

Naloga:

Navodilo: Vaša naloga je da poiščete vse tri genomske zapise za segmente in napišete kako so poimenovani, ter kakšna je njihova dolžina.

Kopiramo ANDV in ga prilepimo v GenBank. Pri filtriranju na genomic DNA/RNA dobimo 68 zadetkov.

Nucleotide [Create alert](#) [Advanced](#)

Species Summary 20 per page Sort by Default order Send to: **File**

Viruses (68) **Items: 1 to 20 of 68**

Customize ...

Molecule types clear genomic DNA/RNA (68) << First < Prev Page 1 of 4 Next > Last >>

mRNA (0) **Filters activated: genomic DNA/RNA [Clear all](#)**

Customize ...

Source databases [Orthohantavirus andesense strain ANDV-LS-CH-2016 ex Chile nucleocapsid protein gene, complete cds](#)

INSDC (GenBank) (68) 1. **1,874 bp linear cRNA**

Customize ... Accession: KY659432.1 GI: 1334385686

Sequence Type [Protein](#) [Taxonomy](#)

Nucleotide (68) [GenBank](#) [FASTA](#) [Graphics](#)

Sequence length [Orthohantavirus andesense strain ANDV-LS-CH-2016 RNA-dependent RNA polymerase gene, complete cds](#)

Custom range... 2. **6,470 bp linear cRNA**

Release date Accession: KY659431.1 GI: 1334385021

Custom range... [Protein](#) [Taxonomy](#)

Revision date [GenBank](#) [FASTA](#) [Graphics](#)

Custom range... [Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment S nucleocapsid protein gene, complete cds](#)

[Clear all](#) 3. **1,793 bp linear cRNA**

[Show additional filters](#) Accession: OR405525.1 GI: 2582768867

[Protein](#) [Taxonomy](#)

[GenBank](#) [FASTA](#) [Graphics](#)

[v/nuccore/KY659432.1](#)

med rezultati identificiramo tri segmente:

- [Protein](#) [Taxonomy](#)
- [GenBank](#) [FASTA](#) [Graphics](#)
- [Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment S nucleocapsid protein gene, complete cds](#)
3. **1,793 bp linear cRNA**
- Accession: OR405525.1 GI: 2582768867
- [Protein](#) [Taxonomy](#)
- [GenBank](#) [FASTA](#) [Graphics](#)
- [Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment M glycoprotein precursor, gene, complete cds](#)
4. **3,606 bp linear cRNA**
- Accession: OR405524.1 GI: 2582768863
- [Protein](#) [Taxonomy](#)
- [GenBank](#) [FASTA](#) [Graphics](#)
- [Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment L RNA-dependent RNA polymerase gene, partial cds](#)
5. **5,924 bp linear cRNA**
- Accession: OR405523.1 GI: 2582768861
- [Protein](#) [Taxonomy](#)
- [GenBank](#) [FASTA](#) [Graphics](#)

Rešitev: S, M, L

Njihove dolžine: ko odpremo zadetke za vsak segment posebej poiščemo njihove dolžine

GenBank [Send to](#)

Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment S nucleocapsid protein gene, complete cds

GenBank: OR405525.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS OR405525 **1793 bp** cRNA linear VRL 02-OCT-2023

DEFINITION Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment S nucleocapsid protein gene, complete cds.

Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment M glycoprotein precursor, gene, complete cds

GenBank: OR405524.1

[FASTA](#) [Graphics](#)

Go to:

LOCUS OR405524 **3606 bp** cRNA linear VRL 02-OCT-2023
DEFINITION Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment M glycoprotein precursor, gene, complete cds.

Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment L RNA-dependent RNA polymerase gene, partial cds

GenBank: OR405523.1

[FASTA](#) [Graphics](#)

Go to:

LOCUS OR405523 **5924 bp** cRNA linear VRL 02-OCT-2023
DEFINITION Orthohantavirus andesense isolate ANDV-CH-LS-2016 segment L

Rešitev: S: 1793 bp, M: 3606 bp, L: 5924 bp

Navodilo: *Zapišite katere proteine kodirajo posamezni segmenti.*

Proteini za katere kodirajo segmenti – razberemo že iz samega imena.

Rešitev:

S: nukleokapsidni protein – protein jedrne kapside

M: prekursor glikoproteina

L: RNA-odvisno RNA polimerazo

Navodilo: *Za protein, ki ga kodira najkrajši segment zapišite UniProt accession code.*

Odpremo UniProt in na advanced search vpišemo ime proteina ter filtriramo organizme.

