**Cancer gene Research Literatures**

Mark Herbert

World Development Institute

39-06 Main Street, Flushing, Queens, New York 11354, USA, ma708090@gmail.com

**Abstract**: Cancer is the general name for a group of more than 100 diseases. Although there are many kinds of cancer, all cancers start because abnormal cells grow out of control. Untreated cancers can cause serious illness and death. The body is made up of trillions of living cells. Normal body cells grow, divide, and die in an orderly fashion. During the early years of a person’s life, normal cells divide faster to allow the person to grow. After the person becomes an adult, most cells divide only to replace worn-out or dying cells or to repair injuries. This article introduces recent research reports as references in the related studies.

**[**Mark Herbert. **Cancer gene Research Literatures.** *Cancer Biology* 2020;10(4):48-120]. ISSN: 2150-1041 (print); ISSN: 2150-105X (online). <http://www.cancerbio.net>. 5. doi:[10.7537/marscbj100420.05](http://www.dx.doi.org/10.7537/marscbj100420.05).

**Key words**: cancer; life; research; literature; cell; gene

**1. Introduction**

Cancer is the general name for a group of more than 100 diseases. Although there are many kinds of cancer, all cancers start because abnormal cells grow out of control. Untreated cancers can cause serious illness and death. The body is made up of trillions of living cells. Normal body cells grow, divide, and die in an orderly fashion. During the early years of a person’s life, normal cells divide faster to allow the person to grow. After the person becomes an adult, most cells divide only to replace worn-out or dying cells or to repair injuries.

The following introduces genes related to cancers as references in the related studies.

# Homo sapiens phosphatase and tensin homolog (PTEN), transcript variant 1, mRNA

NCBI Reference Sequence: NM\_000314.6

LOCUS NM\_000314 8718 bp mRNA linear PRI 18-NOV-2018

DEFINITION Homo sapiens phosphatase and tensin homolog (PTEN), transcript

 variant 1, mRNA.

ACCESSION NM\_000314

VERSION NM\_000314.6

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 8718)

 AUTHORS Li AG, Murphy EC, Culhane AC, Powell E, Wang H, Bronson RT, Von T,

 Giobbie-Hurder A, Gelman RS, Briggs KJ, Piwnica-Worms H, Zhao JJ,

 Kung AL, Kaelin WG Jr and Livingston DM.

 TITLE BRCA1-IRIS promotes human tumor progression through PTEN blockade

 and HIF-1alpha activation

 JOURNAL Proc Natl Acad Sci U S A 115 (41), E9600-E9609 (2018)

 PUBMED [30254159](https://www.ncbi.nlm.nih.gov/pubmed/30254159)

 REMARK GeneRIF: The IRIS-driven metastatic mechanism results from

 IRIS-dependent suppression of phosphatase and tensin homolog (PTEN)

 transcription, which in turn perturbs the PI3K/AKT/GSK-3beta

 pathway leading to prolyl hydroxylase-independent HIF-1alpha

 stabilization and activation in a normoxic environment.

REFERENCE 2 (bases 1 to 8718)

 AUTHORS Jouali F, Marchoudi N, Talbi S, Bilal B, El Khasmi M, Rhaissi H and

 Fekkak J.

 TITLE Detection of PIK3/AKT pathway in Moroccan population with triple

 negative breast cancer

 JOURNAL BMC Cancer 18 (1), 900 (2018)

 PUBMED [30227836](https://www.ncbi.nlm.nih.gov/pubmed/30227836)

 REMARK GeneRIF: In this study, we used the Ion Personal Genome Machine

 (PGM) and Ion Torrent Ampliseq Cancer panel to sequence hotspot

 regions from PIK3CA, AKT and PTEN genes to identify genetic

 mutations in 39 samples of TNBC subtype from Moroccan patients and

 to correlate the results with clinical-pathologic data

 Publication Status: Online-Only

REFERENCE 3 (bases 1 to 8718)

 AUTHORS Li W, Zhang T, Guo L and Huang L.

 TITLE Regulation of PTEN expression by noncoding RNAs

 JOURNAL J Exp Clin Cancer Res 37 (1), 223 (2018)

 PUBMED [30217221](https://www.ncbi.nlm.nih.gov/pubmed/30217221)

 REMARK GeneRIF: we provide a review on current understandings of the

 regulation of PTEN by ncRNAs, which could contribute to the

 development of novel approaches to the diseases with abnormal

 expression of PTEN.

 Review article

 Publication Status: Online-Only

REFERENCE 4 (bases 1 to 8718)

 AUTHORS Xu W, Yang Z, Xie C, Zhu Y, Shu X, Zhang Z, Li N, Chai N, Zhang S,

 Wu K, Nie Y and Lu N.

 TITLE PTEN lipid phosphatase inactivation links the hippo and PI3K/Akt

 pathways to induce gastric tumorigenesis

 JOURNAL J Exp Clin Cancer Res 37 (1), 198 (2018)

 PUBMED [30134988](https://www.ncbi.nlm.nih.gov/pubmed/30134988)

 REMARK GeneRIF: PTEN lipid phosphatase inactivation abolished the

 MOB1-LATS1/2 interaction, decreased YAP phosphorylation and finally

 promoted YAP nuclear translocation, which enhanced the synergistic

 effect of YAP-TEAD, thus inducing cell proliferation and migration.

 Publication Status: Online-Only

REFERENCE 5 (bases 1 to 8718)

 AUTHORS Liang H, He S, Yang J, Jia X, Wang P, Chen X, Zhang Z, Zou X,

 McNutt MA, Shen WH and Yin Y.

 TITLE PTENalpha, a PTEN isoform translated through alternative

 initiation, regulates mitochondrial function and energy metabolism

 JOURNAL Cell Metab 19 (5), 836-848 (2014)

 PUBMED [24768297](https://www.ncbi.nlm.nih.gov/pubmed/24768297)

REFERENCE 6 (bases 1 to 8718)

 AUTHORS Hopkins BD, Fine B, Steinbach N, Dendy M, Rapp Z, Shaw J, Pappas K,

 Yu JS, Hodakoski C, Mense S, Klein J, Pegno S, Sulis ML, Goldstein

 H, Amendolara B, Lei L, Maurer M, Bruce J, Canoll P, Hibshoosh H

 and Parsons R.

 TITLE A secreted PTEN phosphatase that enters cells to alter signaling

 and survival

 JOURNAL Science 341 (6144), 399-402 (2013)

 PUBMED [23744781](https://www.ncbi.nlm.nih.gov/pubmed/23744781)

 REMARK GeneRIF: identified a 576-amino acid translational variant of PTEN,

 PTEN-Long, that arises from an alternative translation start site

 519 base pairs upstream of the ATG initiation sequence; PTEN-Long

 is a membrane-permeable lipid phosphatase that is secreted from

 cells and can enter other cells

REFERENCE 7 (bases 1 to 8718)

 AUTHORS Steck PA, Pershouse MA, Jasser SA, Yung WK, Lin H, Ligon AH,

 Langford LA, Baumgard ML, Hattier T, Davis T, Frye C, Hu R,

 Swedlund B, Teng DH and Tavtigian SV.

 TITLE Identification of a candidate tumour suppressor gene, MMAC1, at

 chromosome 10q23.3 that is mutated in multiple advanced cancers

 JOURNAL Nat Genet 15 (4), 356-362 (1997)

 PUBMED [9090379](https://www.ncbi.nlm.nih.gov/pubmed/9090379)

REFERENCE 8 (bases 1 to 8718)

 AUTHORS Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, Puc J,

 Miliaresis C, Rodgers L, McCombie R, Bigner SH, Giovanella BC,

 Ittmann M, Tycko B, Hibshoosh H, Wigler MH and Parsons R.

 TITLE PTEN, a putative protein tyrosine phosphatase gene mutated in human

 brain, breast, and prostate cancer

 JOURNAL Science 275 (5308), 1943-1947 (1997)

 PUBMED [9072974](https://www.ncbi.nlm.nih.gov/pubmed/9072974)

REFERENCE 9 (bases 1 to 8718)

 AUTHORS Peiffer SL, Herzog TJ, Tribune DJ, Mutch DG, Gersell DJ and

 Goodfellow PJ.

 TITLE Allelic loss of sequences from the long arm of chromosome 10 and

 replication errors in endometrial cancers

 JOURNAL Cancer Res 55 (9), 1922-1926 (1995)

 PUBMED [7728760](https://www.ncbi.nlm.nih.gov/pubmed/7728760)

REFERENCE 10 (bases 1 to 8718)

 AUTHORS Eng,C.

 TITLE PTEN Hamartoma Tumor Syndrome

 JOURNAL (in) Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens

 K and Amemiya A (Eds.);

 GENEREVIEWS((R));

 (1993)

 PUBMED [20301661](https://www.ncbi.nlm.nih.gov/pubmed/20301661)

COMMENT REVIEWED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has been curated by NCBI staff. The

 reference sequence was derived from [U92436.1](https://www.ncbi.nlm.nih.gov/nuccore/U92436.1), [AC063965.8](https://www.ncbi.nlm.nih.gov/nuccore/AC063965.8),

 [BC005821.2](https://www.ncbi.nlm.nih.gov/nuccore/BC005821.2) and [AA836562.1](https://www.ncbi.nlm.nih.gov/nuccore/AA836562.1).

 [WARNING] On Nov 21, 2018 this sequence was replaced by

 [NM\_000314.7](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.7).

 On Mar 25, 2015 this sequence version replaced [NM\_000314.5](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.5).

 Summary: This gene was identified as a tumor suppressor that is

 mutated in a large number of cancers at high frequency. The protein

 encoded by this gene is a phosphatidylinositol-3,4,5-trisphosphate

 3-phosphatase. It contains a tensin like domain as well as a

 catalytic domain similar to that of the dual specificity protein

 tyrosine phosphatases. Unlike most of the protein tyrosine

 phosphatases, this protein preferentially dephosphorylates

 phosphoinositide substrates. It negatively regulates intracellular

 levels of phosphatidylinositol-3,4,5-trisphosphate in cells and

 functions as a tumor suppressor by negatively regulating AKT/PKB

 signaling pathway. The use of a non-canonical (CUG) upstream

 initiation site produces a longer isoform that initiates

 translation with a leucine, and is thought to be preferentially

 associated with the mitochondrial inner membrane. This longer

 isoform may help regulate energy metabolism in the mitochondria. A

 pseudogene of this gene is found on chromosome 9. Alternative

 splicing and the use of multiple translation start codons results

 in multiple transcript variants encoding different isoforms.

 [provided by RefSeq, Feb 2015].

 Transcript Variant: This variant (1) encodes multiple isoforms due

 to the use of alternative translation initiation codons. The

 longest isoform (PTEN-L, PMID:23744781; also known as PTENalpha,

 PMID: 24768297) is derived from the use of an upstream non-AUG

 (CUG) start codon, while two shorter isoforms are derived from

 downstream AUG start codons. The most abundant isoform (PTEN),

 which is derived from the use of the 5'-most AUG start codon, is

 represented in this RefSeq.

 Sequence Note: This RefSeq record was created from transcript and

 genomic sequence data to make the sequence consistent with the

 reference genome assembly. The genomic coordinates used for the

 transcript record were based on transcript alignments.

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

 ##Evidence-Data-START##

 Transcript exon combination :: U92436.1, SRR1660807.191185.1

 [ECO:0000332]

 RNAseq introns :: single sample supports all introns

 SAMEA1965299, SAMEA1966682

 [ECO:0000348]

 ##Evidence-Data-END##

 ##RefSeq-Attributes-START##

 CDS uses downstream in-frame AUG :: experimental evidence

 (PMID:24768297)

 ##RefSeq-Attributes-END##

 COMPLETENESS: complete on the 3' end.

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-128 U92436.1 1-128

 129-129 AC063965.8 185191-185191 c

 130-195 U92436.1 130-195

 196-2253 BC005821.2 1-2058

 2254-8472 AC063965.8 77056-83274 c

 8473-8718 AA836562.1 2-247 c

FEATURES Location/Qualifiers

 source 1..8718

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="10"

 /map="10q23.31"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1&to=8718) 1..8718

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /note="phosphatase and tensin homolog"

 /db\_xref="GeneID:[5728](https://www.ncbi.nlm.nih.gov/gene/5728)"

 /db\_xref="HGNC:[HGNC:9588](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:9588)"

 /db\_xref="MIM:[601728](https://www.ncbi.nlm.nih.gov/omim/601728)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1&to=1110) 1..1110

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=318&to=320) 318..320

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /note="upstream in-frame stop codon"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=513&to=515) 513..515

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /note="alternative non-AUG (CUG) translation initiation

 site used for PTEN-L isoform"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1032&to=2243) 1032..2243

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /EC\_number="[3.1.3.67](https://enzyme.expasy.org/EC/3.1.3.67)"

 /EC\_number="[3.1.3.16](https://enzyme.expasy.org/EC/3.1.3.16)"

 /EC\_number="[3.1.3.48](https://enzyme.expasy.org/EC/3.1.3.48)"

 /note="isoform PTEN is encoded by transcript variant 1;

 phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and

 dual-specificity protein phosphatase PTEN; mitochondrial

 PTENalpha; MMAC1 phosphatase and tensin homolog deleted on

 chromosome 10; mutated in multiple advanced cancers 1;

 phosphatase and tensin-like protein; mitochondrial

 phosphatase and tensin protein alpha; protein tyrosine

 phosphatase"

 /codon\_start=1

 /product="phosphatidylinositol 3,4,5-trisphosphate

 3-phosphatase and dual-specificity protein phosphatase

 PTEN isoform PTEN"

 /protein\_id="[NP\_000305.3](https://www.ncbi.nlm.nih.gov/protein/73765544)"

 /db\_xref="CCDS:[CCDS31238.1](https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?REQUEST=CCDS&DATA=CCDS31238.1)"

 /db\_xref="GeneID:[5728](https://www.ncbi.nlm.nih.gov/gene/5728)"

 /db\_xref="HGNC:[HGNC:9588](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:9588)"

 /db\_xref="MIM:[601728](https://www.ncbi.nlm.nih.gov/omim/601728)"

 /translation="MTAIIKEIVSRNKRRYQEDGFDLDLTYIYPNIIAMGFPAERLEG

 VYRNNIDDVVRFLDSKHKNHYKIYNLCAERHYDTAKFNCRVAQYPFEDHNPPQLELIK

 PFCEDLDQWLSEDDNHVAAIHCKAGKGRTGVMICAYLLHRGKFLKAQEALDFYGEVRT

 RDKKGVTIPSQRRYVYYYSYLLKNHLDYRPVALLFHKMMFETIPMFSGGTCNPQFVVC

 QLKVKIYSSNSGPTRREDKFMYFEFPQPLPVCGDIKVEFFHKQNKMLKKDKMFHFWVN

 TFFIPGPEETSEKVENGSLCDQEIDSICSIERADNDKEYLVLTLTKNDLDKANKDKAN

 RYFSPNFKVKLYFTKTVEEPSNPEASSSTSVTPDVSDNEPDHYRYSDTTDSDPENEPF

 DEDQHTQITKV"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=2&to=2) 1035..1037

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="N-acetylthreonine. {ECO:0000244|PubMed:22814378};

 propagated from UniProtKB/Swiss-Prot (P60484.1); other

 site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=294&to=294) 1911..1913

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphoserine. {ECO:0000250|UniProtKB:O08586};

 propagated from UniProtKB/Swiss-Prot (P60484.1); other

 site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=336&to=336) 2037..2039

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphotyrosine, by FRK.

 {ECO:0000269|PubMed:19345329}; propagated from

 UniProtKB/Swiss-Prot (P60484.1); other site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=338&to=348) 2043..2075

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="propagated from UniProtKB/Swiss-Prot (P60484.1);

 Region: Required for interaction with NOP53.

 {ECO:0000269|PubMed:15355975}"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=366&to=366) 2127..2129

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphothreonine, by GSK3-beta and PLK3.

 {ECO:0000244|PubMed:24275569, ECO:0000269|PubMed:12297295,

 ECO:0000269|PubMed:20940307}; propagated from

 UniProtKB/Swiss-Prot (P60484.1); other site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=370&to=370) 2139..2141

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphoserine, by CK2 and PLK3.

 {ECO:0000269|PubMed:11035045, ECO:0000269|PubMed:12297295,

 ECO:0000269|PubMed:20940307}; propagated from

 UniProtKB/Swiss-Prot (P60484.1); other site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=380&to=380) 2169..2171

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphoserine, by ROCK1 and CK2.

 {ECO:0000269|PubMed:11035045}; propagated from

 UniProtKB/Swiss-Prot (P60484.1); other site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=382&to=382) 2175..2177

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphothreonine, by ROCK1 and CK2.

 {ECO:0000269|PubMed:11035045}; propagated from

 UniProtKB/Swiss-Prot (P60484.1); other site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=383&to=383) 2178..2180

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphothreonine, by ROCK1 and CK2.

 {ECO:0000269|PubMed:11035045}; propagated from

 UniProtKB/Swiss-Prot (P60484.1); other site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=385&to=385) 2184..2186

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphoserine, by CK2.

 {ECO:0000269|PubMed:11035045,

 ECO:0000269|PubMed:12297295}; propagated from

 UniProtKB/Swiss-Prot (P60484.1); other site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=401&to=403) 2232..2240

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="propagated from UniProtKB/Swiss-Prot (P60484.1);

 Region: PDZ domain-binding"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_000305.3?from=401&to=401) 2232..2234

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /experiment="experimental evidence, no additional details

 recorded"

 /note="Phosphothreonine. {ECO:0000269|PubMed:10646847};

 propagated from UniProtKB/Swiss-Prot (P60484.1); other

 site"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1111&to=1195) 1111..1195

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1196&to=1240) 1196..1240

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1241&to=1284) 1241..1284

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1285&to=1523) 1285..1523

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1524&to=1665) 1524..1665

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1666&to=1832) 1666..1832

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=1833&to=2057) 1833..2057

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=2058&to=8701) 2058..8701

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 /inference="alignment:Splign:2.1.0"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=2279&to=2284) 2279..2284

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=2301&to=2301) 2301

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=2505&to=2510) 2505..2510

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=2530&to=2530) 2530

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=3123&to=3128) 3123..3128

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=3147&to=3147) 3147

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=5526&to=5531) 5526..5531

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=5526&to=5526) 5526

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=5546&to=5546) 5546

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=8676&to=8681) 8676..8681

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6?from=8701&to=8701) 8701

 /gene="PTEN"

 /gene\_synonym="10q23del; BZS; CWS1; DEC; GLM2; MHAM;

 MMAC1; PTEN1; PTENbeta; TEP1"

ORIGIN

 1 cctcccctcg cccggcgcgg tcccgtccgc ctctcgctcg cctcccgcct cccctcggtc

 61 ttccgaggcg cccgggctcc cggcgcggcg gcggaggggg cgggcaggcc ggcgggcggt

 121 gatgtggcgg gactctttat gcgctgcggc aggatacgcg ctcggcgctg ggacgcgact

 181 gcgctcagtt ctctcctctc ggaagctgca gccatgatgg aagtttgaga gttgagccgc

 241 tgtgaggcga ggccgggctc aggcgaggga gatgagagac ggcggcggcc gcggcccgga

 301 gcccctctca gcgcctgtga gcagccgcgg gggcagcgcc ctcggggagc cggccggcct

 361 gcggcggcgg cagcggcggc gtttctcgcc tcctcttcgt cttttctaac cgtgcagcct

 421 cttcctcggc ttctcctgaa agggaaggtg gaagccgtgg gctcgggcgg gagccggctg

 481 aggcgcggcg gcggcggcgg cacctcccgc tcctggagcg ggggggagaa gcggcggcgg

 541 cggcggccgc ggcggctgca gctccaggga gggggtctga gtcgcctgtc accatttcca

 601 gggctgggaa cgccggagag ttggtctctc cccttctact gcctccaaca cggcggcggc

 661 ggcggcggca catccaggga cccgggccgg ttttaaacct cccgtccgcc gccgccgcac

 721 cccccgtggc ccgggctccg gaggccgccg gcggaggcag ccgttcggag gattattcgt

 781 cttctcccca ttccgctgcc gccgctgcca ggcctctggc tgctgaggag aagcaggccc

 841 agtcgctgca accatccagc agccgccgca gcagccatta cccggctgcg gtccagagcc

 901 aagcggcggc agagcgaggg gcatcagcta ccgccaagtc cagagccatt tccatcctgc

 961 agaagaagcc ccgccaccag cagcttctgc catctctctc ctcctttttc ttcagccaca

 1021 ggctcccaga catgacagcc atcatcaaag agatcgttag cagaaacaaa aggagatatc

 1081 aagaggatgg attcgactta gacttgacct atatttatcc aaacattatt gctatgggat

 1141 ttcctgcaga aagacttgaa ggcgtataca ggaacaatat tgatgatgta gtaaggtttt

 1201 tggattcaaa gcataaaaac cattacaaga tatacaatct ttgtgctgaa agacattatg

 1261 acaccgccaa atttaattgc agagttgcac aatatccttt tgaagaccat aacccaccac

 1321 agctagaact tatcaaaccc ttttgtgaag atcttgacca atggctaagt gaagatgaca

 1381 atcatgttgc agcaattcac tgtaaagctg gaaagggacg aactggtgta atgatatgtg

 1441 catatttatt acatcggggc aaatttttaa aggcacaaga ggccctagat ttctatgggg

 1501 aagtaaggac cagagacaaa aagggagtaa ctattcccag tcagaggcgc tatgtgtatt

 1561 attatagcta cctgttaaag aatcatctgg attatagacc agtggcactg ttgtttcaca

 1621 agatgatgtt tgaaactatt ccaatgttca gtggcggaac ttgcaatcct cagtttgtgg

 1681 tctgccagct aaaggtgaag atatattcct ccaattcagg acccacacga cgggaagaca

 1741 agttcatgta ctttgagttc cctcagccgt tacctgtgtg tggtgatatc aaagtagagt

 1801 tcttccacaa acagaacaag atgctaaaaa aggacaaaat gtttcacttt tgggtaaata

 1861 cattcttcat accaggacca gaggaaacct cagaaaaagt agaaaatgga agtctatgtg

 1921 atcaagaaat cgatagcatt tgcagtatag agcgtgcaga taatgacaag gaatatctag

 1981 tacttacttt aacaaaaaat gatcttgaca aagcaaataa agacaaagcc aaccgatact

 2041 tttctccaaa ttttaaggtg aagctgtact tcacaaaaac agtagaggag ccgtcaaatc

 2101 cagaggctag cagttcaact tctgtaacac cagatgttag tgacaatgaa cctgatcatt

 2161 atagatattc tgacaccact gactctgatc cagagaatga accttttgat gaagatcagc

 2221 atacacaaat tacaaaagtc tgaatttttt tttatcaaga gggataaaac accatgaaaa

 2281 taaacttgaa taaactgaaa atggaccttt ttttttttaa tggcaatagg acattgtgtc

 2341 agattaccag ttataggaac aattctcttt tcctgaccaa tcttgtttta ccctatacat

 2401 ccacagggtt ttgacacttg ttgtccagtt gaaaaaaggt tgtgtagctg tgtcatgtat

 2461 ataccttttt gtgtcaaaag gacatttaaa attcaattag gattaataaa gatggcactt

 2521 tcccgtttta ttccagtttt ataaaaagtg gagacagact gatgtgtata cgtaggaatt

 2581 ttttcctttt gtgttctgtc accaactgaa gtggctaaag agctttgtga tatactggtt

 2641 cacatcctac ccctttgcac ttgtggcaac agataagttt gcagttggct aagagaggtt

 2701 tccgaagggt tttgctacat tctaatgcat gtattcgggt taggggaatg gagggaatgc

 2761 tcagaaagga aataatttta tgctggactc tggaccatat accatctcca gctatttaca

 2821 cacacctttc tttagcatgc tacagttatt aatctggaca ttcgaggaat tggccgctgt

 2881 cactgcttgt tgtttgcgca ttttttttta aagcatattg gtgctagaaa aggcagctaa

 2941 aggaagtgaa tctgtattgg ggtacaggaa tgaaccttct gcaacatctt aagatccaca

 3001 aatgaaggga tataaaaata atgtcatagg taagaaacac agcaacaatg acttaaccat

 3061 ataaatgtgg aggctatcaa caaagaatgg gcttgaaaca ttataaaaat tgacaatgat

 3121 ttattaaata tgttttctca attgtaacga cttctccatc tcctgtgtaa tcaaggccag

 3181 tgctaaaatt cagatgctgt tagtacctac atcagtcaac aacttacact tattttacta

 3241 gttttcaatc ataatacctg ctgtggatgc ttcatgtgct gcctgcaagc ttcttttttc

 3301 tcattaaata taaaatattt tgtaatgctg cacagaaatt ttcaatttga gattctacag

 3361 taagcgtttt ttttctttga agatttatga tgcacttatt caatagctgt cagccgttcc

 3421 acccttttga ccttacacat tctattacaa tgaattttgc agttttgcac attttttaaa

 3481 tgtcattaac tgttagggaa ttttacttga atactgaata catataatgt ttatattaaa

 3541 aaggacattt gtgttaaaaa ggaaattaga gttgcagtaa actttcaatg ctgcacacaa

 3601 aaaaaagaca tttgattttt cagtagaaat tgtcctacat gtgctttatt gatttgctat

 3661 tgaaagaata gggttttttt tttttttttt tttttttttt ttaaatgtgc agtgttgaat

 3721 catttcttca tagtgctccc ccgagttggg actagggctt caatttcact tcttaaaaaa

 3781 aatcatcata tatttgatat gcccagactg catacgattt taagcggagt acaactacta

 3841 ttgtaaagct aatgtgaaga tattattaaa aaggtttttt tttccagaaa tttggtgtct

 3901 tcaaattata ccttcacctt gacatttgaa tatccagcca ttttgtttct taatggtata

 3961 aaattccatt ttcaataact tattggtgct gaaattgttc actagctgtg gtctgaccta

 4021 gttaatttac aaatacagat tgaataggac ctactagagc agcatttata gagtttgatg

 4081 gcaaatagat taggcagaac ttcatctaaa atattcttag taaataatgt tgacacgttt

 4141 tccatacctt gtcagtttca ttcaacaatt tttaaatttt taacaaagct cttaggattt

 4201 acacatttat atttaaacat tgatatatag agtattgatt gattgctcat aagttaaatt

 4261 ggtaaagtta gagacaacta ttctaacacc tcaccattga aatttatatg ccaccttgtc

 4321 tttcataaaa gctgaaaatt gttacctaaa atgaaaatca acttcatgtt ttgaagatag

 4381 ttataaatat tgttctttgt tacaatttcg ggcaccgcat attaaaacgt aactttattg

 4441 ttccaatatg taacatggag ggccaggtca taaataatga cattataatg ggcttttgca

 4501 ctgttattat ttttcctttg gaatgtgaag gtctgaatga gggttttgat tttgaatgtt

 4561 tcaatgtttt tgagaagcct tgcttacatt ttatggtgta gtcattggaa atggaaaaat

 4621 ggcattatat atattatata tataaatata tattatacat actctcctta ctttatttca

 4681 gttaccatcc ccatagaatt tgacaagaat tgctatgact gaaaggtttt cgagtcctaa

 4741 ttaaaacttt atttatggca gtattcataa ttagcctgaa atgcattctg taggtaatct

 4801 ctgagtttct ggaatatttt cttagacttt ttggatgtgc agcagcttac atgtctgaag

 4861 ttacttgaag gcatcacttt taagaaagct tacagttggg ccctgtacca tcccaagtcc

 4921 tttgtagctc ctcttgaaca tgtttgccat acttttaaaa gggtagttga ataaatagca

 4981 tcaccattct ttgctgtggc acaggttata aacttaagtg gagtttaccg gcagcatcaa

 5041 atgtttcagc tttaaaaaat aaaagtaggg tacaagttta atgtttagtt ctagaaattt

 5101 tgtgcaatat gttcataacg atggctgtgg ttgccacaaa gtgcctcgtt tacctttaaa

 5161 tactgttaat gtgtcatgca tgcagatgga aggggtggaa ctgtgcacta aagtgggggc

 5221 tttaactgta gtatttggca gagttgcctt ctacctgcca gttcaaaagt tcaacctgtt

 5281 ttcatataga atatatatac taaaaaattt cagtctgtta aacagcctta ctctgattca

 5341 gcctcttcag atactcttgt gctgtgcagc agtggctctg tgtgtaaatg ctatgcactg

 5401 aggatacaca aaaataccaa tatgatgtgt acaggataat gcctcatccc aatcagatgt

 5461 ccatttgtta ttgtgtttgt taacaaccct ttatctctta gtgttataaa ctccacttaa

 5521 aactgattaa agtctcattc ttgtcattgt gtgggtgttt tattaaatga gagtttataa

 5581 ttcaaattgc ttaagtccat tgaagtttta attaatgggc agccaaatgt gaatacaaag

 5641 ttttcagttt ttttttttcc tgctgtcctt caaagcctac tgtttaaaaa aaaaaaaaaa

 5701 aaaaaacatg gcctgagagt agagtatctg tctactcatg tttaattaag gaaaaacact

 5761 tatttttagg gctttagtca tcacttcata aattgtataa gcacattaaa tagcgttcta

 5821 gtcctgaaaa agtccaagat tcttagaaaa ttgtgcatat ttttattatg acagatgttt

 5881 gaagataatt ccccagaatg gatttgatac tttagatttc aattttgtgg cttttgtcta

 5941 ttattctgta ctctgccatc agcatatgga aagcttcatt tactcatcat gacttgtgcc

 6001 atataaaaat tgatatttcg gaatagtcta aaggactttt tgtacttgaa tttaatcatg

 6061 ttgtttctaa tattcttaaa agcttgaaga ctaaagcata tcctttcaac aaagcatagt

 6121 aaggtaataa gaaagtgtag tttgtacaag tgttaaaaaa ataaagtaga caatgttaca

 6181 gtgggactta ttatttcaag tttacatttt ctccatgtaa ttttttaaaa agtaaatgaa

 6241 aaaatgtgca ataatgtaaa atatgaagtg tatgtgtaca cacattttat ttttcggtat

 6301 cttgggtata cgtatggttg aaaactatac tggagtctaa aagtattcta atttataaga

 6361 agacattttg gtgatgtttg aaaaatagaa atgtgctagt tttgttttta tatcatgtcc

 6421 tttgtacgtt gtaatatgag ctggcttggt tcagtaaatg ccatcaccat ttccattgag

 6481 aatttaaaac tcaccagtgt ttaatatgca ggcttccaaa ggcttatgaa aaaaatcaag

 6541 acccttaaat ctagttaatt tgctgctaac atgaaactct ttggttcttt tatttttgcc

 6601 agataattag acacacatct aaagcttagt cttaaatggc ttaagtgtag ctattgatta

 6661 gtgctgttgc tagttcagaa agaaatgttt gtgaatggaa acaagaatat tcagtccaaa

 6721 ctgttgtaag gacagtacct gaaaaccagg aaacaggata atggaaaaag tcttttaaag

 6781 atgaaatgtt ggagccaact ttcttataga attaattgta tgtggctata gaaagcctaa

 6841 tgattgttgc ttatttttga gagcatatta ttcttttatg accataatct tgctgttttt

 6901 ccatcttcca aaagatcttc cttctaatat gtatatcaga atgtgggtag ccagtcagac

 6961 aaattcatat tggttggtag ctttaaaaag tttgtaatgt gaagacagga aaggacaaaa

 7021 tagtttgctt tggtggtagt actctggttg ttaagctagg tattttgaga ctacttcccc

 7081 atcacaacaa caataaaata atcactcata atcctatcac ctggagacat agccatcgtt

 7141 aatatgttag tgactataca atcatgtttt cttctgtata tccatgtata ttctttaaaa

 7201 atgaaattta tactgtacct gatctcaaag ctttttagct tagtatatct gtcatgaatt

 7261 tgtaggatgt tccattgcat cagaaaacgg acagtgattt gattactttc taatgccaca

 7321 gatgcagatt acatgtagtt attgagaatc ctttcgaatt cagtggctta atcatgaatg

 7381 tctaaatatt gttgacatta ggatgataca tgtaaattaa agttacattt gtttagcata

 7441 gacaagctta acattgtaga tgtttctctt caaaaatcat cttaaacatt tgcatttgga

 7501 attgtgttaa atagaatgtg tgaaacactg tattagtaaa cttcatcacc tttctacttc

 7561 cttatagttt gaacttttca gtttttgtag ttcccaaaca gttgctcaat ttagagcaaa

 7621 ttaatttaac acctgccaaa aaaaggctgc tgttggctta tcagttgtct ttaaattcaa

 7681 atgctcatgt gacttttatc acatcaaaaa atatttcatt aatgattcac ctttagctct

 7741 gaaaattacc gcgtttagta attatagtgg gcttataaaa acatgcaact ctttttgata

 7801 gttatttgag aattttggtg aaaaatattt agctgagggc agtatagaac ttataaacca

 7861 atatattgat atttttaaaa catttttaca tataagtaaa ctgccatctt tgagcataac

 7921 tacatttaaa aataaagctg catattttta aatcaagtgt ttaacaagaa tttatatttt

 7981 ttatttttta aaattaaaaa taatttatat ttcctctgtt gcatgaggat tctcatctgt

 8041 gcttataatg gttagagatt ttatttgtgt ggaatgaagt gaggcttgta gtcatggttc

 8101 tagtgtttca gtttgccaag tctgtttact gcagtgaaat tcatcaaatg tttcagtgtg

 8161 gttttctgta gcctatcatt tactggctat ttttttatgt acacctttag gattttctgc

 8221 ctactctatc cagttgtcca aatgatatcc tacattttac aaatgccctt tcagtttcta

 8281 ttttcttttt ccattaaatt gccctcatgt cctaatgtgc agtttgtaag tgtgtgtgtg

 8341 tgtgtctgtg tgtgtgtgaa tttgattttc aagagtgcta gacttccaat ttgagagatt

 8401 aaataattta attcaggcaa acatttttca ttggaatttc acagttcatt gtaatgaaaa

 8461 tgttaatcct ggatgacctt tgacatacag taatgaatct tggatattaa tgaatttgtt

 8521 agtagcatct tgatgtgtgt tttaatgagt tattttcaaa gttgtgcatt aaaccaaagt

 8581 tggcatactg gaagtgttta tatcaagttc catttggcta ctgatggaca aaaaatagaa

 8641 atgccttcct atggagagta tttttccttt aaaaaattaa aaaggttaat tattttgact

 8701 aaaaaaaaaa aaaaaaaa

# Homo sapiens mortality factor 4 like 1 (MORF4L1), transcript variant 4, mRNA

NCBI Reference Sequence: NM\_001265604.2

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2%22%20%5Cl%20%22goto1890261236_0)

LOCUS NM\_001265604 2333 bp mRNA linear PRI 16-DEC-2020

DEFINITION Homo sapiens mortality factor 4 like 1 (MORF4L1), transcript

 variant 4, mRNA.

