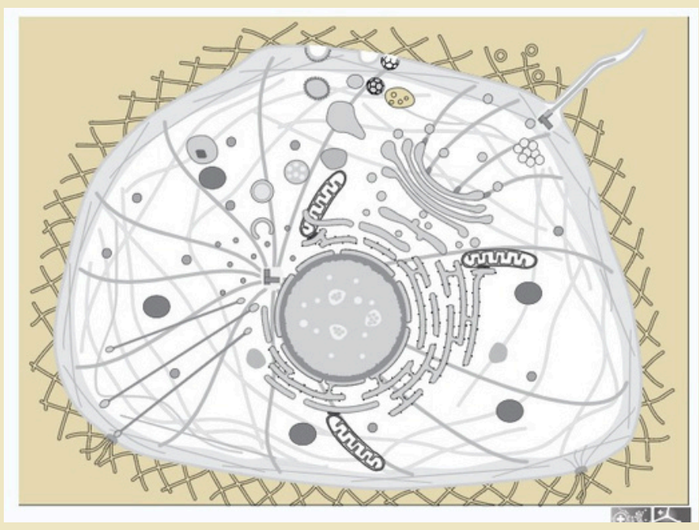
RNA ANALYSIS



RNA FOLD

RNA fold stability refers to the thermodynamic resistance of an RNA molecule to unfolding or degradation, dictated by internal structural features like G-C content and base stacking. It is primarily determined by hydrogen bonding between complementary bases and hydrophobic interactions within the sugar-phosphate backbone. Higher fold stability typically protects RNA from nucleolytic degradation, as seen in the stem-loops of 3' UTRs or ribosomal RNAs. Conversely, low fold stability is crucial for regulatory RNAs, such as noncoding RNAs or specific mRNAs, allowing them to rapidly respond to environmental stimuli or undergo degradation. Stability is highly sensitive to external factors like temperature, pH, and the presence of divalent cations like Mg2+.

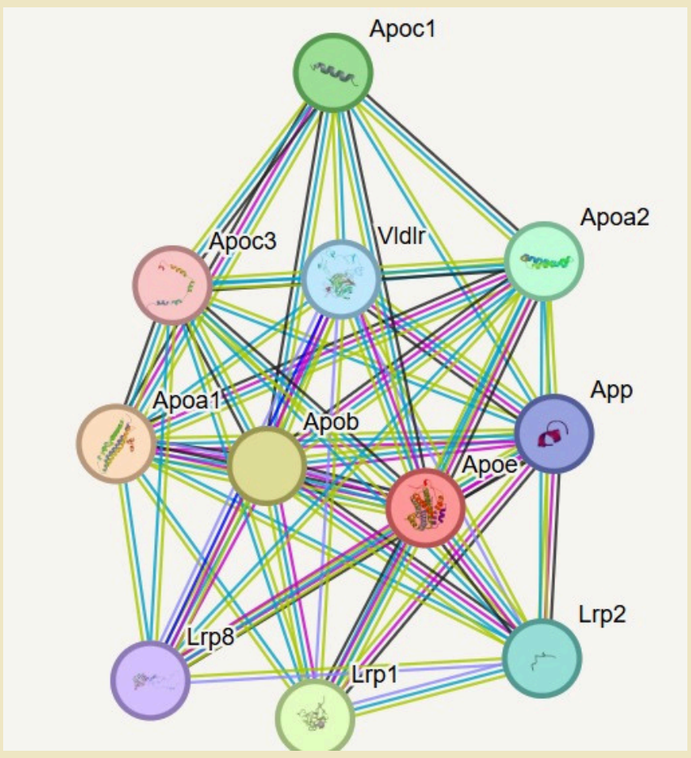
SUBCELLULAR LOCALIZATION



LOCALIZATION	RELIABILITY
Extracellular	0.288
PlasmaMembrane	0.937
PlasmaMembrane	0.840
PlasmaMembrane	0.798
Extracellular	0.522

Apolipoprotein E (ApoE) has a complex localization, found both extracellularly (plasma, brain parenchyma/vessels) and intracellularly in various compartments like the endoplasmic reticulum (ER), Golgi, endosomes, lysosomes, and even the nucleus and mitochondria (especially MAMs), acting in lipid transport but also having diverse cellular roles, including signaling and mitochondrial function. Its location shifts depending on cell type (astrocytes, neurons) and disease state (Alzheimer's), often near amyloid-beta (Aβ) deposits.

PROTEIN PROTEIN INTERACTION



Apolipoprotein E (ApoE) serves as a critical bridge between lipoprotein metabolism and cellular signaling, functioning as a key ligand for receptors (LDL-R, LRP1) to clear cholesterol-rich particles. Its interactions with other apolipoproteins, such as ApoC-II, ApoC-III, and ApoA-I, are vital for regulating triglyceride lipolysis, remnant clearance, and HDL metabolism.
Opposing Action to ApoC-III
Displacement of ApoC-II
Interaction with ApoA-I in HDL Metabolism
ApoE4-Specific Interactions
Signaling and Inflammation