Advanced Search

Searching in UniProtKB

Protein Name [DE] nucleocapsid protein Remove

Organism [OS] Andes orthohantavirus (ANDV/Andes virus) [1980456] Remove

Add Field Cancel Search

Type * in the search box to search for all values for the selected field.

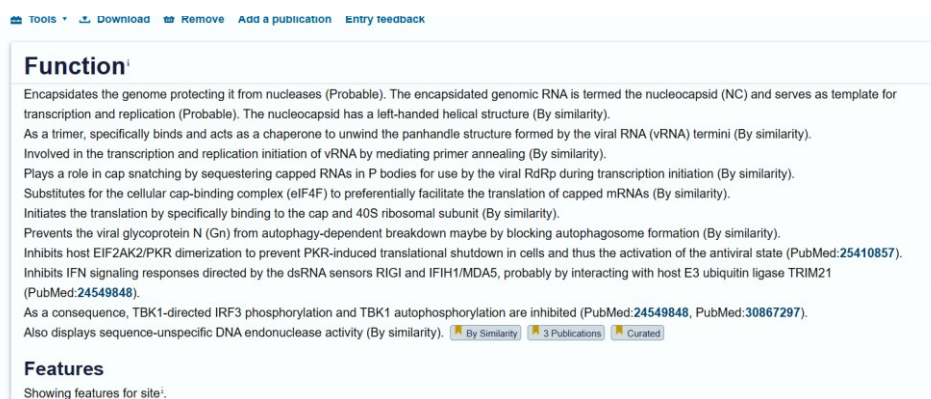
UniProtKB 52 results or expand search to taxon ID "1980456" to include lower taxonomic ranks or restrict search to reference proteome UP000204348

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
O36307	NCAP_ANDV	Nucleoprotein[...]	N	Andes orthohantavirus (ANDV) (Andes virus)	428 AA
COIM21	COIM21_ANDV	Nucleoprotein[...]		Andes orthohantavirus (ANDV) (Andes virus)	132 AA
COIM31	COIM31_ANDV	Nucleoprotein[...]		Andes orthohantavirus (ANDV) (Andes virus)	132 AA
COIM37	COIM37_ANDV	Nucleoprotein[...]		Andes orthohantavirus (ANDV) (Andes virus)	132 AA

Rešitev: Uniprot accession code: O36307

Navodilo: *Navedite 3 funkcije tega proteina.*

Odpremo ustrezen zadetek in si pod razdelkom Function preberemo funkcije proteina.



The screenshot shows the UniProt 'Function' section for protein O36307. At the top, there are navigation links: Tools, Download, Remove, Add a publication, and Entry feedback. The 'Function' section contains several lines of text describing the protein's role: 'Encapsidates the genome protecting it from nucleases (Probable). The encapsidated genomic RNA is termed the nucleocapsid (NC) and serves as template for transcription and replication (Probable). The nucleocapsid has a left-handed helical structure (By similarity). As a trimer, specifically binds and acts as a chaperone to unwind the panhandle structure formed by the viral RNA (vRNA) termini (By similarity). Involved in the transcription and replication initiation of vRNA by mediating primer annealing (By similarity). Plays a role in cap snatching by sequestering capped RNAs in P bodies for use by the viral RdRp during transcription initiation (By similarity). Substitutes for the cellular cap-binding complex (eIF4F) to preferentially facilitate the translation of capped mRNAs (By similarity). Initiates the translation by specifically binding to the cap and 40S ribosomal subunit (By similarity). Prevents the viral glycoprotein N (Gn) from autophagy-dependent breakdown maybe by blocking autophagosome formation (By similarity). Inhibits host EIF2AK2/PKR dimerization to prevent PKR-induced translational shutdown in cells and thus the activation of the antiviral state (PubMed:25410857). Inhibits IFN signaling responses directed by the dsRNA sensors RIGI and IFIH1/MDA5, probably by interacting with host E3 ubiquitin ligase TRIM21 (PubMed:24549848). As a consequence, TBK1-directed IRF3 phosphorylation and TBK1 autophosphorylation are inhibited (PubMed:24549848, PubMed:30867297). Also displays sequence-unspecific DNA endonuclease activity (By similarity). Below the text are three small icons: 'By Similarity', '3 Publications', and 'Curated'. The 'Features' section below shows 'Showing features for site:'.