ACCESSION NM\_001265604

VERSION NM\_001265604.2

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2333)

 AUTHORS Fragoza R, Das J, Wierbowski SD, Liang J, Tran TN, Liang S, Beltran

 JF, Rivera-Erick CA, Ye K, Wang TY, Yao L, Mort M, Stenson PD,

 Cooper DN, Wei X, Keinan A, Schimenti JC, Clark AG and Yu H.

 TITLE Extensive disruption of protein interactions by genetic variants

 across the allele frequency spectrum in human populations

 JOURNAL Nat Commun 10 (1), 4141 (2019)

 PUBMED [31515488](https://www.ncbi.nlm.nih.gov/pubmed/31515488)

 REMARK Publication Status: Online-Only

REFERENCE 2 (bases 1 to 2333)

 AUTHORS Hou P, Huang C, Liu CP, Yang N, Yu T, Yin Y, Zhu B and Xu RM.

 TITLE Structural Insights into Stimulation of Ash1L's H3K36

 Methyltransferase Activity through Mrg15 Binding

 JOURNAL Structure 27 (5), 837-845 (2019)

 PUBMED [30827843](https://www.ncbi.nlm.nih.gov/pubmed/30827843)

 REMARK GeneRIF: Ash1L stimulates H3K36 methyltransferase activity through

 Mrg15 binding

REFERENCE 3 (bases 1 to 2333)

 AUTHORS Lee Y, Yoon E, Cho S, Schmahling S, Muller J and Song JJ.

 TITLE Structural Basis of MRG15-Mediated Activation of the ASH1L Histone

 Methyltransferase by Releasing an Autoinhibitory Loop

 JOURNAL Structure 27 (5), 846-852 (2019)

 PUBMED [30827841](https://www.ncbi.nlm.nih.gov/pubmed/30827841)

 REMARK GeneRIF: ASH1L activation by MRG15 represents a delicate regulatory

 mechanism for how a cofactor activates an SET domain HMTase by

 releasing autoinhibition

REFERENCE 4 (bases 1 to 2333)

 AUTHORS Bleuyard JY, Fournier M, Nakato R, Couturier AM, Katou Y, Ralf C,

 Hester SS, Dominguez D, Rhodes D, Humphrey TC, Shirahige K and

 Esashi F.

 TITLE MRG15-mediated tethering of PALB2 to unperturbed chromatin protects

 active genes from genotoxic stress

 JOURNAL Proc Natl Acad Sci U S A 114 (29), 7671-7676 (2017)

 PUBMED [28673974](https://www.ncbi.nlm.nih.gov/pubmed/28673974)

 REMARK GeneRIF: PALB2 associates with active genes through its major

 binding partner, MRG15, which recognizes histone H3 trimethylated

 at lysine 36 (H3K36me3) by the SETD2 methyltransferase

REFERENCE 5 (bases 1 to 2333)

 AUTHORS Marcon E, Ni Z, Pu S, Turinsky AL, Trimble SS, Olsen JB,

 Silverman-Gavrila R, Silverman-Gavrila L, Phanse S, Guo H, Zhong G,

 Guo X, Young P, Bailey S, Roudeva D, Zhao D, Hewel J, Li J,

 Graslund S, Paduch M, Kossiakoff AA, Lupien M, Emili A, Wodak SJ

 and Greenblatt J.

 TITLE Human-chromatin-related protein interactions identify a demethylase

 complex required for chromosome segregation

 JOURNAL Cell Rep 8 (1), 297-310 (2014)

 PUBMED [24981860](https://www.ncbi.nlm.nih.gov/pubmed/24981860)

REFERENCE 6 (bases 1 to 2333)

 AUTHORS Cai Y, Jin J, Tomomori-Sato C, Sato S, Sorokina I, Parmely TJ,

 Conaway RC and Conaway JW.

 TITLE Identification of new subunits of the multiprotein mammalian

 TRRAP/TIP60-containing histone acetyltransferase complex

 JOURNAL J Biol Chem 278 (44), 42733-42736 (2003)

 PUBMED [12963728](https://www.ncbi.nlm.nih.gov/pubmed/12963728)

REFERENCE 7 (bases 1 to 2333)

 AUTHORS Pardo PS, Leung JK, Lucchesi JC and Pereira-Smith OM.

 TITLE MRG15, a novel chromodomain protein, is present in two distinct

 multiprotein complexes involved in transcriptional activation

 JOURNAL J Biol Chem 277 (52), 50860-50866 (2002)

 PUBMED [12397079](https://www.ncbi.nlm.nih.gov/pubmed/12397079)

 REMARK GeneRIF: a novel chromodomain protein that is present in two

 distinct multiprotein complexes involved in transcriptional

 activation

REFERENCE 8 (bases 1 to 2333)

 AUTHORS Yochum GS and Ayer DE.

 TITLE Role for the mortality factors MORF4, MRGX, and MRG15 in

 transcriptional repression via associations with Pf1, mSin3A, and

 Transducin-Like Enhancer of Split

 JOURNAL Mol Cell Biol 22 (22), 7868-7876 (2002)

 PUBMED [12391155](https://www.ncbi.nlm.nih.gov/pubmed/12391155)

REFERENCE 9 (bases 1 to 2333)

 AUTHORS Leung JK, Berube N, Venable S, Ahmed S, Timchenko N and

 Pereira-Smith OM.

 TITLE MRG15 activates the B-myb promoter through formation of a nuclear

 complex with the retinoblastoma protein and the novel protein PAM14

 JOURNAL J Biol Chem 276 (42), 39171-39178 (2001)

 PUBMED [11500496](https://www.ncbi.nlm.nih.gov/pubmed/11500496)

REFERENCE 10 (bases 1 to 2333)

 AUTHORS Bertram MJ, Berube NG, Hang-Swanson X, Ran Q, Leung JK, Bryce S,

 Spurgers K, Bick RJ, Baldini A, Ning Y, Clark LJ, Parkinson EK,

 Barrett JC, Smith JR and Pereira-Smith OM.

 TITLE Identification of a gene that reverses the immortal phenotype of a

 subset of cells and is a member of a novel family of transcription

 factor-like genes

 JOURNAL Mol Cell Biol 19 (2), 1479-1485 (1999)

 PUBMED [9891081](https://www.ncbi.nlm.nih.gov/pubmed/9891081)

COMMENT VALIDATED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has undergone validation or

 preliminary review. The reference sequence was derived from

 [AC103975.9](https://www.ncbi.nlm.nih.gov/nuccore/AC103975.9), [DB455429.1](https://www.ncbi.nlm.nih.gov/nuccore/DB455429.1), [AK300789.1](https://www.ncbi.nlm.nih.gov/nuccore/AK300789.1), [AY148481.1](https://www.ncbi.nlm.nih.gov/nuccore/AY148481.1), [BM996530.1](https://www.ncbi.nlm.nih.gov/nuccore/BM996530.1) and

 [AC011944.12](https://www.ncbi.nlm.nih.gov/nuccore/AC011944.12).

 On Aug 13, 2020 this sequence version replaced [NM\_001265604.1](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.1).

 Transcript Variant: This variant (4) differs in the 5' UTR and

 initiates translation at a downstream, in-frame start codon,

 compared to variant 1. Variants 3, 4 and 5 encode the same isoform

 (3), which has a shorter N-terminus compared to isoform 1.

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

 ##Evidence-Data-START##

 Transcript exon combination :: AK300789.1, SRR1660805.249535.1

 [ECO:0000332]

 ##Evidence-Data-END##

 COMPLETENESS: complete on the 3' end.

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-99 AC103975.9 199546-199644

 100-115 DB455429.1 1-16

 116-1449 AK300789.1 1-1334

 1450-1924 AY148481.1 1409-1883

 1925-1933 BM996530.1 14-22 c

 1934-2333 AC011944.12 128215-128614 c

FEATURES Location/Qualifiers

 source 1..2333

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="15"

 /map="15q25.1"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1&to=2333) 1..2333

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="mortality factor 4 like 1"

 /db\_xref="GeneID:[10933](https://www.ncbi.nlm.nih.gov/gene/10933)"

 /db\_xref="HGNC:[HGNC:16989](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:16989)"

 /db\_xref="MIM:[607303](https://www.ncbi.nlm.nih.gov/omim/607303)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1&to=329) 1..329

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=272&to=274) 272..274

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="upstream in-frame stop codon"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=330&to=376) 330..376

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=377&to=444) 377..444

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=445&to=531) 445..531

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=532&to=612) 532..612

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=554&to=1261) 554..1261

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="isoform 3 is encoded by transcript variant 4;

 MORF-related gene on chromosome 15; Esa1p-associated

 factor 3 homolog; protein MSL3-1; MORF-related gene 15

 protein; transcription factor-like protein MRG15;

 mortality factor 4-like protein 1"

 /codon\_start=1

 /product="mortality factor 4-like protein 1 isoform 3"

 /protein\_id="[NP\_001252533.1](https://www.ncbi.nlm.nih.gov/protein/388240814)"

 /db\_xref="CCDS:[CCDS58393.1](https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?REQUEST=CCDS&DATA=CCDS58393.1)"

 /db\_xref="GeneID:[10933](https://www.ncbi.nlm.nih.gov/gene/10933)"

 /db\_xref="HGNC:[HGNC:16989](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:16989)"

 /db\_xref="MIM:[607303](https://www.ncbi.nlm.nih.gov/omim/607303)"

 /translation="MRGAAPGKKTSGLQQKNVEVKTKKNKQKTPGNGDGGSTSETPQP

 PRKKRARVDPTVENEETFMNRVEVKVKIPEELKPWLVDDWDLITRQKQLFYLPAKKNV

 DSILEDYANYKKSRGNTDNKEYAVNEVVAGIKEYFNVMLGTQLLYKFERPQYAEILAD

 HPDAPMSQVYGAPHLLRLFVRIGAMLAYTPLDEKSLALLLNYLHDFLKYLAKNSATLF

 SASDYEVAPPEYHRKAV"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=613&to=638) 613..638

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=639&to=727) 639..727

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=728&to=829) 728..829

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=830&to=918) 830..918

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=919&to=1091) 919..1091

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1092&to=1176) 1092..1176

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1177&to=2333) 1177..2333

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /inference="alignment:Splign:2.1.0"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1404&to=1409) 1404..1409

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1424&to=1424) 1424

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="major polyA site"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1902&to=1907) 1902..1907

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=1933&to=1933) 1933

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="major polyA site"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=2303&to=2308) 2303..2308

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

 /note="hexamer: ATTAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001265604.2?from=2333&to=2333) 2333

 /gene="MORF4L1"

 /gene\_synonym="Eaf3; FWP006; HsT17725; MEAF3; MORFRG15;

 MRG15; S863-6"

ORIGIN

 1 aaatccggcc caggatgtag agctggcagt gcctgacggc gcgtctgacg cggagttggg

 61 tggggtagag agtagggggc ggtagtcggg ggtggtggga gaaggaggag gcggcgaatc

 121 acttataaat ggcgccgaag caggacccga agcctaaatt ccaggagggt gagtgtgcgc

 181 ctttgggaaa aaggcaccta acggcgcagg agatagaggc gggctcgagg tgattgaggc

 241 ttgagggccg ggggcggcgc gggctgcgcc ctgagaaggc ggcggtcagt gctttgtgcg

 301 gggaaagcgc attgcggatg gcaattccgg tgagcgagtg ctgtgctttc atgggcctct

 361 tctttatgaa gcaaagtgtg taaaggttgc cataaaggac aaacaagtga aatacttcat

 421 acattacagt ggttggaata aaaattggga tgaatgggtt ccggagagca gagtactcaa

 481 atacgtggac accaatttgc agaaacagcg agaacttcaa aaagccaatc aggagcagta

 541 tgcagagggg aagatgagag gggctgcccc aggaaagaag acatctggtc tgcaacagaa

 601 aaatgttgaa gtgaaaacga aaaagaacaa acagaaaaca cctggaaatg gagatggtgg

 661 cagtaccagt gagacccctc agcctcctcg gaagaaaagg gcccgggtag atcctactgt

 721 tgaaaatgag gaaacattca tgaacagagt tgaagttaaa gtaaagattc ctgaagagct

 781 aaaaccgtgg cttgttgatg actgggactt aattaccagg caaaaacagc tcttttatct

 841 tcctgccaag aagaatgtgg attccattct tgaggattat gcaaattaca agaaatctcg

 901 tggaaacaca gataataagg agtatgcggt taatgaagtt gtggcaggga taaaagaata

 961 cttcaacgta atgttgggta cccagctact ctataaattt gagagaccac agtatgctga

 1021 aattcttgca gatcatcccg atgcacccat gtcccaggtg tatggagcgc cacatctcct

 1081 gagattattt gtacgaattg gagcaatgtt ggcttataca cctctggatg agaagagcct

 1141 tgctttatta ctcaattatc ttcacgattt cctaaagtac ctggcaaaga attctgcaac

 1201 tttgttcagt gccagcgatt atgaagtggc tcctcctgag taccatcgga aagctgtgtg

 1261 agaggcactc tcactcactt atgtttggat ctccgtaaac acatttttgt tcttagtcta

 1321 tctcttgtac aaacgatgtg ctttgaagat gttagtgtat aacaattgat gtttgttttc

 1381 tgtttgattt taaacagaga aaaaataaaa gggggtaata gctccttttt tcttctttct

 1441 tttttttttt catttcaaaa ttgctgccag tgttttcaat gatggacaac agagggatat

 1501 gctgtagagt gttttattgc ctagttgaca aagctgcttt tgaatgctgg tggttctatt

 1561 cctttgacac tacgcacttt tataatacat gttaatgcta tatgacaaaa tgctctgatt

 1621 cctagtgcca aaggttcaat tcagtgtata taactgaaca cactcatcca tttgtgcttt

 1681 tgtttttttt tatggtgctt aaagtaaaga gcccatcctt tgcaagtcat ccatgttgtt

 1741 acttaggcat tttatcttgg ctcaaattgt tgaagaatgg tggcttgttt catggttttt

 1801 gtatttgtgt ctaatgcacg ttttaacatg atagacgcaa tgcattgtgt agctagtttt

 1861 ctggaaaagt caatctttta ggaattgttt ttcagatctt caataaattt tttctttaaa

 1921 tttcaaagaa caatgtgctt gtgttgatgc cttacaaaaa ccattgtata tttgtgtatt

 1981 ccttcttgta tttagacagt ggtttttcag gtgcgtgctt tgttttctgg tatggccttt

 2041 atggaatgag acgctttagc tttggtacgt agcgctaatc catagcagct ttggcagttt

 2101 gctgtcttga gtcttagcta aaaagttaga agtttacatg actgtttttt ttattttccc

 2161 taaattatta cttactctga gcattaatta agggcatttt cacctgtgta aaattatggt

 2221 cagctttttt ctgtctataa ttgtttactt ttgtgggttt actctagaaa catgagccaa

 2281 aaatgtcaat agacaacaca gtattaaaat aacccaaaag ttgtaaaggg caa

# Homo sapiens suppressor of cytokine signaling 3 (SOCS3), transcript variant 3, mRNA

NCBI Reference Sequence: NM\_001378933.1

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1%22%20%5Cl%20%22goto1820735736_0)

LOCUS NM\_001378933 2476 bp mRNA linear PRI 15-DEC-2020

DEFINITION Homo sapiens suppressor of cytokine signaling 3 (SOCS3), transcript

 variant 3, mRNA.

ACCESSION NM\_001378933

VERSION NM\_001378933.1

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2476)

 AUTHORS Johnson HM, Lewin AS and Ahmed CM.

 TITLE SOCS, Intrinsic Virulence Factors, and Treatment of COVID-19

 JOURNAL Front Immunol 11, 582102 (2020)

 PUBMED [33193390](https://www.ncbi.nlm.nih.gov/pubmed/33193390)

 REMARK GeneRIF: SOCS, Intrinsic Virulence Factors, and Treatment of

 COVID-19.

 Review article

 Publication Status: Online-Only

REFERENCE 2 (bases 1 to 2476)

 AUTHORS Jiang M, Zhang W, Zhang R, Liu P, Ye Y, Yu W, Guo X and Yu J.

 TITLE Cancer exosome-derived miR-9 and miR-181a promote the development

 of early-stage MDSCs via interfering with SOCS3 and PIAS3

 respectively in breast cancer

 JOURNAL Oncogene 39 (24), 4681-4694 (2020)

 PUBMED [32398867](https://www.ncbi.nlm.nih.gov/pubmed/32398867)

 REMARK GeneRIF: Cancer exosome-derived miR-9 and miR-181a promote the

 development of early-stage MDSCs via interfering with SOCS3 and

 PIAS3 respectively in breast cancer.

REFERENCE 3 (bases 1 to 2476)

 AUTHORS Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, Brignall

 R, Cafarelli T, Campos-Laborie FJ, Charloteaux B, Choi D, Cote AG,

 Daley M, Deimling S, Desbuleux A, Dricot A, Gebbia M, Hardy MF,

 Kishore N, Knapp JJ, Kovacs IA, Lemmens I, Mee MW, Mellor JC,

 Pollis C, Pons C, Richardson AD, Schlabach S, Teeking B, Yadav A,

 Babor M, Balcha D, Basha O, Bowman-Colin C, Chin SF, Choi SG,

 Colabella C, Coppin G, D'Amata C, De Ridder D, De Rouck S,

 Duran-Frigola M, Ennajdaoui H, Goebels F, Goehring L, Gopal A,

 Haddad G, Hatchi E, Helmy M, Jacob Y, Kassa Y, Landini S, Li R, van

 Lieshout N, MacWilliams A, Markey D, Paulson JN, Rangarajan S,

 Rasla J, Rayhan A, Rolland T, San-Miguel A, Shen Y, Sheykhkarimli

 D, Sheynkman GM, Simonovsky E, Tasan M, Tejeda A, Tropepe V,

 Twizere JC, Wang Y, Weatheritt RJ, Weile J, Xia Y, Yang X,

 Yeger-Lotem E, Zhong Q, Aloy P, Bader GD, De Las Rivas J, Gaudet S,

 Hao T, Rak J, Tavernier J, Hill DE, Vidal M, Roth FP and Calderwood

 MA.

 TITLE A reference map of the human binary protein interactome

 JOURNAL Nature 580 (7803), 402-408 (2020)

 PUBMED [32296183](https://www.ncbi.nlm.nih.gov/pubmed/32296183)

REFERENCE 4 (bases 1 to 2476)

 AUTHORS Yang Z, Zhu H, Zhang L, Wei Q, Zhou L, Xu X, Song P, Liu J, Xie H

 and Zheng S.

 TITLE DNA methylation of SOCS1/2/3 predicts hepatocellular carcinoma

 recurrence after liver transplantation

 JOURNAL Mol Biol Rep 47 (3), 1773-1782 (2020)

 PUBMED [32006198](https://www.ncbi.nlm.nih.gov/pubmed/32006198)

 REMARK GeneRIF: DNA methylation of SOCS1/2/3 predicts hepatocellular

 carcinoma recurrence after liver transplantation.

REFERENCE 5 (bases 1 to 2476)

 AUTHORS Sun Y, Ju XL, Li D, Zhou PP, Li X and Luo RH.

 TITLE RETRACTED: miR-1290 promotes proliferation and suppresses apoptosis

 in acute myeloid leukemia by targeting FOXG1/SOCS3

 JOURNAL J Biol Regul Homeost Agents 33 (6) (2019)

 PUBMED [31960662](https://www.ncbi.nlm.nih.gov/pubmed/31960662)

 REMARK GeneRIF: miR-1290 promoted proliferation and suppressed apoptosis

 in acute myeloid leukemia by targeting FOXG1 and SOCS3

REFERENCE 6 (bases 1 to 2476)

 AUTHORS Marine JC, McKay C, Wang D, Topham DJ, Parganas E, Nakajima H,

 Pendeville H, Yasukawa H, Sasaki A, Yoshimura A and Ihle JN.

 TITLE SOCS3 is essential in the regulation of fetal liver erythropoiesis

 JOURNAL Cell 98 (5), 617-627 (1999)

 PUBMED [10490101](https://www.ncbi.nlm.nih.gov/pubmed/10490101)

REFERENCE 7 (bases 1 to 2476)

 AUTHORS Sasaki A, Yasukawa H, Suzuki A, Kamizono S, Syoda T, Kinjyo I,

 Sasaki M, Johnston JA and Yoshimura A.

 TITLE Cytokine-inducible SH2 protein-3 (CIS3/SOCS3) inhibits Janus

 tyrosine kinase by binding through the N-terminal kinase inhibitory

 region as well as SH2 domain

 JOURNAL Genes Cells 4 (6), 339-351 (1999)

 PUBMED [10421843](https://www.ncbi.nlm.nih.gov/pubmed/10421843)

REFERENCE 8 (bases 1 to 2476)

 AUTHORS Zhang JG, Farley A, Nicholson SE, Willson TA, Zugaro LM, Simpson

 RJ, Moritz RL, Cary D, Richardson R, Hausmann G, Kile BT, Kent SB,

 Alexander WS, Metcalf D, Hilton DJ, Nicola NA and Baca M.

 TITLE The conserved SOCS box motif in suppressors of cytokine signaling

 binds to elongins B and C and may couple bound proteins to

 proteasomal degradation

 JOURNAL Proc Natl Acad Sci U S A 96 (5), 2071-2076 (1999)

 PUBMED [10051596](https://www.ncbi.nlm.nih.gov/pubmed/10051596)

 REMARK Erratum:[Proc Natl Acad Sci U S A. 2015 Jun 2;112(22):E2979. Kile,

 B J [corrected to Kile, Benjamin T]. PMID: 25956176]

REFERENCE 9 (bases 1 to 2476)

 AUTHORS Masuhara M, Sakamoto H, Matsumoto A, Suzuki R, Yasukawa H, Mitsui

 K, Wakioka T, Tanimura S, Sasaki A, Misawa H, Yokouchi M, Ohtsubo M

 and Yoshimura A.

 TITLE Cloning and characterization of novel CIS family genes

 JOURNAL Biochem Biophys Res Commun 239 (2), 439-446 (1997)

 PUBMED [9344848](https://www.ncbi.nlm.nih.gov/pubmed/9344848)

REFERENCE 10 (bases 1 to 2476)

 AUTHORS Minamoto S, Ikegame K, Ueno K, Narazaki M, Naka T, Yamamoto H,

 Matsumoto T, Saito H, Hosoe S and Kishimoto T.

 TITLE Cloning and functional analysis of new members of STAT induced STAT

 inhibitor (SSI) family: SSI-2 and SSI-3

 JOURNAL Biochem Biophys Res Commun 237 (1), 79-83 (1997)

 PUBMED [9266833](https://www.ncbi.nlm.nih.gov/pubmed/9266833)

COMMENT REVIEWED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has been curated by NCBI staff. The

 reference sequence was derived from [AC061992.11](https://www.ncbi.nlm.nih.gov/nuccore/AC061992.11).

 Summary: This gene encodes a member of the STAT-induced STAT

 inhibitor (SSI), also known as suppressor of cytokine signaling

 (SOCS), family. SSI family members are cytokine-inducible negative

 regulators of cytokine signaling. The expression of this gene is

 induced by various cytokines, including IL6, IL10, and interferon

 (IFN)-gamma. The protein encoded by this gene can bind to JAK2

 kinase, and inhibit the activity of JAK2 kinase. Studies of the

 mouse counterpart of this gene suggested the roles of this gene in

 the negative regulation of fetal liver hematopoiesis, and placental

 development. [provided by RefSeq, Jul 2008].

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-70 AC061992.11 23095-23164 c

 71-2476 AC061992.11 19017-21422 c

FEATURES Location/Qualifiers

 source 1..2476

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="17"

 /map="17q25.3"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?from=1&to=2476) 1..2476

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /note="suppressor of cytokine signaling 3"

 /db\_xref="GeneID:[9021](https://www.ncbi.nlm.nih.gov/gene/9021)"

 /db\_xref="HGNC:[HGNC:19391](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:19391)"

 /db\_xref="MIM:[604176](https://www.ncbi.nlm.nih.gov/omim/604176)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?from=1&to=70) 1..70

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?from=71&to=2476) 71..2476

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /inference="alignment:Splign:2.1.0"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?from=159&to=836) 159..836

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /note="STAT-induced STAT inhibitor 3; cytokine-inducible

 SH2 protein 3"

 /codon\_start=1

 /product="suppressor of cytokine signaling 3"

 /protein\_id="[NP\_001365862.1](https://www.ncbi.nlm.nih.gov/protein/1820735737)"

 /db\_xref="GeneID:[9021](https://www.ncbi.nlm.nih.gov/gene/9021)"

 /db\_xref="HGNC:[HGNC:19391](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:19391)"

 /db\_xref="MIM:[604176](https://www.ncbi.nlm.nih.gov/omim/604176)"

 /translation="MVTHSKFPAAGMSRPLDTSLRLKTFSSKSEYQLVVNAVRKLQES

 GFYWSAVTGGEANLLLSAEPAGTFLIRDSSDQRHFFTLSVKTQSGTKNLRIQCEGGSF

 SLQSDPRSTQPVPRFDCVLKLVHHYMPPPGAPSFPSPPTEPSSEVPEQPSAQPLPGSP

 PRRAYYIYSGGEKIPLVLSRPLSSNVATLQHLCRKTVNGHLDSYEKVTQLPGPIREFL

 DQYDAPL"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001365862.1?from=22&to=33) 222..257

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /note="propagated from UniProtKB/Swiss-Prot (O14543.1);

 Region: Kinase inhibitory region (KIR)"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001365862.1?from=34&to=45) 258..293

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /note="propagated from UniProtKB/Swiss-Prot (O14543.1);

 Region: Extended SH2 subdomain (ESS)"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?from=2448&to=2453) 2448..2453

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1?from=2476&to=2476) 2476

 /gene="SOCS3"

 /gene\_synonym="ATOD4; CIS3; Cish3; SOCS-3; SSI-3; SSI3"

 /note="major polyA site"

ORIGIN

 1 gccatcgttc cacctcgagc tctccagccc ggccgagagc acctggccca acccacaccg

 61 gaccaaccag ccccccggga tgcggtagcg gccgctgtgc ggaggccgcg aagcagctgc

 121 agccgccgcc gcgcagatcc acgctggctc cgtgcgccat ggtcacccac agcaagtttc

 181 ccgccgccgg gatgagccgc cccctggaca ccagcctgcg cctcaagacc ttcagctcca

 241 agagcgagta ccagctggtg gtgaacgcag tgcgcaagct gcaggagagc ggcttctact

 301 ggagcgcagt gaccggcggc gaggcgaacc tgctgctcag tgccgagccc gccggcacct

 361 ttctgatccg cgacagctcg gaccagcgcc acttcttcac gctcagcgtc aagacccagt

 421 ctgggaccaa gaacctgcgc atccagtgtg aggggggcag cttctctctg cagagcgatc

 481 cccggagcac gcagcccgtg ccccgcttcg actgcgtgct caagctggtg caccactaca

 541 tgccgccccc tggagccccc tccttcccct cgccacctac tgaaccctcc tccgaggtgc

 601 ccgagcagcc gtctgcccag ccactccctg ggagtccccc cagaagagcc tattacatct

 661 actccggggg cgagaagatc cccctggtgt tgagccggcc cctctcctcc aacgtggcca

 721 ctcttcagca tctctgtcgg aagaccgtca acggccacct ggactcctat gagaaagtca

 781 cccagctgcc ggggcccatt cgggagttcc tggaccagta cgatgccccg ctttaagggg

 841 taaagggcgc aaagggcatg ggtcgggaga ggggacgcag gcccctctcc tccgtggcac

 901 atggcacaag cacaagaagc caaccaggag agagtcctgt agctctgggg ggaaagaggg

 961 cggacaggcc cctccctctg ccctctccct gcagaatgtg gcaggcggac ctggaatgtg

 1021 ttggagggaa gggggagtac cacctgagtc tccagcttct ccggaggagc cagctgtcct

 1081 ggtgggacga tagcaaccac aagtggattc tccttcaatt cctcagcttc ccctctgcct

 1141 ccaaacaggg gacacttcgg gaatgctgaa ctaatgagaa ctgccaggga atcttcaaac

 1201 tttccaacgg aacttgtttg ctctttgatt tggtttaaac ctgagctggt tgtggagcct

 1261 gggaaaggtg gaagagagag aggtcctgag ggccccaggg ctgcgggctg gcgaaggaaa

 1321 tggtcacacc ccccgcccac cccaggcgag gatcctggtg acatgctcct ctccctggct

 1381 ccggggagaa gggcttgggg tgacctgaag ggaaccatcc tggtacccca catcctctcc

 1441 tccgggacag tcaccgaaaa cacaggttcc aaagtctacc tggtgcctga gagcccaggg

 1501 cccttcctcc gttttaaggg ggaagcaaca tttggagggg atggatgggc tggtcagctg

 1561 gtctcctttt cctactcata ctataccttc ctgtacctgg gtggatggag cgggaggatg

 1621 gaggagacgg gacatctttc acctcaggct cctggtagag aagacagggg attctactct

 1681 gtgcctcctg actatgtctg gctaagagat tcgccttaaa tgctccctgt cccatggaga

 1741 gggacccagc ataggaaagc cacatactca gcctggatgg gtggagaggc tgagggactc

 1801 actggagggc accaagccag cccacagcca gggaagtggg gagggggggc ggaaacccat

 1861 gcctcccagc tgagcactgg gaatgtcagc ccagtaagta ttggccagtc aggcgcctcg

 1921 tggtcagagc agagccacca ggtcccactg ccccgagccc tgcacagccc tccctcctgc

 1981 ctgggtgggg gaggctggag gtcattggag aggctggact gctgccaccc cgggtgctcc

 2041 cgctctgcca tagcactgat cagtgacaat ttacaggaat gtagcagcga tggaattacc

 2101 tggaacagtt ttttgttttt gtttttgttt ttgtttttgt gggggggggc aactaaacaa

 2161 acacaaagta ttctgtgtca ggtattgggc tggacagggc agttgtgtgt tggggtggtt

 2221 tttttctcta tttttttgtt tgtttcttgt tttttaataa tgtttacaat ctgcctcaat

 2281 cactctgtct tttataaaga ttccacctcc agtcctctct cctcccccct actcaggccc

 2341 ttgaggctat taggagatgc ttgaagaact caacaaaatc ccaatccaag tcaaactttg

 2401 cacatattta tatttatatt cagaaaagaa acatttcagt aatttataat aaagagcact

 2461 attttttaat gaaaaa

# Homo sapiens coagulation factor III, tissue factor (F3), transcript variant 2, mRNA

NCBI Reference Sequence: NM\_001178096.2

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2%22%20%5Cl%20%22goto1890283268_0)

LOCUS NM\_001178096 2138 bp mRNA linear PRI 16-DEC-2020

DEFINITION Homo sapiens coagulation factor III, tissue factor (F3), transcript

 variant 2, mRNA.

ACCESSION NM\_001178096

VERSION NM\_001178096.2

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2138)

 AUTHORS Skendros P, Mitsios A, Chrysanthopoulou A, Mastellos DC, Metallidis

 S, Rafailidis P, Ntinopoulou M, Sertaridou E, Tsironidou V,

 Tsigalou C, Tektonidou M, Konstantinidis T, Papagoras C, Mitroulis

 I, Germanidis G, Lambris JD and Ritis K.

