Rešitve:

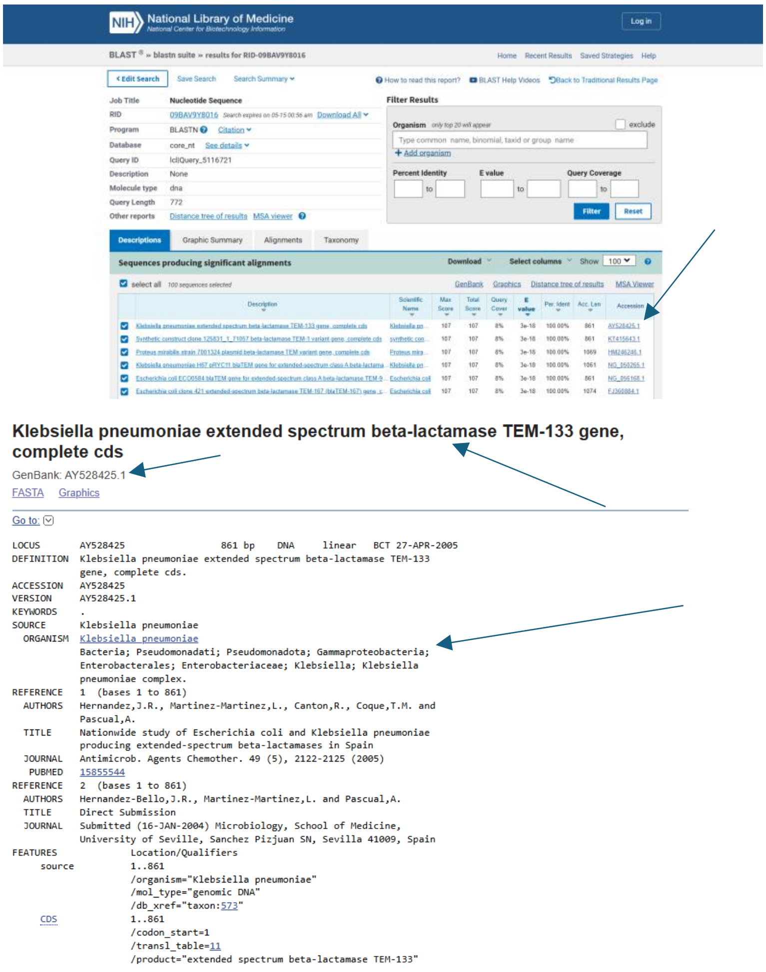
## 1. Naloga

Dani sta nukleotidni zaporedji. Vsako zaporedje posebej kopiram in prilepim v BLASTn, kjer uporabim funkcijo Program selection: optimize for: more dissimilar sequences.

Odgovori za vsako zaporedje:

Zaporedje A: Vidim, da se prvi zadetek zelo dobro ujema, zato uporabim tega. Čisto desno je napisan accession code: AY528425. Kliknem nanj in vidim da gre za bakterijo Klebsiella pneumoniae. K. pneumoniae spada v rod Klebsiella, družina Enterobacteriaceae.

Na isti strani vidim da gre za zapis za protein: beta-lactamase.



| | |
|---|---|
| Job Title | Nucleotide Sequence |
| RID | 09BAV9Y8016 Search expires on 05-15 00:56 am Download All |
| Program | BLASTN Citation |
| Database | core_nt See details |
| Query ID | lcl\|Query_5116721 |
| Description | None |
| Molecule type | dna |
| Query Length | 772 |
| Other reports | Distance tree of results MSA viewer |

Sequences producing significant alignments

☑ select all 100 sequences selected

| | Description | Scientific Name | Max Score | Total Score | Query Cover | E value | Per. Ident | Acc. Len | Accession |
|---|---|---|---|---|---|---|---|---|---|
| ☑ | Klebsiella pneumoniae extended spectrum beta-lactamase TEM-133 gene, complete cds | Klebsiella pn... | 107 | 107 | 8% | 3e-18 | 100.00% | 861 | AY528425.1 |
| ☑ | Synthetic construct clone 125831_1_71057 beta-lactamase TEM-1 variant gene, complete cds | synthetic con... | 107 | 107 | 8% | 3e-18 | 100.00% | 861 | KT415643.1 |
| ☑ | Proteus mirabilis strain 7001324 plasmid beta-lactamase TEM variant gene, complete cds | Proteus mira... | 107 | 107 | 8% | 3e-18 | 100.00% | 1069 | HM246246.1 |
| ☑ | Klebsiella pneumoniae H67 pRYC11 blaTEM gene for extended-spectrum class A beta-lactama... | Klebsiella pn... | 107 | 107 | 8% | 3e-18 | 100.00% | 1061 | NG_050265.1 |
| ☑ | Escherichia coli ECO0584 blaTEM gene for extended-spectrum class A beta-lactamase TEM-9... | Escherichia coli | 107 | 107 | 8% | 3e-18 | 100.00% | 861 | NG_056168.1 |
| ☑ | Escherichia coli clone 421 extended-spectrum beta-lactamase TEM-167 (blaTEM-167) gene, c... | Escherichia coli | 107 | 107 | 8% | 3e-18 | 100.00% | 1074 | FJ360884.1 |

**Klebsiella pneumoniae extended spectrum beta-lactamase TEM-133 gene, complete cds**

GenBank: AY528425.1

```
LOCUS       AY528425                 861 bp    DNA     linear   BCT 27-APR-2005
DEFINITION  Klebsiella pneumoniae extended spectrum beta-lactamase TEM-133
            gene, complete cds.
ACCESSION   AY528425
VERSION     AY528425.1
KEYWORDS    .
SOURCE      Klebsiella pneumoniae
  ORGANISM  Klebsiella pneumoniae
            Bacteria; Pseudomonadati; Pseudomonadota; Gammaproteobacteria;
            Enterobacterales; Enterobacteriaceae; Klebsiella; Klebsiella
            pneumoniae complex.
REFERENCE   1  (bases 1 to 861)
  AUTHORS   Hernandez,J.R., Martinez-Martinez,L., Canton,R., Coque,T.M. and
            Pascual,A.
  TITLE     Nationwide study of Escherichia coli and Klebsiella pneumoniae
            producing extended-spectrum beta-lactamases in Spain
  JOURNAL   Antimicrob. Agents Chemother. 49 (5), 2122-2125 (2005)
   PUBMED   15855544
REFERENCE   2  (bases 1 to 861)
  AUTHORS   Hernandez-Bello,J.R., Martinez-Martinez,L. and Pascual,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-2004) Microbiology, School of Medicine,
            University of Seville, Sanchez Pizjuan SN, Sevilla 41009, Spain
FEATURES             Location/Qualifiers
     source          1..861
                     /organism="Klebsiella pneumoniae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:573"
     CDS             1..861
                     /codon_start=1
                     /transl_table=11
                     /product="extended spectrum beta-lactamase TEM-133"
```