Rešitev: 3 funkcije proteina:

vse možne rešitve:

- zaščita virusnega RNA genoma pred razgradnjo z nukleazami,
- tvorba nukleokapside okoli virusne RNA,
- sodelovanje pri transkripciji in replikaciji virusne RNA,
- odvijanje »panhandle« strukture virusne RNA,
- pomoč pri začetku replikacije/transkripcije z vezavo začetnih oligonukleotidov,
- sodelovanje pri »cap snatching« procesu,
- omogočanje translacije virusnih mRNA,
- vezava na kapo mRNA in 40S ribosomsko podenoto,
- zaščita virusnega glikoproteina Gn pred avtofagijsko razgradnjo,
- zaviranje PKR/EIF2AK2 signalne poti in preprečevanje zaustavitve translacije,
- zaviranje interferonskega protivirusnega odziva preko vpliva na RIG-I/MDA5 signalizacijo.

Navodilo: *Poiščite ustrezen protein in ga primerjajte s proteinom iz Andskega virusa. Ali bi ocenili, da sta si zaporedji med seboj podobni? (torej proteina iz *Andes virus* in SARS-CoV-2)*

V UniProtu vpišemo: Nucleoprotein ter filtriramo rezultate za SARS-CoV-2. Najdemo protein:

on **PO2TC9 · NCAP_SARS2**

s & Taxonomy
 ilular Location
 types & Variants
 Processing
 ision
 ction
 ure
 & Domains
 nce
 Proteins
 logs

Protein¹ Nucleoprotein
Gene¹ N
Status¹ UniProtKB reviewed (Swiss-Prot)
Organism¹ Severe acute respiratory syndrome coronavirus 2 (2019-nCoV) (SARS-CoV-2)

Amino acids 419 (go to sequence)
Protein existence² Evidence at protein level
Annotation score³

Entry Variant viewer Feature viewer Genomic coordinates Publications External links History

Tools Download Remove Community curated (2) Add a publication Entry feedback

Function¹
 Packages the positive strand viral genome RNA into a helical ribonucleocapsid (RNP) and plays a fundamental role during virion assembly through its interactions with the viral genome and membrane protein M (PubMed:33264373).
 Plays an important role in enhancing the efficiency of subgenomic viral RNA transcription as well as viral replication. Attenuates the stress granules formation by

Proteina nato primerjamo z uporabo EMBOSS Needle – primerjava pokaže da si zaporedji med seboj nista precej podobni.

```

*****
#=====
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 510
# Identity: 91/510 (17.8%)
# Similarity: 147/510 (28.8%)
# Gaps: 173/510 (33.9%)
# Score: 45.5
#
#
#=====
EMBOSS_001 1 MSTLQELQENITAHEQQLVTARQKLKDAEKAVEVDPDDVVKSTLQSRRAA 50
          :.:| ..:..:|:.....|:.....|:..
EMBOSS_001 1 -----MSDN-----GPQNRNAPRITFGGSDSTGSDNGERSG 34
FMROSS_001 51 VSTI FTXI GFI KRDI ADI VAAOKI ATK-----PVDPTGI FPDDH 89
  
```

Navodilo: *Kateri izmed virusov je bolj prilagojen translacijskemu sistemu človeka? Razloži.*