 TITLE Complement and tissue factor-enriched neutrophil extracellular

 traps are key drivers in COVID-19 immunothrombosis

 JOURNAL J Clin Invest 130 (11), 6151-6157 (2020)

 PUBMED [32759504](https://www.ncbi.nlm.nih.gov/pubmed/32759504)

 REMARK GeneRIF: Complement and tissue factor-enriched neutrophil

 extracellular traps are key drivers in COVID-19 immunothrombosis.

REFERENCE 2 (bases 1 to 2138)

 AUTHORS Hottz ED, Azevedo-Quintanilha IG, Palhinha L, Teixeira L, Barreto

 EA, Pao CRR, Righy C, Franco S, Souza TML, Kurtz P, Bozza FA and

 Bozza PT.

 TITLE Platelet activation and platelet-monocyte aggregate formation

 trigger tissue factor expression in patients with severe COVID-19

 JOURNAL Blood 136 (11), 1330-1341 (2020)

 PUBMED [32678428](https://www.ncbi.nlm.nih.gov/pubmed/32678428)

REFERENCE 3 (bases 1 to 2138)

 AUTHORS Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, Brignall

 R, Cafarelli T, Campos-Laborie FJ, Charloteaux B, Choi D, Cote AG,

 Daley M, Deimling S, Desbuleux A, Dricot A, Gebbia M, Hardy MF,

 Kishore N, Knapp JJ, Kovacs IA, Lemmens I, Mee MW, Mellor JC,

 Pollis C, Pons C, Richardson AD, Schlabach S, Teeking B, Yadav A,

 Babor M, Balcha D, Basha O, Bowman-Colin C, Chin SF, Choi SG,

 Colabella C, Coppin G, D'Amata C, De Ridder D, De Rouck S,

 Duran-Frigola M, Ennajdaoui H, Goebels F, Goehring L, Gopal A,

 Haddad G, Hatchi E, Helmy M, Jacob Y, Kassa Y, Landini S, Li R, van

 Lieshout N, MacWilliams A, Markey D, Paulson JN, Rangarajan S,

 Rasla J, Rayhan A, Rolland T, San-Miguel A, Shen Y, Sheykhkarimli

 D, Sheynkman GM, Simonovsky E, Tasan M, Tejeda A, Tropepe V,

 Twizere JC, Wang Y, Weatheritt RJ, Weile J, Xia Y, Yang X,

 Yeger-Lotem E, Zhong Q, Aloy P, Bader GD, De Las Rivas J, Gaudet S,

 Hao T, Rak J, Tavernier J, Hill DE, Vidal M, Roth FP and Calderwood

 MA.

 TITLE A reference map of the human binary protein interactome

 JOURNAL Nature 580 (7803), 402-408 (2020)

 PUBMED [32296183](https://www.ncbi.nlm.nih.gov/pubmed/32296183)

REFERENCE 4 (bases 1 to 2138)

 AUTHORS Zioncheck TF, Roy S and Vehar GA.

 TITLE The cytoplasmic domain of tissue factor is phosphorylated by a

 protein kinase C-dependent mechanism

 JOURNAL J Biol Chem 267 (6), 3561-3564 (1992)

 PUBMED [1740409](https://www.ncbi.nlm.nih.gov/pubmed/1740409)

REFERENCE 5 (bases 1 to 2138)

 AUTHORS Broze GJ Jr, Girard TJ and Novotny WF.

 TITLE Regulation of coagulation by a multivalent Kunitz-type inhibitor

 JOURNAL Biochemistry 29 (33), 7539-7546 (1990)

 PUBMED [2271516](https://www.ncbi.nlm.nih.gov/pubmed/2271516)

 REMARK Review article

REFERENCE 6 (bases 1 to 2138)

 AUTHORS Mackman N, Morrissey JH, Fowler B and Edgington TS.

 TITLE Complete sequence of the human tissue factor gene, a highly

 regulated cellular receptor that initiates the coagulation protease

 cascade

 JOURNAL Biochemistry 28 (4), 1755-1762 (1989)

 PUBMED [2719931](https://www.ncbi.nlm.nih.gov/pubmed/2719931)

REFERENCE 7 (bases 1 to 2138)

 AUTHORS Kao FT, Hartz J, Horton R, Nemerson Y and Carson SD.

 TITLE Regional assignment of human tissue factor gene (F3) to chromosome

 1p21-p22

 JOURNAL Somat Cell Mol Genet 14 (4), 407-410 (1988)

 PUBMED [3399965](https://www.ncbi.nlm.nih.gov/pubmed/3399965)

REFERENCE 8 (bases 1 to 2138)

 AUTHORS Scarpati EM, Wen D, Broze GJ Jr, Miletich JP, Flandermeyer RR,

 Siegel NR and Sadler JE.

 TITLE Human tissue factor: cDNA sequence and chromosome localization of

 the gene

 JOURNAL Biochemistry 26 (17), 5234-5238 (1987)

 PUBMED [2823875](https://www.ncbi.nlm.nih.gov/pubmed/2823875)

REFERENCE 9 (bases 1 to 2138)

 AUTHORS Spicer,E.K., Horton,R., Bloem,L., Bach,R., Williams,K.R., Guha,A.,

 Kraus,J., Lin,T.C., Nemerson,Y. and Konigsberg,W.H.

 TITLE Isolation of cDNA clones coding for human tissue factor: primary

 structure of the protein and cDNA

 JOURNAL Proc Natl Acad Sci U S A 84 (15), 5148-5152 (1987)

 PUBMED [3037536](https://www.ncbi.nlm.nih.gov/pubmed/3037536)

REFERENCE 10 (bases 1 to 2138)

 AUTHORS Gouault-Helimann,M. and Josso,F.

 TITLE [Initiation in vivo of blood coagulation. The role of white blood

 cells and tissue factor (author's transl)]

 JOURNAL Nouv Presse Med 8 (40), 3249-3253 (1979)

 PUBMED [392457](https://www.ncbi.nlm.nih.gov/pubmed/392457)

 REMARK Review article

COMMENT REVIEWED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has been curated by NCBI staff. The

 reference sequence was derived from [DB499288.1](https://www.ncbi.nlm.nih.gov/nuccore/DB499288.1), [AF487337.1](https://www.ncbi.nlm.nih.gov/nuccore/AF487337.1) and

 [AC093117.2](https://www.ncbi.nlm.nih.gov/nuccore/AC093117.2).

 On Aug 13, 2020 this sequence version replaced [NM\_001178096.1](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.1).

 Summary: This gene encodes coagulation factor III which is a cell

 surface glycoprotein. This factor enables cells to initiate the

 blood coagulation cascades, and it functions as the high-affinity

 receptor for the coagulation factor VII. The resulting complex

 provides a catalytic event that is responsible for initiation of

 the coagulation protease cascades by specific limited proteolysis.

 Unlike the other cofactors of these protease cascades, which

 circulate as nonfunctional precursors, this factor is a potent

 initiator that is fully functional when expressed on cell surfaces,

 for example, on monocytes. There are 3 distinct domains of this

 factor: extracellular, transmembrane, and cytoplasmic. Platelets

 and monocytes have been shown to express this coagulation factor

 under procoagulatory and proinflammatory stimuli, and a major role

 in HIV-associated coagulopathy has been described.

 Platelet-dependent monocyte expression of coagulation factor III

 has been described to be associated with Coronavirus Disease 2019

 (COVID-19) severity and mortality. This protein is the only one in

 the coagulation pathway for which a congenital deficiency has not

 been described. Alternate splicing results in multiple transcript

 variants.[provided by RefSeq, Aug 2020].

 Transcript Variant: This variant (2) lacks an exon in the coding

 region, which results in a frameshift and an early stop codon,

 compared to variant 1. The encoded isoform (2) is shorter and has a

 distinct C-terminus, compared to isoform 1.

 Sequence Note: This RefSeq record was created from transcript and

 genomic sequence data to make the sequence consistent with the

 reference genome assembly. The genomic coordinates used for the

 transcript record were based on transcript alignments.

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

 ##Evidence-Data-START##

 Transcript exon combination :: AF497569.1, AF497570.1 [ECO:0000332]

 ##Evidence-Data-END##

 ##RefSeq-Attributes-START##

 coronavirus related :: relevant for disease process

 ##RefSeq-Attributes-END##

 COMPLETENESS: full length.

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-122 DB499288.1 101-222

 123-852 AF487337.1 1-730

 853-2138 AC093117.2 14463-15748 c

FEATURES Location/Qualifiers

 source 1..2138

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="1"

 /map="1p21.3"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=1&to=2138) 1..2138

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="coagulation factor III, tissue factor"

 /db\_xref="GeneID:[2152](https://www.ncbi.nlm.nih.gov/gene/2152)"

 /db\_xref="HGNC:[HGNC:3541](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:3541)"

 /db\_xref="MIM:[134390](https://www.ncbi.nlm.nih.gov/omim/134390)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=1&to=223) 1..223

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /inference="alignment:Splign:2.1.0"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=124&to=840) 124..840

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="isoform 2 precursor is encoded by transcript

 variant 2; coagulation factor III (thromboplastin, tissue

 factor)"

 /codon\_start=1

 /product="tissue factor isoform 2 precursor"

 /protein\_id="[NP\_001171567.1](https://www.ncbi.nlm.nih.gov/protein/296010912)"

 /db\_xref="CCDS:[CCDS53345.1](https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?REQUEST=CCDS&DATA=CCDS53345.1)"

 /db\_xref="GeneID:[2152](https://www.ncbi.nlm.nih.gov/gene/2152)"

 /db\_xref="HGNC:[HGNC:3541](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:3541)"

 /db\_xref="MIM:[134390](https://www.ncbi.nlm.nih.gov/omim/134390)"

 /translation="METPAWPRVPRPETAVARTLLLGWVFAQVAGASGTTNTVAAYNL

 TWKSTNFKTILEWEPKPVNQVYTVQISTKSGDWKSKCFYTTDTECDLTDEIVKDVKQT

 YLARVFSYPAGNVESTGSAGEPLYENSPEFTPYLETNLGQPTIQSFEQVGTKVNVTVE

 DERTLVRRNNTFLSLRDVFGKDLIYTLYYWKSSSSGKKYSTSLELWYLWSSSLSSSWL

 YLYTSVERQEWGRAGRRTPH"

 [sig\_peptide](https://www.ncbi.nlm.nih.gov/protein/NP_001171567.1?from=1&to=32) 124..219

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/NP_001171567.1?from=33&to=238) 220..837

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /product="tissue factor isoform 2"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001171567.1?from=46&to=48) 259..267

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="propagated from UniProtKB/Swiss-Prot (P13726.1);

 Region: WKS motif"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001171567.1?from=77&to=79) 352..360

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="propagated from UniProtKB/Swiss-Prot (P13726.1);

 Region: WKS motif"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001171567.1?from=156&to=156) 589..591

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="N-linked (GlcNAc...) asparagine.

 /evidence=ECO:0000255; propagated from

 UniProtKB/Swiss-Prot (P13726.1); glycosylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001171567.1?from=169&to=169) 628..630

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="N-linked (GlcNAc...) asparagine.

 /evidence=ECO:0000255; propagated from

 UniProtKB/Swiss-Prot (P13726.1); glycosylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001171567.1?from=190&to=192) 691..699

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="propagated from UniProtKB/Swiss-Prot (P13726.1);

 Region: WKS motif"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=224&to=335) 224..335

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=336&to=535) 336..535

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=536&to=714) 536..714

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=715&to=2138) 715..2138

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /inference="alignment:Splign:2.1.0"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=1972&to=1977) 1972..1977

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=1995&to=1995) 1995

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="major polyA site"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=2064&to=2069) 2064..2069

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="hexamer: AGTAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=2086&to=2086) 2086

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=2119&to=2124) 2119..2124

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

 /note="hexamer: AATACA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2?from=2138&to=2138) 2138

 /gene="F3"

 /gene\_synonym="CD142; TF; TFA"

ORIGIN

 1 aagactgcga gctccccgca ccccctcgca ctccctctgg ccggcccagg gcgccttcag

 61 cccaacctcc ccagccccac gggcgccacg gaacccgctc gatctcgccg ccaactggta

 121 gacatggaga cccctgcctg gccccgggtc ccgcgccccg agaccgccgt cgctcggacg

 181 ctcctgctcg gctgggtctt cgcccaggtg gccggcgctt caggcactac aaatactgtg

 241 gcagcatata atttaacttg gaaatcaact aatttcaaga caattttgga gtgggaaccc

 301 aaacccgtca atcaagtcta cactgttcaa ataagcacta agtcaggaga ttggaaaagc

 361 aaatgctttt acacaacaga cacagagtgt gacctcaccg acgagattgt gaaggatgtg

 421 aagcagacgt acttggcacg ggtcttctcc tacccggcag ggaatgtgga gagcaccggt

 481 tctgctgggg agcctctgta tgagaactcc ccagagttca caccttacct ggagacaaac

 541 ctcggacagc caacaattca gagttttgaa caggtgggaa caaaagtgaa tgtgaccgta

 601 gaagatgaac ggactttagt cagaaggaac aacactttcc taagcctccg ggatgttttt

 661 ggcaaggact taatttatac actttattat tggaaatctt caagttcagg aaagaaatat

 721 tctacatcat tggagctgtg gtatttgtgg tcatcatcct tgtcatcatc ctggctatat

 781 ctctacacaa gtgtagaaag gcaggagtgg ggcagagctg gaaggagaac tccccactga

 841 atgtttcata aaggaagcac tgttggagct actgcaaatg ctatattgca ctgtgaccga

 901 gaacttttaa gaggatagaa tacatggaaa cgcaaatgag tatttcggag catgaagacc

 961 ctggagttca aaaaactctt gatatgacct gttattacca ttagcattct ggttttgaca

 1021 tcagcattag tcactttgaa atgtaacgaa tggtactaca accaattcca agttttaatt

 1081 tttaacacca tggcaccttt tgcacataac atgctttaga ttatatattc cgcactcaag

 1141 gagtaaccag gtcgtccaag caaaaacaaa tgggaaaatg tcttaaaaaa tcctgggtgg

 1201 acttttgaaa agcttttttt tttttttttt tttttttgag acggagtctt gctctgttgc

 1261 ccaggctgga gtgcagtagc acgatctcgg ctcactgcac cctccgtctc tcgggttcaa

 1321 gcaattgtct gcctcagcct cccgagtagc tgggattaca ggtgcgcact accacgccaa

 1381 gctaattttt gtatttttta gtagagatgg ggtttcacca tcttggccag gctggtcttg

 1441 aattcctgac ctcaggtgat ccacccacct tggcctccca aagtgctagt attatgggcg

 1501 tgaaccacca tgcccagccg aaaagctttt gaggggctga cttcaatcca tgtaggaaag

 1561 taaaatggaa ggaaattggg tgcatttcta ggacttttct aacatatgtc tataatatag

 1621 tgtttaggtt cttttttttt tcaggaatac atttggaaat tcaaaacaat tggcaaactt

 1681 tgtattaatg tgttaagtgc aggagacatt ggtattctgg gcaccttcct aatatgcttt

 1741 acaatctgca ctttaactga cttaagtggc attaaacatt tgagagctaa ctatattttt

 1801 ataagactac tatacaaact acagagttta tgatttaagg tacttaaagc ttctatggtt

 1861 gacattgtat atataatttt ttaaaaaggt tttctatatg gggattttct atttatgtag

 1921 gtaatattgt tctatttgta tatattgaga taatttattt aatatacttt aaataaaggt

 1981 gactgggaat tgttactgtt gtacttattc tatcttccat ttattattta tgtacaattt

 2041 ggtgtttgta ttagctctac tacagtaaat gactgtaaaa ttgtcagtgg cttacaacaa

 2101 cgtatctttt tcgcttataa tacattttgg tgactgta

# Homo sapiens TNF superfamily member 14 (TNFSF14), transcript variant 2, mRNA

NCBI Reference Sequence: NM\_172014.3

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3%22%20%5Cl%20%22goto1015809705_0)

LOCUS NM\_172014 4406 bp mRNA linear PRI 13-DEC-2020

DEFINITION Homo sapiens TNF superfamily member 14 (TNFSF14), transcript

 variant 2, mRNA.

ACCESSION NM\_172014 XM\_005259670

VERSION NM\_172014.3

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4406)

 AUTHORS Perlin DS, Zafir-Lavie I, Roadcap L, Raines S, Ware CF and Neil GA.

 TITLE Levels of the TNF-Related Cytokine LIGHT Increase in Hospitalized

 COVID-19 Patients with Cytokine Release Syndrome and ARDS

 JOURNAL mSphere 5 (4), e00699-20 (2020)

 PUBMED [32817460](https://www.ncbi.nlm.nih.gov/pubmed/32817460)

 REMARK GeneRIF: Levels of the TNF-Related Cytokine LIGHT Increase in

 Hospitalized COVID-19 Patients with Cytokine Release Syndrome and

 ARDS.

 Publication Status: Online-Only

REFERENCE 2 (bases 1 to 4406)

 AUTHORS Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, Brignall

 R, Cafarelli T, Campos-Laborie FJ, Charloteaux B, Choi D, Cote AG,

 Daley M, Deimling S, Desbuleux A, Dricot A, Gebbia M, Hardy MF,

 Kishore N, Knapp JJ, Kovacs IA, Lemmens I, Mee MW, Mellor JC,

 Pollis C, Pons C, Richardson AD, Schlabach S, Teeking B, Yadav A,

 Babor M, Balcha D, Basha O, Bowman-Colin C, Chin SF, Choi SG,

 Colabella C, Coppin G, D'Amata C, De Ridder D, De Rouck S,

 Duran-Frigola M, Ennajdaoui H, Goebels F, Goehring L, Gopal A,

 Haddad G, Hatchi E, Helmy M, Jacob Y, Kassa Y, Landini S, Li R, van

 Lieshout N, MacWilliams A, Markey D, Paulson JN, Rangarajan S,

 Rasla J, Rayhan A, Rolland T, San-Miguel A, Shen Y, Sheykhkarimli

 D, Sheynkman GM, Simonovsky E, Tasan M, Tejeda A, Tropepe V,

 Twizere JC, Wang Y, Weatheritt RJ, Weile J, Xia Y, Yang X,

 Yeger-Lotem E, Zhong Q, Aloy P, Bader GD, De Las Rivas J, Gaudet S,

 Hao T, Rak J, Tavernier J, Hill DE, Vidal M, Roth FP and Calderwood

 MA.

 TITLE A reference map of the human binary protein interactome

 JOURNAL Nature 580 (7803), 402-408 (2020)

 PUBMED [32296183](https://www.ncbi.nlm.nih.gov/pubmed/32296183)

REFERENCE 3 (bases 1 to 4406)

 AUTHORS Brunetti G, Storlino G, Oranger A, Colaianni G, Faienza MF,

 Ingravallo G, Di Comite M, Reseland JE, Celi M, Tarantino U,

 Passeri G, Ware CF, Grano M and Colucci S.

 TITLE LIGHT/TNFSF14 regulates estrogen deficiency-induced bone loss

 JOURNAL J Pathol 250 (4), 440-451 (2020)

 PUBMED [31990039](https://www.ncbi.nlm.nih.gov/pubmed/31990039)

 REMARK GeneRIF: TNFSF14 regulates estrogen deficiency-induced bone loss

REFERENCE 4 (bases 1 to 4406)

 AUTHORS Iriyama T, Wang G, Yoshikawa M, Mimura N, Matsui H, Sayama S,

 Kumasawa K, Nagamatsu T, Koga K, Kotani T, Niimi K, Yamamoto E,

 Kellems RE, Xia Y, Osuga Y and Fujii T.

 TITLE Increased LIGHT leading to sFlt-1 elevation underlies the

 pathogenic link between hydatidiform mole and preeclampsia

 JOURNAL Sci Rep 9 (1), 10107 (2019)

 PUBMED [31300808](https://www.ncbi.nlm.nih.gov/pubmed/31300808)

 REMARK GeneRIF: Increased LIGHT leading to sFlt-1 elevation underlies the

 pathogenic link between hydatidiform mole and preeclampsia.

 Publication Status: Online-Only

REFERENCE 5 (bases 1 to 4406)

 AUTHORS Hsu CY, Tseng WK, Wu YW, Lin TH, Yeh HI, Chang KC, Wang JH, Chou

 RH, Huang CY, Huang PH, Leu HB, Yin WH, Wu CC, Lin SJ and Chen JW.

 TITLE Circulating TNFSF14 (Tumor Necrosis Factor Superfamily 14) Predicts

 Clinical Outcome in Patients With Stable Coronary Artery Disease

 JOURNAL Arterioscler Thromb Vasc Biol 39 (6), 1240-1252 (2019)

 PUBMED [30943772](https://www.ncbi.nlm.nih.gov/pubmed/30943772)

 REMARK GeneRIF: Increased TNFSF14 levels were independently associated

 with the occurrence of cardiovascular events in patients with

 stable coronary artery disease.

REFERENCE 6 (bases 1 to 4406)

 AUTHORS Yu KY, Kwon B, Ni J, Zhai Y, Ebner R and Kwon BS.

 TITLE A newly identified member of tumor necrosis factor receptor

 superfamily (TR6) suppresses LIGHT-mediated apoptosis

 JOURNAL J Biol Chem 274 (20), 13733-13736 (1999)

 PUBMED [10318773](https://www.ncbi.nlm.nih.gov/pubmed/10318773)

REFERENCE 7 (bases 1 to 4406)

 AUTHORS Harrop JA, McDonnell PC, Brigham-Burke M, Lyn SD, Minton J, Tan KB,

 Dede K, Spampanato J, Silverman C, Hensley P, DiPrinzio R, Emery

 JG, Deen K, Eichman C, Chabot-Fletcher M, Truneh A and Young PR.

 TITLE Herpesvirus entry mediator ligand (HVEM-L), a novel ligand for

 HVEM/TR2, stimulates proliferation of T cells and inhibits HT29

 cell growth

 JOURNAL J Biol Chem 273 (42), 27548-27556 (1998)

 PUBMED [9765287](https://www.ncbi.nlm.nih.gov/pubmed/9765287)

REFERENCE 8 (bases 1 to 4406)

 AUTHORS Zhai Y, Guo R, Hsu TL, Yu GL, Ni J, Kwon BS, Jiang GW, Lu J, Tan J,

 Ugustus M, Carter K, Rojas L, Zhu F, Lincoln C, Endress G, Xing L,

 Wang S, Oh KO, Gentz R, Ruben S, Lippman ME, Hsieh SL and Yang D.

 TITLE LIGHT, a novel ligand for lymphotoxin beta receptor and TR2/HVEM

 induces apoptosis and suppresses in vivo tumor formation via gene

 transfer

 JOURNAL J Clin Invest 102 (6), 1142-1151 (1998)

 PUBMED [9739048](https://www.ncbi.nlm.nih.gov/pubmed/9739048)

REFERENCE 9 (bases 1 to 4406)

 AUTHORS Marsters SA, Sheridan JP, Pitti RM, Brush J, Goddard A and

 Ashkenazi A.

 TITLE Identification of a ligand for the death-domain-containing receptor

 Apo3

 JOURNAL Curr Biol 8 (9), 525-528 (1998)

 PUBMED [9560343](https://www.ncbi.nlm.nih.gov/pubmed/9560343)

REFERENCE 10 (bases 1 to 4406)

 AUTHORS Mauri DN, Ebner R, Montgomery RI, Kochel KD, Cheung TC, Yu GL,

 Ruben S, Murphy M, Eisenberg RJ, Cohen GH, Spear PG and Ware CF.

 TITLE LIGHT, a new member of the TNF superfamily, and lymphotoxin alpha

 are ligands for herpesvirus entry mediator

 JOURNAL Immunity 8 (1), 21-30 (1998)

 PUBMED [9462508](https://www.ncbi.nlm.nih.gov/pubmed/9462508)

COMMENT REVIEWED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has been curated by NCBI staff. The

 reference sequence was derived from [AF064090.1](https://www.ncbi.nlm.nih.gov/nuccore/AF064090.1), [AY028261.1](https://www.ncbi.nlm.nih.gov/nuccore/AY028261.1),

 [AC008760.7](https://www.ncbi.nlm.nih.gov/nuccore/AC008760.7) and [AK026704.1](https://www.ncbi.nlm.nih.gov/nuccore/AK026704.1).

 On or before Apr 7, 2016 this sequence version replaced

 [XM\_005259670.2](https://www.ncbi.nlm.nih.gov/nuccore/XM_005259670.2), [NM\_172014.2](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.2).

 Summary: The protein encoded by this gene is a member of the tumor

 necrosis factor (TNF) ligand family. This protein is a ligand for

 TNFRSF14, which is a member of the tumor necrosis factor receptor

 superfamily, and which is also known as a herpesvirus entry

 mediator (HVEM). This protein may function as a costimulatory

 factor for the activation of lymphoid cells and as a deterrent to

 infection by herpesvirus. This protein has been shown to stimulate

 the proliferation of T cells, and trigger apoptosis of various

 tumor cells. This protein is also reported to prevent tumor

 necrosis factor alpha mediated apoptosis in primary hepatocyte. Two

 alternatively spliced transcript variant encoding distinct isoforms

 have been reported. [provided by RefSeq, Jul 2008].

 Sequence Note: This RefSeq record was created from transcript and

 genomic sequence data to make the sequence consistent with the

 reference genome assembly. The genomic coordinates used for the

 transcript record were based on transcript alignments.

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

 ##Evidence-Data-START##

 Transcript exon combination :: SRR1163658.48979.1, AY028261.1

 [ECO:0000332]

 RNAseq introns :: single sample supports all introns

 SAMEA2142586, SAMEA2144333

 [ECO:0000348]

 ##Evidence-Data-END##

 COMPLETENESS: complete on the 3' end.

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-80 AF064090.1 294-373

 81-620 AY028261.1 40-579

 621-2615 AC008760.7 98701-100695 c

 2616-4091 AK026704.1 442-1917

 4092-4268 AC008760.7 97048-97224 c

 4269-4406 AK026704.1 2095-2232

FEATURES Location/Qualifiers

 source 1..4406

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="19"

 /map="19p13.3"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=1&to=4406) 1..4406

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /note="TNF superfamily member 14"

 /db\_xref="GeneID:[8740](https://www.ncbi.nlm.nih.gov/gene/8740)"

 /db\_xref="HGNC:[HGNC:11930](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:11930)"

 /db\_xref="MIM:[604520](https://www.ncbi.nlm.nih.gov/omim/604520)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=1&to=200) 1..200

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /inference="alignment:Splign:2.1.0"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=21&to=23) 21..23

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /note="upstream in-frame stop codon"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=90&to=704) 90..704

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /note="isoform 2 is encoded by transcript variant 2; tumor

 necrosis factor (ligand) superfamily, member 14;

 herpesvirus entry mediator ligand; tumor necrosis factor

 superfamily member 14; tumor necrosis factor ligand 1D"

 /codon\_start=1

 /product="tumor necrosis factor ligand superfamily member

 14 isoform 2"

 /protein\_id="[NP\_742011.2](https://www.ncbi.nlm.nih.gov/protein/291045244)"

 /db\_xref="CCDS:[CCDS45939.1](https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?REQUEST=CCDS&DATA=CCDS45939.1)"

 /db\_xref="GeneID:[8740](https://www.ncbi.nlm.nih.gov/gene/8740)"

 /db\_xref="HGNC:[HGNC:11930](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:11930)"

 /db\_xref="MIM:[604520](https://www.ncbi.nlm.nih.gov/omim/604520)"

 /translation="MEESVVRPSVFVVDGQTDIPFTRLGRSHRRQSCSVARDGPAGSW

 EQLIQERRSHEVNPAAHLTGANSSLTGSGGPLLWETQLGLAFLRGLSYHDGALVVTKA

 GYYYIYSKVQLGGVGCPLGLASTITHGLYKRTPRYPEELELLVSQQSPCGRATSSSRV

 WWDSSFLGGVVHLEAGEKVVVRVLDERLVRLRDGTRSYFGAFMV"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=201&to=237) 201..237

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=238&to=279) 238..279

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=280&to=4377) 280..4377

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /inference="alignment:Splign:2.1.0"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=2467&to=2472) 2467..2472

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=2493&to=2493) 2493

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /note="major polyA site"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=4352&to=4357) 4352..4357

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3?from=4377&to=4377) 4377

 /gene="TNFSF14"

 /gene\_synonym="CD258; HVEML; LIGHT; LTg"

ORIGIN

 1 cgtgcacagc ccaggagtgt tgagcaattt cggtttcctc tgaggttgaa ggacccaggc

 61 gtgtcagccc tgctccagac accttgggca tggaggagag tgtcgtacgg ccctcagtgt

 121 ttgtggtgga tggacagacc gacatcccat tcacgaggct gggacgaagc caccggagac

 181 agtcgtgcag tgtggcccgg gacggacctg caggctcctg ggagcagctg atacaagagc

 241 gaaggtctca cgaggtcaac ccagcagcgc atctcacagg ggccaactcc agcttgaccg

 301 gcagcggggg gccgctgtta tgggagactc agctgggcct ggccttcctg aggggcctca

 361 gctaccacga tggggccctt gtggtcacca aagctggcta ctactacatc tactccaagg

 421 tgcagctggg cggtgtgggc tgcccgctgg gcctggccag caccatcacc cacggcctct

 481 acaagcgcac accccgctac cccgaggagc tggagctgtt ggtcagccag cagtcaccct

 541 gcggacgggc caccagcagc tcccgggtct ggtgggacag cagcttcctg ggtggtgtgg

 601 tacacctgga ggctggggag aaggtggtcg tccgtgtgct ggatgaacgc ctggttcgac

 661 tgcgtgatgg tacccggtct tacttcgggg ctttcatggt gtgaaggaag gagcgtggtg

 721 cattggacat gggtctgaca cgtggagaac tcagagggtg cctcagggga aagaaaactc

 781 acgaagcaga ggctgggcgt ggtggctctc gcctgtaatc ccagcacttt gggaggccaa

 841 ggcaggcgga tcacctgagg tcaggagttc gagaccagcc tggctaacat ggcaaaaccc

 901 catctctact aaaaatacaa aaattagccg gacgtggtgg tgcctgcctg taatccagct

 961 actcaggagg ctgaggcagg ataattttgc ttaaacccgg gaggcggagg ttgcagtgag

 1021 ccgagatcac accactgcac tccaacctgg gaaacgcagt gagactgtgc ctcaaaaaaa

 1081 agaaaggaag aaaaaagaaa actcaggaaa cagatcttgg gggacactcc agggaaccca

 1141 aaactcaaag gcggagagct cagtgggcac caccaaggcg agatgaagcc ccagcaggca

 1201 ccttcagaag acccacgtag actgcagacc ctgccacgga caatactaag gacaaaaacc

 1261 cagagacttg gggtctgtgg gcccccaaac atggggtaaa gttgatttgc ctgatattca

 1321 ggaagaaggg gtgaggggtg ggtatttatg cttttgattc agaagaaagt ggggcttggg

 1381 attccaggga cttggctggg ggtgggaaac ttcatccact tccctactct catcatgagt

 1441 acggacaggg tgggcgggag actgatcatc gggactcatc atgaagagcc cagccccacc

 1501 ccacatactc agatcccacc cacagactgg tggccacacc tcagcctggt cacaaagagt

 1561 tacactcaga tacatgagca cggcagcgtg ctcataactg tttaacaacc agctgtcctg

 1621 ggagggggac agctttgtaa tgtttgccaa tttccatggt gtaaatgcta ccaccatggc

 1681 tgatttcatc actgccaagc atagacatcc ctaataggac accacggatc tgtccccggc

 1741 atccggccca gggcctggca caaagcatgc tctagggaaa tgcttgctga ttgaaaggaa

 1801 ggaagaatga ctctacagtc acacctatgg catcccacaa aatctgtcac atggctgcat

 1861 aatctcagcc actctttcac aactatagac tcatacacgc gaagtgccag attcatgcac

 1921 aaccacacaa tcacatggaa gtcacagacg gcatcacaga cagtcacagc actgtgtgta

 1981 tgttataaca caagcacaca aaactcagac agcatcccag ctacacagcc actcccagag

 2041 gtgtcaccgt cacacttggt aattaatact cattacatta gacacagaca gaccaagtta

 2101 tagtcagacc tggttacaca catacacaca cacaatatca ccatgacaaa tacacattac

 2161 acacacacaa catcacaatg acaaacacac attacacaca caacatcacg atgacaaaca

 2221 cacattacac acacaacatc acgatgacaa acacacatta cacacacatc acaatgacaa

 2281 acacaacatt acacacacac aacatcacaa tgacacacac atcacacaca catcacaatg

 2341 acaaacacac aacattacac acatatacac acagcctgag ggccctcccc agcccagact

 2401 aacacatctc ggggtgagga ccagaccttg ttcataaccc tgggcctctt aaccactgat

 2461 ctttgaaata aatggcaaat agttgtacct ggatctgtct agttcttagg ggaacaaact

 2521 gaagaagggt ggagaggaat tgtcaggcct aaagagcccc acagggaaag ggaggagtcg

 2581 gatggggggc aaccatcagc aacaagtggt ggctcctaga ggcagaggga tggaggtaat

 2641 gacccatgga ggtcattcta cagatgagga acctggaccc agttggctca agtccatgca

 2701 ggaaatgtgg gggaaaccag agacctcacg tctggatctg gcttcctctc caatccacaa

 2761 ttcctgagga agtagaggct acatcccgca agacgccctt attagacaca tccaggacag

 2821 aatgacaatc cgccaagcca gctggaagca taaaacacag ggagctggtg ggttgggtgg

 2881 gggcagataa tgatatgcat acaaattaga gggtctatgc aaatgagcat tgctgcagtg

 2941 tggctggagg gaatccttag ttcctaggat tctaggatat gggtttcgac cccagaggtg

 3001 aatgtattgt tattattgtt ttgttgttgt tgtgaatgac aagtcaaaat ttgtgggtta

 3061 ttgttgttat cgccaatagt attcttgtca ttgttgcaca gtacagagat gaaggaaaca

 3121 gattttgcaa tcagatgatc ctgggttctg agtccactct gccactcacc agctatatga

 3181 cctccagcaa tttccatcac ctctcaatgc ttcagtttcc ccatcggcaa gatggttgtg

 3241 gggggagagg aacaacagta cagattcacc atcccaaatt caaaatgctc caaaatctag

 3301 gccgggcgtg gtggctcata cctgtaatcc cagcactttg ggaggtcaaa gtggacggat

 3361 aacctgaggt caggagctcc agaccagcct ggccaacatg gcgaaacccc atctctacta

 3421 aaaatacaaa aaattacctg ggtgtggtgg ggggcacctg taaccccagc tactcgggag

 3481 gctgaggcag gaaccctgga ggttgaggtt gcagtgagct gagatcacac cactgcactc

 3541 cagcctgggt gacagagcaa ggctcccatc tcaaaaaaca aaaaaacatg ctccaaaatc

 3601 tgaaactctt tgagccccag tgtgatgcca caagtgggaa attccacaac tcatcacatg

 3661 tgatagattg cagtggaaat gcaggcacac accacgaagt ttactcagca tcctcaaagg

 3721 aaatccccgt cagtagctat atatcatttt ctcacatgcc agataggtat ctctcatctt

 3781 ttactgttag gtacttctgt gttgaatagg tggaggaaaa tgattgctgg ttagtagtat

 3841 ataaattcag agtcaggaag gatggtgatg tcggctgggt gcagtggctc atgcctgtaa

 3901 ttccaatgtg ataccctacc ttgtgtttaa cgtgattgac tctcccttag ctgagagggc

 3961 caggcagact ctattttggc ttcttcgctt gcagtctctc acccaccccc cttcctcaag

 4021 gacttaagct gactcccagc acatccaaga atgcgattac tgataagata ctgtgacaag

 4081 ctatatccac aattcccagg aattcgtccg gttgatagca cccaaagccc ccgcgtctat

 4141 caccttgtga tagatttaaa gcccctgcac ctggaactgt ttgtttttct gttaccattt

 4201 atctttttca ctttcttgcc tgttttgctt ctgtaaaatt gcttcagctc ggctccctct

 4261 tccccttcta aaccaaggta taaaaagaaa cctagcccct tctttggggt ggagagaatt

 4321 ttgagcgcta gccgtctctc agtcgccggc taataaagga ctcctgaatt agtctaaaaa

 4381 aaaaaaaaaa aaaaaaaaaa aaaaaa

# Homo sapiens troponin I1, slow skeletal type (TNNI1), mRNA

NCBI Reference Sequence: NM\_003281.4

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4%22%20%5Cl%20%22goto1519243850_0)

LOCUS NM\_003281 6110 bp mRNA linear PRI 13-OCT-2020

DEFINITION Homo sapiens troponin I1, slow skeletal type (TNNI1), mRNA.