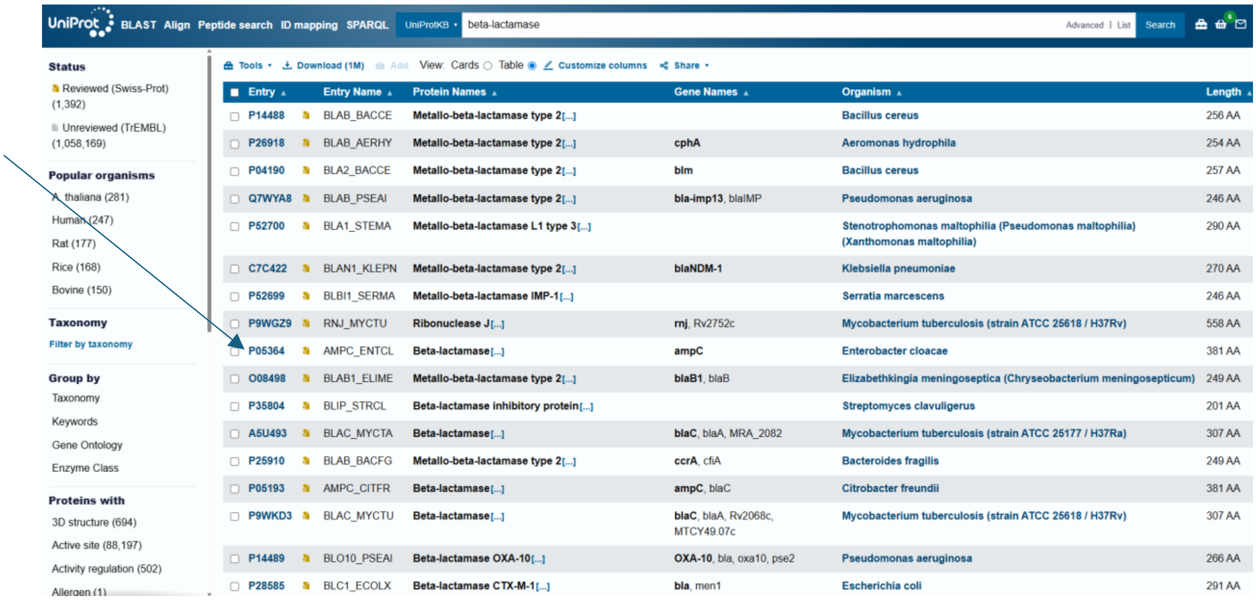
Zaporedje B: Tudi v tem primeru se zaporedje zelo dobro ujema s prvim zadetkom. Ponovim enak postopek in odčitam accession code: AF148850. Zaporedje se nahaja v Escherichia Coli in prav tako kodira protein beta laktamazo. E. coli spada v rod Escherichia, prav tako družina Enterobacteriaceae

Beta laktamaza je encim, ki ga proizvajajo nekatere bakterije kot mehanizem obrambe pred beta-laktamskimi antibiotiki (kot je na primer penicilin)

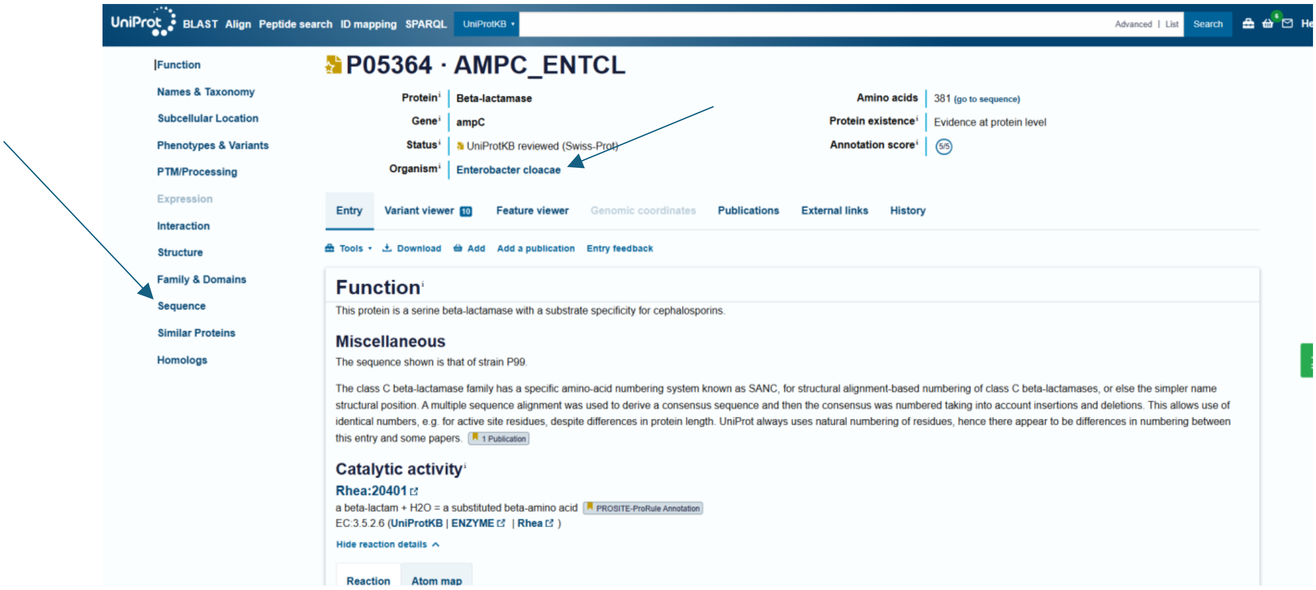
## 1. Naloga

Zdaj moram poiskati 5 homolognih zaporedij. Grem v uniprot in vpišem beta-lactamase. Tako poiščem aminokislinska zaporedja za 5 različnih bakterij in jih zapišem v FASTA obliki.

Pojasnim na primeru:



| ☐ | Entry | Entry Name | Protein Names | Gene Names | Organism | Length |
|---|---|---|---|---|---|---|
| ☐ | P14488 | BLAB_BACCE | Metallo-beta-lactamase type 2[...] | | Bacillus cereus | 256 AA |
| ☐ | P26918 | BLAB_AERHY | Metallo-beta-lactamase type 2[...] | cphA | Aeromonas hydrophila | 254 AA |
| ☐ | P04190 | BLA2_BACCE | Metallo-beta-lactamase type 2[...] | blm | Bacillus cereus | 257 AA |
| ☐ | Q7WYA8 | BLAB_PSEAI | Metallo-beta-lactamase type 2[...] | bla-imp13, blaIMP | Pseudomonas aeruginosa | 246 AA |
| ☐ | P52700 | BLA1_STEMA | Metallo-beta-lactamase L1 type 3[...] | | Stenotrophomonas maltophilia (Pseudomonas maltophilia) (Xanthomonas maltophilia) | 290 AA |
| ☐ | C7C422 | BLAN1_KLEPN | Metallo-beta-lactamase type 2[...] | blaNDM-1 | Klebsiella pneumoniae | 270 AA |
| ☐ | P52699 | BLBI1_SERMA | Metallo-beta-lactamase IMP-1[...] | | Serratia marcescens | 246 AA |
| ☐ | P9WGZ9 | RNJ_MYCTU | Ribonuclease J[...] | rnj, Rv2752c | Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) | 558 AA |
| ☐ | P05364 | AMPC_ENTCL | Beta-lactamase[...] | ampC | Enterobacter cloacae | 381 AA |
| ☐ | O08498 | BLAB1_ELIME | Metallo-beta-lactamase type 2[...] | blaB1, blaB | Elizabethkingia meningoseptica (Chryseobacterium meningosepticum) | 249 AA |
| ☐ | P35804 | BLIP_STRCL | Beta-lactamase inhibitory protein[...] | | Streptomyces clavuligerus | 201 AA |
| ☐ | A5U493 | BLAC_MYCTA | Beta-lactamase[...] | blaC, blaA, MRA_2082 | Mycobacterium tuberculosis (strain ATCC 25177 / H37Ra) | 307 AA |
| ☐ | P25910 | BLAB_BACFG | Metallo-beta-lactamase type 2[...] | ccrA, cfiA | Bacteroides fragilis | 249 AA |
| ☐ | P05193 | AMPC_CITFR | Beta-lactamase[...] | ampC, blaC | Citrobacter freundii | 381 AA |
| ☐ | P9WKD3 | BLAC_MYCTU | Beta-lactamase[...] | blaC, blaA, Rv2068c, MTCY49.07c | Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) | 307 AA |
| ☐ | P14489 | BLO10_PSEAI | Beta-lactamase OXA-10[...] | OXA-10, bla, oxa10, pse2 | Pseudomonas aeruginosa | 266 AA |
| ☐ | P28585 | BLC1_ECOLX | Beta-lactamase CTX-M-1[...] | bla, men1 | Escherichia coli | 291 AA |



P05364 · AMPC_ENTCL

Protein: Beta-lactamase
Gene: ampC
Status: UniProtKB reviewed (Swiss-Prot)
Organism: Enterobacter cloacae
Amino acids: 381 (go to sequence)
Protein existence: Evidence at protein level
Annotation score: 5/5

Function

This protein is a serine beta-lactamase with a substrate specificity for cephalosporins.