Za nukleokapsidni protein SARS-CoV-2 poiščemo kodirajoče zaporedje in ga kopiramo.

```

27901 LLLTGLLL LLLGGMLL LSLGSLG LGGTGLLL LSLGSLG LGGTGLLL
27961 agtcatgtac tcaacatcaa ccatatgtag ttgatgacc gtgtcttatt cacttcatt
28021 ctaaatggta tattagatga ggagtagaaa aatcagcacc ttaattgaa ttgtcgtgg
28081 atgagctgg ttctaaatca cccattcagt acatgatat cgtaattat acagttcct
28141 gtttaccctt tacaattaat tcccaggaac ctaattggg tagtcttga gtgcgtgtt
28201 cgttctatga agactttta gattataag agtctcgtg tgttttagat ttactctaaa
28261 cgaacaact aaatgtctg ataatggac ccaaatctg cgaatgac ctcgattca
28321 ttatggaga caactcgtt cactctgac taacagatg ggaagatca gtaagctca
28381 atcaaaacaa cgtcgcgcc aaggtttacc caataactt cgtcttgg taaccgctt
28441 cactcaact ggaaggaag accttaatt cctctgaga caagcgttc caatcaac
28501 caatpcagt ccaatgacc aaattgcta ctaccpaga gctaccapac gaattctgg
28561 tggtagcgt aaatgaaag atctcagtc aagatggtat ttctactacc taggaactgg
28621 cccagaact ggaactcct atggtctaa caaagagac atcatatgg ttgcaactga
28681 gggagcttg aatacaccan aagatcaat tgcaccgcc aatcctgta acaatgctg
28741 aatgttca caactcttc aaagatcaa atggcaaaa gctttctag caaagatga
28801 caagagcgc agtcaagct ctctctgtc ctctacagt atccaacca ptccaagaa
28861 atcaatcca ggaagcaga ggaagactt tctctgata atggctgca atggcgtga
28921 tctgtcctt gctttctgc tctgtcag atgaaacag ctgagagca aaatgtctg
28981 aaagcccaa caaaccaag gcaaacctt caatagaaa tctgtctg aggtcttaa
29041 aaagctcgg caaaaacga ctcccactaa gcatataat gtaacacaag cttcggcag
29101 acgtgctca gaacaaacc aaggaattt tggagaccg gaactaaca gacaaggaac
29161 gpatcaaaa catggcggc aaatgcaaa attgccccc agcgtctcg cttctctgg
29221 aatgtctg atggcagtg aatctcaac ttgaaaga gtttcaact caaagctg
29281 atcaaatg atgacaag atcaaatit caaagatca gkatcttg taaatgaca
29341 katgagca tacaacact tcccacac agagctcaa aagacaana gaaagaagc
29401 gpatgaaat caagcttac cgcagagca gaagaaacag caactgtga ctctctcc
29461 tctcagat ttgatgatt tctcaaca atgcaaca tcatgagca gtctgactt
29521 aactcagcc taactctat gacaccac aagcagat gcatataaa agtcttcc
29581 ttttcgctt agatataa gctactctt gtagcaat gattctga actcatagc
29641 caagtagat gtagtacc ttaactcac atagcaact ttaactag tgaacatta
29701 ggaagact gaagagca ccaacttc accagacca cggagatc gatcaggtt
29761 acagtaca atgctagga gactgcta tatgaaag cctatgtg taaattat
29821 tttatgtg ctatcccat gtattttaa tagttctta ggaagatg aaaaaaaaaa
29881 aaaaaaaaaa aaaaaaaaaa aaa
//
  
```

```

28274..29533
/gene="N"
/note="structural protein"
/codon_start=1
/product="nucleocapsid phosphoprotein"
/protein_id="090332"
/translation="MSDLPQNRNAPRITFGGSDSTGSDNGERSGARSQRPPQG
LPNITASNFALTQNGKEDLKFRRQGVPTINTSSPDQIGYRRATRRIRGGDGLVQ
DLSPRMYFYLLGTGFEAGLPGMKDGIINAVTEGALNTPKDMITGRNPNAAIVLQ
LPLQGTLLPKGYAEGSRGQASRRSSRSRNSRNSRTPGSSRGTSPARMAGGDA
LALLLLDRMLQESKMSGKQQQQQTVTKSAEAASKPRQRITKAYNVTFQGR
RGPETQGNFQDQLIRQGTIDYKHPQIAQFAPSASAFSGMSRIGHEVPSGTHLTYT
GATKLDKDRPHKQVILLNHIDYAKTFPPTPKDKKKKADQALPQRKQQQTV
TLLPAADLDFSKLQSSHSADSTQA"
  
```

Po navodilih v nalogi odpremo spletno orodje CAI calculator in nastavimo eksperimentalni sistem. Prilepimo kodirajoče zaporedje v polje Target DNA Sequence in pritisnemo Calculate Codon Adaptation Index Now. Za analizo

kaj pomeni Codon Adaptation Index Value uporabimo Google, ki nam pove, da bližje kot je CAI 1, boljša je prilagojenost ekspresijskemu sistemu.

Home » Bioinformatics Center » Codon Adaptation Index Calculator

Codon Adaptation Index Calculator

Introduction: CAI (Codon Adaptation Index) is an effective measure of synonymous codon usage bias. It may give an approximate indication of the likely success of the heterologous gene expression. This online tool calculates CAI according to the relative synonymous codon usage of a reference sequence or existing expression host organisms.

[\[Recommended\] BiologicsCorp Guarantee Protein Service \(\\$1,550\)
Codon Optimization + Gene Synthesis + 5mg purified protein](#)
[Click to read more in New window](#)

Codon Adaptation Index Value: 0.71
Reference Set: the codon usage table of *Homo sapiens* (Mamalian/HEK)
Codon Usage of Reference Set
Format: fields: [triplet] [amino acid] [fraction] [rscu] [rscu weight] [frequency: per thousand] [(number)]