ACCESSION NM\_003281

VERSION NM\_003281.4

KEYWORDS RefSeq; MANE Select.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 6110)

 AUTHORS Nie SF, Yu M, Xie T, Yang F, Wang HB, Wang ZH, Li M, Gao XL, Lv BJ,

 Wang SJ, Zhang XB, He SL, Qiu ZH, Liao YH, Zhou ZH and Cheng X.

 TITLE Cardiac Troponin I Is an Independent Predictor for Mortality in

 Hospitalized Patients With COVID-19

 JOURNAL Circulation 142 (6), 608-610 (2020)

 PUBMED [32539541](https://www.ncbi.nlm.nih.gov/pubmed/32539541)

 REMARK GeneRIF: Cardiac Troponin I Is an Independent Predictor for

 Mortality in Hospitalized Patients With COVID-19.

REFERENCE 2 (bases 1 to 6110)

 AUTHORS Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, Brignall

 R, Cafarelli T, Campos-Laborie FJ, Charloteaux B, Choi D, Cote AG,

 Daley M, Deimling S, Desbuleux A, Dricot A, Gebbia M, Hardy MF,

 Kishore N, Knapp JJ, Kovacs IA, Lemmens I, Mee MW, Mellor JC,

 Pollis C, Pons C, Richardson AD, Schlabach S, Teeking B, Yadav A,

 Babor M, Balcha D, Basha O, Bowman-Colin C, Chin SF, Choi SG,

 Colabella C, Coppin G, D'Amata C, De Ridder D, De Rouck S,

 Duran-Frigola M, Ennajdaoui H, Goebels F, Goehring L, Gopal A,

 Haddad G, Hatchi E, Helmy M, Jacob Y, Kassa Y, Landini S, Li R, van

 Lieshout N, MacWilliams A, Markey D, Paulson JN, Rangarajan S,

 Rasla J, Rayhan A, Rolland T, San-Miguel A, Shen Y, Sheykhkarimli

 D, Sheynkman GM, Simonovsky E, Tasan M, Tejeda A, Tropepe V,

 Twizere JC, Wang Y, Weatheritt RJ, Weile J, Xia Y, Yang X,

 Yeger-Lotem E, Zhong Q, Aloy P, Bader GD, De Las Rivas J, Gaudet S,

 Hao T, Rak J, Tavernier J, Hill DE, Vidal M, Roth FP and Calderwood

 MA.

 TITLE A reference map of the human binary protein interactome

 JOURNAL Nature 580 (7803), 402-408 (2020)

 PUBMED [32296183](https://www.ncbi.nlm.nih.gov/pubmed/32296183)

REFERENCE 3 (bases 1 to 6110)

 AUTHORS de Almeida Thiengo D, Strogoff-de-Matos JP, Lugon JR and Graciano

 ML.

 TITLE Troponin I at admission in the intensive care unit predicts the

 need of dialysis in septic patients

 JOURNAL BMC Nephrol 19 (1), 329 (2018)

 PUBMED [30453890](https://www.ncbi.nlm.nih.gov/pubmed/30453890)

 REMARK GeneRIF: Troponin I levels at intensive care unit admission are a

 strong independent predictor of dialysis needs in sepsis.

 Publication Status: Online-Only

REFERENCE 4 (bases 1 to 6110)

 AUTHORS Kaess BM, de Las Heras Gala T, Zierer A, Meisinger C, Wahl S,

 Peters A, Todd J, Herder C, Huth C, Thorand B and Koenig W.

 TITLE Ultra-sensitive troponin I is an independent predictor of incident

 coronary heart disease in the general population

 JOURNAL Eur J Epidemiol 32 (7), 583-591 (2017)

 PUBMED [28585121](https://www.ncbi.nlm.nih.gov/pubmed/28585121)

 REMARK GeneRIF: Ultrasensitive troponin I was detectable in almost all

 individuals of a study sample reflecting middle-aged to elderly

 European general population. Ultrasensitive troponin concentrations

 exhibit an independent, graded, positive relation with incident CHD

 [coronary heart disease].

REFERENCE 5 (bases 1 to 6110)

 AUTHORS Shafi A, Siddiqui N, Imtiaz S and Din Sajid MU.

 TITLE Left Ventricular Systolic Dysfunction Predicted By Early Troponin I

 Release After Anthracycline Based Chemotherapy In Breast Cancer

 Patients

 JOURNAL J Ayub Med Coll Abbottabad 29 (2), 266-269 (2017)

 PUBMED [28718245](https://www.ncbi.nlm.nih.gov/pubmed/28718245)

 REMARK GeneRIF: Studied use of serum levels of Troponin I as a predictive

 biomarker for diagnosis of left ventricular systolic dysfunction

 after anthracycline treatment in breast cancer.

REFERENCE 6 (bases 1 to 6110)

 AUTHORS Corin SJ, Juhasz O, Zhu L, Conley P, Kedes L and Wade R.

 TITLE Structure and expression of the human slow twitch skeletal muscle

 troponin I gene

 JOURNAL J Biol Chem 269 (14), 10651-10659 (1994)

 PUBMED [8144655](https://www.ncbi.nlm.nih.gov/pubmed/8144655)

REFERENCE 7 (bases 1 to 6110)

 AUTHORS Bhavsar PK, Dhoot GK, Cumming DV, Butler-Browne GS, Yacoub MH and

 Barton PJ.

 TITLE Developmental expression of troponin I isoforms in fetal human

 heart

 JOURNAL FEBS Lett 292 (1-2), 5-8 (1991)

 PUBMED [1959627](https://www.ncbi.nlm.nih.gov/pubmed/1959627)

REFERENCE 8 (bases 1 to 6110)

 AUTHORS Hunkeler NM, Kullman J and Murphy AM.

 TITLE Troponin I isoform expression in human heart

 JOURNAL Circ Res 69 (5), 1409-1414 (1991)

 PUBMED [1934363](https://www.ncbi.nlm.nih.gov/pubmed/1934363)

REFERENCE 9 (bases 1 to 6110)

 AUTHORS Wade R, Eddy R, Shows TB and Kedes L.

 TITLE cDNA sequence, tissue-specific expression, and chromosomal mapping

 of the human slow-twitch skeletal muscle isoform of troponin I

 JOURNAL Genomics 7 (3), 346-357 (1990)

 PUBMED [2365354](https://www.ncbi.nlm.nih.gov/pubmed/2365354)

REFERENCE 10 (bases 1 to 6110)

 AUTHORS Suzuki H, Kawarabayasi Y, Kondo J, Abe T, Nishikawa K, Kimura S,

 Hashimoto T and Yamamoto T.

 TITLE Structure and regulation of rat long-chain acyl-CoA synthetase

 JOURNAL J Biol Chem 265 (15), 8681-8685 (1990)

 PUBMED [2341402](https://www.ncbi.nlm.nih.gov/pubmed/2341402)

COMMENT REVIEWED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has been curated by NCBI staff. The

 reference sequence was derived from [BC012600.1](https://www.ncbi.nlm.nih.gov/nuccore/BC012600.1), [AC096677.2](https://www.ncbi.nlm.nih.gov/nuccore/AC096677.2) and

 [AL832006.2](https://www.ncbi.nlm.nih.gov/nuccore/AL832006.2).

 This sequence is a reference standard in the [RefSeqGene](https://www.ncbi.nlm.nih.gov/refseq/rsg/) project.

 On Nov 22, 2018 this sequence version replaced [NM\_003281.3](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.3).

 Summary: Troponin proteins associate with tropomyosin and regulate

 the calcium sensitivity of the myofibril contractile apparatus of

 striated muscles. Troponin I (TnI), along with troponin T (TnT) and

 troponin C (TnC), is one of 3 subunits that form the troponin

 complex of the thin filaments of striated muscle. TnI is the

 inhibitory subunit; blocking actin-myosin interactions and thereby

 mediating striated muscle relaxation. The TnI subfamily contains

 three genes: TnI-skeletal-fast-twitch, TnI-skeletal-slow-twitch,

 and TnI-cardiac. The TnI-fast and TnI-slow genes are expressed in

 fast-twitch and slow-twitch skeletal muscle fibers, respectively,

 while the TnI-cardiac gene is expressed exclusively in cardiac

 muscle tissue. This gene encodes the

 Troponin-I-skeletal-slow-twitch protein. This gene is expressed in

 cardiac and skeletal muscle during early development but is

 restricted to slow-twitch skeletal muscle fibers in adults. The

 encoded protein prevents muscle contraction by inhibiting

 calcium-mediated conformational changes in actin-myosin complexes.

 [provided by RefSeq, Jul 2008].

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

 ##Evidence-Data-START##

 Transcript exon combination :: AL831975.1, BX510903.2 [ECO:0000332]

 RNAseq introns :: single sample supports all introns

 SAMEA2162946, SAMN03267751

 [ECO:0000348]

 ##Evidence-Data-END##

 ##RefSeq-Attributes-START##

 MANE Ensembl match :: ENST00000361379.9/ ENSP00000354488.4

 RefSeq Select criteria :: based on single protein-coding transcript

 ##RefSeq-Attributes-END##

 COMPLETENESS: full length.

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-1029 BC012600.1 5-1033

 1030-5522 AC096677.2 2370-6862 c

 5523-6110 AL832006.2 4239-4826

FEATURES Location/Qualifiers

 source 1..6110

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="1"

 /map="1q32.1"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=1&to=6110) 1..6110

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="troponin I1, slow skeletal type"

 /db\_xref="GeneID:[7135](https://www.ncbi.nlm.nih.gov/gene/7135)"

 /db\_xref="HGNC:[HGNC:11945](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:11945)"

 /db\_xref="MIM:[191042](https://www.ncbi.nlm.nih.gov/omim/191042)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=1&to=58) 1..58

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=51&to=53) 51..53

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="upstream in-frame stop codon"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=59&to=88) 59..88

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=78&to=641) 78..641

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="troponin I type 1 (skeletal, slow)"

 /codon\_start=1

 /product="troponin I, slow skeletal muscle"

 /protein\_id="[NP\_003272.3](https://www.ncbi.nlm.nih.gov/protein/56682969)"

 /db\_xref="CCDS:[CCDS1411.1](https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?REQUEST=CCDS&DATA=CCDS1411.1)"

 /db\_xref="GeneID:[7135](https://www.ncbi.nlm.nih.gov/gene/7135)"

 /db\_xref="HGNC:[HGNC:11945](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:11945)"

 /db\_xref="MIM:[191042](https://www.ncbi.nlm.nih.gov/omim/191042)"

 /translation="MPEVERKPKITASRKLLLKSLMLAKAKECWEQEHEEREAEKVRY

 LAERIPTLQTRGLSLSALQDLCRELHAKVEVVDEERYDIEAKCLHNTREIKDLKLKVM

 DLRGKFKRPPLRRVRVSADAMLRALLGSKHKVSMDLRANLKSVKKEDTEKERPVEVGD

 WRKNVEAMSGMEGRKKMFDAAKSPTSQ"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_003272.3?from=2&to=48) 81..221

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="propagated from UniProtKB/Swiss-Prot (P19237.3);

 Region: Involved in binding TNC"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_003272.3?from=2&to=2) 81..83

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="N-acetylproline.

 /evidence=ECO:0000250|UniProtKB:P02645; propagated from

 UniProtKB/Swiss-Prot (P19237.3); acetylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_003272.3?from=58&to=58) 249..251

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="Phosphoserine.

 /evidence=ECO:0000250|UniProtKB:Q9WUZ5; propagated from

 UniProtKB/Swiss-Prot (P19237.3); phosphorylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_003272.3?from=97&to=118) 366..431

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="propagated from UniProtKB/Swiss-Prot (P19237.3);

 Region: Involved in binding TNC and actin"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=89&to=92) 89..92

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=93&to=134) 93..134

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=135&to=266) 135..266

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=267&to=356) 267..356

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=357&to=533) 357..533

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=534&to=643) 534..643

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=644&to=6110) 644..6110

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /inference="alignment:Splign:2.1.0"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=1009&to=1014) 1009..1014

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="hexamer: ATTAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=1032&to=1032) 1032

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="major polyA site"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=3555&to=3560) 3555..3560

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="hexamer: AATAAT"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=3579&to=3579) 3579

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=5377&to=5382) 5377..5382

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=5397&to=5397) 5397

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=6090&to=6095) 6090..6095

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

 /note="hexamer: AATACA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4?from=6110&to=6110) 6110

 /gene="TNNI1"

 /gene\_synonym="SSTNI; TNN1"

ORIGIN

 1 acagtctgca gtctacggcg aggcacaggc cagcccagct ccacgaggac tgaacaaggt

 61 gctgtctcac tgccaccatg ccggaagtcg agagaaaacc caagatcact gcctcccgca

 121 aactcttgct gaagagcctg atgctggcca aggccaagga atgctgggag caggagcacg

 181 aggagcgcga ggctgagaag gtgcgctacc tggcagagcg catccccacg ctgcagaccc

 241 gtggcctgtc cctcagtgcc ctgcaggacc tgtgccggga gctgcacgcc aaggtggagg

 301 tggtggatga ggagcgatac gacattgagg ccaaatgcct ccacaacacc agggagatta

 361 aggacctgaa gctgaaggtg atggacctcc gtgggaagtt caagcgcccg cccctgcgtc

 421 gagtccgtgt ctcggctgac gccatgctcc gggccctgct gggctccaag cacaaggtgt

 481 ccatggatct gcgggccaac ctcaagtctg tgaagaagga agacacagag aaggagcggc

 541 ctgtggaggt gggtgactgg aggaagaacg tggaggccat gtctggcatg gaaggccgga

 601 agaagatgtt tgatgccgcc aagtctccga cctcacaata gaggccagct tgctgtgctg

 661 cgctctgagc tcctgcttca tgcttcttct ccaacccagc tcactcacct ctctgcctgt

 721 gtctggagca tcccttccca cctctccccc acttcttccc tccagcctgc atgccctcct

 781 ctggaactgg gattaaacag atacccaaga ggcaggaaaa acgcgagtgt ctgaagacct

 841 caccctggca gggtgggctt ctatccaccc atgaggcacc tggggcttct cagccacccc

 901 gtggtgtaac aagtggagga ggaggagggg gtgagaacac ctttcctcca gccctgtgcc

 961 acactccttc tatgtgtgta cagtgtccac aggcatcaca ccttcttcat taaaaaccca

 1021 gtttcctggt catttctgca gctctgactg tgagccctga gaagaaaggg aagacttggg

 1081 gcttgggaag agagggctgt gctgggctcc agtcaccctg tctgcagtcc ggagctcggg

 1141 atctgaaggg agggctaagg aattggtggc tgcttgccag ggtcccctgg ggccccttct

 1201 gcccagtccc tgtgacgtca cactgcacac actcaccacc caggatgctc agctcaccag

 1261 agccctgacc tccgagaaat gaccccgtcc tgcccctgac tctggccatc agaagaggcc

 1321 tctccattaa gagccgcttc tcctgccggc tttcatactc tccttcctct gcaaacctgg

 1381 attaactctc cctgatgtgt gctgggttga gaggtcaggg cactttcaat ccctcaccta

 1441 gtcctgctgc acagggcagg aggcctcaca gaagagatac ttgcagttcc aggcctggcc

 1501 accaggtgac gccttccttc tctcctgggc tctcctgtcc ctccggcctc tgctctcccc

 1561 acgttcaccc accctcctcc caattgtttc cccttcacac tcctgcactg gtcctatccc

 1621 ccaggactga ttctttcccg gagatggggt ggccacttat ctgtcctctg gatcaggtcc

 1681 tttctccatc ccactacctc ctgggtccaa agcctggtga gagcataggg ccttctgggg

 1741 tggaaggggg gtctggggca gcctctcttc ccccaacata ctcactagag tggctcctgt

 1801 gaagcagaga ccaaagttct gacttttctt tttccttttt tttttttttt tgagatggag

 1861 tctcgctgtg tcacccaggc tggagtgcag tgacgcgatc tcgactcact gcaacctccg

 1921 cctcccaggt tcaagcgatt ctcctgcttc aacctcctga gtagctggga ttacaggtgc

 1981 gtgccaccac accccgctaa tttttgtatt tttagtagag acagggtttc accatgttgg

 2041 tcaggctggt ctcgaactcc tgacctcgta atctgcccac ctcggcctcc caaagttctg

 2101 ggattacagg tgtgagcgac cacatccggt caagttctga cttttcaggg cagtgggatg

 2161 gtgacttctg agctaagcct gcaagggaag tggctgttcc cagacctggg gtcccttggc

 2221 ttttcgacca ccctagcact ggccctagga ggggctgagt tatcagagca gcaacccttc

 2281 tacccaccct actgagtccc cattctacct tcgaatccca tgtattatct tccataagtg

 2341 ccagaaattg gtacaaagtg cagtgccagc tgctccccag ggcatggtct ggctctctgg

 2401 gctgctgggc tgccagcccc tttcagaaag catggcagtg ggaaccagca acttgggcta

 2461 ccttagggga agcagcactc tgcccagggt tcttgctata aataggtcag tgggcagaga

 2521 ggcagatgtg tgcaaaaagc acaaccccaa tcacacacag aaacccgtag gcacggacac

 2581 acccagtctc aatgcacgaa gttggagaac ctgacagggc tggtccaggc cacagaggat

 2641 ggggttaggc ttcatgcctt catgcccagg cagccacaga aacccactct tccccgcctc

 2701 acagggctcc tccggttggc ctggatccaa agcgatctcc ttgtttcacc tctttccagc

 2761 tcttccctag agcctggccc caagttccag agtcaaatcc cagttttgca gtcaactatt

 2821 aacctggggc cctggcagat tggggtatcc acccagtcat gtaaaggcag gagacctgct

 2881 ctgggcctga ctcaaagcaa acacgtgggt gcagaagacc catgggagga ggaggtgggg

 2941 cccactcaac acacctaagc cagtggccct ctaggagggg cagaagagcc aaccagaggg

 3001 ttcctggagg tcatcgggtt cctccttctg cttccaggca gcaacagacc taattcatct

 3061 cctgtgagga tggagcaagg gcgaatatag gcatagacat atctggagcg caaggattga

 3121 gaggccctgt ggacaagggg aaggagagct tgggggtggg agagggatct atggaagggg

 3181 ctgttgtagc caagaatggg gcctgggctg ggggcctggc ctctggagag ccctgattct

 3241 actccagtct tctctcacca tggagacaac accagtagag ctgggagctg gggcctgcca

 3301 ctgagtaaca gctcacacat gtagcacact tatgcgtcag gcaccagcat gtacacctta

 3361 atttatttaa ctttcacaac aaccctagga gaaaggtaca gctactattc tcatggtaca

 3421 gtggggaaaa ctgaggcaca gagaggttag ggacttttca tgatcacaca gctaataagt

 3481 ggtggtccca ggattcaacc caagaagtct gttagaatcc atactcttat cgactatttt

 3541 aataacagca tttcaataat agtaatatgt gatgttaaat gctttatctg agttatatca

 3601 ttgcattctc acaaaaatgg tacatttgtt agctcaaatt tacagatgag gaaatggagg

 3661 cctggagagt tcaagtagct cactctaagt tacagggttt gctagctgca gactggagac

 3721 tagatctcag gtcactctgc cttggggttg cccttttctg agatcagact ggccttaaca

 3781 gagcccaact ttcttggaag gacacagctg ggggtcgagg gtcctgcata tgcacctggg

 3841 gagtccccac acacccgcac caaaggggaa acattcaaag gagatgcagg tgacatacag

 3901 atgggctaag gtgaggcctt cctccaacat ggggccctgg ctgagagtgg agggataggg

 3961 tgctacttgg gggatatcta ttactaatca gcacacaata gctgggtggg caccccctct

 4021 ctctgcagct ggtggagatg ggcctggtgc tcctctctgc ttggtaggaa tgaccagggc

 4081 aggatgccac ctgctacact cttgcctgtc ccagagccct tgggagacca gctgttgctc

 4141 ccacctttcc ccacagggtg cccggtcccc acggcgcaag gaggccaagc tcagctgtgt

 4201 gtctgagccc actcagcagc aggcatgatc agggcagggc ccagcgctaa gggctgctgg

 4261 ggctaaagtg ccttgctgaa ggaggggcag agttaaacat tagccacaat ggccccccgc

 4321 gaagtgagtc atactcccca caagtctctg ggctgggcag ccgggggcgg gggggttagc

 4381 ttcgttccag gtgggtgacc aagctgcgga agcccaggac agttcgctgg gaccagtgct

 4441 aatggcagct tctgctcctg acccccatca ggtctctttc ctcagacatc cccaaagctt

 4501 gtctgcctgt gggtctttca ttaaaggtga cgggcaaggg ttttctcttc cttccaagac

 4561 acacttctag gctggcttgt gggtcccctg agacgaaggc tgagcagggt gagcctcgaa

 4621 aaccccatag cactggccca gagcactgcg gtcaatgtcc gaaggaccag agagggaggg

 4681 tgacttgctc aaggtcacaa acctaactgg agctaggaag taggtcccca gactcccagt

 4741 ctgaggaaat tggatgtgga cagggctttg gtctcctggc cagaggcaaa cgcccaggga

 4801 tctgaacagg ctgcccatct ctggcctcaa ctcagcaccc caactggctc cagtgccaac

 4861 cattggcctc tctcactggt tggcacaggc tgggtgagcc tggccctttg ccctgggttc

 4921 aggagagagg tatggctttg cctggcagtc aggggaggtg ttccaacctg ggagcccctg

 4981 ggggaggagc gtgtgtgtga acgcttgtgc ctgttcatca gaaggggtga gcatggtgtg

 5041 tatgcgtggg catgcagaat gcagggtgat gacctgtctg cccgccctga gtgtcacaga

 5101 cacctgtgtg tcccctgtac tagtgcctcc cagatgctct tctcctgact ggctctttgt

 5161 gacctaaatt tcaactgctc caccccacaa ggaggaaacc agggccacaa atgtagtgag

 5221 attttcataa agctaaatgt attcagtgga aacgaccttc ctctgatcta aaagcatgca

 5281 tgcctttgtt ttattactgg agtattaaat tagcccttac ttatggaata ggaggacagt

 5341 aggtggcaat tgttttctaa tgctcttact tcacaaaata aaatggtggc agcacaatgt

 5401 cagtatttat attacacaat attattagtt tgtaaaaacc gaaagccagg gagcttactg

 5461 ccctggattc tttagggcgt gggtcttatc acctcattca tttcttccct tcccctaaga

 5521 gtgccacaag catggcagtt gctcagctga gacacatgaa tgggaacaga actgaaatat

 5581 caggatgtga cctcgagcca gctggttggg gctggggatc cccctggatc tccactctgt

 5641 cctgtccgag aaaggggaag ctgccctatg agtctttttt ctgggggatt ccagaggttg

 5701 aacgtctgtt ctgaggctcc tcagaacctc gagggagggg acgagagaga tgaccccttg

 5761 cccaaccact ggctgaaaaa tcttgaaatt ccctgacttt gaccagggaa ttcccagttg

 5821 tgctaagccc tgcctggtgg aatgtccctg gaggaagtaa ccaggacacg catacgtttt

 5881 taatgtacgt ataaagaatc gtaaagtgaa cacccatctg ccaaccactc attttaagaa

 5941 ctaggatatg gcacatactt ttgaagccct tatgcgtccc tccccaatca caaccccttc

 6001 ctctccacaa aggtaaatac aatcctcaat tttgtgtttg ttattctctt ggttctcttc

 6061 ataatgttac tacgtattta tgaattccta atacatcatt cagttttgga

# Homo sapiens solute carrier family 6 member 20 (SLC6A20), transcript variant 1, mRNA

NCBI Reference Sequence: NM\_020208.4

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4%22%20%5Cl%20%22goto1519314525_0)

LOCUS NM\_020208 5425 bp mRNA linear PRI 12-DEC-2020

DEFINITION Homo sapiens solute carrier family 6 member 20 (SLC6A20),

 transcript variant 1, mRNA.

ACCESSION NM\_020208

VERSION NM\_020208.4

KEYWORDS RefSeq; MANE Select.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5425)

 AUTHORS Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A,

 Invernizzi P, Fernandez J, Prati D, Baselli G, Asselta R, Grimsrud

 MM, Milani C, Aziz F, Kassens J, May S, Wendorff M, Wienbrandt L,

 Uellendahl-Werth F, Zheng T, Yi X, de Pablo R, Chercoles AG, Palom

 A, Garcia-Fernandez AE, Rodriguez-Frias F, Zanella A, Bandera A,

 Protti A, Aghemo A, Lleo A, Biondi A, Caballero-Garralda A, Gori A,

 Tanck A, Carreras Nolla A, Latiano A, Fracanzani AL, Peschuck A,

 Julia A, Pesenti A, Voza A, Jimenez D, Mateos B, Nafria Jimenez B,

 Quereda C, Paccapelo C, Gassner C, Angelini C, Cea C, Solier A,

 Pestana D, Muniz-Diaz E, Sandoval E, Paraboschi EM, Navas E, Garcia

 Sanchez F, Ceriotti F, Martinelli-Boneschi F, Peyvandi F, Blasi F,

 Tellez L, Blanco-Grau A, Hemmrich-Stanisak G, Grasselli G,

 Costantino G, Cardamone G, Foti G, Aneli S, Kurihara H, ElAbd H, My

 I, Galvan-Femenia I, Martin J, Erdmann J, Ferrusquia-Acosta J,

 Garcia-Etxebarria K, Izquierdo-Sanchez L, Bettini LR, Sumoy L,

 Terranova L, Moreira L, Santoro L, Scudeller L, Mesonero F, Roade

 L, Ruhlemann MC, Schaefer M, Carrabba M, Riveiro-Barciela M,

 Figuera Basso ME, Valsecchi MG, Hernandez-Tejero M, Acosta-Herrera

 M, D'Angio M, Baldini M, Cazzaniga M, Schulzky M, Cecconi M, Wittig

 M, Ciccarelli M, Rodriguez-Gandia M, Bocciolone M, Miozzo M,

 Montano N, Braun N, Sacchi N, Martinez N, Ozer O, Palmieri O,

 Faverio P, Preatoni P, Bonfanti P, Omodei P, Tentorio P, Castro P,

 Rodrigues PM, Blandino Ortiz A, de Cid R, Ferrer R, Gualtierotti R,

 Nieto R, Goerg S, Badalamenti S, Marsal S, Matullo G, Pelusi S,

 Juzenas S, Aliberti S, Monzani V, Moreno V, Wesse T, Lenz TL,

 Pumarola T, Rimoldi V, Bosari S, Albrecht W, Peter W, Romero-Gomez

 M, D'Amato M, Duga S, Banales JM, Hov JR, Folseraas T, Valenti L,

 Franke A and Karlsen TH.

 CONSRTM Severe Covid-19 GWAS Group

 TITLE Genomewide Association Study of Severe Covid-19 with Respiratory

 Failure

 JOURNAL N Engl J Med 383 (16), 1522-1534 (2020)

 PUBMED [32558485](https://www.ncbi.nlm.nih.gov/pubmed/32558485)

REFERENCE 2 (bases 1 to 5425)

 AUTHORS Xie X, He Q, Huang L, Li L, Yao Y, Xia H, Zhao J, Zhong W and Zhang

 Y.

 TITLE Associations of SLC6A20 genetic polymorphisms with Hirschsprung's

 disease in a Southern Chinese population

 JOURNAL Biosci Rep 39 (8) (2019)

 PUBMED [31358688](https://www.ncbi.nlm.nih.gov/pubmed/31358688)

 REMARK GeneRIF: Associations of SLC6A20 genetic polymorphisms with

 Hirschsprung's disease in a Southern Chinese population.

 Publication Status: Online-Only

REFERENCE 3 (bases 1 to 5425)

 AUTHORS Sweeney MD, Zhao Z, Montagne A, Nelson AR and Zlokovic BV.

 TITLE Blood-Brain Barrier: From Physiology to Disease and Back

 JOURNAL Physiol Rev 99 (1), 21-78 (2019)

 PUBMED [30280653](https://www.ncbi.nlm.nih.gov/pubmed/30280653)

 REMARK Review article

REFERENCE 4 (bases 1 to 5425)

 AUTHORS Meier C, Camargo SM, Hunziker S, Moehrlen U, Gros SJ, Bode P, Leu

 S, Meuli M, Holland-Cunz S, Verrey F and Vuille-Dit-Bille RN.

 TITLE Intestinal IMINO transporter SIT1 is not expressed in human

 newborns

 JOURNAL Am J Physiol Gastrointest Liver Physiol 315 (5), G887-G895 (2018)

 PUBMED [30160974](https://www.ncbi.nlm.nih.gov/pubmed/30160974)

 REMARK GeneRIF: SIT1 is not expressed in small intestine of human

 newborns.

REFERENCE 5 (bases 1 to 5425)

 AUTHORS Lee JS, Oh JT, Kim JH, Seo JM, Kim DY, Park KW, Kim HY, Jung K,

 Park BL, Koh I and Shin HD.

 TITLE Association Analysis of SLC6A20 Polymorphisms With Hirschsprung

 Disease

 JOURNAL J Pediatr Gastroenterol Nutr 62 (1), 64-70 (2016)

 PUBMED [26049783](https://www.ncbi.nlm.nih.gov/pubmed/26049783)

 REMARK GeneRIF: Imputed meta-analysis revealed that 13 SLC6A20 SNPs were

 significantly associated with Hirschsprung disease. In further

 subgroup analysis, SLC6A20 polymorphisms appeared to have increased

 associations with Long-Segment Hirschsprung disease.