Miscellaneous

The sequence shown is that of strain P99.

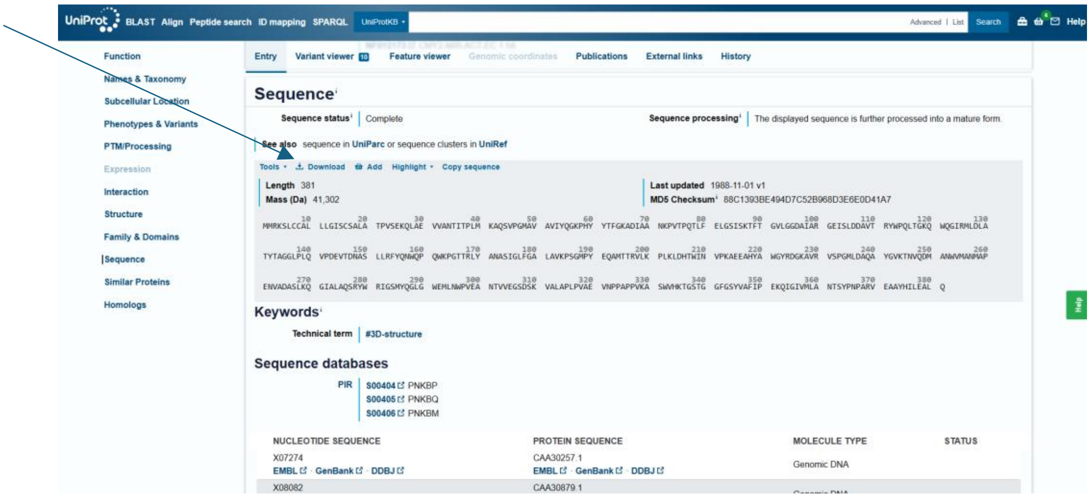
The class C beta-lactamase family has a specific amino-acid numbering system known as SANC, for structural alignment-based numbering of class C beta-lactamases, or else the simpler name structural position. A multiple sequence alignment was used to derive a consensus sequence and then the consensus was numbered taking into account insertions and deletions. This allows use of identical numbers, e.g. for active site residues, despite differences in protein length. UniProt always uses natural numbering of residues, hence there appear to be differences in numbering between this entry and some papers. 1 Publication

Catalytic activity

Rhea:20401

a beta-lactam + $H_2O$ = a substituted beta-amino acid PROSITE-ProRule Annotation

EC:3.5.2.6 (UniProtKB | ENZYME | Rhea )



## Sequence

**Sequence status** Complete

**Sequence processing** The displayed sequence is further processed into a mature form.

**See also** sequence in UniParc or sequence clusters in UniRef

**Length** 381
**Mass (Da)** 41,302

**Last updated** 1988-11-01 v1
**MD5 Checksum** 88C1393BE494D7C52B968D3E6E0D41A7

```
        10         20         30         40         50         60         70         80         90        100        110        120        130
MMRKSLCCAL LLGISCSALA TPVSEKQLAE VVANTITPLM KAQSVPGMAV AVIYQGKPHY YTFGKADIAA NKPVTPQTLF ELGSISKTFT GVLGGDAIAR GEISLDDAVT RYWPQLTGKQ WQGIRMLDLA

       140        150        160        170        180        190        200        210        220        230        240        250        260
TYTAGGLPLQ VPDEVTDNAS LLRFYQNWQP QWKPGTTRLY ANASIGLFGA LAVKPSGMPY EQAMTTRVLK PLKLDHTWIN VPKAEEAHYA WGYRDGKAVR VSPGMLDAQA YGVKTNVQDM ANWVMANMAP

       270        280        290        300        310        320        330        340        350        360        370        380
ENVADASLKQ GIALAQSRYW RIGSMYQGLG WEMLNWPVEA NTVVEGSDSK VALAPLPVAE VNPPAPPVKA SWVHKTGSTG GFGSYVAFIP EKQIGIVMLA NTSYPNPARV EAAYHILEAL Q
```

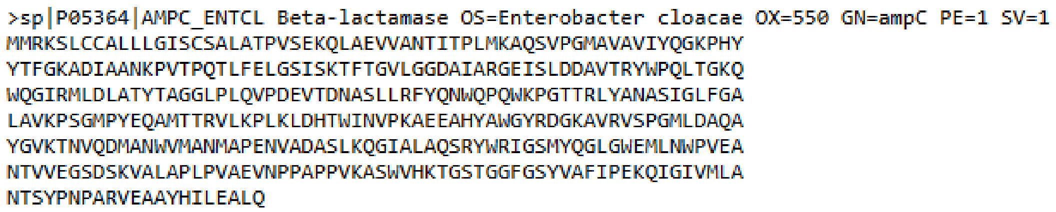
## Keywords

**Technical term** #3D-structure

## Sequence databases

**PIR** S00404 PNKBP
S00405 PNKBQ
S00406 PNKBM

| NUCLEOTIDE SEQUENCE | PROTEIN SEQUENCE | MOLECULE TYPE | STATUS |
|---|---|---|---|
| X07274<br>EMBL · GenBank · DDBJ | CAA30257.1<br>EMBL · GenBank · DDBJ | Genomic DNA | |
| X08082 | CAA30879.1 | | |



```
>sp|P05364|AMPC_ENTCL Beta-lactamase OS=Enterobacter cloacae OX=550 GN=ampC PE=1 SV=1
MMRKSLCCALLLGISCSALATPVSEKQLAEVVANTITPLMKAQSVPGMAVAVIYQGKPHY
YTFGKADIAANKPVTPQTLFELGSISKTFTGVLGGDAIARGEISLDDAVTRYWPQLTGKQ
WQGIRMLDLATYTAGGLPLQVPDEVTDNASLLRFYQNWQPQWKPGTTRLYANASIGLFGA
LAVKPSGMPYEQAMTTRVLKPLKLDHTWINVPKAEEAHYAWGYRDGKAVRVSPGMLDAQA
YGVKTNVQDMANWVMANMAPENVADASLKQGIALAQSRYWRIGSMYQGLGWEMLNWPVEA
NTVVEGSDSKVALAPLPVAEVNPPAPPVKASWVHKTGSTGGFGSYVAFIPEKQIGIVMLA
NTSYPNPARVEAAYHILEALQ
```

> Enterobacter_cloacae

MMRKSLCCALLLGISCSALATPVSEKQLAEVVANTITPLMKAQSVPGMAVAVIYQGKPHYYTFGKADIAAN
KPVTPQTLFELGSISKTFTGVLGGDAIARGEISLDDAVTRYWPQLTGKQWQGIRMLDLATYTAGGLPLQVP
DEVTDNASLLRFYQNWQPQWKPGTTRLYANASIGLFGALAVKPSGMPYEQAMTTRVLKPLKLDHTWINV
PKAEEAHYAWGYRDGKAVRVSPGMLDAQAYGVKTNVQDMANWVMANMAPENVADASLKQGIALAQS
RYWRIGSMYQGLGWEMLNWPVEANTVVEGSDSKVALAPLPVAEVNPPAPPVKASWVHKTGSTGGFGS
YVAFIPEKQIGIVMLANTSYPNPARVEAAYHILEALQ

Enako ponovim še za 4 druge bakterije in dobim:

>Streptomyces_clavuligerus

MRTVGIGAGVRRLGRAVVMAAAVGGLVLGSAGASNAAGVMTGAKFTQIQFGMTRQQVLDIAGAENCET
GGSFGDSIHCRGHAAGDYYAYATFGFTSAAADAKVDSKSQEKLLAPSAPTLTLAKFNQVTVGMTRAQVLA
TVGQGSCTTWSEYYPAYPSTAGVTLSLSCFDVDGYSSTGFYRGSAHLWFTDGVLQGKRQWDLV