TTT F	0.46	0.93	0.866	17.6	(714298)	TTA L	0.08	0.46	0.193	7.7	(311881)
TTG L(s)	0.13	0.77	0.326	12.9	(525688)	TTC F	0.54	1.07	1.000	20.3	(824692)
TAT Y	0.44	0.89	0.796	12.2	(495699)	TAA *	0.30	0.89	0.637	1.0	(40285)
TAG *	0.24	0.71	0.508	0.8	(32109)	TAC Y	0.56	1.11	1.000	15.3	(622407)
TGT C	0.46	0.91	0.839	10.6	(430311)	TGA *	0.47	1.40	1.000	1.6	(63237)
TGG W	1.00	1.00	1.000	13.2	(535595)	TGC C	0.54	1.09	1.000	12.6	(513028)
TCT S	0.19	1.13	0.782	15.2	(618711)	TCA S	0.15	0.91	0.627	12.2	(496448)
TCG S	0.05	0.33	0.227	4.4	(179419)	TCC S	0.22	1.31	0.908	17.7	(718892)
ATT I	0.36	1.08	0.768	16.0	(650473)	ATA I	0.17	0.51	0.360	7.5	(304565)
ATG M(s)	1.00	1.00	1.000	22.0	(896905)	ATC I	0.47	1.41	1.000	20.8	(845466)

Orthohantavirus andesense

Codon Optimization + Gene Synthesis + 5mg purified protein
[Click to read more in New window](#)

Codon Adaptation Index Value: 0.71
Reference Set: the codon usage table of *Homo sapiens* (Mamalian/HEK)
Codon Usage of Reference Set
Format: fields: [triplet] [amino acid] [fraction] [rscu] [rscu weight] [frequency: per thousand] [(number)]

TTT F	0.46	0.93	0.866	17.6	(714298)	TTA L	0.08	0.46	0.193	7.7	(311881)
TTG L(s)	0.13	0.77	0.326	12.9	(525688)	TTC F	0.54	1.07	1.000	20.3	(824692)
TAT Y	0.44	0.89	0.796	12.2	(495699)	TAA *	0.30	0.89	0.637	1.0	(40285)
TAG *	0.24	0.71	0.508	0.8	(32109)	TAC Y	0.56	1.11	1.000	15.3	(622407)
TGT C	0.46	0.91	0.839	10.6	(430311)	TGA *	0.47	1.40	1.000	1.6	(63237)
TGG W	1.00	1.00	1.000	13.2	(535595)	TGC C	0.54	1.09	1.000	12.6	(513028)
TCT S	0.19	1.13	0.782	15.2	(618711)	TCA S	0.15	0.91	0.627	12.2	(496448)
TCG S	0.05	0.33	0.227	4.4	(179419)	TCC S	0.22	1.31	0.908	17.7	(718892)
ATT I	0.36	1.08	0.768	16.0	(650473)	ATA I	0.17	0.51	0.360	7.5	(304565)
ATG M(s)	1.00	1.00	1.000	22.0	(896905)	ATC I	0.47	1.41	1.000	20.8	(845466)
AAT N	0.47	0.94	0.808	17.0	(689701)	AAA K	0.43	0.87	0.767	24.4	(993621)
AAG K	0.57	1.13	1.000	31.9	(1295568)	AAC N	0.53	1.06	1.000	19.1	(776603)
AGT S	0.15	0.90	0.624	12.1	(493429)	AGA R	0.21	1.29	1.000	12.2	(494682)
AGG R	0.21	1.27	0.983	12.0	(486463)	AGC S	0.24	1.44	1.000	19.5	(791383)
ACT T	0.25	0.99	0.695	13.1	(533609)	ACA T	0.28	1.14	0.800	15.1	(614523)
ACG T	0.11	0.46	0.320	6.0	(246105)	ACC T	0.35	1.42	1.000	18.9	(768147)
GTT V	0.18	0.73	0.392	11.0	(448607)	GTA V	0.12	0.47	0.252	7.1	(287712)
GTG V	0.46	1.85	1.000	28.1	(1143534)	GTC V	0.24	0.95	0.514	14.5	(588130)
GAT D	0.47	0.93	0.868	21.8	(885429)	GAA E	0.42	0.84	0.731	29.0	(1177632)
GAG E	0.58	1.16	1.000	39.6	(1609975)	GAC D	0.54	1.07	1.000	25.1	(1020595)
GGT G	0.16	0.65	0.484	10.8	(437126)	GGA G	0.25	1.00	0.741	16.5	(669873)
GGG G	0.25	1.00	0.741	16.5	(669768)	GGC G	0.34	1.35	1.000	22.2	(903565)

SARS-CoV-2

Rešitev: vidimo da je rezultat za obe primerjavi enak kar pomeni da oba virusa podobno učinkovito izkoriščata človeške tRNA.