REFERENCE 6 (bases 1 to 5425)

 AUTHORS Vuille-dit-Bille RN, Camargo SM, Emmenegger L, Sasse T, Kummer E,

 Jando J, Hamie QM, Meier CF, Hunziker S, Forras-Kaufmann Z, Kuyumcu

 S, Fox M, Schwizer W, Fried M, Lindenmeyer M, Gotze O and Verrey F.

 TITLE Human intestine luminal ACE2 and amino acid transporter expression

 increased by ACE-inhibitors

 JOURNAL Amino Acids 47 (4), 693-705 (2015)

 PUBMED [25534429](https://www.ncbi.nlm.nih.gov/pubmed/25534429)

 REMARK GeneRIF: SIT1, B(0)AT1 and ACE2 were co-localized in the

 brush-border membrane of small intestine enterocytes.

REFERENCE 7 (bases 1 to 5425)

 AUTHORS Kanei-Ishii C, Nomura T, Tanikawa J, Ichikawa-Iwata E and Ishii S.

 TITLE Differential sensitivity of v-Myb and c-Myb to Wnt-1-induced

 protein degradation

 JOURNAL J Biol Chem 279 (43), 44582-44589 (2004)

 PUBMED [15308626](https://www.ncbi.nlm.nih.gov/pubmed/15308626)

REFERENCE 8 (bases 1 to 5425)

 AUTHORS Kiss H, Kedra D, Kiss C, Kost-Alimova M, Yang Y, Klein G, Imreh S

 and Dumanski JP.

 TITLE The LZTFL1 gene is a part of a transcriptional map covering 250 kb

 within the common eliminated region 1 (C3CER1) in 3p21.3

 JOURNAL Genomics 73 (1), 10-19 (2001)

 PUBMED [11352561](https://www.ncbi.nlm.nih.gov/pubmed/11352561)

REFERENCE 9 (bases 1 to 5425)

 AUTHORS Nash SR, Giros B, Kingsmore SF, Kim KM, el-Mestikawy S, Dong Q,

 Fumagalli F, Seldin MF and Caron MG.

 TITLE Cloning, gene structure and genomic localization of an orphan

 transporter from mouse kidney with six alternatively-spliced

 isoforms

 JOURNAL Recept Channels 6 (2), 113-128 (1998)

 PUBMED [9932288](https://www.ncbi.nlm.nih.gov/pubmed/9932288)

REFERENCE 10 (bases 1 to 5425)

 AUTHORS Stevens,B.R. and Wright,E.M.

 TITLE Kinetics of the intestinal brush border proline (Imino) carrier

 JOURNAL J Biol Chem 262 (14), 6546-6551 (1987)

 PUBMED [3571270](https://www.ncbi.nlm.nih.gov/pubmed/3571270)

 REMARK GeneRIF: Characterization and substrate specificity of the Na+

 coupled IMINO transport system in apical brush border membranes of

 epithelial cells. Identically found in kidney proximal tubule.

COMMENT REVIEWED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has been curated by NCBI staff. The

 reference sequence was derived from [DR006419.1](https://www.ncbi.nlm.nih.gov/nuccore/DR006419.1), [AJ276207.1](https://www.ncbi.nlm.nih.gov/nuccore/AJ276207.1),

 [AC098476.2](https://www.ncbi.nlm.nih.gov/nuccore/AC098476.2) and [EL949345.1](https://www.ncbi.nlm.nih.gov/nuccore/EL949345.1).

 This sequence is a reference standard in the [RefSeqGene](https://www.ncbi.nlm.nih.gov/refseq/rsg/) project.

 On Nov 23, 2018 this sequence version replaced [NM\_020208.3](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.3).

 Summary: Transport of small hydrophilic substances across cell

 membranes is mediated by substrate-specific transporter proteins

 which have been classified into several families of related genes.

 The protein encoded by this gene belongs to the

 sodium:neurotransmitter symporter (SNF) family and functions as a

 proline transporter expressed in kidney and small intestine.

 Mutations in this gene are associated with Hyperglycinuria and

 Iminoglycinuria. [provided by RefSeq, Jul 2020].

 Transcript Variant: This variant (1) encodes isoform 1.

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

 ##Evidence-Data-START##

 Transcript exon combination :: AJ276207.1, BC126197.1 [ECO:0000332]

 RNAseq introns :: single sample supports all introns

 SAMEA2144835 [ECO:0000348]

 ##Evidence-Data-END##

 ##RefSeq-Attributes-START##

 MANE Ensembl match :: ENST00000358525.9/ ENSP00000346298.4

 RefSeq Select criteria :: based on conservation, expression

 ##RefSeq-Attributes-END##

 COMPLETENESS: full length.

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-66 DR006419.1 15-80

 67-533 AJ276207.1 1-467

 534-534 AC098476.2 126599-126599 c

 535-2494 AJ276207.1 469-2428

 2495-4904 AC098476.2 106643-109052 c

 4905-5425 EL949345.1 9-529

FEATURES Location/Qualifiers

 source 1..5425

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="3"

 /map="3p21.31"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=1&to=5425) 1..5425

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="solute carrier family 6 member 20"

 /db\_xref="GeneID:[54716](https://www.ncbi.nlm.nih.gov/gene/54716)"

 /db\_xref="HGNC:[HGNC:30927](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:30927)"

 /db\_xref="MIM:[605616](https://www.ncbi.nlm.nih.gov/omim/605616)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=1&to=238) 1..238

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=118&to=1896) 118..1896

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="isoform 1 is encoded by transcript variant 1; X

 transporter protein 3; orphan transporter XT3;

 neurotransmitter transporter RB21A; solute carrier family

 6 (neurotransmitter transporter), member 20; solute

 carrier family 6 (proline IMINO transporter), member 20;

 sodium/imino-acid transporter 1; transporter rB21A

 homolog; sodium- and chloride-dependent transporter XTRP3;

 sodium-dependent imino acid transporter 1"

 /codon\_start=1

 /product="sodium- and chloride-dependent transporter XTRP3

 isoform 1"

 /protein\_id="[NP\_064593.1](https://www.ncbi.nlm.nih.gov/protein/11181770)"

 /db\_xref="CCDS:[CCDS43077.1](https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?REQUEST=CCDS&DATA=CCDS43077.1)"

 /db\_xref="GeneID:[54716](https://www.ncbi.nlm.nih.gov/gene/54716)"

 /db\_xref="HGNC:[HGNC:30927](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:30927)"

 /db\_xref="MIM:[605616](https://www.ncbi.nlm.nih.gov/omim/605616)"

 /translation="MEKARPLWANSLQFVFACISYAVGLGNVWRFPYLCQMYGGGSFL

 VPYIIMLIVEGMPLLYLELAVGQRMRQGSIGAWRTISPYLSGVGVASVVVSFFLSMYY

 NVINAWAFWYLFHSFQDPLPWSVCPLNGNHTGYDEECEKASSTQYFWYRKTLNISPSL

 QENGGVQWEPALCLLLAWLVVYLCILRGTESTGKVVYFTASLPYCVLIIYLIRGLTLH

 GATNGLMYMFTPKIEQLANPKAWINAATQIFFSLGLGFGSLIAFASYNEPSNNCQKHA

 IIVSLINSFTSIFASIVTFSIYGFKATFNYENCLKKVSLLLTNTFDLEDGFLTASNLE

 QVKGYLASAYPSKYSEMFPQIKNCSLESELDTAVQGTGLAFIVYTEAIKNMEVSQLWS

 VLYFFMLLMLGIGSMLGNTAAILTPLTDSKIISSHLPKEAISGLVCLVNCAIGMVFTM

 EAGNYWFDIFNDYAATLSLLLIVLVETIAVCYVYGLRRFESDLKAMTGRAVSWYWKVM

 WAGVSPLLIVSLFVFYLSDYILTGTLKYQAWDASQGQLVTKDYPAYALAVIGLLVASS

 TMCIPLAALGTFVQRRLKRGDADPVA"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=6&to=26) 133..195

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=43&to=63) 244..306

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=80&to=100) 355..417

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=131&to=131) 508..510

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="N-linked (GlcNAc...) asparagine.

 /evidence=ECO:0000255; propagated from

 UniProtKB/Swiss-Prot (Q9NP91.1); glycosylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=166&to=186) 613..675

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=195&to=215) 700..762

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=242&to=262) 841..903

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=277&to=297) 946..1008

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=357&to=357) 1186..1188

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="N-linked (GlcNAc...) asparagine.

 /evidence=ECO:0000255; propagated from

 UniProtKB/Swiss-Prot (Q9NP91.1); glycosylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=390&to=410) 1285..1347

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=432&to=452) 1411..1473

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=466&to=486) 1513..1575

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=505&to=525) 1630..1692

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_064593.1?from=555&to=575) 1780..1842

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="propagated from UniProtKB/Swiss-Prot (Q9NP91.1);

 transmembrane region"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=239&to=379) 239..379

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=380&to=471) 380..471

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=472&to=699) 472..699

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=700&to=810) 700..810

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=811&to=1052) 811..1052

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=1053&to=1215) 1053..1215

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=1216&to=1420) 1216..1420

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=1421&to=1580) 1421..1580

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=1581&to=1746) 1581..1746

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=1747&to=5425) 1747..5425

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /inference="alignment:Splign:2.1.0"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=4051&to=4051) 4051

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="major polyA site"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=5398&to=5403) 5398..5403

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4?from=5425&to=5425) 5425

 /gene="SLC6A20"

 /gene\_synonym="IMINO; SIT1; XT3; Xtrp3"

 /note="major polyA site"

ORIGIN

 1 ggaggcgcgg ggccaccccc aagccgagcc ggggtgggac ggcgcgcctg gcgaccgcgc

 61 actgagactg ccgtgccgga cccccgagcc ggagccgagc gcgccgaggc cggggccatg

 121 gagaaagcgc ggccgctgtg ggccaactcg ctacagttcg tgttcgcctg catctcgtac

 181 gccgtgggcc tgggcaacgt gtggcgattc ccgtacctgt gccagatgta cggcggaggt

 241 agtttcctgg tcccctacat catcatgctt atcgtggagg gaatgccgct cttgtacctg

 301 gaactggctg tggggcagcg catgcggcag ggcagcatcg gcgcctggag gaccatcagc

 361 ccgtacctca gtggtgtcgg ggtcgccagc gtggtggtct ctttcttcct ctccatgtac

 421 tacaacgtca tcaacgcctg ggccttctgg tacctcttcc actccttcca ggatcccctg

 481 ccgtggtctg tctgcccact gaatggtaac cacacgggct acgatgagga gtgtgagaag

 541 gcgtcctcca cacagtactt ctggtacagg aaaaccctca atatctcgcc gtccctccag

 601 gagaacgggg gtgtgcagtg ggagccggcg ctgtgcctcc tcctggcctg gctggtggtg

 661 tacctgtgca tcctgcgtgg caccgagtcc actggcaagg tggtgtattt cacggcgtca

 721 ctgccctatt gcgtgctcat catctacctc atcaggggcc tcacgctcca cggagccacc

 781 aatggcctca tgtacatgtt cactcccaag atagagcagc tggccaaccc caaggcctgg

 841 atcaatgcag ccacccagat cttcttctca cttggcctgg gcttcggcag cctgatcgcc

 901 ttcgccagct acaatgagcc atccaacaac tgccagaagc acgccatcat cgtgtccctc

 961 atcaacagct tcacctccat atttgccagc attgtcacct tctccatcta tggcttcaag

 1021 gccaccttca attatgaaaa ctgcttgaag aaggtgagtc tgctgctgac caacactttt

 1081 gaccttgaag atggcttttt gacagccagc aacctggagc aggtgaaggg ctacctcgca

 1141 tctgcctacc caagcaaata cagcgagatg ttcccgcaaa tcaaaaactg cagcttggaa

 1201 tcggagctag acacggccgt ccagggcact ggcctggcat tcatcgtcta cacagaggcc

 1261 attaaaaaca tggaggtgtc ccagctgtgg tcggtgctct acttcttcat gctgctgatg

 1321 ctgggcattg ggagcatgct ggggaacaca gcggccatcc tcacccctct gacagacagc

 1381 aagatcatct ccagccacct gcccaaggag gccatctcag gtctggtgtg ccttgtcaac

 1441 tgtgccattg gcatggtgtt cacgatggag gctgggaact actggtttga catattcaac

 1501 gactacgcgg ccacactgtc cctgctgctc atcgtgctgg tggagacgat tgccgtgtgc

 1561 tacgtgtacg ggctgaggag atttgaaagt gaccttaagg ccatgaccgg ccgagctgtg

 1621 agctggtact ggaaggtgat gtgggctggc gtaagcccac tgctgattgt cagcctcttt

 1681 gtcttctacc tgagcgacta catcctcacg gggaccctga agtatcaagc ctgggacgcc

 1741 tcccagggcc agctcgtgac caaagattac ccggcctatg cactggctgt catcgggctg

 1801 cttgtggcct cctccaccat gtgcatcccc ctggcggccc tggggacttt tgttcagcgt

 1861 cgcctcaaga ggggagacgc agaccccgtg gcctgagatg tgggcttccc agccgctcac

 1921 ggttttacag atactattta caggcggaaa ctcctcggct gctttttcaa atgcttaagc

 1981 caggagtgct cagcccatca acttcctgag tgtctaaaga agatgaggaa ggtgtgcagg

 2041 aagaaaactc ccttgggaga acgcacaccc tcccgtggtg gctgttcctc cctgtcacct

 2101 gcctcctcat catggaaggg ggtgggctat gaaagccggt ctcaaagata actgcatcct

 2161 tcattccagg aaagccctag aattagggca cattgcaaac tgaaatatga ctataattct

 2221 tatgggacca aatttaagca atttttgttt ttggctgaag agacaccaaa atattagagg

 2281 acaaatattt ttagatccat ttaaggagtt ttgaagtgcc taagatgacc tatttgtcag

 2341 tggtgcaaaa ttaattctct tcttttttga gttgtagtga atatgcaatt tctgtgttcc

 2401 ccttccaccc tttaaatctt aggatgacaa gttataaaga aagaagatct ttgtctggga

 2461 cccccaaagg gatcctttct ctaaggtctc tgacggtggg tccaggacca gacctctcta

 2521 caaaaaattg ccccaactac agtttgcaac cccaaaccac attagaagtc tgtgcagaca

 2581 tccctccgtg gtgtgtgtct tggtgcattg gaaaaggagt caggagccac tgtgaggtga

 2641 gaatgaaagt ggatctcagc tgggcacggt ggctcacgcc tgtaatccta gcaccttggg

 2701 ggtcaaggtg ggtggatcac ttgaggtcag gagtttgagg ccagcctggc caacatggcg

 2761 aaaccccatt tctactaaaa atacaaaaaa attagctggg agtggtggca taggcctgta

 2821 attccagcta ctctggctgc tgaggcacaa gaatcatttg aacctaggag gtggcggttg

 2881 ctgtgagcca agatcatgcc actccactcc agtctgggcg acagagcaag actctgactc

 2941 aaaaaaaaaa aaaaaaaaaa gtggctctga ctttcccata ggcttaatat tacatttgga

 3001 catcgtggcc aactctatca agaccacaaa taccacttac aattaatttt aaattattca

 3061 tctgtacata gttttctaaa atgtatataa ttcaaacaga gcatcttgta actgaagaca

 3121 caccatatct atgatatcgc attagtccat gtggggaaaa gaaagatcag attgttactg

 3181 tgtctgtgta gaaaaggaag acataagaaa ctccattttg atctgtacta agaaaaattg

 3241 tttctgcttt gagatgttgt tagcctataa ctttagcccc aactctgtgc tcacagaaac

 3301 atgcgctgta atgaatcaaa gtttaatgga tttagggctg tgcaggatgt gccttgttaa

 3361 caatatgttt gcaggctgta tgcttggtag aagtcatcgc cattctccat tctcgattaa

 3421 ccagggacac aatgcacagc agaaagctgc agggacctct gcccaagaaa gcctgggtat

 3481 tgtccaaggt ttccccccca ctgagacagc ctgagatatg gcctcatggg aagggaaaga

 3541 cctgactgtc ccccagcccg acacctgtaa agggtcggtg ctgaggagga atagtgaagg

 3601 agggaggcct ctttgcagtt gagataagag gaaggcttct gtctcctgct tgtccctggt

 3661 aatggaatgt ctcggtgtaa aaagctgacc attcccattc gttctattct gagataggag

 3721 aaaaccgccc tgtggctgga ggtgagatat gctggcagca atactgctct gttactcctt

 3781 gctacactga gatgtttggg taaagagaaa cataaatcta gcctacgtgc acatctgggc

 3841 acagtacctt tccttgaact tattcgtgat acagattcct ttgctcacat gtttccctgc

 3901 tgaccttctt cccacctgtt gccctgctac actcccctcg ctaagacagt aaaaataatg

 3961 atcaataaat actgagggaa ctcagaggcc agcgccggtg cgggtcctcc acatgctgag

 4021 cgccggtcct gggcccactg ttctttctct atactttgtc tctgtgtctt atttcttttc

 4081 tcaaaaaaat ttttcttttt ttaactgcaa taaagtaggt cgttacccct aaagtgagcg

 4141 agtcccttcc tgctctaacg tccagtgctt ccttgagatc atggaattct ggagagtttc

 4201 tgtggatata tttgcacaga tatctttgaa gacacaaaca ctcacctggg acgttggccg

 4261 gagatgtctt tttatttttg tgctgtaaaa ttctcttaca gcaaaaatag gctttagaaa

 4321 ggtcttctac tgtcttcagc aaccatctca tcttccagct tcacctgatt gtccagttat

 4381 catacatttg actttcaaat gtatgaacca gcatgtaccc catggattta atcttatcta

 4441 ccccgtggat tcaatcttct tatcagaagg ttcttttatg tcaaaaaacc tgctgtcaag

 4501 gcttgaagag ccggcacact caatggcaaa cacagcaccg agtctgctct gaatcctgga

 4561 ggatctggcc ctcctctcaa tccccactca cagtcaccgt cttacaactc agggccacct

 4621 gggatcagtc atcagtcagg gtgcgtaagc cttgaatacc aggtagcctc aggagtgaaa

 4681 agataaatgt cctagatcat taccttattc agtgtcccca ccttgcagcg cattccaacc

 4741 acctgggagc atttaaaact ccagatgccc acaccacacc ctggggccac ccatcagacc

 4801 ttctggaagc aagacctggg cctccatggc cccaaaaact ccctaggtga tccgatgtgc

 4861 agccaaatct gagaggcccc atttaaaaaa gaaagaacat gggtggtaca ttgaggagta

 4921 tttacatttt ataaaatgac ttaaaaattt gaaggcattt ttgagcattt ccaattatat

 4981 ggaagagtta cttctacgga atagtttttg ctcatggaac tcaaacagat gaagcaccac

 5041 tgttacagaa taatgtgctc cagatgaaaa tgtctcgttt ctgtgaattt catgaagagc

 5101 agaacatttc tcaagaatcc tcttgagcca gtaatcaatc ctgtctcaaa aaatgttctt

 5161 tgccatttcc tagatactgc acaaaagtgg ccatgtcgac atttgtccac ccaccctcca

 5221 ataagctgga gcgacaaagg gacattccat ccctgtaccc ttagtggtag ccatgacacg

 5281 atggccagat catggactcc ggaaagcttt ctgtttttac tggaaacata gcaaaccttg

 5341 atttagctcc aagaaattga gtagggaaat atttgttttt tagcaattgt catagtaaat

 5401 aaaactctaa gcaagtgtat ctaca

# Homo sapiens FYVE and coiled-coil domain autophagy adaptor 1 (FYCO1), transcript variant 10, mRNA

NCBI Reference Sequence: NM\_001386421.1

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1%22%20%5Cl%20%22goto1896085764_0)

LOCUS NM\_001386421 8598 bp mRNA linear PRI 17-DEC-2020

DEFINITION Homo sapiens FYVE and coiled-coil domain autophagy adaptor 1

 (FYCO1), transcript variant 10, mRNA.

ACCESSION NM\_001386421 XM\_011534111

VERSION NM\_001386421.1

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

 ORGANISM [Homo sapiens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)

 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 8598)

 AUTHORS Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A,

 Invernizzi P, Fernandez J, Prati D, Baselli G, Asselta R, Grimsrud

 MM, Milani C, Aziz F, Kassens J, May S, Wendorff M, Wienbrandt L,

 Uellendahl-Werth F, Zheng T, Yi X, de Pablo R, Chercoles AG, Palom

 A, Garcia-Fernandez AE, Rodriguez-Frias F, Zanella A, Bandera A,

 Protti A, Aghemo A, Lleo A, Biondi A, Caballero-Garralda A, Gori A,

 Tanck A, Carreras Nolla A, Latiano A, Fracanzani AL, Peschuck A,

 Julia A, Pesenti A, Voza A, Jimenez D, Mateos B, Nafria Jimenez B,

 Quereda C, Paccapelo C, Gassner C, Angelini C, Cea C, Solier A,

 Pestana D, Muniz-Diaz E, Sandoval E, Paraboschi EM, Navas E, Garcia

 Sanchez F, Ceriotti F, Martinelli-Boneschi F, Peyvandi F, Blasi F,

 Tellez L, Blanco-Grau A, Hemmrich-Stanisak G, Grasselli G,

 Costantino G, Cardamone G, Foti G, Aneli S, Kurihara H, ElAbd H, My

 I, Galvan-Femenia I, Martin J, Erdmann J, Ferrusquia-Acosta J,

 Garcia-Etxebarria K, Izquierdo-Sanchez L, Bettini LR, Sumoy L,

 Terranova L, Moreira L, Santoro L, Scudeller L, Mesonero F, Roade

 L, Ruhlemann MC, Schaefer M, Carrabba M, Riveiro-Barciela M,

 Figuera Basso ME, Valsecchi MG, Hernandez-Tejero M, Acosta-Herrera

 M, D'Angio M, Baldini M, Cazzaniga M, Schulzky M, Cecconi M, Wittig

 M, Ciccarelli M, Rodriguez-Gandia M, Bocciolone M, Miozzo M,

 Montano N, Braun N, Sacchi N, Martinez N, Ozer O, Palmieri O,

 Faverio P, Preatoni P, Bonfanti P, Omodei P, Tentorio P, Castro P,

 Rodrigues PM, Blandino Ortiz A, de Cid R, Ferrer R, Gualtierotti R,

 Nieto R, Goerg S, Badalamenti S, Marsal S, Matullo G, Pelusi S,

 Juzenas S, Aliberti S, Monzani V, Moreno V, Wesse T, Lenz TL,

 Pumarola T, Rimoldi V, Bosari S, Albrecht W, Peter W, Romero-Gomez

 M, D'Amato M, Duga S, Banales JM, Hov JR, Folseraas T, Valenti L,

 Franke A and Karlsen TH.

 CONSRTM Severe Covid-19 GWAS Group

 TITLE Genomewide Association Study of Severe Covid-19 with Respiratory

 Failure

 JOURNAL N Engl J Med 383 (16), 1522-1534 (2020)

 PUBMED [32558485](https://www.ncbi.nlm.nih.gov/pubmed/32558485)

REFERENCE 2 (bases 1 to 8598)

 AUTHORS Sakurai S, Shimizu T and Ohto U.

 TITLE Crystal structure of the FYCO1 RUN domain suggests possible

 interfaces with small GTPases

 JOURNAL Acta Crystallogr F Struct Biol Commun 76 (Pt 8), 326-333 (2020)

 PUBMED [32744243](https://www.ncbi.nlm.nih.gov/pubmed/32744243)

REFERENCE 3 (bases 1 to 8598)

 AUTHORS Thavarajah T, Dos Santos CC, Slutsky AS, Marshall JC, Bowden P,

 Romaschin A and Marshall JG.

 TITLE The plasma peptides of sepsis

 JOURNAL Clin Proteomics 17, 26 (2020)

 PUBMED [32636717](https://www.ncbi.nlm.nih.gov/pubmed/32636717)

 REMARK Publication Status: Online-Only

REFERENCE 4 (bases 1 to 8598)

 AUTHORS Iqbal H, Khan SY, Zhou L, Irum B, Ali M, Ahmed MR, Shahzad M, Ali

 MH, Naeem MA, Riazuddin S, Hejtmancik JF and Riazuddin SA.

 TITLE Mutations in FYCO1 identified in families with congenital cataracts

 JOURNAL Mol Vis 26, 334-344 (2020)

 PUBMED [32355443](https://www.ncbi.nlm.nih.gov/pubmed/32355443)

 REMARK Publication Status: Online-Only

REFERENCE 5 (bases 1 to 8598)

 AUTHORS Rothwell S, Lilleker JB and Lamb JA.

 TITLE Genetics in inclusion body myositis

 JOURNAL Curr Opin Rheumatol 29 (6), 639-644 (2017)

 PUBMED [28777108](https://www.ncbi.nlm.nih.gov/pubmed/28777108)

 REMARK Review article

REFERENCE 6 (bases 1 to 8598)

 AUTHORS Chen J, Ma Z, Jiao X, Fariss R, Kantorow WL, Kantorow M, Pras E,

 Frydman M, Pras E, Riazuddin S, Riazuddin SA and Hejtmancik JF.

 TITLE Mutations in FYCO1 cause autosomal-recessive congenital cataracts

 JOURNAL Am J Hum Genet 88 (6), 827-838 (2011)

 PUBMED [21636066](https://www.ncbi.nlm.nih.gov/pubmed/21636066)

 REMARK GeneRIF: FYCO1 is involved in lens development and transparency in

 humans, and mutations in this gene are one of the most common

 causes of autosomal-recessive congenital cataracts in the Pakistani

 population.

REFERENCE 7 (bases 1 to 8598)

 AUTHORS Pankiv S and Johansen T.

 TITLE FYCO1: linking autophagosomes to microtubule plus end-directing

 molecular motors

 JOURNAL Autophagy 6 (4), 550-552 (2010)

 PUBMED [20364109](https://www.ncbi.nlm.nih.gov/pubmed/20364109)

REFERENCE 8 (bases 1 to 8598)

 AUTHORS Pankiv S, Alemu EA, Brech A, Bruun JA, Lamark T, Overvatn A,

 Bjorkoy G and Johansen T.

 TITLE FYCO1 is a Rab7 effector that binds to LC3 and PI3P to mediate

 microtubule plus end-directed vesicle transport

 JOURNAL J Cell Biol 188 (2), 253-269 (2010)

 PUBMED [20100911](https://www.ncbi.nlm.nih.gov/pubmed/20100911)

 REMARK GeneRIF: We have characterized the LC3-, Rab7-, and

 phosphatidylinositol-3-phosphate-binding domains in FYCO1 and

 mapped part of the CC region essential for MT plus end-directed

 transport.

REFERENCE 9 (bases 1 to 8598)

 AUTHORS Ghosh D, Lippert D, Krokhin O, Cortens JP and Wilkins JA.

 TITLE Defining the membrane proteome of NK cells

 JOURNAL J Mass Spectrom 45 (1), 1-25 (2010)

 PUBMED [19946888](https://www.ncbi.nlm.nih.gov/pubmed/19946888)

REFERENCE 10 (bases 1 to 8598)

 AUTHORS Kiss H, Yang Y, Kiss C, Andersson K, Klein G, Imreh S and Dumanski

 JP.

 TITLE The transcriptional map of the common eliminated region 1 (C3CER1)

 in 3p21.3

 JOURNAL Eur J Hum Genet 10 (1), 52-61 (2002)

 PUBMED [11896456](https://www.ncbi.nlm.nih.gov/pubmed/11896456)

 REMARK GeneRIF: Maps to a region of chromosome 3p21.3 which is frequently

 deleted in tumors.

COMMENT REVIEWED [REFSEQ](https://www.ncbi.nlm.nih.gov/RefSeq/): This record has been curated by NCBI staff. The

 reference sequence was derived from [AC099782.2](https://www.ncbi.nlm.nih.gov/nuccore/AC099782.2).

 On Aug 25, 2020 this sequence version replaced [XM\_011534111.3](https://www.ncbi.nlm.nih.gov/nuccore/XM_011534111.3).

 Summary: The gene encodes a Rab7 adapter protein that is implicated

 in the microtubule transport of autophagosomes. The encoded protein

 contains a RUN domain, a FYVE-type zinc finger domain, and Golgi

 dynamics (GOLD) domain. The encoded protein plays a role in

 microtubule plus end-directed transport of autophagic vesicles

 through interactions with the small GTPase Rab7,

 phosphatidylinositol-3-phosphate (PI3P), the autophagosome marker

 LC3, and the kinesin KIF5. Mutations in this gene are associated

 with inclusion body myositis (IBM) and autosomal recessive

 congenital cataracts (CATC2). [provided by RefSeq, Aug 2020].

 Publication Note: This RefSeq record includes a subset of the

 publications that are available for this gene. Please see the Gene

 record to access additional publications.

 ##Evidence-Data-START##

 Transcript exon combination :: SRR7346977.2471468.1,

 DRR138518.449027.1 [ECO:0000332]

 ##Evidence-Data-END##

PRIMARY REFSEQ\_SPAN PRIMARY\_IDENTIFIER PRIMARY\_SPAN COMP

 1-103 AC099782.2 181647-181749 c

 104-187 AC099782.2 171077-171160 c

 188-354 AC099782.2 170781-170947 c

 355-461 AC099782.2 167495-167601 c

 462-587 AC099782.2 165630-165755 c

 588-694 AC099782.2 161164-161270 c

 695-838 AC099782.2 159013-159156 c

 839-929 AC099782.2 155600-155690 c

 930-3356 AC099782.2 152202-154628 c

 3357-3449 AC099782.2 150958-151050 c

 3450-3568 AC099782.2 150261-150379 c

 3569-3736 AC099782.2 148150-148317 c

 3737-3886 AC099782.2 145318-145467 c

 3887-4098 AC099782.2 144333-144544 c

 4099-4243 AC099782.2 141174-141318 c

 4244-4339 AC099782.2 122373-122468 c

 4340-4550 AC099782.2 116996-117206 c

 4551-4660 AC099782.2 109581-109690 c

 4661-8598 AC099782.2 103828-107765 c

FEATURES Location/Qualifiers

 source 1..8598

 /organism="Homo sapiens"

 /mol\_type="mRNA"

 /db\_xref="taxon:[9606](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=9606)"

 /chromosome="3"

 /map="3p21.31"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=1&to=8598) 1..8598

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="FYVE and coiled-coil domain autophagy adaptor 1"

 /db\_xref="GeneID:[79443](https://www.ncbi.nlm.nih.gov/gene/79443)"

 /db\_xref="HGNC:[HGNC:14673](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:14673)"

 /db\_xref="MIM:[607182](https://www.ncbi.nlm.nih.gov/omim/607182)"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=1&to=103) 1..103

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=104&to=187) 104..187

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=188&to=354) 188..354

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=204&to=206) 204..206

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="upstream in-frame stop codon"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=300&to=4736) 300..4736

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="isoform 1 is encoded by transcript variant 10; FYVE

 and coiled-coil domain containing 1; RUN and FYVE domain

 containing 3; zinc finger FYVE domain-containing protein

 7; FYVE and coiled-coil domain-containing protein 1"

 /codon\_start=1

 /product="FYVE and coiled-coil domain-containing protein 1

 isoform 1"

 /protein\_id="[NP\_001373350.1](https://www.ncbi.nlm.nih.gov/protein/1896085765)"

 /db\_xref="GeneID:[79443](https://www.ncbi.nlm.nih.gov/gene/79443)"

 /db\_xref="HGNC:[HGNC:14673](https://www.genenames.org/data/gene-symbol-report/#!/hgnc_id/HGNC:14673)"

 /db\_xref="MIM:[607182](https://www.ncbi.nlm.nih.gov/omim/607182)"

 /translation="MASTNAESQLQRIIRDLQDAVTELSKEFQEAGEPITDDSTSLHK

 FSYKLEYLLQFDQKEKATLLGNKKDYWDYFCACLAKVKGANDGIRFVKSISELRTSLG

 KGRAFIRYSLVHQRLADTLQQCFMNTKVTSDWYYARSPFLQPKLSSDIVGQLYELTEV

 QFDLASRGFDLDAAWPTFARRTLTTGSSAYLWKPPSRSSSMSSLVSSYLQTQEMVSNF

 DLNSPLNNEALEGFDEMRLELDQLEVREKQLRERMQQLDRENQELRAAVSQQGEQLQT

 ERERGRTAAEDNVRLTCLVAELQKQWEVTQATQNTVKELQTCLQGLELGAAEKEEDYH

 TALRRLESMLQPLAQELEATRDSLDKKNQHLASFPGWLAMAQQKADTASDTKGRQEPI

 PSDAAQEMQELGEKLQALERERTKVEEVNRQQSAQLEQLVKELQLKEDARASLERLVK

 EMAPLQEELSGKGQEADQLWRRLQELLAHTSSWEEELAELRREKKQQQEEKELLEQEV

 RSLTRQLQFLETQLAQVSQHVSDLEEQKKQLIQDKDHLSQQVGMLERLAGPPGPELPV

 AGEKNEALVPVNSSLQEAWGKPEEEQRGLQEAQLDDTKVQEGSQEEELRQANRELEKE

 LQNVVGRNQLLEGKLQALQADYQALQQRESAIQGSLASLEAEQASIRHLGDQMEASLL

 AVRKAKEAMKAQMAEKEAILQSKEGECQQLREEVEQCQQLAEARHRELRALESQCQQQ

 TQLIEVLTAEKGQQGVGPPTDNEARELAAQLALSQAQLEVHQGEVQRLQAQVVDLQAK

 MRAALDDQDKVQSQLSMAEAVLREHKTLVQQLKEQNEALNRAHVQELLQCSEREGALQ

 EERADEAQQREEELRALQEELSQAKCSSEEAQLEHAELQEQLHRANTDTAELGIQVCA

 LTVEKERVEEALACAVQELQDAKEAASREREGLERQVAGLQQEKESLQEKLKAAKAAA

 GSLPGLQAQLAQAEQRAQSLQEAAHQELNTLKFQLSAEIMDYQSRLKNAGEECKSLRG

 QLEEQGRQLQAAEEAVEKLKATQADMGEKLSCTSNHLAECQAAMLRKDKEGAALREDL

 ERTQKELEKATTKIQEYYNKLCQEVTNRERNDQKMLADLDDLNRTKKYLEERLIELLR

 DKDALWQKSDALEFQQKLSAEERWLGDTEANHCLDCKREFSWMVRRHHCRICGRIFCY

 YCCNNYVLSKHGGKKERCCRACFQKLSEGPGSPDSSGSGTSQGEPSPALSPASPGPQA

 TGGQGANTDYRPPDDAVFDIITDEELCQIQESGSSLPETPTETDSLDPNAAEQDTTST

 SLTPEDTEDMPVGQDSEICLLKSGELMIKVPLTVDEIASFGEGSRELFVRSSTYSLIP

 ITVAEAGLTISWVFSSDPKSISFSVVFQEAEDTPLDQCKVLIPTTRCNSHKENIQGQL

 KVRTPGIYMLIFDNTFSRFVSKKVFYHLTVDRPVIYDGSDFL"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001373350.1?from=2&to=2) 303..305

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="N-acetylalanine.