> Citrobacter_amalonaticus

MMRHRVKRMMLMTTACISLLLGSAPLYAQANDVQQKLAALEKSSGGRLGVALIDTADNAQTLYRADERFA
MCSTSKVMAAAAVLKQSETQKKVLSQKVEIKSSDLINYNPITEKHVNGTMTLAELSAAALQYSDNTAMNKL
IAHLGGPDKVTAFARAIGDNTFRLDRTEPTLNTAIPGDPRDTTTPLAMAQTLRNLTLGSALGETQRAQLVTW

LKGNTTGAASIQAGLPTSWVVGDKTGSGDYGTTNDIAVIWPEGRAPLILVTYFTQPEQKAESRRDVLAAAA
KIVTDGY

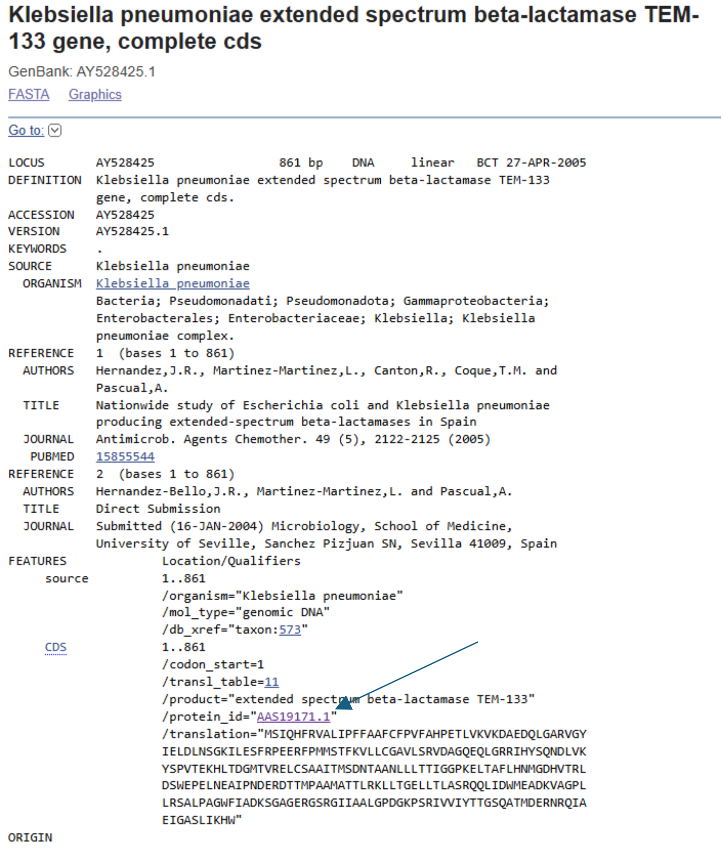
> Mycobacterium_tuberculosis

MRNRGFGRRELLVAMAMLVSVTGCARHASGARPASTTLPAGADLADRFAELERRYDARLGVYVPATGTTA
AIEYRADERFAFCSTFKAPLVAAVLHQNPLTHLDKLITYTSDDIRSISPVAQQHVQTGMTIGQLCDAAIRYSD
GTAANLLLADLGGPGGGTAAFTGYLRSLGDTVSRLDAEEPELNRDPPGDERDTTTPHAIALVLQQLVLGN
ALPPDKRALLTDWMARNTTGAKRIRAGFPADWKVIDKTGTGDYGRANDIAVVWSPTGVPYVVAVMSDRA
GGGYDAEPREALLAEAATCVAGVLA

> Nocardia_farcinica

MPGVDISFLKKSGRRTMAAAAVIALLGGCGADAGSEPATTAASTTAPSAATDAATAEFAALEQRFGARLGV
YAVDTTSGAVVAYRADERFGMASTFKGLACGALLREHPLSSGYFDQVVRYSREEVVSYSPVTETRVDTGM
TVAELCHATITVSDNTAGNQILKLLGGPAGFTAFLRSLGDEVSRLDRWETELNEVPPGEERDTTTPAAVAA
NYRALVLGDVLAEPERAQLRDWLVANTTGDQRIRAGVPAGWTVGDKTGGGSHGGNNDVAVAWTETGD
PIVIALLSHRTDPAAKADNALLAEATRAVVTALR

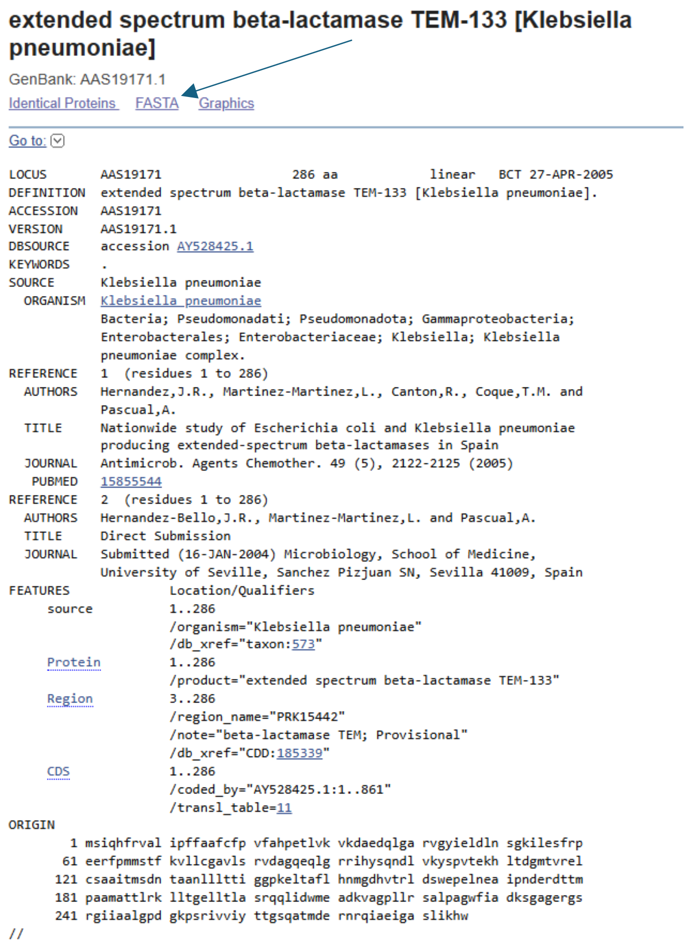
Da pridem pa do aminokislinskega zaporedja Klebsiella pneumoniae in Escherichia Coli izhajam iz datoteke iz prve naloge.



Klebsiella pneumoniae extended spectrum beta-lactamase TEM-133 gene, complete cds

GenBank: AY528425.1

```
LOCUS       AY528425                 861 bp    DNA     linear   BCT 27-APR-2005
DEFINITION  Klebsiella pneumoniae extended spectrum beta-lactamase TEM-133
            gene, complete cds.
ACCESSION   AY528425
VERSION     AY528425.1
KEYWORDS    .
SOURCE      Klebsiella pneumoniae
  ORGANISM  Klebsiella pneumoniae
            Bacteria; Pseudomonadati; Pseudomonadota; Gammaproteobacteria;
            Enterobacterales; Enterobacteriaceae; Klebsiella; Klebsiella
            pneumoniae complex.
REFERENCE   1  (bases 1 to 861)
  AUTHORS   Hernandez,J.R., Martinez-Martinez,L., Canton,R., Coque,T.M. and
            Pascual,A.
  TITLE     Nationwide study of Escherichia coli and Klebsiella pneumoniae
            producing extended-spectrum beta-lactamases in Spain
  JOURNAL   Antimicrob. Agents Chemother. 49 (5), 2122-2125 (2005)
   PUBMED   15855544
REFERENCE   2  (bases 1 to 861)
  AUTHORS   Hernandez-Bello,J.R., Martinez-Martinez,L. and Pascual,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-2004) Microbiology, School of Medicine,
            University of Seville, Sanchez Pizjuan SN, Sevilla 41009, Spain
FEATURES             Location/Qualifiers
     source          1..861
                     /organism="Klebsiella pneumoniae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:573"
     CDS             1..861
                     /codon_start=1
                     /transl_table=11
                     /product="extended spectrum beta-lactamase TEM-133"
                     /protein_id="AAS19171.1"
                     /translation="MSIQHFRVALIPFFAAFCFPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRVDAGQEQLGRRIHYSQNDLVK
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DSWEPELNEAIPNDERDTTMPAAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
ORIGIN
```