 /evidence=ECO:0000244|PubMed:19413330,

 ECO:0000244|PubMed:22223895, ECO:0000244|PubMed:22814378;

 propagated from UniProtKB/Swiss-Prot (Q9BQS8.3);

 acetylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001373350.1?from=196&to=196) 885..887

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="Phosphoserine.

 /evidence=ECO:0000244|PubMed:23186163; propagated from

 UniProtKB/Swiss-Prot (Q9BQS8.3); phosphorylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001373350.1?from=342&to=342) 1323..1325

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="Phosphoserine.

 /evidence=ECO:0000244|PubMed:24275569; propagated from

 UniProtKB/Swiss-Prot (Q9BQS8.3); phosphorylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001373350.1?from=381&to=381) 1440..1442

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="Phosphothreonine.

 /evidence=ECO:0000250|UniProtKB:Q8VDC1; propagated from

 UniProtKB/Swiss-Prot (Q9BQS8.3); phosphorylation site"

 [misc\_feature](https://www.ncbi.nlm.nih.gov/protein/NP_001373350.1?from=878&to=878) 2931..2933

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="Phosphoserine.

 /evidence=ECO:0000244|PubMed:17525332; propagated from

 UniProtKB/Swiss-Prot (Q9BQS8.3); phosphorylation site"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=355&to=461) 355..461

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=462&to=587) 462..587

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=588&to=694) 588..694

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=695&to=838) 695..838

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=839&to=929) 839..929

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=930&to=3356) 930..3356

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=3357&to=3449) 3357..3449

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=3450&to=3568) 3450..3568

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=3569&to=3736) 3569..3736

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=3737&to=3886) 3737..3886

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=3887&to=4098) 3887..4098

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=4099&to=4243) 4099..4243

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=4244&to=4339) 4244..4339

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=4340&to=4550) 4340..4550

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=4551&to=4660) 4551..4660

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [exon](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=4661&to=8598) 4661..8598

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /inference="alignment:Splign:2.1.0"

 [regulatory](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=8574&to=8579) 8574..8579

 /regulatory\_class="polyA\_signal\_sequence"

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="hexamer: AATAAA"

 [polyA\_site](https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1?from=8598&to=8598) 8598

 /gene="FYCO1"

 /gene\_synonym="CATC2; CTRCT18; RUFY3; ZFYVE7"

 /note="major polyA site"

ORIGIN

 1 accctcttgg tgtgcgcgtc atggccgtca gcaccgcgtt cccgtcctct tccgcttggc

 61 cccagaaagt ttcggttctg cccggcggtg gacccacgag cgcccccaaa ccaggtcact

 121 gtgtctgaat agaagtgtgt cagaaacaaa ccagagagtg cccacagaac aaggcccaca

 181 cagaaaggtg ccaccatgga gtctgaccac tgctgagcag acagccaccg agggccgaaa

 241 ttctgagcct tcctctggac ccaggcagga gacatacaga caagaaaggc aaactcacca

 301 tggcctccac caatgcagag agccagctcc agagaatcat ccgagacttg caagatgctg

 361 tgacagaact aagcaaagaa tttcaggaag caggggaacc catcacggat gacagcacca

 421 gcttgcataa attttcttat aaacttgagt atctcctgca atttgatcag aaagagaagg

 481 ccaccctcct gggcaacaag aaggactact gggattactt ctgtgcctgc ctggccaagg

 541 tgaaaggagc caatgatggg atccgctttg tcaagtctat ctcagagctc cgaacatcct

 601 tggggaaagg aagagcattt attcgctact ccttggtgca ccagaggttg gcagacacct

 661 tacagcagtg cttcatgaac accaaagtga ccagtgactg gtactatgca agaagcccct

 721 ttctgcagcc aaagctgagc tcggacattg tgggccaact ctatgagctg actgaggttc

 781 agtttgacct ggcgtcgagg ggctttgact tggatgctgc ctggccaaca tttgccagga

 841 ggacgctgac cactggctct tctgcttacc tgtggaaacc ccctagccgc agctccagca

 901 tgagcagctt ggtgagcagc tacctgcaga ctcaagagat ggtgtccaac tttgacctga

 961 acagccccct aaacaacgag gcattggagg gctttgatga gatgcgacta gagctggacc

 1021 agttggaggt gcgggagaag cagctacggg agcgcatgca gcagctggac agagagaacc

 1081 aggagctgag ggcagctgtc agccagcaag gggagcaact gcagacagag agggagaggg

 1141 ggcgcactgc agcggaggac aacgttcgcc tcacttgctt ggtagctgag ctccagaagc

 1201 agtgggaggt cacccaggcc acccagaaca ctgtgaagga gctgcagaca tgcctgcagg

 1261 gcctggagct aggagcagca gagaaggagg aggactacca cacagccctg cggcggctgg

 1321 agtccatgct gcagcccttg gcacaggagc ttgaggccac acgggactca ctggacaaga

 1381 aaaaccagca tttagccagc ttcccaggct ggctagccat ggctcagcag aaggcagata

 1441 cggcatcaga cacaaagggc cggcaagaac ctattcccag tgatgcggcc caggagatgc

 1501 aggagctagg ggagaagctt caagccctag aaagggagag aaccaaggtc gaggaggtca

 1561 acagacagca gagtgcccaa ctggaacagc tggtcaagga gcttcagctg aaagaggatg

 1621 cccgggccag cctggagcgc ctggtgaagg agatggcccc actccaggag gagttgtctg

 1681 ggaagggaca ggaggcagac cagctctggc gacggctgca ggagttgctg gcccacacga

 1741 gctcctggga ggaggagcta gcagagttga ggcgggagaa aaaacagcaa caggaggaga

 1801 aggagctgct ggagcaggag gtcaggtctc tgacccggca gctgcagttc ctggagaccc

 1861 agctggcaca ggtgagccaa catgtgagtg acctggagga gcagaagaag cagctcattc

 1921 aggacaaaga ccacctcagc cagcaggtgg gtatgctcga gcggcttgct gggccgcctg

 1981 gcccagaact gccagtggca ggtgagaaga atgaggccct ggtccctgtg aactccagtc

 2041 tgcaagaggc ctgggggaag ccagaggagg agcagagggg cctgcaggag gcacagttag

 2101 acgataccaa ggtgcaagag ggcagccagg aggaagagct ccggcaggcc aacagggagc

 2161 tggagaagga gctacagaat gtggtcgggc gtaaccagct cctggagggc aagctgcaag

 2221 ccctgcaggc cgattaccag gctttgcagc agcgggaatc agccatccag ggctccttgg

 2281 cctccctgga ggccgagcag gccagcatcc ggcacttggg tgaccagatg gaggcgagct

 2341 tgctggctgt aaggaaggcc aaggaggcca tgaaagccca gatggcagag aaggaggcca

 2401 ttctacagag caaggagggc gagtgtcagc agctgcggga ggaggtggag cagtgccagc

 2461 aactggcaga agcccggcac agagagctta gggctctcga gagccagtgc cagcagcaga

 2521 cccagctgat tgaggtcctc acagcagaga aaggccaaca gggagttggc ccacccactg

 2581 acaatgaagc ccgtgagctg gctgcccagc tagccctgtc tcaggcgcag ctggaagtcc

 2641 atcaggggga ggtccaacgg ctgcaggctc aggtggtgga cctccaggcc aagatgcggg

 2701 cagccctgga tgaccaggac aaggtgcaga gccagctaag catggctgag gccgtcctga

 2761 gggagcacaa aacccttgtg cagcagctga aggagcagaa tgaagccctt aacagagccc

 2821 atgtccagga gctgctgcaa tgctcggagc gtgaaggggc actgcaggag gagagggccg

 2881 atgaggccca gcagagggag gaggagctgc gggccctgca ggaggagctg tcccaggcca

 2941 aatgcagctc cgaggaagca cagctggagc acgctgagct gcaagagcag ctgcaccggg

 3001 ccaacacaga cacagctgag ctgggcatcc aggtttgcgc actgaccgtg gaaaaggagc

 3061 gagtggagga ggcactggcc tgtgctgtcc aggagctcca ggacgccaaa gaggcagcct

 3121 caagggagcg agagggcctg gagcgccaag tagctgggct gcagcaagag aaggagagct

 3181 tgcaggagaa gctgaaggcg gccaaggcag cagccggctc actgcctggc ctgcaggccc

 3241 agctcgccca ggcagagcag cgggcccaga gcctccaaga ggctgcacac caggagctca

 3301 acaccctcaa gttccagctg agtgctgaaa tcatggacta ccagagcaga cttaagaatg

 3361 ctggtgaaga gtgcaagagc ctcaggggcc agcttgagga gcaaggccgg cagctgcagg

 3421 ctgctgagga agctgtggag aagctgaagg ccacccaagc agacatggga gagaagctga

 3481 gctgcactag caaccatctt gcagagtgcc aggcggccat gctgaggaag gacaaggagg

 3541 gggctgccct gcgtgaagac ctagaaagga cccagaagga actcgaaaaa gccacaacaa

 3601 aaatccaaga gtattacaac aaactctgcc aggaggtgac aaatcgtgag aggaatgacc

 3661 agaagatgct tgctgacctg gatgacctca acagaaccaa gaagtatctc gaggagcggc

 3721 tgatagagct gctcagggac aaggatgctc tctggcagaa gtcagatgcc ctggaattcc

 3781 agcagaagct cagtgctgag gagagatggc tcggagacac agaggcaaac cactgcctcg

 3841 actgtaagcg ggagttcagc tggatggtgc ggcggcacca ctgcaggata tgtggccgca

 3901 tcttctgtta ctactgctgc aacaactacg tcctgagcaa gcacggtggc aaaaaggagc

 3961 gctgctgccg agcctgtttc cagaagctca gtgaaggccc tggctcccct gatagcagtg

 4021 gctcaggcac tagccaggga gagcccagcc ctgcactgtc accagcctca cctgggcccc

 4081 aggccacagg aggccaagga gcaaatacag actacaggcc accggacgac gctgtgtttg

 4141 atatcatcac agatgaggaa ttgtgccaga tacaggagtc cggctcctct ttgcctgaaa

 4201 cacccactga aactgattct ctcgacccaa atgcggctga acaggatact acatcaacct

 4261 cgctaacgcc tgaggacact gaagacatgc ccgtggggca ggattcggaa atctgcctgc

 4321 tgaagtctgg agaactgatg atcaaagtac ccctcacagt ggatgagatc gccagcttcg

 4381 gggagggtag cagggagctg tttgtgaggt ccagcaccta cagcctgatc cccatcactg

 4441 tggccgaggc aggcctcacc atcagctggg tcttctcctc tgaccccaag agcatctcct

 4501 tcagtgtggt cttccaggag gccgaggaca caccgctgga tcagtgtaag gtcctcattc

 4561 ccacgacccg atgcaactcc cacaaggaga acatccaggg ccagctcaag gttcgcacac

 4621 ccggcatcta catgctcatc ttcgacaata ccttctcaag gtttgtctct aaaaaggtat

 4681 tttatcactt gacggttgat cggcctgtga tctacgatgg aagtgatttc ctgtagcttc

 4741 agcacctcag taacttcact tcatccacag gaaacactgc tcttcctcac ctgtcacata

 4801 aagcattttt ttaaaaagtc agctgctcca aaatcatcaa ctcagcccct gggctgcccc

 4861 tcagaggcgg tgtctgggga ggactttgtg ctcagcactc tgcaccggcc actcttagcc

 4921 cccgaggcgt tgaagggctc aggcaatgtt tccattaagt agagactcag ctgttgtcac

 4981 acccaaaggg atgctctgcc aaaggtttaa acacccagga gaccatcagc ctctcctggg

 5041 agcacagttg gctacaggcc tcttgtggag agtttcacgg gcaggggtga ttccaacttc

 5101 tgcctgtgga gagattttct gccctgcccc accagggccc tgcatgttgg agactgagct

 5161 gggtgcactg gccatgccct gtgaatcctc aggctgtgac gccctcaggt actcctggga

 5221 aaaggaggta cacagccatc atgcgagtcg gtgccagggg accccccgga gatcctgacc

 5281 agctcctcca gtcatgctct tgtccctcac tgccccagta agctggaggc tgctccagaa

 5341 ctcagcagtg ttggaggggc ctctaagctg cactctcttt ctggcccttt tgtctgggtg

 5401 attctgtcct caaataaagc ccttcactca gccagacctc tccacagctc aaagcattgc

 5461 cctaagaatc agaagtaaag ataatccaag agcaaaaccc actgtacttg gggcctgcaa

 5521 tggctgtgtg tacactacat ctaatgccca aatgccagcc agtgtggatg ttgtgaccac

 5581 agagcaggat tgtgcattgg ctttagagct actcctcagc tgatggccca cttttgttta

 5641 tataaataag agcttctgcc ccacctgcag acatgtttac taatgatcat agccaggatt

 5701 agaaccactt tcaaacattg gggccttctt aacaaaagtc tttgataact taagaaccaa

 5761 agtaacagag taaacagagg catgatggat ccctgggccc cactcccctc ctgacaggtt

 5821 ccccaacagc ccatttgccc acttcccact gctcagccca caccagacct ccaggagaca

 5881 tccccccttg aggcagagag atcctgttcc ctattcccag acaagaatta tttaatcttc

 5941 cctgttctct gtggtccttt tcttccccaa caacagatag ctcaccttgg acagctcttc

 6001 gtcccttgtt catggaacca gctgcctgca gtcaggcccc aggttcttcc atgggtgaac

 6061 agagcatctg acaaaaggtc ccagtttggc caggggtgag ggagagagca ccagacaggc

 6121 tatccgagaa tctgagagct gggcccggca gttcctccag ctacccttgt gacctaagtc

 6181 cagtcacaca tttcccaaag tttctctttg tcataaccct ggtctggctg gttttgaggg

 6241 cttgagaatg ggtcagggac tccaggccaa gtccaacaga gaccccaaac ccaccacaca

 6301 ccagcagcca caacctcacc accaacaaag aggacttttg tggggccaca agtaagaggt

 6361 catttctgga atggactcag acctttaaac aggagagttg agcacttcca gtcagttttt

 6421 aagcaaggca tggggaacag ggaatagaac ctttcaaaga ggttgcccag agaaaagctg

 6481 ggcctcttgc attcggcttc cttggagcag cctcttctgg cagaaagcca tcaggtgctc

 6541 aatcatcttc tcctggccaa ggctctgacc atgcttagta ctggaataga ggtggccagg

 6601 cccccagcga ctcttcttgg cctgatgttt gtcctcacag gcatgccacg tggcctgaga

 6661 tgattcagaa caaatcatgc taactttgaa tccatccagc cacttgcaaa tgataatcag

 6721 aagtcagctt gttcactgtt agaaagaaac taacaaaaga gaacccagag caatctagaa

 6781 tctttgagtg cttggctttc caaggatact gcggagactc tggccaagct gatgaccttc

 6841 tgaagtgtca ctggcaccat atgcaacaag aaccaccatt cactgagtag ctaatgggtt

 6901 tggggcctgg gacattccat ctgaggtcct tcctgaacat gtcactccac agcagaggac

 6961 cggttgcagc ttacccagaa ccactcctcc aggagagctg gatgttttgc gtgcaacacc

 7021 ttgagcactg actgctattg ttcaaaaaaa gcctttgctg cattcggagg actgccccgt

 7081 gccctgaggt gacttcctaa ctatgtggtt tcattagcga atttattttt tgtgctgggt

 7141 ggacatttgt attttgttag gttgctgttt aagctcaagt ttgctgtgct ctctgcagct

 7201 acaaaacatc ttggcatatt taagagtggc ttttataaat agctttattc tgatattaat

 7261 cagattccca actttactga gaattaagga ctggggtact ttaaagaaat gcaaatagca

 7321 attgaagaac cactgctgca ggtggtagcc ctggctagac tgaattacac tagaaatcag

 7381 ccagaaggaa gcgtccttgg gatcccagat cactcttttt tttttttttt aaaaggggca

 7441 gccccttgat ggctcatctc tctgaataac agttacgtct tcatatcgat accagatgcc

 7501 ttcttcatca tgccactgaa gccactcacc accttcaaga acatgccaac ctctgtcaga

 7561 ttcacttacc cacaaacaag gaggcacgtt tggcacaaag tgttgtcctc caggtccaag

 7621 tggactctac agagtgcttg acctcaacac actggattcc aggtggactg gaccaagagc

 7681 aggcaaagac acgggaactg aaaaactcca cagggtttgg agaatagaaa tgaaaagcca

 7741 cgtcatataa ctcaagaata aatggtgttt tggaaatttt aaaattatca tcgaaggtgg

 7801 tgaaactatt tcaggcccaa atgaaaggaa atcgccagtt ggggatgaaa tcacagagcc

 7861 tgtgttttat gatatggttg gatgtccact gatgaaattt taaaggagtt tcagttttaa

 7921 aagtgcgcat gattctacat atgagaattc tttaggccaa gaaactgtcc ttggctcaga

 7981 ggtgttggga attaaagcag agagaagcca ttcgtgatgc ttagaaccaa ggatggtcat

 8041 gtacacaaag accatcgaga cggccattct tgtttacaaa acacttacca agaaagcact

 8101 ttgtagggga actttagtaa gttcttctca tttcattatg tttcttccaa ggaaacagga

 8161 gagactgaat taataattct ctctttcctc ttaagcactt ttaaaataat aaagtacatc

 8221 ttgaaatttg gggaggcatc tctgatttaa aaaaagaaaa aggctgcttg atgtatgtta

 8281 tgcagagaca ctctgcctct ggtggctgca gagcaatacc caagcctcat ttggaaggct

 8341 caacatttgg aattgcactt taattgatta atcctcaatt catgtggcct tacgggatgg

 8401 tgggtctggg accccaattc attcttatct gccaaagaat tatctagaag cacatcaaat

 8461 accagcaccc cacctgctca atgggggtgg aaaacttttg tatccctaag catattattt

 8521 tatagtgtct gccatgccat gtggaaatac tttattttta acctcaggat ttaaataaag

 8581 taaacactat gacattta

# Mutant Severe acute respiratory syndrome coronavirus 2 clone SARS-CoV-2-MA10, complete genome

GenBank: MT952602.1

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1%22%20%5Cl%20%22goto1898953378_0)

LOCUS MT952602 29882 bp RNA linear SYN 19-OCT-2020

DEFINITION Mutant Severe acute respiratory syndrome coronavirus 2 clone

 SARS-CoV-2-MA10, complete genome.

ACCESSION MT952602

VERSION MT952602.1

KEYWORDS .

SOURCE Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

 ORGANISM [Severe acute respiratory syndrome coronavirus 2](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2697049)

 Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes;

 Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae;

 Betacoronavirus; Sarbecovirus.

REFERENCE 1 (bases 1 to 29882)

 AUTHORS Leist,S.R., Dinnon,K.H. III, Schafer,A., Tse,L.V., Okuda,K.,

 Hou,Y.J., West,A., Edwards,C.E., Sanders,W., Fritch,E.J.,

 Gully,K.L., Scobey,T., Brown,A.J., Sheahan,T.P., Moorman,N.J.,

 Boucher,R.C., Gralinski,L.E., Montgomery,S.A. and Baric,R.S.

 TITLE A Mouse-Adapted SARS-CoV-2 Induces Acute Lung Injury and Mortality

 in Standard Laboratory Mice

 JOURNAL Cell (2020) In press

 PUBMED [33031744](https://www.ncbi.nlm.nih.gov/pubmed/33031744)

 REMARK Publication Status: Available-Online prior to print

REFERENCE 2 (bases 1 to 29882)

 AUTHORS Leist,S.R., Dinnon,K.H. III, Schafer,A., Tse,L.V., Okuda,K.,

 Hou,Y.J., West,A., Edwards,C.E., Sanders,W., Fritch,E.J.,

 Gully,K.L., Scobey,T., Brown,A.J., Sheahan,T.P., Moorman,N.J.,

 Boucher,R.C., Gralinski,L.E., Montgomery,S.A. and Baric,R.S.

 TITLE Direct Submission

 JOURNAL Submitted (31-AUG-2020) Epidemiology, University of North Carolina

 at Chapel Hill, 2107 McGavran-Greenberg CB7435, Chapel Hill, NC

 27599, USA

COMMENT ##Assembly-Data-START##

 Sequencing Technology :: Sanger dideoxy sequencing

 ##Assembly-Data-END##

FEATURES Location/Qualifiers

 source 1..29882

 /organism="Severe acute respiratory syndrome coronavirus

 2"

 /mol\_type="genomic RNA"

 /db\_xref="taxon:[2697049](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2697049)"

 /clone="SARS-CoV-2-MA10"

 /note="mouse-adapted; passage 10"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=266&to=21555) 266..21555

 /gene="ORF1ab"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?location=266:13468,13468:21555) join(266..13468,13468..21555)

 /gene="ORF1ab"

 /ribosomal\_slippage

 /codon\_start=1

 /product="ORF1ab polyprotein"

 /protein\_id="[QNN93351.1](https://www.ncbi.nlm.nih.gov/protein/1898953379)"

 /translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ

 HLKDGTCGLVEVEKGVLPQLEQPYVFIKRSDARTAPHGHVMVELVAELEGIQYGRSGE

 TLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHSYGADLKSFDLGDELGTDPYEDFQEN

 WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ

 LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP

 LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTG

 DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESG

 LKTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNL

 LEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGN

 FKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAA

 ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL

 KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLV

 NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII

 FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEK

 YCALAPNMMVTNNTFTLKGGAPTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEK

 CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDESGEF

 KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPE

 EEQEEDWLDDDSQQTVGQQDGSEDNQTTTIQTIVEVQPQLEMELTPVVQTIEVNSFSG

 YLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD

 DYIATNGPLKVGGSCVLSGHNLAKHCLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLA

 PLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIA

 EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTLEETKFLTENLLLYIDINGN

 LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKV

 PTDNYITTYPGQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM

 LAHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLIN

 TLNDLNETLVTMPLGYVTHGLNLEEAARYMRSLKVPATVSVSSPDAVTAYNGYLTSSS

 KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVIT

 FDNLKTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPH

 NSHEGKTFYVLPNDDTLRVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIK

 WADNNCYLATALLTLQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELG

 DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSYEQFKKGVQ

 IPCTCGKQATKYLVQQESPFVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK

 ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIKPVTYKLDGVVCTEIDPKLDNYY

 KKDNSYFTEQPIDLVPNQPYPNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVT

 FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNTWCIRCLWST

 KPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD

 IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS

 VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRI

 KASMPTTIAKNTVKSVGKFCLEASFNYLKSPNFSKLINIIIWFLLLSVCLGSLIYSTA

 ALGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSLDTYPSLET

 IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL

 MWLIINLVQMAPISAMVRMYIFFASFYYVWKSYVHVVDGCNSSTCMMCYKRNRATRVE

 CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRP

 INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSLPI

 NVIVFDGKSKCEESSAKSASVYYSQLMCQPILLLDQALVSDVGDSAEVAVKMFDAYVN

 TFSSTFNVPMEKLKTLVATAEAELAKNVSLDNVLSTFISAARQGFVDSDVETKDVVEC

 LKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI

 WNVKDFMSLSEQLRKQIRSAAKKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL

 KQLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFA

 NKHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFVVPGLPGTILRTTNGDFLHFLP

 RVFSAVGNICYTPSKLIEYTDFATSACVLAAECTIFKDASGKPVPYCYDTNVLEGSVA

 YESLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSG

 RWVLNNDYYRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDISASIVAGGIVAIVVICL

 AYYFMRFRRAFGEYSHVVAFNTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTN

 DVSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFE

 EAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAACCHL

 AKALNDFSNSGSDVLYQPPQTSITSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNG

 LWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL

 KLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSC

 GSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN

 VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAV

 LDMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHW

 LLLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFL

 LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTAR

 TVYDDGARRVWTLMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTTVMFLAR

 GIVFMCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYL

 VSTQEFRYMNSQGLLPPKNSIDAFKLNIKLLGVGGKPCIKVATVQSRMSDVKCTSVVL

 LSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKL

 CEEMLDNRATLQAIASEFSSLPSYAAFATAQEAYGQAVANGDSEVVLKKLKKSLNVAK

 SEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALN

 NIINNARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDA

 DSKIVQLSEISMDNSPNLAWPLIVTALRANSAVKLQNNELSPVALRQMSCAAGTTQTA

 CTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP

 KGPKVKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAK

 AYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDH

 PNPKGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSA

 DAQSFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKD

 EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQR

 LTKYTMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYA

 NLGERVRQALLKTVQFCDAMRNAGIVGVLTLDNQDLNGNWYDFGDFIQTTPGSGVPVV

 DSYYSLLMPILTLTRALTAESHVDTDLTKPYIKWDLLKYDFTEERLKLFDRYFKYWDQ

 TYHPNCVNCLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRE

 LGVVHNQDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQ

 TVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDI

 RQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQ

 DALFAYTKRNVIPTITQMNLKYAISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAAT

 RGATVVIGTSKFYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLAR

 KHTTCCSLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQ

 AVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVDTDFVNEFYAYLRKHFSMM

 ILSDDAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGPHEFCS

 QHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKH

 PNQEYADVFHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTV

 LQAVGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCD

 VTDVTQLYLGGMSYYCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDW

 TNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRELHLSWEVGKP

 RPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSH

 TVMPLSAPTLVPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGK

 SHFAIGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKCSRIIPARARVECFDKF

 KVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQ

 LPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLK

 AHKDKSAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNA

 VASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNRFNVAITRAKVGILCIMSD

 RDLYDKLQFTSLEIPRRNVATLQAENVTGLFKDCSKVITGLHPTQAPTHLSVDTKFKT

 EGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEG

 CHATREAVGTNLPLQLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAKPPPGDQFKHLIP

 LMYKGLPWNVVRIKIVQMLSDTLKNLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCL

 CDRRATCFSTASDTYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNA

 HVASCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACRKVQHMVVKAALLADK

 FPVLHDIGNPKAIKCVPQADVEWKFYDAQPCSDKAYKIEELFYSYATHSDKFTDGVCL

 FWNCNVDRYPANSIVCRFDTRVLSNLNLPGCDGGSLYVNKHAFHTPAFDKSAFVNLKQ

 LPFFYYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAVCRHHANEYRLYLDAYN

 MMISAGFSLWVYKQFDTYNLWNTFTRLQSLENVAFNVVNKGHFDGQQGEVPVSIINNT

 VYTKVDGVDVELFENKTTLPVNVAFELWAKRNIKPVPEVKILNNLGVDIAANTVIWDY

 KRDAPAHISTIGVCSMTDIAKKPTETICAPLTVFFDGRVDGQVDLFRNARNGVLITEG

 SVKGLQPSVGPKQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFK

 PRSQMEIDFLELAMDEFIERYKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESP

 FELEDFIPMDSTVKNYFITDAQTGSSKCVCSVIDLLLDDFVEIIKSQDLSVVSKVVKV

 TIDYTEISFMLWCKDGHVETFYPKLQSSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGD

 SATLPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLP

 TGTLLVDSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMYDPKTKNVTKENDSKEG

 FFTYICGFIQQKLALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSSEAFL

 IGCNYLGKPREQIDGYVMHANYIFWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKE

 GQINDMILSLLSKGRLIIRENNRVVISSDVLVNN"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=1&to=180) 266..805

 /gene="ORF1ab"

 /product="leader protein"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=181&to=818) 806..2719

 /gene="ORF1ab"

 /product="nsp2"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=819&to=2763) 2720..8554

 /gene="ORF1ab"

 /product="nsp3"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=2764&to=3263) 8555..10054

 /gene="ORF1ab"

 /product="nsp4"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=3264&to=3569) 10055..10972

 /gene="ORF1ab"

 /product="3C-like proteinase"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=3570&to=3859) 10973..11842

 /gene="ORF1ab"

 /product="nsp6"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=3860&to=3942) 11843..12091

 /gene="ORF1ab"

 /product="nsp7"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=3943&to=4140) 12092..12685

 /gene="ORF1ab"

 /product="nsp8"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=4141&to=4253) 12686..13024

 /gene="ORF1ab"

 /product="nsp9"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=4254&to=4392) 13025..13441

 /gene="ORF1ab"

 /product="nsp10"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=4393&to=5324) join(13442..13468,13468..16236)

 /gene="ORF1ab"

 /product="RNA-dependent RNA polymerase"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=5325&to=5925) 16237..18039

 /gene="ORF1ab"

 /product="helicase"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=5926&to=6452) 18040..19620

 /gene="ORF1ab"

 /product="3'-to-5' exonuclease"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=6453&to=6798) 19621..20658

 /gene="ORF1ab"

 /product="endoRNAse"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93351.1?from=6799&to=7096) 20659..21552

 /gene="ORF1ab"

 /product="2'-O-ribose methyltransferase"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=266&to=13483) 266..13483

 /gene="ORF1ab"

 /codon\_start=1

 /product="ORF1a polyprotein"

 /protein\_id="[QNN93352.1](https://www.ncbi.nlm.nih.gov/protein/1898953380)"

 /translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ

 HLKDGTCGLVEVEKGVLPQLEQPYVFIKRSDARTAPHGHVMVELVAELEGIQYGRSGE

 TLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHSYGADLKSFDLGDELGTDPYEDFQEN

 WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ

 LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP

 LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTG

 DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESG

 LKTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNL

 LEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGN

 FKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAA

 ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL

 KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLV

 NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII

 FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEK

 YCALAPNMMVTNNTFTLKGGAPTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEK

 CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDESGEF

 KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPE

 EEQEEDWLDDDSQQTVGQQDGSEDNQTTTIQTIVEVQPQLEMELTPVVQTIEVNSFSG

 YLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD

 DYIATNGPLKVGGSCVLSGHNLAKHCLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLA

 PLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIA

 EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTLEETKFLTENLLLYIDINGN

 LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKV

 PTDNYITTYPGQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM

 LAHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLIN

 TLNDLNETLVTMPLGYVTHGLNLEEAARYMRSLKVPATVSVSSPDAVTAYNGYLTSSS

 KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVIT

 FDNLKTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPH

 NSHEGKTFYVLPNDDTLRVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIK

 WADNNCYLATALLTLQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELG

 DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSYEQFKKGVQ

 IPCTCGKQATKYLVQQESPFVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK

 ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIKPVTYKLDGVVCTEIDPKLDNYY

 KKDNSYFTEQPIDLVPNQPYPNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVT

 FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNTWCIRCLWST

 KPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD

 IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS

 VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRI

 KASMPTTIAKNTVKSVGKFCLEASFNYLKSPNFSKLINIIIWFLLLSVCLGSLIYSTA

 ALGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSLDTYPSLET

 IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL

 MWLIINLVQMAPISAMVRMYIFFASFYYVWKSYVHVVDGCNSSTCMMCYKRNRATRVE

 CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRP

 INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSLPI

 NVIVFDGKSKCEESSAKSASVYYSQLMCQPILLLDQALVSDVGDSAEVAVKMFDAYVN

 TFSSTFNVPMEKLKTLVATAEAELAKNVSLDNVLSTFISAARQGFVDSDVETKDVVEC

 LKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI

 WNVKDFMSLSEQLRKQIRSAAKKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL

 KQLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFA

 NKHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFVVPGLPGTILRTTNGDFLHFLP

 RVFSAVGNICYTPSKLIEYTDFATSACVLAAECTIFKDASGKPVPYCYDTNVLEGSVA

 YESLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSG

 RWVLNNDYYRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDISASIVAGGIVAIVVICL