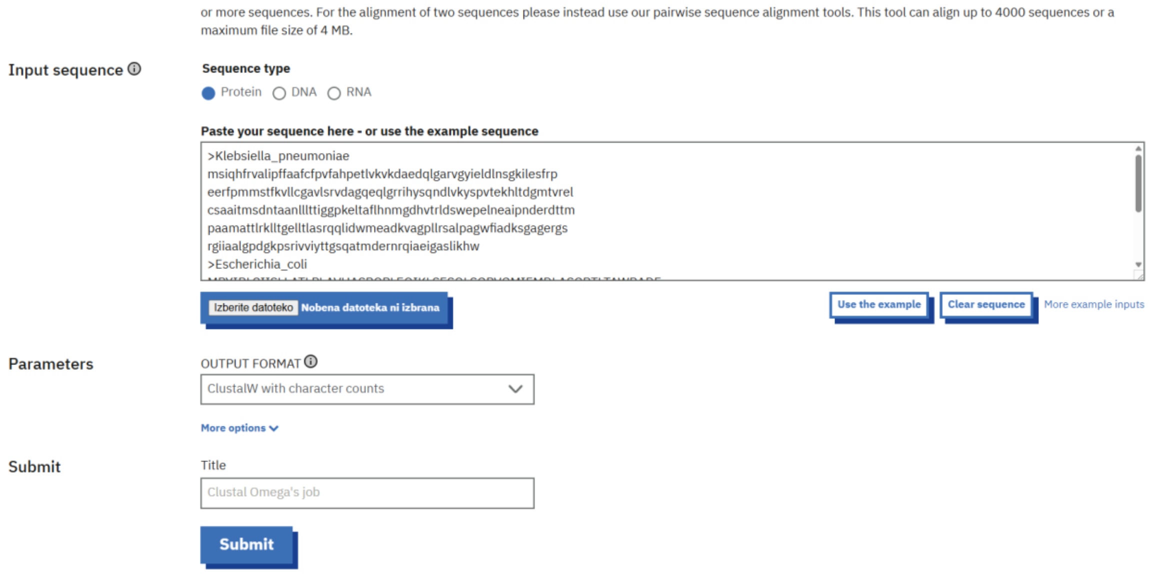


**extended spectrum beta-lactamase TEM-133 [Klebsiella pneumoniae]**

GenBank: AAS19171.1

```
LOCUS       AAS19171                 286 aa            linear   BCT 27-APR-2005
DEFINITION  extended spectrum beta-lactamase TEM-133 [Klebsiella pneumoniae].
ACCESSION   AAS19171
VERSION     AAS19171.1
DBSOURCE    accession AY528425.1
KEYWORDS    .
SOURCE      Klebsiella pneumoniae
  ORGANISM  Klebsiella pneumoniae
            Bacteria; Pseudomonadati; Pseudomonadota; Gammaproteobacteria;
            Enterobacterales; Enterobacteriaceae; Klebsiella; Klebsiella
            pneumoniae complex.
REFERENCE   1  (residues 1 to 286)
  AUTHORS   Hernandez,J.R., Martinez-Martinez,L., Canton,R., Coque,T.M. and
            Pascual,A.
  TITLE     Nationwide study of Escherichia coli and Klebsiella pneumoniae
            producing extended-spectrum beta-lactamases in Spain
  JOURNAL   Antimicrob. Agents Chemother. 49 (5), 2122-2125 (2005)
   PUBMED   15855544
REFERENCE   2  (residues 1 to 286)
  AUTHORS   Hernandez-Bello,J.R., Martinez-Martinez,L. and Pascual,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-2004) Microbiology, School of Medicine,
            University of Seville, Sanchez Pizjuan SN, Sevilla 41009, Spain
FEATURES             Location/Qualifiers
     source          1..286
                     /organism="Klebsiella pneumoniae"
                     /db_xref="taxon:573"
     Protein         1..286
                     /product="extended spectrum beta-lactamase TEM-133"
     Region          3..286
                     /region_name="PRK15442"
                     /note="beta-lactamase TEM; Provisional"
                     /db_xref="CDD:185339"
     CDS             1..286
                     /coded_by="AY528425.1:1..861"
                     /transl_table=11
ORIGIN
        1 msiqhfrval ipffaafcfp vfahpetlvk vkdaedqlga rvgyieldln sgkilesfrp
       61 eerfpmmstf kvllcgavls rvdagqeqlg rrihysqndl vkyspvtekh ltdgmtvrel
      121 csaaitmsdn taanllltti ggpkeltafl hnmgdhvtrl dswepelnea ipnderdttm
      181 paamattlrk lltgelltla srqqlidwme adkvagpllr salpagwfia dksgagergs
      241 rgiiaalgpd gkpsrivviy ttgsqatmde rnrqiaeiga slikhw
//
```

Za poravnave aminokislinskih zaporedij bom uporabila Clustal omega. Najprej primerjam aminokislinski zaporedji iskanih proteinov.



or more sequences. For the alignment of two sequences please instead use our pairwise sequence alignment tools. This tool can align up to 4000 sequences or a maximum file size of 4 MB.

**Input sequence**

**Sequence type**

● Protein ○ DNA ○ RNA

**Paste your sequence here - or use the example sequence**

```
>Klebsiella_pneumoniae
msiqhfrvalipffaafcfpvfahpetlvkvkdaedqlgarvgyieldlnsgkilesfrp
eerfpmmstfkvllcgavlsrvdagqeqlgrrihysqndlvkyspvtekhltdgmtvrel
csaaitmsdntaanllttiggpkeltaflhnmgdhvtrldswepelneaipnderdttm
paamattlrklltgelltlasrqqlidwmeadkvagpllrsalpagwfiadksgagergs
rgiiaalgpdgkpsrivviyttgsqatmdernrqiaeigaslikhw
>Escherichia_coli
```

**Parameters**

OUTPUT FORMAT: ClustalW with character counts

**Submit**

Title: Clustal Omega's job

Zaporedji prilepim v obliki fasta formata in zaženem poravnavo.

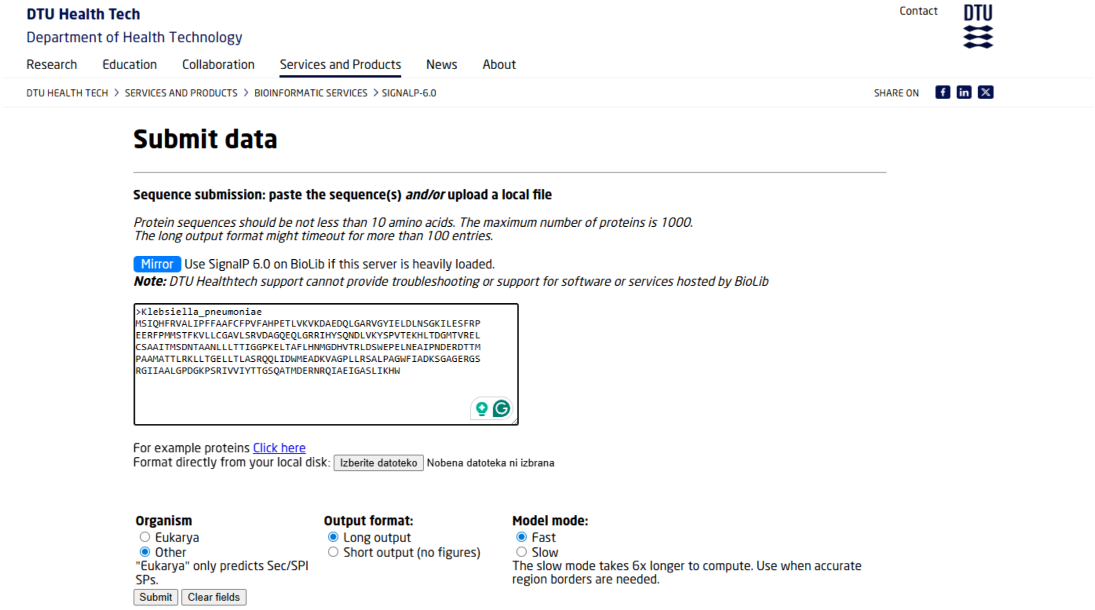
```
CLUSTAL O(1.2.4) multiple sequence alignment


Klebsiella_pneumoniae      msiqhfrvalipffaafcfpvfahpetlvkvkdaedqlgarvgyieldlnsgkilesfrp 60
Escherichia_coli           --MRYIRLCIISLLATLPLAVHASPQPLEQIKLSESQLSGRVGMIEMDLASGRTLTAWRA 58
                             ::::*:.:* ::*:: : *.* *: * ::* :*.**..*** **:** **: * ::*

Klebsiella_pneumoniae      eerfpmmstfkvllcgavlsrvdagqeqlgrrihysqndlvkyspvtekhltdgmtvrel 120
Escherichia_coli           DERFPMMSTFKVVLCGAVLARVDAGDEQLERKIHYRQQDLVDYSPVSEKHLADGMTVGEL 118
                           :**********:******:*****:*** *:*** *:***.****:****:***** **

Klebsiella_pneumoniae      csaaitmsdntaanllltiggpkeltaflhnmgdhvtrldswepelneaipnderdttm 180
Escherichia_coli           CAAAITMSDNSAANLLLATVGGPAGLTAFLRQIGDNVTRLDRWETELNEALPGDARDTTT 178
                           *:********:******:*:***  *****:::**:***** ** *****:*.* ****

Klebsiella_pneumoniae      paamattlrklltgelltlasrqqlidwmeadkvagpllrsalpagwfiadksgagergs 240
Escherichia_coli           PASMAATLRKLLTSQRLSARSQRQLLQWMVDDRVAGPLIRSVLPAGWFIADKTGAGERGA 238
                           **:**:*******.: *:  *::**::**  *:*****:**.*********:******:

Klebsiella_pneumoniae      rgiiaalgpdgkpsrivviyttgsqatmdernrqiaeigaslikhw--     286
Escherichia_coli           RGIVALLGPNNKAERIVVIYLRDTPASMAERNQQIAGIGAALIEHWQR     286
                           ***:* ***:.* .******  .: *:* ***:*** ***:**:**
```

Odgovori na vprašanja:

1. Ker sta zaporedji enako dolgi in sorodni sem uporabila globalno poravnavo.
2. Da
3. Zaporedji sta si bolj podobne na C koncu kot na N. Verjetno zato, ker je lahko na N koncu signalni peptid, ki pa je pri vsaki bakteriji malo drugačen.

## 3. Naloga

Uporabim spletno stran v navodilih in prilepim fasta zaporedje vzorca A.



**Submit data**

**Sequence submission: paste the sequence(s) *and/or* upload a local file**

*Protein sequences should be not less than 10 amino acids. The maximum number of proteins is 1000.*
*The long output format might timeout for more than 100 entries.*

Mirror Use SignalP 6.0 on BioLib if this server is heavily loaded.
***Note:*** *DTU Healthtech support cannot provide troubleshooting or support for software or services hosted by BioLib*

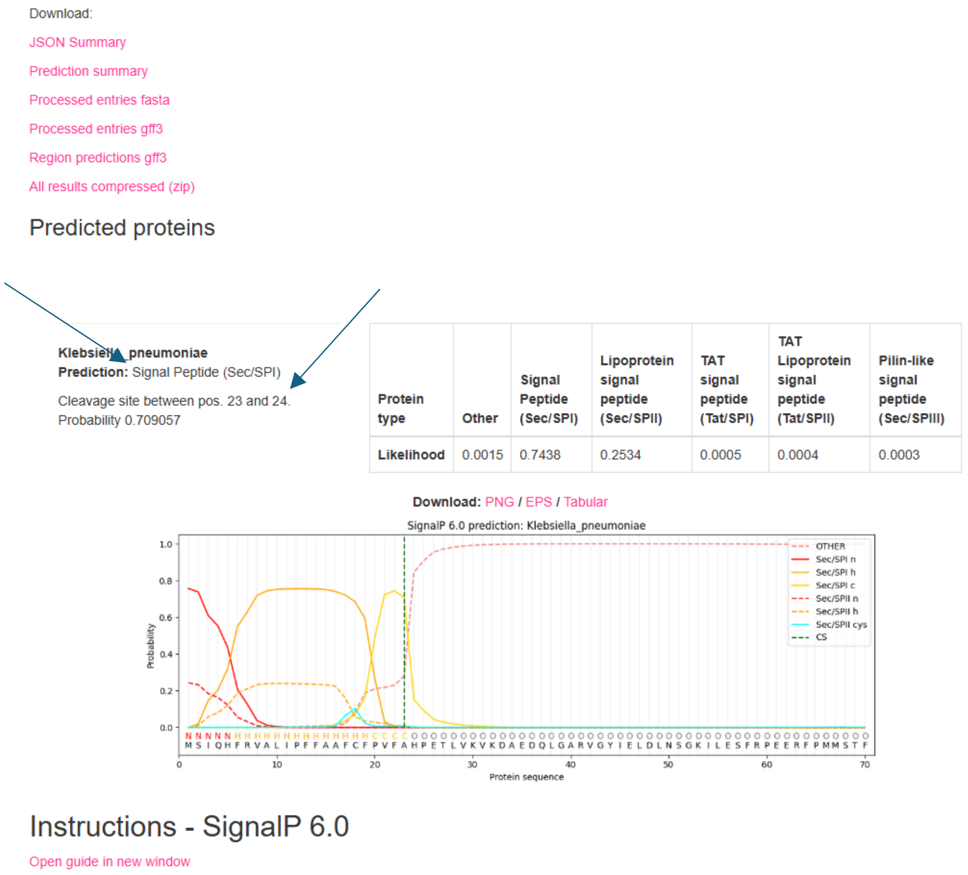
```
>Klebsiella_pneumoniae
MSIQHFRVALIPFFAAFCFPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRP
EERFPMMSTFKVLLCGAVLSRVDAGQEQLGRRIHYSQNDLVKYSPVTEKHLTDGMTVREL
CSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDSWEPELNEAIPNDERDTTM
PAAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAGERGS
RGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
```

For example proteins Click here
Format directly from your local disk: Izberite datoteko Nobena datoteka ni izbrana

**Organism**
○ Eukarya
● Other
"Eukarya" only predicts Sec/SPI SPs.

**Output format:**
● Long output
○ Short output (no figures)

**Model mode:**
● Fast
○ Slow
The slow mode takes 6x longer to compute. Use when accurate region borders are needed.


Download:

JSON Summary

Prediction summary

Processed entries fasta

Processed entries gff3

Region predictions gff3

All results compressed (zip)

Predicted proteins

**Klebsiella_pneumoniae**
**Prediction:** Signal Peptide (Sec/SPI)

Cleavage site between pos. 23 and 24.
Probability 0.709057

| Protein type | Other | Signal Peptide (Sec/SPI) | Lipoprotein signal peptide (Sec/SPII) | TAT signal peptide (Tat/SPI) | TAT Lipoprotein signal peptide (Tat/SPII) | Pilin-like signal peptide (Sec/SPIII) |
|---|---|---|---|---|---|---|
| **Likelihood** | 0.0015 | 0.7438 | 0.2534 | 0.0005 | 0.0004 | 0.0003 |