 AYYFMRFRRAFGEYSHVVAFNTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTN

 DVSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFE

 EAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAACCHL

 AKALNDFSNSGSDVLYQPPQTSITSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNG

 LWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL

 KLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSC

 GSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN

 VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAV

 LDMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHW

 LLLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFL

 LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTAR

 TVYDDGARRVWTLMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTTVMFLAR

 GIVFMCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYL

 VSTQEFRYMNSQGLLPPKNSIDAFKLNIKLLGVGGKPCIKVATVQSRMSDVKCTSVVL

 LSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKL

 CEEMLDNRATLQAIASEFSSLPSYAAFATAQEAYGQAVANGDSEVVLKKLKKSLNVAK

 SEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALN

 NIINNARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDA

 DSKIVQLSEISMDNSPNLAWPLIVTALRANSAVKLQNNELSPVALRQMSCAAGTTQTA

 CTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP

 KGPKVKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAK

 AYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDH

 PNPKGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSA

 DAQSFLNGFAV"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=1&to=180) 266..805

 /gene="ORF1ab"

 /product="leader protein"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=181&to=818) 806..2719

 /gene="ORF1ab"

 /product="nsp2"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=819&to=2763) 2720..8554

 /gene="ORF1ab"

 /product="nsp3"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=2764&to=3263) 8555..10054

 /gene="ORF1ab"

 /product="nsp4"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=3264&to=3569) 10055..10972

 /gene="ORF1ab"

 /product="3C-like proteinase"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=3570&to=3859) 10973..11842

 /gene="ORF1ab"

 /product="nsp6"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=3860&to=3942) 11843..12091

 /gene="ORF1ab"

 /product="nsp7"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=3943&to=4140) 12092..12685

 /gene="ORF1ab"

 /product="nsp8"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=4141&to=4253) 12686..13024

 /gene="ORF1ab"

 /product="nsp9"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=4254&to=4392) 13025..13441

 /gene="ORF1ab"

 /product="nsp10"

 [mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/QNN93352.1?from=4393&to=4405) 13442..13480

 /gene="ORF1ab"

 /product="nsp11"

 [stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=13476&to=13503) 13476..13503

 /gene="ORF1ab"

 /note="Coronavirus frameshifting stimulation element

 stem-loop 1"

 [stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=13488&to=13542) 13488..13542

 /gene="ORF1ab"

 /note="Coronavirus frameshifting stimulation element

 stem-loop 2"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=21563&to=25384) 21563..25384

 /gene="S"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=21563&to=25384) 21563..25384

 /gene="S"

 /codon\_start=1

 /product="surface glycoprotein"

 /protein\_id="[QNN93353.1](https://www.ncbi.nlm.nih.gov/protein/1898953381)"

 /translation="MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFR

 SSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR

 GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVY

 SSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQ

 GFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFL

 LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN

 LCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCF

 TNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN

 YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLKSYGFYTTNGVGYQPY

 RVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFG

 RDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAI

 HADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPR

 RARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTM

 YICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG

 GFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFN

 GLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQN

 VLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGA

 ISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMS

 ECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAH

 FPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELD

 SFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG

 KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSE

 PVLKGVKLHYT"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=25393&to=26220) 25393..26220

 /gene="ORF3a"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=25393&to=26220) 25393..26220

 /gene="ORF3a"

 /codon\_start=1

 /product="ORF3a protein"

 /protein\_id="[QNN93354.1](https://www.ncbi.nlm.nih.gov/protein/1898953382)"

 /translation="MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATATIPIQASLPFG

 WLIVGVALLAVFQSASKIITLKKRWQLALSKGVHFVCNLLLLFVTVYSHLLLVAAGLE

 APFLYLYALVYFLQSINFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIPY

 NSVTSSIVITSGDGTTSPISEHDYQIGGYTEKWESGVKDCVVLHSYFTSDYYQLYSTQ

 LSTDTGVEHVTFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYDEPTTTTSVPL"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=26245&to=26472) 26245..26472

 /gene="E"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=26245&to=26472) 26245..26472

 /gene="E"

 /codon\_start=1

 /product="envelope protein"

 /protein\_id="[QNN93355.1](https://www.ncbi.nlm.nih.gov/protein/1898953383)"

 /translation="MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCC

 NIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=26523&to=27191) 26523..27191

 /gene="M"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=26523&to=27191) 26523..27191

 /gene="M"

 /codon\_start=1

 /product="membrane glycoprotein"

 /protein\_id="[QNN93356.1](https://www.ncbi.nlm.nih.gov/protein/1898953384)"

 /translation="MADSNGTITVEELKKLLEQWNLVIGFLFLTWICLLQFAYANRNR

 FLYIIKLIFLWLLWPVTLACFVLAAVYRINWITGGIAIAMACLVGLMWLSYFIASFRL

 FARTRSMWSFNPETNILLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCD

 IKDLPKEITVATSRTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIA

 LLVQ"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27202&to=27387) 27202..27387

 /gene="ORF6"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27202&to=27387) 27202..27387

 /gene="ORF6"

 /codon\_start=1

 /product="ORF6 protein"

 /protein\_id="[QNN93357.1](https://www.ncbi.nlm.nih.gov/protein/1898953385)"

 /translation="MFHLVDSQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKSL

 TENKYSQLDEEQPMEID"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27394&to=27759) 27394..27759

 /gene="ORF7a"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27394&to=27759) 27394..27759

 /gene="ORF7a"

 /codon\_start=1

 /product="ORF7a protein"

 /protein\_id="[QNN93358.1](https://www.ncbi.nlm.nih.gov/protein/1898953386)"

 /translation="MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNS

 PFHPLADNKFALTCFSTQFAFACPDGVKHVYQLRARSVSPKLFIRQEEVQELYSPIFL

 IVAAIVFITLCFTLKRKTE"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27756&to=27887) 27756..27887

 /gene="ORF7b"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27756&to=27887) 27756..27887

 /gene="ORF7b"

 /codon\_start=1

 /product="ORF7b"

 /protein\_id="[QNN93359.1](https://www.ncbi.nlm.nih.gov/protein/1898953387)"

 /translation="MIELSLIDFYLCFLAFLLFLVLIMLIIFWFSLELQDHNETCHA"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27894&to=28259) 27894..28259

 /gene="ORF8"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=27894&to=28259) 27894..28259

 /gene="ORF8"

 /codon\_start=1

 /product="ORF8 protein"

 /protein\_id="[QNN93360.1](https://www.ncbi.nlm.nih.gov/protein/1898953388)"

 /translation="MKFLVFLGIITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSK

 WYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCSPFTINCQEPKLGSLVVRC

 SFYEDFLEYHDVRVVLDFI"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=28274&to=29533) 28274..29533

 /gene="N"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=28274&to=29533) 28274..29533

 /gene="N"

 /codon\_start=1

 /product="nucleocapsid phosphoprotein"

 /protein\_id="[QNN93361.1](https://www.ncbi.nlm.nih.gov/protein/1898953389)"

 /translation="MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQG

 LPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMK

 DLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQ

 LPQGTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAA

 LALLLLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGR

 RGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYT

 GAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTV

 TLLPAADLDDFSKQLQQSMSSADSTQA"

 [gene](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=29558&to=29674) 29558..29674

 /gene="ORF10"

 [CDS](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=29558&to=29674) 29558..29674

 /gene="ORF10"

 /codon\_start=1

 /product="ORF10 protein"

 /protein\_id="[QNN93362.1](https://www.ncbi.nlm.nih.gov/protein/1898953390)"

 /translation="MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT"

 [stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=29609&to=29644) 29609..29644

 /gene="ORF10"

 /note="Coronavirus 3' UTR pseudoknot stem-loop 1"

 [stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=29629&to=29657) 29629..29657

 /gene="ORF10"

 /note="Coronavirus 3' UTR pseudoknot stem-loop 2"

 [stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1?from=29728&to=29768) 29728..29768

 /note="Coronavirus 3' stem-loop II-like motif (s2m)"