**Download:** PNG / EPS / Tabular

SignalP 6.0 prediction: Klebsiella_pneumoniae

Instructions - SignalP 6.0

Open guide in new window

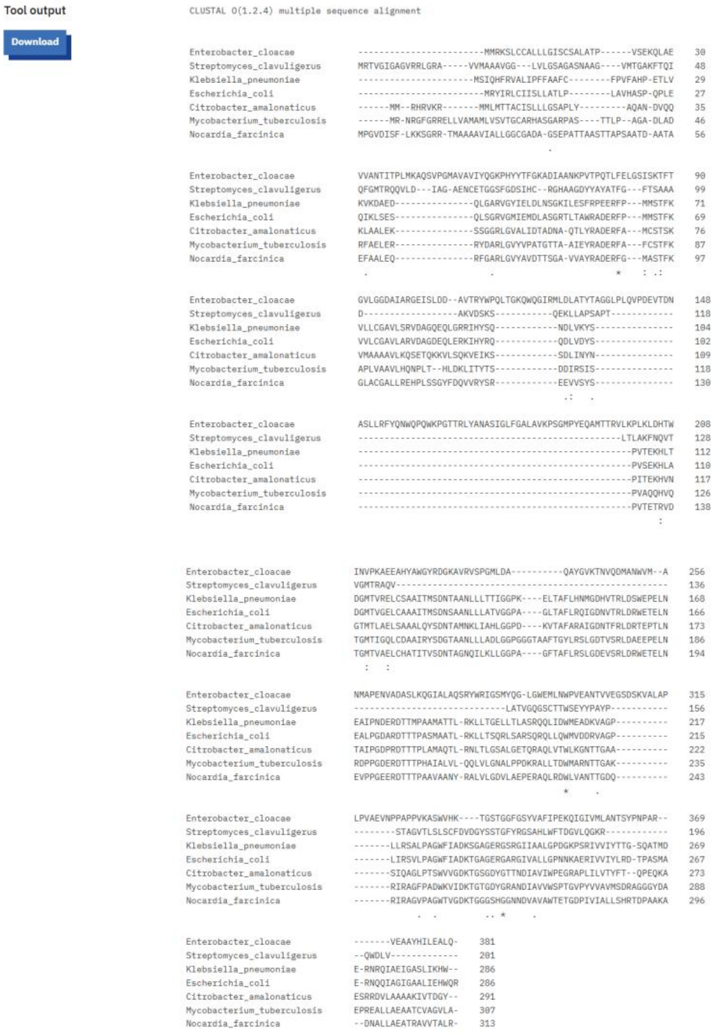
Enak postopek uporabim za vzorec A.

Odgovori:

1. Da.
2. Pri Klebsiella pneumoniae se signalni peptid odcepi med 23. in 24. aminokislinskim ostankom. Pri E. coli pa med 21. in 22.
3. E. coli ima signalni peptid dolg 21 aminokislinskih ostankov in je krajši od signalnega peptida Klebsiella pneumoniae, ki je dolg 23 aminokislinskih ostankov. Oba vzorca imata enak tip signalnega peptida (Sec/SPI), kar nakazuje to, da imata oba proteina podobno pot v periplazmatski prostor..

## 4. Naloga

Zdaj ponovno uporabim Clustal omega, tokrat dodam še 5 homologov.



```
CLUSTAL O(1.2.4) multiple sequence alignment


Enterobacter_cloacae            ------------------------MMRKSLCCALLLGISCSALATP------VSEKQLAE    30
Streptomyces_clavuligerus       MRTVGIGAGVRRLGRA-----VVMAAAVGG---LVLGSAGASNAAG----VMTGAKFTQI    48
Klebsiella_pneumoniae           ----------------------MSIQHFRVALIPFFAAFC--------FPVFAHP-ETLV    29
Escherichia_coli                ------------------------MRYIRLCIISLLATLP--------LAVHASP-QPLE    27
Citrobacter_amalonaticus        ------MM--RHRVKR-------MMLMTTACISLLLGSAPLY---------AQAN-DVQQ    35
Mycobacterium_tuberculosis      ------MR-NRGFGRRELLVAMAMLVSVTGCARHASGARPAS----TTLP--AGA-DLAD    46
Nocardia_farcinica              MPGVDISF-LKKSGRR-TMAAAAVIALLGGCGADA-GSEPATTAASTTAPSAATD-AATA    56
                                                                    .

Enterobacter_cloacae            VVANTITPLMKAQSVPGMAVAVIYQGKPHYYTFGKADIAANKPVTPQTLFELGSISKTFT    90
Streptomyces_clavuligerus       QFGMTRQQVLD---IAG-AENCETGGSFGDSIHC--RGHAAGDYYAYATFG---FTSAAA    99
Klebsiella_pneumoniae           KVKDAED---------------QLGARVGYIELDLNSGKILESFRPEERFP---MMSTFK    71
Escherichia_coli                QIKLSES---------------QLSGRVGMIEMDLASGRTLTAWRADERFP---MMSTFK    69
Citrobacter_amalonaticus        KLAALEK---------------SSGGRLGVALIDTADNA-QTLYRADERFA---MCSTSK    76
Mycobacterium_tuberculosis      RFAELER---------------RYDARLGVYVPATGTTA-AIEYRADERFA---FCSTFK    87
Nocardia_farcinica              EFAALEQ---------------RFGARLGVYAVDTTSGA-VVAYRADERFG---MASTFK    97
                                 .                       .                       *    : .:

Enterobacter_cloacae            GVLGGDAIARGEISLDD--AVTRYWPQLTGKQWQGIRMLDLATYTAGGLPLQVPDEVTDN    148
Streptomyces_clavuligerus       D------------------AKVDSKS-----------QEKLLAPSAPT------------    118
Klebsiella_pneumoniae           VLLCGAVLSRVDAGQEQLGRRIHYSQ------------NDLVKYS---------------    104
Escherichia_coli                VVLCGAVLARVDAGDEQLERKIHYRQ------------QDLVDYS---------------    102
Citrobacter_amalonaticus        VMAAAAVLKQSETQKKVLSQKVEIKS------------SDLINYN---------------    109
Mycobacterium_tuberculosis      APLVAAVLHQNPLT--HLDKLITYTS------------DDIRSIS---------------    118
Nocardia_farcinica              GLACGALLREHPLSSGYFDQVVRYSR------------EEVVSYS---------------    130
                                                                       .:   .

Enterobacter_cloacae            ASLLRFYQNWQPQWKPGTTRLYANASIGLFGALAVKPSGMPYEQAMTTRVLKPLKLDHTW    208
Streptomyces_clavuligerus       --------------------------------------------------LTLAKFNQVT    128
Klebsiella_pneumoniae           ----------------------------------------------------PVTEKHLT    112
Escherichia_coli                ----------------------------------------------------PVSEKHLA    110
Citrobacter_amalonaticus        ----------------------------------------------------PITEKHVN    117
Mycobacterium_tuberculosis      ----------------------------------------------------PVAQQHVQ    126
Nocardia_farcinica              ----------------------------------------------------PVTETRVD    138
                                                                                         :


Enterobacter_cloacae            INVPKAEEAHYAWGYRDGKAVRVSPGMLDA----------QAYGVKTNVQDMANWVM--A    256
Streptomyces_clavuligerus       VGMTRAQV----------------------------------------------------    136
Klebsiella_pneumoniae           DGMTVRELCSAAITMSDNTAANLLLTTIGGPK----ELTAFLHNMGDHVTRLDSWEPELN    168
Escherichia_coli                DGMTVGELCAAAITMSDNSAANLLLATVGGPA----GLTAFLRQIGDNVTRLDRWETELN    166
Citrobacter_amalonaticus        GTMTLAELSAAALQYSDNTAMNKLIAHLGGPD----KVTAFARAIGDNTFRLDRTEPTLN    173
Mycobacterium_tuberculosis      TGMTIGQLCDAAIRYSDGTAANLLLADLGGPGGGTAAFTGYLRSLGDTVSRLDAEEPELN    186
Nocardia_farcinica              TGMTVAELCHATITVSDNTAGNQILKLLGGPA----GFTAFLRSLGDEVSRLDRWETELN    194
                                  :   :

Enterobacter_cloacae            NMAPENVADASLKQGIALAQSRYWRIGSMYQG-LGWEMLNWPVEANTVVEGSDSKVALAP    315
Streptomyces_clavuligerus       -----------------------------LATVGQGSCTTWSEYYPAYP-----------    156
Klebsiella_pneumoniae           EAIPNDERDTTMPAAMATTL-RKLLTGELLTLASRQQLIDWMEADKVAGP----------    217
Escherichia_coli                EALPGDARDTTTPASMAATL-RKLLTSQRLSARSQRQLLQWMVDDRVAGP----------    215
Citrobacter_amalonaticus        TAIPGDPRDTTTPLAMAQTL-RNLTLGSALGETQRAQLVTWLKGNTTGAA----------    222
Mycobacterium_tuberculosis      RDPPGDERDTTTPHAIALVL-QQLVLGNALPPDKRALLTDWMARNTTGAK----------    235
Nocardia_farcinica              EVPPGEERDTTTPAAVAANY-RALVLGDVLAEPERAQLRDWLVANTTGDQ----------    243
                                                                        *     .

Enterobacter_cloacae            LPVAEVNPPAPPVKASWVHK----TGSTGGFGSYVAFIPEKQIGIVMLANTSYPNPAR--    369
Streptomyces_clavuligerus       --------STAGVTLSLSCFDVDGYSSTGFYRGSAHLWFTDGVLQGKR------------    196
Klebsiella_pneumoniae           -------LLRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTG-SQATMD    269
Escherichia_coli                -------LIRSVLPAGWFIADKTGAGERGARGIVALLGPNNKAERIVVIYLRD-TPASMA    267
Citrobacter_amalonaticus        -------SIQAGLPTSWVVGDKTGSGDYGTTNDIAVIWPEGRAPLILVTYFT--QPEQKA    273
Mycobacterium_tuberculosis      -------RIRAGFPADWKVIDKTGTGDYGRANDIAVVWSPTGVPYVVAVMSDRAGGGYDA    288
Nocardia_farcinica              -------RIRAGVPAGWTVGDKTGGGSHGGNNDVAVAWTETGDPIVIALLSHRTDPAAKA    296
                                            .  .         .. *     .

Enterobacter_cloacae            -------VEAAYHILEALQ-    381
Streptomyces_clavuligerus       --QWDLV-------------    201
Klebsiella_pneumoniae           E-RNRQIAEIGASLIKHW--    286
Escherichia_coli                E-RNQQIAGIGAALIEHWQR    286
Citrobacter_amalonaticus        ESRRDVLAAAAKIVTDGY--    291
Mycobacterium_tuberculosis      EPREALLAEAATCVAGVLA-    307
Nocardia_farcinica              --DNALLAEATRAVVTALR-    313
```