ORIGIN

 1 attaaaggtt tataccttcc caggtaacaa accaaccaac tttcgatctc ttgtagatct

 61 gttctctaaa cgaactttaa aatctgtgtg gctgtcactc ggctgcatgc ttagtgcact

 121 cacgcagtat aattaataac taattactgt cgttgacagg acacgagtaa ctcgtctatc

 181 ttctgcaggc tgcttacggt ttcgtccgtg ttgcagccga tcatcagcac atctaggttt

 241 cgtccgggtg tgaccgaaag gtaagatgga gagccttgtc cctggtttca acgagaaaac

 301 acacgtccaa ctcagtttgc ctgttttaca ggttcgcgac gtgctcgtac gtggctttgg

 361 agactccgtg gaggaggtct tatcagaggc acgtcaacat cttaaagatg gcacttgtgg

 421 cttagtagaa gttgaaaaag gcgttttgcc tcaacttgaa cagccctatg tgttcatcaa

 481 acgttcggat gctcgaactg cacctcatgg tcatgttatg gttgagctgg tagcagaact

 541 cgaaggcatt cagtacggtc gtagtggtga gacacttggt gtccttgtcc ctcatgtggg

 601 cgaaatacca gtggcttacc gcaaggttct tcttcgtaag aacggtaata aaggagctgg

 661 tggccatagt tacggcgccg atctaaagtc atttgactta ggcgacgagc ttggcactga

 721 tccttatgaa gattttcaag aaaactggaa cactaaacat agcagtggtg ttacccgtga

 781 actcatgcgt gagcttaacg gaggggcata cactcgctat gtcgataaca acttctgtgg

 841 ccctgatggc taccctcttg agtgcattaa agaccttcta gcacgtgctg gtaaagcttc

 901 atgcactttg tccgaacaac tggactttat tgacactaag aggggtgtat actgctgccg

 961 tgaacatgag catgaaattg cttggtacac ggaacgttct gaaaagagct atgaattgca

 1021 gacacctttt gaaattaaat tggcaaagaa atttgacacc ttcaatgggg aatgtccaaa

 1081 ttttgtattt cccttaaatt ccataatcaa gactattcaa ccaagggttg aaaagaaaaa

 1141 gcttgatggc tttatgggta gaattcgatc tgtctatcca gttgcgtcac caaatgaatg

 1201 caaccaaatg tgcctttcaa ctctcatgaa gtgtgatcat tgtggtgaaa cttcatggca

 1261 gacgggcgat tttgttaaag ccacttgcga attttgtggc actgagaatt tgactaaaga

 1321 aggtgccact acttgtggtt acttacccca aaatgctgtt gttaaaattt attgtccagc

 1381 atgtcacaat tcagaagtag gacctgagca tagtcttgcc gaataccata atgaatctgg

 1441 cttgaaaacc attcttcgta agggtggtcg cactattgcc tttggaggct gtgtgttctc

 1501 ttatgttggt tgccataaca agtgtgccta ttgggttcca cgtgctagcg ctaacatagg

 1561 ttgtaaccat acaggtgttg ttggagaagg ttccgaaggt cttaatgaca accttcttga

 1621 aatactccaa aaagagaaag tcaacatcaa tattgttggt gactttaaac ttaatgaaga

 1681 gatcgccatt attttggcat ctttttctgc ttccacaagt gcttttgtgg aaactgtgaa

 1741 aggtttggat tataaagcat tcaaacaaat tgttgaatcc tgtggtaatt ttaaagttac

 1801 aaaaggaaaa gctaaaaaag gtgcctggaa tattggtgaa cagaaatcaa tactgagtcc

 1861 tctttatgca tttgcatcag aggctgctcg tgttgtacga tcaattttct cccgcactct

 1921 tgaaactgct caaaattctg tgcgtgtttt acagaaggcc gctataacaa tactagatgg

 1981 aatttcacag tattcactga gactcattga tgctatgatg ttcacatctg atttggctac

 2041 taacaatcta gttgtaatgg cctacattac aggtggtgtt gttcagttga cttcgcagtg

 2101 gctaactaac atctttggca ctgtttatga aaaactcaaa cccgtccttg attggcttga

 2161 agagaagttt aaggaaggtg tagagtttct tagagacggt tgggaaattg ttaaatttat

 2221 ctcaacctgt gcttgtgaaa ttgtcggtgg acaaattgtc acctgtgcaa aggaaattaa

 2281 ggagagtgtt cagacattct ttaagcttgt aaataaattt ttggctttgt gtgctgactc

 2341 tatcattatt ggtggagcta aacttaaagc cttgaattta ggtgaaacat ttgtcacgca

 2401 ctcaaaggga ttgtacagaa agtgtgttaa atccagagaa gaaactggcc tactcatgcc

 2461 tctaaaagcc ccaaaagaaa ttatcttctt agagggagaa acacttccca cagaagtgtt

 2521 aacagaggaa gttgtcttga aaactggtga tttacaacca ttagaacaac ctactagtga

 2581 agctgttgaa gctccattgg ttggtacacc agtttgtatt aacgggctta tgttgctcga

 2641 aatcaaagac acagaaaagt actgtgccct tgcacctaat atgatggtaa caaacaatac

 2701 cttcacactc aaaggcggtg caccaacaaa ggttactttt ggtgatgaca ctgtgataga

 2761 agtgcaaggt tacaagagtg tgaatatcac ttttgaactt gatgaaagga ttgataaagt

 2821 acttaatgag aagtgctctg cctatacagt tgaactcggt acagaagtaa atgagttcgc

 2881 ctgtgttgtg gcagatgctg tcataaaaac tttgcaacca gtatctgaat tacttacacc

 2941 actgggcatt gatttagatg agtggagtat ggctacatac tacttatttg atgagtctgg

 3001 tgagtttaaa ttggcttcac atatgtattg ttctttctac cctccagatg aggatgaaga

 3061 agaaggtgat tgtgaagaag aagagtttga gccatcaact caatatgagt atggtactga

 3121 agatgattac caaggtaaac ctttggaatt tggtgccact tctgctgctc ttcaacctga

 3181 agaagagcaa gaagaagatt ggttagatga tgatagtcaa caaactgttg gtcaacaaga

 3241 cggcagtgag gacaatcaga caactactat tcaaacaatt gttgaggttc aacctcaatt

 3301 agagatggaa cttacaccag ttgttcagac tattgaagtg aatagtttta gtggttattt

 3361 aaaacttact gacaatgtat acattaaaaa tgcagacatt gtggaagaag ctaaaaaggt

 3421 aaaaccaaca gtggttgtta atgcagccaa tgtttacctt aaacatggag gaggtgttgc

 3481 aggagcctta aataaggcta ctaacaatgc catgcaagtt gaatctgatg attacatagc

 3541 tactaatgga ccacttaaag tgggtggtag ttgtgtttta agcggacaca atcttgctaa

 3601 acactgtctt catgttgtcg gcccaaatgt taacaaaggt gaagacattc aacttcttaa

 3661 gagtgcttat gaaaatttta atcagcacga agttctactt gcaccattat tatcagctgg

 3721 tatttttggt gctgacccta tacattcttt aagagtttgt gtagatactg ttcgcacaaa

 3781 tgtctactta gctgtctttg ataaaaatct ctatgacaaa cttgtttcaa gctttttgga

 3841 aatgaagagt gaaaagcaag ttgaacaaaa gatcgctgag attcctaaag aggaagttaa

 3901 gccatttata actgaaagta aaccttcagt tgaacagaga aaacaagatg ataagaaaat

 3961 caaagcttgt gttgaagaag ttacaacaac tctggaagaa actaagttcc tcacagaaaa

 4021 cttgttactt tatattgaca ttaatggcaa tcttcatcca gattctgcca ctcttgttag

 4081 tgacattgac atcactttct taaagaaaga tgctccatat atagtgggtg atgttgttca

 4141 agagggtgtt ttaactgctg tggttatacc tactaaaaag gctggtggca ctactgaaat

 4201 gctagcgaaa gctttgagaa aagtgccaac agacaattat ataaccactt acccgggtca

 4261 gggtttaaat ggttacactg tagaggaggc aaagacagtg cttaaaaagt gtaaaagtgc

 4321 cttttacatt ctaccatcta ttatctctaa tgagaagcaa gaaattcttg gaactgtttc

 4381 ttggaatttg cgagaaatgc ttgcacatgc agaagaaaca cgcaaattaa tgcctgtctg

 4441 tgtggaaact aaagccatag tttcaactat acagcgtaaa tataagggta ttaaaataca

 4501 agagggtgtg gttgattatg gtgctagatt ttacttttac accagtaaaa caactgtagc

 4561 gtcacttatc aacacactta acgatctaaa tgaaactctt gttacaatgc cacttggcta

 4621 tgtaacacat ggcttaaatt tggaagaagc tgctcggtat atgagatctc tcaaagtgcc

 4681 agctacagtt tctgtttctt cacctgatgc tgttacagcg tataatggtt atcttacttc

 4741 ttcttctaaa acacctgaag aacattttat tgaaaccatc tcacttgctg gttcctataa

 4801 agattggtcc tattctggac aatctacaca actaggtata gaatttctta agagaggtga

 4861 taaaagtgta tattacacta gtaatcctac cacattccac ctagatggtg aagttatcac

 4921 ctttgacaat cttaagacac ttctttcttt gagagaagtg aggactatta aggtgtttac

 4981 aacagtagac aacattaacc tccacacgca agttgtggac atgtcaatga catatggaca

 5041 acagtttggt ccaacttatt tggatggagc tgatgttact aaaataaaac ctcataattc

 5101 acatgaaggt aaaacatttt atgttttacc taatgatgac actctacgtg ttgaggcttt

 5161 tgagtactac cacacaactg atcctagttt tctgggtagg tacatgtcag cattaaatca

 5221 cactaaaaag tggaaatacc cacaagttaa tggtttaact tctattaaat gggcagataa

 5281 caactgttat cttgccactg cattgttaac actccaacaa atagagttga agtttaatcc

 5341 acctgctcta caagatgctt attacagagc aagggctggt gaagctgcta acttttgtgc

 5401 acttatctta gcctactgta ataagacagt aggtgagtta ggtgatgtta gagaaacaat

 5461 gagttacttg tttcaacatg ccaatttaga ttcttgcaaa agagtcttga acgtggtgtg

 5521 taaaacttgt ggacaacagc agacaaccct taagggtgta gaagctgtta tgtacatggg

 5581 cacactttct tatgaacaat ttaagaaagg tgttcagata ccttgtacgt gtggtaaaca

 5641 agctacaaaa tatctagtac aacaggagtc accttttgtt atgatgtcag caccacctgc

 5701 tcagtatgaa cttaagcatg gtacatttac ttgtgctagt gagtacactg gtaattacca

 5761 gtgtggtcac tataaacata taacttctaa agaaactttg tattgcatag acggtgcttt

 5821 acttacaaag tcctcagaat acaaaggtcc tattacggat gttttctaca aagaaaacag

 5881 ttacacaaca accataaaac cagttactta taaattggat ggtgttgttt gtacagaaat

 5941 tgaccctaag ttggacaatt attataagaa agacaattct tatttcacag agcaaccaat

 6001 tgatcttgta ccaaaccaac catatccaaa cgcaagcttc gataatttta agtttgtatg

 6061 tgataatatc aaatttgctg atgatttaaa ccagttaact ggttataaga aacctgcttc

 6121 aagagagctt aaagttacat ttttccctga cttaaatggt gatgtggtgg ctattgatta

 6181 taaacactac acaccctctt ttaagaaagg agctaaattg ttacataaac ctattgtttg

 6241 gcatgttaac aatgcaacta ataaagccac gtataaacca aatacctggt gtatacgttg

 6301 tctttggagc acaaaaccag ttgaaacatc aaattcgttt gatgtactga agtcagagga

 6361 cgcgcaggga atggataatc ttgcctgcga agatctaaaa ccagtctctg aagaagtagt

 6421 ggaaaatcct accatacaga aagacgttct tgagtgtaat gtgaaaacta ccgaagttgt

 6481 aggagacatt atacttaaac cagcaaataa tagtttaaaa attacagaag aggttggcca

 6541 cacagatcta atggctgctt atgtagacaa ttctagtctt actattaaga aacctaatga

 6601 attatctaga gtattaggtt tgaaaaccct tgctactcat ggtttagctg ctgttaatag

 6661 tgtcccttgg gatactatag ctaattatgc taagcctttt cttaacaaag ttgttagtac

 6721 aactactaac atagttacac ggtgtttaaa ccgtgtttgt actaattata tgccttattt

 6781 ctttacttta ttgctacaat tgtgtacttt tactagaagt acaaattcta gaattaaagc

 6841 atctatgccg actactatag caaagaatac tgttaagagt gtcggtaaat tttgtctaga

 6901 ggcttcattt aattatttga agtcacctaa tttttctaaa ctgataaata ttataatttg

 6961 gtttttacta ttaagtgttt gcctaggttc tttaatctac tcaaccgctg ctttaggtgt

 7021 tttaatgtct aatttaggca tgccttctta ctgtactggt tacagagaag gctatttgaa

 7081 ctctactaat gtcactattg caacctactg tactggttct ataccttgta gtgtttgtct

 7141 tagtggttta gattctttag acacctatcc ttctttagaa actatacaaa ttaccatttc

 7201 atcttttaaa tgggatttaa ctgcttttgg cttagttgca gagtggtttt tggcatatat

 7261 tcttttcact aggtttttct atgtacttgg attggctgca atcatgcaat tgtttttcag

 7321 ctattttgca gtacatttta ttagtaattc ttggcttatg tggttaataa ttaatcttgt

 7381 acaaatggcc ccgatttcag ctatggttag aatgtacatc ttctttgcat cattttatta

 7441 tgtatggaaa agttatgtgc atgttgtaga cggttgtaat tcatcaactt gtatgatgtg

 7501 ttacaaacgt aatagagcaa caagagtcga atgtacaact attgttaatg gtgttagaag

 7561 gtccttttat gtctatgcta atggaggtaa aggcttttgc aaactacaca attggaattg

 7621 tgttaattgt gatacattct gtgctggtag tacatttatt agtgatgaag ttgcgagaga

 7681 cttgtcacta cagtttaaaa gaccaataaa tcctactgac cagtcttctt acatcgttga

 7741 tagtgttaca gtgaagaatg gttccatcca tctttacttt gataaagctg gtcaaaagac

 7801 ttatgaaaga cattctctct ctcattttgt taacttagac aacctgagag ctaataacac

 7861 taaaggttca ttgcctatta atgttatagt ttttgatggt aaatcaaaat gtgaagaatc

 7921 atctgcaaaa tcagcgtctg tttactacag tcagcttatg tgtcaaccta tactgttact

 7981 agatcaggca ttagtgtctg atgttggtga tagtgcggaa gttgcagtta aaatgtttga

 8041 tgcttacgtt aatacgtttt catcaacttt taacgtacca atggaaaaac tcaaaacact

 8101 agttgcaact gcagaagctg aacttgcaaa gaatgtgtcc ttagacaatg tcttatctac

 8161 ttttatttca gcagctcggc aagggtttgt tgattcagat gtagaaacta aagatgttgt

 8221 tgaatgtctt aaattgtcac atcaatctga catagaagtt actggcgata gttgtaataa

 8281 ctatatgctc acctataaca aagttgaaaa catgacaccc cgtgaccttg gtgcttgtat

 8341 tgactgtagt gcgcgtcata ttaatgcgca ggtagcaaaa agtcacaaca ttgctttgat

 8401 atggaacgtt aaagatttca tgtcattgtc tgaacaacta cgaaaacaaa tacgtagtgc

 8461 tgctaaaaag aataacttac cttttaagtt gacatgtgca actactagac aagttgttaa

 8521 tgttgtaaca acaaagatag cacttaaggg tggtaaaatt gttaataatt ggttgaagca

 8581 gttaattaaa gttacacttg tgttcctttt tgttgctgct attttctatt taataacacc

 8641 tgttcatgtc atgtctaaac atactgactt ttcaagtgaa atcataggat acaaggctat

 8701 tgatggtggt gtcactcgtg acatagcatc tacagatact tgttttgcta acaaacatgc

 8761 tgattttgac acatggttta gccagcgtgg tggtagttat actaatgaca aagcttgccc

 8821 attgattgct gcagtcataa caagagaagt gggttttgtc gtgcctggtt tgcctggcac

 8881 gatattacgc acaactaatg gtgacttttt gcatttctta cctagagttt ttagtgcagt

 8941 tggtaacatc tgttacacac catcaaaact tatagagtac actgactttg caacatcagc

 9001 ttgtgttttg gctgctgaat gtacaatttt taaagatgct tctggtaagc cagtaccata

 9061 ttgttatgat accaatgtac tagaaggttc tgttgcttat gaaagtttac gccctgacac

 9121 acgttatgtg ctcatggatg gctctattat tcaatttcct aacacctacc ttgaaggttc

 9181 tgttagagtg gtaacaactt ttgattctga gtactgtagg cacggcactt gtgaaagatc

 9241 agaagctggt gtttgtgtat ctactagtgg tagatgggta cttaacaatg attattacag

 9301 atctttacca ggagttttct gtggtgtaga tgctgtaaat ttacttacta atatgtttac

 9361 accactaatt caacctattg gtgctttgga catatcagca tctatagtag ctggtggtat

 9421 tgtagctatc gtagtaatat gccttgccta ctattttatg aggtttagaa gagcttttgg

 9481 tgaatacagt catgtagttg cctttaatac tttactattc cttatgtcat tcactgtact

 9541 ctgtttaaca ccagtttact cattcttacc tggtgtttat tctgttattt acttgtactt

 9601 gacattttat cttactaatg atgtttcttt tttagcacat attcagtgga tggttatgtt

 9661 cacaccttta gtacctttct ggataacaat tgcttatatc atttgtattt ccacaaagca

 9721 tttctattgg ttctttagta attacctaaa gagacgtgta gtctttaatg gtgtttcctt

 9781 tagtactttt gaagaagctg cgctgtgcac ctttttgtta aataaagaaa tgtatctaaa

 9841 gttgcgtagt gatgtgctat tacctcttac gcaatataat agatacttag ctctttataa

 9901 taagtacaag tattttagtg gagcaatgga tacaactagc tacagagaag ctgcttgttg

 9961 tcatctcgca aaggctctca atgacttcag taactcaggt tctgatgttc tttaccaacc

 10021 accacaaacc tctatcacct cagctgtttt gcagagtggt tttagaaaaa tggcattccc

 10081 atctggtaaa gttgagggtt gtatggtaca agtaacttgt ggtacaacta cacttaacgg

 10141 tctttggctt gatgacgtag tttactgtcc aagacatgtg atctgcacct ctgaagacat

 10201 gcttaaccct aattatgaag atttactcat tcgtaagtct aatcataatt tcttggtaca

 10261 ggctggtaat gttcaactca gggttattgg acattctatg caaaattgtg tacttaagct

 10321 taaggttgat acagccaatc ctaagacacc taagtataag tttgttcgca ttcaaccagg

 10381 acagactttt tcagtgttag cttgttacaa tggttcacca tctggtgttt accaatgtgc

 10441 tatgaggccc aatttcacta ttaagggttc attccttaat ggttcatgtg gtagtgttgg

 10501 ttttaacata gattatgact gtgtctcttt ttgttacatg caccatatgg aattaccaac

 10561 tggagttcat gctggcacag acttagaagg taacttttat ggaccttttg ttgacaggca

 10621 aacagcacaa gcagctggta cggacacaac tattacagtt aatgttttag cttggttgta

 10681 cgctgctgtt ataaatggag acaggtggtt tctcaatcga tttaccacaa ctcttaatga

 10741 ctttaacctt gtggctatga agtacaatta tgaacctcta acacaagacc atgttgacat

 10801 actaggacct ctttctgctc aaactggaat tgccgtttta gatatgtgtg cttcattaaa

 10861 agaattactg caaaatggta tgaatggacg taccatattg ggtagtgctt tattagaaga

 10921 tgaatttaca ccttttgatg ttgttagaca atgctcaggt gttactttcc aaagtgcagt

 10981 gaaaagaaca atcaagggta cacaccactg gttgttactc acaattttga cttcactttt

 11041 agttttagtc cagagtactc aatggtcttt gttctttttt ttgtatgaaa atgccttttt

 11101 accttttgct atgggtatta ttgctatgtc tgcttttgca atgatgtttg tcaaacataa

 11161 gcatgcattt ctctgtttgt ttttgttacc ttctcttgcc actgtagctt attttaatat

 11221 ggtctatatg cctgctagtt gggtgatgcg tattatgaca tggttggata tggttgatac

 11281 tagtttgtct ggttttaagc taaaagactg tgttatgtat gcatcagctg tagtgttact

 11341 aatccttatg acagcaagaa ctgtgtatga tgatggtgct aggagagtgt ggacacttat

 11401 gaatgtcttg acactcgttt ataaagttta ttatggtaat gctttagatc aagccatttc

 11461 catgtgggct cttataatct ctgttacttc taactactca ggtgtagtta caactgtcat

 11521 gtttttggcc agaggtattg tttttatgtg tgttgagtat tgccctattt tcttcataac

 11581 tggtaataca cttcagtgta taatgctagt ttattgtttc ttaggctatt tttgtacttg

 11641 ttactttggc ctcttttgtt tactcaaccg ctactttaga ctgactcttg gtgtttatga

 11701 ttacttagtt tctacacagg agtttagata tatgaattca cagggactac tcccacccaa

 11761 gaatagcata gatgccttca aactcaacat taaattgttg ggtgttggtg gcaaaccttg

 11821 tatcaaagta gccactgtac agtctagaat gtcagatgta aagtgcacat cagtagtctt

 11881 actctcagtt ttgcaacaac tcagagtaga atcatcatct aaattgtggg ctcaatgtgt

 11941 ccagttacac aatgacattc tcttagctaa agatactact gaagcctttg aaaaaatggt

 12001 ttcactactt tctgttttgc tttccatgca gggtgctgta gacataaaca agctttgtga

 12061 agaaatgctg gacaacaggg caaccttaca agctatagcc tcagagttta gttcccttcc

 12121 atcatatgca gcttttgcta ctgctcaaga agcttatggg caggctgttg ctaatggtga

 12181 ttctgaagtt gttcttaaaa agttgaagaa gtctttgaat gtggctaaat ctgaatttga

 12241 ccgtgatgca gccatgcaac gtaagttgga aaagatggct gatcaagcta tgacccaaat

 12301 gtataaacag gctagatctg aggacaagag ggcaaaagtt actagtgcta tgcagacaat

 12361 gcttttcact atgcttagaa agttggataa tgatgcactc aacaacatta tcaacaatgc

 12421 aagagatggt tgtgttccct tgaacataat acctcttaca acagcagcca aactaatggt

 12481 tgtcatacca gactataaca catataaaaa tacgtgtgat ggtacaacat ttacttatgc

 12541 atcagcattg tgggaaatcc aacaggttgt agatgcagat agtaaaattg ttcaacttag

 12601 tgaaattagt atggacaatt cacctaattt agcatggcct cttattgtaa cagctttaag

 12661 ggccaattct gctgtcaaat tacagaataa tgagcttagt cctgttgcac tacgacagat

 12721 gtcttgtgct gccggtacta cacaaactgc ttgcactgat gacaatgcgt tagcttacta

 12781 caacacaaca aagggaggta ggtttgtact tgcactgtta tccgatttac aggatttgaa

 12841 atgggctaga ttccctaaga gtgatggaac tggtactatc tatacagaac tggaaccacc

 12901 ttgtaggttt gttacagaca cacctaaagg tcctaaagtg aagtatttat actttattaa

 12961 aggattaaac aacctaaata gaggtatggt acttggtagt ttagctgcca cagtacgtct

 13021 acaagctggt aatgcaacag aagtgcctgc caattcaact gtattatctt tctgtgcttt

 13081 tgctgtagat gctgctaaag cttacaaaga ttatctagct agtgggggac aaccaatcac

 13141 taattgtgtt aagatgttgt gtacacacac tggtactggt caggcaataa cagttacacc

 13201 ggaagccaat atggatcaag aatcctttgg tggtgcatcg tgttgtctgt actgccgttg

 13261 ccacatagat catccaaatc ctaaaggatt ttgtgactta aaaggtaagt atgtacaaat

 13321 acctacaact tgtgctaatg accctgtggg ttttacactt aaaaacacag tctgtaccgt

 13381 ctgcggtatg tggaaaggtt atggctgtag ttgtgatcaa ctccgcgaac ccatgcttca

 13441 gtcagctgat gcacaatcgt ttttaaacgg gtttgcggtg taagtgcagc ccgtcttaca

 13501 ccgtgcggca caggcactag tactgatgtc gtatacaggg cttttgacat ctacaatgat

 13561 aaagtagctg gttttgctaa attcctaaaa actaattgtt gtcgcttcca agaaaaggac

 13621 gaagatgaca atttaattga ttcttacttt gtagttaaga gacacacttt ctctaactac

 13681 caacatgaag aaacaattta taatttactt aaggattgtc cagctgttgc taaacatgac

 13741 ttctttaagt ttagaataga cggtgacatg gtaccacata tatcacgtca acgtcttact

 13801 aaatacacaa tggcagacct cgtctatgct ttaaggcatt ttgatgaagg taattgtgac

 13861 acattaaaag aaatacttgt cacatacaat tgttgtgatg atgattattt caataaaaag

 13921 gactggtatg attttgtaga aaacccagat atattacgcg tatacgccaa cttaggtgaa

 13981 cgtgtacgcc aagctttgtt aaaaacagta caattctgtg atgccatgcg aaatgctggt

 14041 attgttggtg tactgacatt agataatcaa gatctcaatg gtaactggta tgatttcggt

 14101 gatttcatac aaaccacgcc aggtagtgga gttcctgttg tagattctta ttattcattg

 14161 ttaatgccta tattaacctt gaccagggct ttaactgcag agtcacatgt tgacactgac

 14221 ttaacaaagc cttacattaa gtgggatttg ttaaaatatg acttcacgga agagaggtta

 14281 aaactctttg accgttattt taaatattgg gatcagacat accacccaaa ttgtgttaac

 14341 tgtttggatg acagatgcat tctgcattgt gcaaacttta atgttttatt ctctacagtg

 14401 ttcccaccta caagttttgg accactagtg agaaaaatat ttgttgatgg tgttccattt

 14461 gtagtttcaa ctggatacca cttcagagag ctaggtgttg tacataatca ggatgtaaac

 14521 ttacatagct ctagacttag ttttaaggaa ttacttgtgt atgctgctga ccctgctatg

 14581 cacgctgctt ctggtaatct attactagat aaacgcacta cgtgcttttc agtagctgca

 14641 cttactaaca atgttgcttt tcaaactgtc aaacccggta attttaacaa agacttctat

 14701 gactttgctg tgtctaaggg tttctttaag gaaggaagtt ctgttgaatt aaaacacttc

 14761 ttctttgctc aggatggtaa tgctgctatc agcgattatg actactatcg ttataatcta

 14821 ccaacaatgt gtgatatcag acaactacta tttgtagttg aagttgttga taagtacttt

 14881 gattgttacg atggtggctg tattaatgct aaccaagtca tcgtcaacaa cctagacaaa

 14941 tcagctggtt ttccatttaa taaatggggt aaggctagac tttattatga ttcaatgagt

 15001 tatgaggatc aagatgcact tttcgcatat acaaaacgta atgtcatccc tactataact

 15061 caaatgaatc ttaagtatgc cattagtgca aagaatagag cacgcaccgt agctggtgtc

 15121 tctatctgta gtactatgac caatagacag tttcatcaaa aattattgaa atcaatagcc

 15181 gccactagag gagctactgt agtaattgga acaagcaaat tctatggtgg ttggcacaac

 15241 atgttaaaaa ctgtttatag tgatgtagaa aaccctcacc ttatgggttg ggattatcct

 15301 aaatgtgata gagccatgcc taacatgctt agaattatgg cctcacttgt tcttgctcgc

 15361 aaacatacaa cgtgttgtag cttgtcacac cgtttctata gattagctaa tgagtgtgct

 15421 caagtattga gtgaaatggt catgtgtggc ggttcactat atgttaaacc aggtggaacc

 15481 tcatcaggag atgccacaac tgcttatgct aatagtgttt ttaacatttg tcaagctgtc

 15541 acggccaatg ttaatgcact tttatctact gatggtaaca aaattgccga taagtatgtc

 15601 cgcaatttac aacacagact ttatgagtgt ctctatagaa atagagatgt tgacacagac

 15661 tttgtgaatg agttttacgc atatttgcgt aaacatttct caatgatgat actctctgac

 15721 gatgctgttg tgtgtttcaa tagcacttat gcatctcaag gtctagtggc tagcataaag

 15781 aactttaagt cagttcttta ttatcaaaac aatgttttta tgtctgaagc aaaatgttgg

 15841 actgagactg accttactaa aggacctcat gaattttgct ctcaacatac aatgctagtt

 15901 aaacagggtg atgattatgt gtaccttcct tacccagatc catcaagaat cctaggggcc

 15961 ggctgttttg tagatgatat cgtaaaaaca gatggtacac ttatgattga acggttcgtg

 16021 tctttagcta tagatgctta cccacttact aaacatccta atcaggagta tgctgatgtc

 16081 tttcatttgt acttacaata cataagaaag ctacatgatg agttaacagg acacatgtta

 16141 gacatgtatt ctgttatgct tactaatgat aacacttcaa ggtattggga acctgagttt

 16201 tatgaggcta tgtacacacc gcatacagtc ttacaggctg ttggggcttg tgttctttgc

 16261 aattcacaga cttcattaag atgtggtgct tgcatacgta gaccattctt atgttgtaaa

 16321 tgctgttacg accatgtcat atcaacatca cataaattag tcttgtctgt taatccgtat

 16381 gtttgcaatg ctccaggttg tgatgtcaca gatgtgactc aactttactt aggaggtatg

 16441 agctattatt gtaaatcaca taaaccaccc attagttttc cattgtgtgc taatggacaa

 16501 gtttttggtt tatataaaaa tacatgtgtt ggtagcgata atgttactga ctttaatgca

 16561 attgcaacat gtgactggac aaatgctggt gattacattt tagctaacac ctgtactgaa

 16621 agactcaagc tttttgcagc agaaacgctc aaagctactg aggagacatt taaactgtct

 16681 tatggtattg ctactgtacg tgaagtgctg tctgacagag aattacatct ttcatgggaa

 16741 gttggtaaac ctagaccacc acttaaccga aattatgtct ttactggtta tcgtgtaact

 16801 aaaaacagta aagtacaaat aggagagtac acctttgaaa aaggtgacta tggtgatgct

 16861 gttgtttacc gaggtacaac aacttacaaa ttaaatgttg gtgattattt tgtgctgaca

 16921 tcacatacag taatgccatt aagtgcacct acactagtgc cacaagagca ctatgttaga

 16981 attactggct tatacccaac actcaatatc tcagatgagt tttctagcaa tgttgcaaat

 17041 tatcaaaagg ttggtatgca aaagtattct acactccagg gaccacctgg tactggtaag

 17101 agtcattttg ctattggcct agctctctac tacccttctg ctcgcatagt gtatacagct

 17161 tgctctcatg ccgctgttga tgcactatgt gagaaggcat taaaatattt gcctatagat

 17221 aaatgtagta gaattatacc tgcacgtgct cgtgtagagt gttttgataa attcaaagtg

 17281 aattcaacat tagaacagta tgtcttttgt actgtaaatg cattgcctga gacgacagca

 17341 gatatagttg tctttgatga aatttcaatg gccacaaatt atgatttgag tgttgtcaat

 17401 gccagattac gtgctaagca ctatgtgtac attggcgacc ctgctcaatt acctgcacca

 17461 cgcacattgc taactaaggg cacactagaa ccagaatatt tcaattcagt gtgtagactt

 17521 atgaaaacta taggtccaga catgttcctc ggaacttgtc ggcgttgtcc tgctgaaatt

 17581 gttgacactg tgagtgcttt ggtttatgat aataagctta aagcacataa agacaaatca

 17641 gctcaatgct ttaaaatgtt ttataagggt gttatcacgc atgatgtttc atctgcaatt

 17701 aacaggccac aaataggcgt ggtaagagaa ttccttacac gtaaccctgc ttggagaaaa

 17761 gctgtcttta tttcacctta taattcacag aatgctgtag cctcaaagat tttgggacta

 17821 ccaactcaaa ctgttgattc atcacagggc tcagaatatg actatgtcat attcactcaa

 17881 accactgaaa cagctcactc ttgtaatgta aacagattta atgttgctat taccagagca

 17941 aaagtaggca tactttgcat aatgtctgat agagaccttt atgacaagtt gcaatttaca

 18001 agtcttgaaa ttccacgtag gaatgtggca actttacaag ctgaaaatgt aacaggactt

 18061 tttaaagatt gtagtaaggt aatcactggg ttacatccta cacaggcacc tacacacctc

 18121 agtgttgaca ctaaattcaa aactgaaggt ttatgtgttg acatacctgg catacctaag

 18181 gacatgacct atagaagact catctctatg atgggtttta aaatgaatta tcaagttaat

 18241 ggttacccta acatgtttat cacccgcgaa gaagctataa gacatgtacg tgcatggatt

 18301 ggcttcgatg tcgaggggtg tcatgctact agagaagctg ttggtaccaa tttaccttta

 18361 cagctaggtt tttctacagg tgttaaccta gttgctgtac ctacaggtta tgttgataca

 18421 cctaataata cagatttttc cagagttagt gctaaaccac cgcctggaga tcaatttaaa

 18481 cacctcatac cacttatgta caaaggactt ccttggaatg tagtgcgtat aaagattgta

 18541 caaatgttaa gtgacacact taaaaatctc tctgacagag tcgtatttgt cttatgggca

 18601 catggctttg agttgacatc tatgaagtat tttgtgaaaa taggacctga gcgcacctgt

 18661 tgtctatgtg atagacgtgc cacatgcttt tccactgctt cagacactta tgcctgttgg

 18721 catcattcta ttggatttga ttacgtctat aatccgttta tgattgatgt tcaacaatgg

 18781 ggttttacag gtaacctaca aagcaaccat gatctgtatt gtcaagtcca tggtaatgca

 18841 catgtagcta gttgtgatgc aatcatgact aggtgtctag ctgtccacga gtgctttgtt

 18901 aagcgtgttg actggactat tgaatatcct ataattggtg atgaactgaa gattaatgcg

 18961 gcttgtagaa aggttcaaca catggttgtt aaagctgcat tattagcaga caaattccca

 19021 gttcttcacg acattggtaa ccctaaagct attaagtgtg tacctcaagc tgatgtagaa

 19081 tggaagttct atgatgcaca gccttgtagt gacaaagctt ataaaataga agaattattc

 19141 tattcttatg ccacacattc tgacaaattc acagatggtg tatgcctatt ttggaattgc

 19201 aatgtcgata gatatcctgc taattccatt gtttgtagat ttgacactag agtgctatct

 19261 aaccttaact tgcctggttg tgatggtggc agtttgtatg taaataaaca tgcattccac

 19321 acaccagctt ttgataaaag tgcttttgtt aatttaaaac aattaccatt tttctattac

 19381 tctgacagtc catgtgagtc tcatggaaaa caagtagtgt cagatataga ttatgtacca

 19441 ctaaagtctg ctacgtgtat aacacgttgc aatttaggtg gtgctgtctg tagacatcat

 19501 gctaatgagt acagattgta tctcgatgct tataacatga tgatctcagc tggctttagc

 19561 ttgtgggttt acaaacaatt tgatacttat aacctctgga acacttttac aagacttcag

 19621 agtttagaaa atgtggcttt taatgttgta aataagggac actttgatgg acaacagggt

 19681 gaagtaccag tttctatcat taataacact gtttacacaa aagttgatgg tgttgatgta

 19741 gaattgtttg aaaataaaac aacattacct gttaatgtag catttgagct ttgggctaag

 19801 cgcaacatta aaccagtacc agaggtgaaa atactcaata atttgggtgt ggacattgct

 19861 gctaatactg tgatctggga ctacaaaaga gatgctccag cacatatatc tactattggt

 19921 gtttgttcta tgactgacat agccaagaaa ccaactgaaa cgatttgtgc accactcact

 19981 gtcttttttg atggtagagt tgatggtcaa gtagacttat ttagaaatgc ccgtaatggt

 20041 gttcttatta cagaaggtag tgttaaaggt ttacaaccat ctgtaggtcc caaacaagct

 20101 agtcttaatg gagtcacatt aattggagaa gccgtaaaaa cacagttcaa ttattataag

 20161 aaagttgatg gtgttgtcca acaattacct gaaacttact ttactcagag tagaaattta

 20221 caagaattta aacccaggag tcaaatggaa attgatttct tagaattagc tatggatgaa

 20281 ttcattgaac ggtataaatt agaaggctat gccttcgaac atatcgttta tggagatttt

 20341 agtcatagtc agttaggtgg tttacatcta ctgattggac tagctaaacg ttttaaggaa

 20401 tcaccttttg aattagaaga ttttattcct atggacagta cagttaaaaa ctatttcata

 20461 acagatgcgc aaacaggttc atctaagtgt gtgtgttctg ttattgattt attacttgat

 20521 gattttgttg aaataataaa atcccaagat ttatctgtag tttctaaggt tgtcaaagtg

 20581 actattgact atacagaaat ttcatttatg ctttggtgta aagatggcca tgtagaaaca

 20641 ttttacccaa aattacaatc tagtcaagcg tggcaaccgg gtgttgctat gcctaatctt

 20701 tacaaaatgc aaagaatgct attagaaaag tgtgaccttc aaaattatgg tgatagtgca

 20761 acattaccta aaggcataat gatgaatgtc gcaaaatata ctcaactgtg tcaatattta

 20821 aacacattaa cattagctgt accctataat atgagagtta tacattttgg tgctggttct

 20881 gataaaggag ttgcaccagg tacagctgtt ttaagacagt ggttgcctac gggtacgctg

 20941 cttgtcgatt cagatcttaa tgactttgtc tctgatgcag attcaacttt gattggtgat

 21001 tgtgcaactg tacatacagc taataaatgg gatctcatta ttagtgatat gtacgaccct

 21061 aagactaaaa atgttacaaa agaaaatgac tctaaagagg gttttttcac ttacatttgt

 21121 gggtttatac aacaaaagct agctcttgga ggttccgtgg ctataaagat aacagaacat

 21181 tcttggaatg ctgatcttta taagctcatg ggacacttcg catggtggac agcctttgtt

 21241 actaatgtga atgcgtcatc atctgaagca tttttaattg gatgtaatta tcttggcaaa

 21301 ccacgcgaac aaatagatgg ttatgtcatg catgcaaatt acatattttg gaggaataca

 21361 aatccaattc agttgtcttc ctattcttta tttgacatga gtaaatttcc ccttaaatta

 21421 aggggtactg ctgttatgtc tttaaaagaa ggtcaaatca atgatatgat tttatctctt

 21481 cttagtaaag gtagacttat aattagagaa aacaacagag ttgttatttc tagtgatgtt

 21541 cttgttaaca actaaacgaa caatgtttgt ttttcttgtt ttattgccac tagtctctag

 21601 tcagtgtgtt aatcttacaa ccagaactca attaccccct gcatacacta attctttcac

 21661 acgtggtgtt tattaccctg acaaagtttt cagatcctca gttttacatt caactcagga

 21721 cttgttctta cctttctttt ccaatgttac ttggttccat gctatacatg tctctgggac

 21781 caatggtact aagaggtttg ataaccctgt cctaccattt aatgatggtg tttattttgc

 21841 ttccactgag aagtctaaca taataagagg ctggattttt ggtactactt tagattcgaa

 21901 gacccagtcc ctacttattg ttaataacgc tactaatgtt gttattaaag tctgtgaatt

 21961 tcaattttgt aatgatccat ttttgggtgt ttattaccac aaaaacaaca aaagttggat

 22021 ggaaagtgag ttcagagttt attctagtgc gaataattgc acttttgaat atgtctctca

 22081 gccttttctt atggaccttg aaggaaaaca gggtaatttc aaaaatctta gggaatttgt

 22141 gtttaagaat attgatggtt attttaaaat atattctaag cacacgccta ttaatttagt

 22201 gcgtgatctc cctcagggtt tttcggcttt agaaccattg gtagatttgc caataggtat

 22261 taacatcact aggtttcaaa ctttacttgc tttacataga agttatttga ctcctggtga

 22321 ttcttcttca ggttggacag ctggtgctgc agcttattat gtgggttatc ttcaacctag

 22381 gacttttcta ttaaaatata atgaaaatgg aaccattaca gatgctgtag actgtgcact

 22441 tgaccctctc tcagaaacaa agtgtacgtt gaaatccttc actgtagaaa aaggaatcta

 22501 tcaaacttct aactttagag tccaaccaac agaatctatt gttagatttc ctaatattac

 22561 aaacttgtgc ccttttggtg aagtttttaa cgccaccaga tttgcatctg tttatgcttg

 22621 gaacaggaag agaatcagca actgtgttgc tgattattct gtcctatata attccgcatc

 22681 attttccact tttaagtgtt atggagtgtc tcctactaaa ttaaatgatc tctgctttac

 22741 taatgtctat gcagattcat ttgtaattag aggtgatgaa gtcagacaaa tcgctccagg

 22801 gcaaactgga aagattgctg attataatta taaattacca gatgatttta caggctgcgt

 22861 tatagcttgg aattctaaca atcttgattc taaggttggt ggtaattata attacctgta

 22921 tagattgttt aggaagtcta atctcaaacc ttttgagaga gatatttcaa ctgaaatcta

 22981 tcaggccggt agcacacctt gtaatggtgt tgaaggtttt aattgttact ttcctttaaa

 23041 atcatatggt ttctacacga ctaatggtgt tggttaccaa ccatacagag tagtagtact

 23101 ttcttttgaa cttctacatg caccagcaac tgtttgtgga cctaaaaagt ctactaattt

 23161 ggttaaaaac aaatgtgtca atttcaactt caatggttta acaggcacag gtgttcttac

 23221 tgagtctaac aaaaagtttc tgcctttcca acaatttggc agagacattg ctgacactac

 23281 tgatgctgtc cgtgatccac agacacttga gattcttgac attacaccat gttcttttgg

 23341 tggtgtcagt gttataacac caggaacaaa tacttctaac caggttgctg ttctttatca

 23401 ggatgttaac tgcacagaag tccctgttgc tattcatgca gatcaactta ctcctacttg

 23461 gcgtgtttat tctacaggtt ctaatgtttt tcaaacacgt gcaggctgtt taataggggc

 23521 tgaacatgtc aacaactcat atgagtgtga catacccatt ggtgcaggta tatgcgctag

 23581 ttatcagact cagactaatt ctcctcggcg ggcacgtagt gtagctagtc aatccatcat

 23641 tgcctacact atgtcacttg gtgcagaaaa ttcagttgct tactctaata actctattgc

 23701 catacccaca aattttacta ttagtgttac cacagaaatt ctaccagtgt ctatgaccaa

 23761 gacatcagta gattgtacaa tgtacatttg tggtgattca actgaatgca gcaatctttt

 23821 gttgcaatat ggcagttttt gtacacaatt aaaccgtgct ttaactggaa tagctgttga

 23881 acaagacaaa aacacccaag aagtttttgc acaagtcaaa caaatttaca aaacaccacc

 23941 aattaaagat tttggtggtt ttaatttttc acaaatatta ccagatccat caaaaccaag

 24001 caagaggtca tttattgaag atctactttt caacaaagtg acacttgcag atgctggctt

 24061 catcaaacaa tatggtgatt gccttggtga tattgctgct agagacctca tttgtgcaca

 24121 aaagtttaac ggccttactg ttttgccacc tttgctcaca gatgaaatga ttgctcaata

 24181 cacttctgca ctgttagcgg gtacaatcac ttctggttgg acctttggtg caggtgctgc

 24241 attacaaata ccatttgcta tgcaaatggc ttataggttt aatggtattg gagttacaca

 24301 gaatgttctc tatgagaacc aaaaattgat tgccaaccaa tttaatagtg ctattggcaa

 24361 aattcaagac tcactttctt ccacagcaag tgcacttgga aaacttcaag atgtggtcaa

 24421 ccaaaatgca caagctttaa acacgcttgt taaacaactt agctccaatt ttggtgcaat

 24481 ttcaagtgtt ttaaatgata tcctttcacg tcttgacaaa gttgaggctg aagtgcaaat

 24541 tgataggttg atcacaggca gacttcaaag tttgcagaca tatgtgactc aacaattaat

 24601 tagagctgca gaaatcagag cttctgctaa tcttgctgct actaaaatgt cagagtgtgt

 24661 acttggacaa tcaaaaagag ttgatttttg tggaaagggc tatcatctta tgtccttccc

 24721 tcagtcagca cctcatggtg tagtcttctt gcatgtgact tatgtccctg cacaagaaaa

 24781 gaacttcaca actgctcctg ccatttgtca tgatggaaaa gcacactttc ctcgtgaagg

 24841 tgtctttgtt tcaaatggca cacactggtt tgtaacacaa aggaattttt atgaaccaca

 24901 aatcattact acagacaaca catttgtgtc tggtaactgt gatgttgtaa taggaattgt

 24961 caacaacaca gtttatgatc ctttgcaacc tgaattagac tcattcaagg aggagttaga

 25021 taaatatttt aagaatcata catcaccaga tgttgattta ggtgacatct ctggcattaa

 25081 tgcttcagtt gtaaacattc aaaaagaaat tgaccgcctc aatgaggttg ccaagaattt

 25141 aaatgaatct ctcatcgatc tccaagaact tggaaagtat gagcagtata taaaatggcc

 25201 atggtacatt tggctaggtt ttatagctgg cttgattgcc atagtaatgg tgacaattat

 25261 gctttgctgt atgaccagtt gctgtagttg tctcaagggc tgttgttctt gtggatcctg

 25321 ctgcaaattt gatgaagacg actctgagcc agtgctcaaa ggagtcaaat tacattacac

 25381 ataaacgaac ttatggattt gtttatgaga atcttcacaa ttggaactgt aactttgaag

 25441 caaggtgaaa tcaaggatgc tactccttca gattttgttc gcgctactgc aacgataccg

 25501 atacaagcct cactcccttt cggatggctt attgttggcg ttgcacttct tgctgttttt

 25561 cagagcgctt ccaaaatcat aaccctcaaa aagagatggc aactagcact ctccaagggt

 25621 gttcactttg tttgcaactt gctgttgttg tttgtaacag tttactcaca ccttttgctc

 25681 gttgctgctg gccttgaagc cccttttctc tatctttatg ctttagtcta cttcttgcag

 25741 agtataaact ttgtaagaat aataatgagg ctttggcttt gctggaaatg ccgttccaaa

 25801 aacccattac tttatgatgc caactatttt ctttgctggc atactaattg ttacgactat

 25861 tgtatacctt acaatagtgt aacttcttca attgtcatta cttcaggtga tggcacaaca

 25921 agtcctattt ctgaacatga ctaccagatt ggtggttata ctgaaaaatg ggaatctgga

 25981 gtaaaagact gtgttgtatt acacagttac ttcacttcag actattacca gctgtactca

 26041 actcaattga gtacagacac tggtgttgaa catgttacct tcttcatcta caataaaatt

 26101 gttgatgagc ctgaagaaca tgtccaaatt cacacaatcg acggttcatc cggagttgtt

 26161 aatccagtaa tggaaccaat ttatgatgaa ccgacgacga ctactagcgt gcctttgtaa

 26221 gcacaagctg atgagtacga acttatgtac tcattcgttt cggaagagac aggtacgtta

 26281 atagttaata gcgtacttct ttttcttgct ttcgtggtat tcttgctagt tacactagcc

 26341 atccttactg cgcttcgatt gtgtgcgtac tgctgcaata ttgttaacgt gagtcttgta

 26401 aaaccttctt tttacgttta ctctcgtgtt aaaaatctga attcttctag agttcctgat

 26461 cttctggtct aaacgaacta aatattatat tagtttttct gtttggaact ttaattttag

 26521 ccatggcaga ttccaacggt actattaccg ttgaagagct taaaaagctc cttgaacaat

 26581 ggaacctagt aataggtttc ctattcctta catggatttg tcttctacaa tttgcctatg

 26641 ccaacaggaa taggtttttg tatataatta agttaatttt cctctggctg ttatggccag

 26701 taactttagc ttgttttgtg cttgctgctg tttacagaat aaattggatc accggtggaa

 26761 ttgctatcgc aatggcttgt cttgtaggct tgatgtggct cagctacttc attgcttctt

 26821 tcagactgtt tgcgcgtacg cgttccatgt ggtcattcaa tccagaaact aacattcttc

 26881 tcaacgtgcc actccatggc actattctga ccagaccgct tctagaaagt gaactcgtaa

 26941 tcggagctgt gatccttcgt ggacatcttc gtattgctgg acaccatcta ggacgctgtg

 27001 acatcaagga cctgcctaaa gaaatcactg ttgctacatc acgaacgctt tcttattaca

 27061 aattgggagc ttcgcagcgt gtagcaggtg actcaggttt tgctgcatac agtcgctaca

 27121 ggattggcaa ctataaatta aacacagacc attccagtag cagtgacaat attgctttgc

 27181 ttgtacagta agtgacaaca gatgtttcat ctcgttgact ctcaggttac tatagcagag

 27241 atattactaa ttattatgag gacttttaaa gtttccattt ggaatcttga ttacatcata

 27301 aacctcataa ttaaaaattt atctaagtca ctaactgaga ataaatattc tcaattagat

 27361 gaagagcaac caatggagat tgattaaacg aacatgaaaa ttattctttt cttggcactg

 27421 ataacactcg ctacttgtga gctttatcac taccaagagt gtgttagagg tacaacagta

 27481 cttttaaaag aaccttgctc ttctggaaca tacgagggca attcaccatt tcatcctcta

 27541 gctgataaca aatttgcact gacttgcttt agcactcaat ttgcttttgc ttgtcctgac

 27601 ggcgtaaaac acgtctatca gttacgtgcc agatcagttt cacctaaact gttcatcaga

 27661 caagaggaag ttcaagaact ttactctcca atttttctta ttgttgcggc aatagtgttt

 27721 ataacacttt gcttcacact caaaagaaag acagaatgat tgaactttca ttaattgact

 27781 tctatttgtg ctttttagcc tttctgctat tccttgtttt aattatgctt attatctttt

 27841 ggttctcact tgaactgcaa gatcataatg aaacttgtca cgcctaaacg aacatgaaat

 27901 ttcttgtttt cttaggaatc atcacaactg tagctgcatt tcaccaagaa tgtagtttac

 27961 agtcatgtac tcaacatcaa ccatatgtag ttgatgaccc gtgtcctatt cacttctatt

 28021 ctaaatggta tattagagta ggagctagaa aatcagcacc tttaattgaa ttgtgcgtgg

 28081 atgaggctgg ttctaaatca cccattcagt acatcgatat cggtaattat acagtttcct

 28141 gttcaccttt tacaattaat tgccaggaac ctaaattggg tagtcttgta gtgcgttgtt

 28201 cgttctatga agacttttta gagtatcatg acgttcgtgt tgttttagat ttcatctaaa

 28261 cgaacaaact aaaatgtctg ataatggacc ccaaaatcag cgaaatgcac cccgcattac

 28321 gtttggtgga ccctcagatt caactggcag taaccagaat ggagaacgca gtggggcgcg

 28381 atcaaaacaa cgtcggcccc aaggtttacc caataatact gcgtcttggt tcaccgctct

 28441 cactcaacat ggcaaggaag accttaaatt ccctcgagga caaggcgttc caattaacac

 28501 caatagcagt ccagatgacc aaattggcta ctaccgaaga gctaccagac gaattcgtgg

 28561 tggtgacggt aaaatgaaag atctcagtcc aagatggtat ttctactacc taggaactgg

 28621 gccagaagct ggacttccct atggtgctaa caaagacggc atcatatggg ttgcaactga

 28681 gggagccttg aatacaccaa aagatcacat tggcacccgc aatcctgcta acaatgctgc

 28741 aatcgtgcta caacttcctc aaggaacaac attgccaaaa ggcttctacg cagaagggag

 28801 cagaggcggc agtcaagcct cttctcgttc ctcatcacgt agtcgcaaca gttcaagaaa

 28861 ttcaactcca ggcagcagta ggggaacttc tcctgctaga atggctggca atggcggtga

 28921 tgctgctctt gctttgctgc tgcttgacag attgaaccag cttgagagca aaatgtctgg

 28981 taaaggccaa caacaacaag gccaaactgt cactaagaaa tctgctgctg aggcttctaa

 29041 gaagcctcgg caaaaacgta ctgccactaa agcatacaat gtaacacaag ctttcggcag

 29101 acgtggtcca gaacaaaccc aaggaaattt tggggaccag gaactaatca gacaaggaac

 29161 tgattacaaa cattggccgc aaattgcaca atttgccccc agcgcttcag cgttcttcgg

 29221 aatgtcgcgc attggcatgg aagtcacacc ttcgggaacg tggttgacct acacaggtgc

 29281 catcaaattg gatgacaaag atccaaattt caaagatcaa gtcattttgc tgaataagca

 29341 tattgacgca tacaaaacat tcccaccaac agagcctaaa aaggacaaaa agaagaaggc

 29401 tgatgaaact caagccttac cgcagagaca gaagaaacag caaactgtga ctcttcttcc

 29461 tgctgcagat ttggatgatt tctccaaaca attgcaacaa tccatgagca gtgctgactc

 29521 aactcaggcc taaactcatg cagaccacac aaggcagatg ggctatataa acgttttcgc

 29581 ttttccgttt acgatatata gtctactctt gtgcagaatg aattctcgta actacatagc

 29641 acaagtagat gtagttaact ttaatctcac atagcaatct ttaatcagtg tgtaacatta

 29701 gggaggactt gaaagagcca ccacattttc accgaggcca cgcggagtac gatcgagtgt

 29761 acagtgaaca atgctaggga gagctgccta tatggaagag ccctaatgtg taaaattaat

 29821 tttagtagtg ctatccccat gtgattttaa tagcttctta ggagaatgac aaaaaaaaaa

 29881 aa

The above contents are the collected information from Internet PubMed (<https://www.ncbi.nlm.nih.gov/nuccore>) to offer to the people for the convenient reading and information disseminating and sharing.

**References**

1. Baidu. <http://www.baidu.com>. 2019.
2. Cancer Biology. <http://www.cancerbio.net>. 2019.
3. Google. <http://www.google.com>. 2019.
4. <https://www.ncbi.nlm.nih.gov/nuccore/MT952602.1>.
5. <https://www.ncbi.nlm.nih.gov/nuccore/NM_000314.6>.
6. <https://www.ncbi.nlm.nih.gov/nuccore/NM_001178096.2>.
7. <https://www.ncbi.nlm.nih.gov/nuccore/NM_001378933.1>.
8. <https://www.ncbi.nlm.nih.gov/nuccore/NM_001386421.1>.
9. <https://www.ncbi.nlm.nih.gov/nuccore/NM_003281.4>,.
10. <https://www.ncbi.nlm.nih.gov/nuccore/NM_020208.4>.
11. <https://www.ncbi.nlm.nih.gov/nuccore/NM_172014.3>.
12. Journal of American Science. <http://www.jofamericanscience.org>. 2019.
13. Life Science Journal. <http://www.lifesciencesite.com>. 2019.
14. Ma H, Chen G. Stem cell. The Journal of American Science 2005;1(2):90-92. doi:[10.7537/marsjas010205.14](http://www.dx.doi.org/10.7537/marsjas010205.14). <http://www.jofamericanscience.org/journals/am-sci/0102/14-mahongbao.pdf>.
15. Ma H, Cherng S. Eternal Life and Stem Cell. Nature and Science. 2007;5(1):81-96. doi:[10.7537/marsnsj050107.10](http://www.dx.doi.org/10.7537/marsnsj050107.10). <http://www.sciencepub.net/nature/0501/10-0247-mahongbao-eternal-ns.pdf>.
16. Ma H, Cherng S. Nature of Life. Life Science Journal 2005;2(1):7-15. doi:[10.7537/marslsj020105.03](http://www.dx.doi.org/10.7537/marslsj020105.03). <http://www.lifesciencesite.com/lsj/life0201/life-0201-03.pdf>.
17. Ma H, Yang Y. Turritopsis nutricula. Nature and Science 2010;8(2):15-20. doi:[10.7537/marsnsj080210.03](http://www.dx.doi.org/10.7537/marsnsj080210.03). <http://www.sciencepub.net/nature/ns0802/03_1279_hongbao_turritopsis_ns0802_15_20.pdf>.
18. Ma H. The Nature of Time and Space. Nature and science 2003;1(1):1-11. doi:[10.7537/marsnsj010103.01](http://www.dx.doi.org/10.7537/marsnsj010103.01). <http://www.sciencepub.net/nature/0101/01-ma.pdf>.
19. Marsland Press. <http://www.sciencepub.net>. 2019; <http://www.sciencepub.org>. 2019.
20. National Center for Biotechnology Information, U.S. National Library of Medicine. <http://www.ncbi.nlm.nih.gov/pubmed>. 2019.
21. Nature and Science. <http://www.sciencepub.net/nature>. 2019.
22. Stem Cell. <http://www.sciencepub.net/stem>. 2019.
23. Wikipedia. The free encyclopedia. <http://en.wikipedia.org>. 2019.

12/18/2020