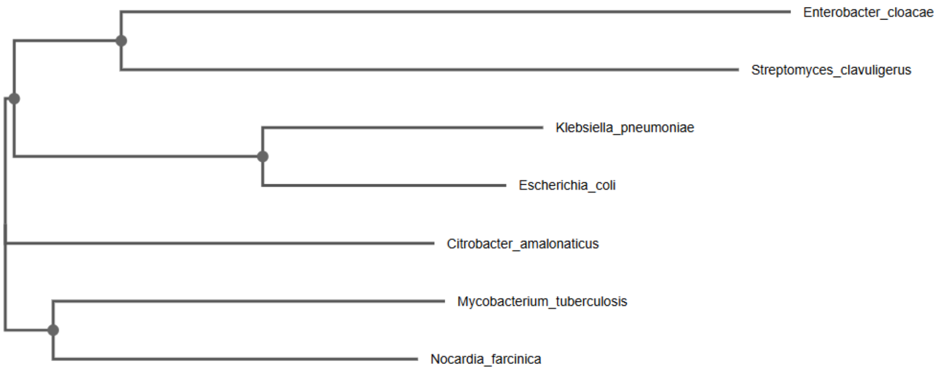
Odgovori na vprašanja:

1. Fenilalanin, triptofan in glicin.
2. Najbolj podobni sta si ravno zaporedji A in B. Imata zelo dobro ujemanje aminokislin in tudi vrzeli.

## 5. Naloga

Filogenetsko drevo dobim tako, da kopiram zapis v clustal omega in ga odprem v Phylo.io

```
(
(
(
Enterobacter_cloacae:0.44983,
Streptomyces_clavuligerus:0.41488)
:0.07194,
(
Klebsiella_pneumoniae:0.18847,
Escherichia_coli:0.16364)
:0.16697)
:0.00601,
Citrobacter_amalonaticus:0.28830,
(
Mycobacterium_tuberculosis:0.26333,
Nocardia_farcinica:0.24503)
:0.03217);
```

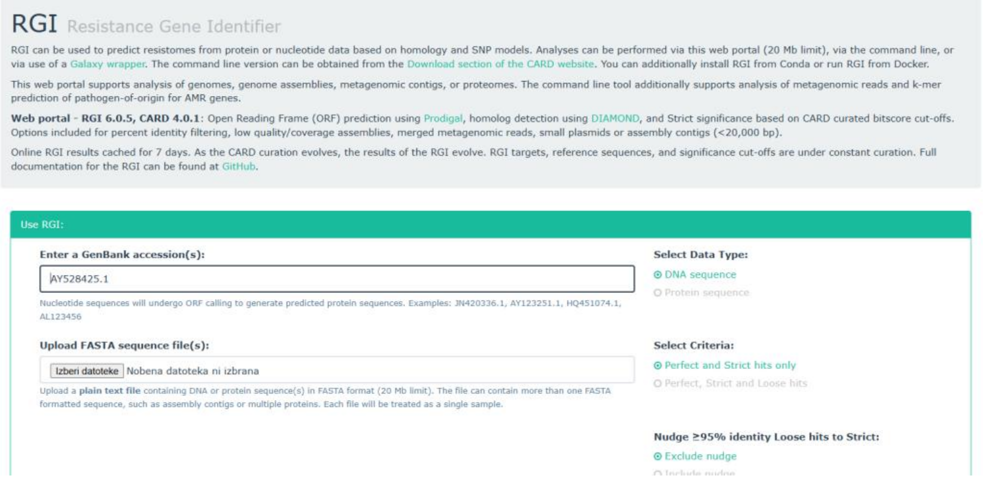


Odgovori:

1. Najbližji sta si Klebsiella pneumoniae in Escherichia coli. To je skladno s pričakovanji, saj obe bakteriji spadata v družino Enterobacteriaceae. Ujema se tudi z opažanji iz 4. naloge. Prav tako sta si blizu Mycobacterium tuberculosis in Nocardia farcinica, ki pa sta obe aktinobakteriji.
2. Ker sta K. pneumoniae in E. coli na isti veji, jima je evolucijsko najbližje Citrobacter amalonaticus, ki je prav tako enterobakterija.

### 6. Naloga

Pri zadnji nalogi uporabim link v navodilih in vanj prilepim accession code vzorca A.



RGI Resistance Gene Identifier

RGI can be used to predict resistomes from protein or nucleotide data based on homology and SNP models. Analyses can be performed via this web portal (20 Mb limit), via the command line, or via use of a Galaxy wrapper. The command line version can be obtained from the Download section of the CARD website. You can additionally install RGI from Conda or run RGI from Docker.

This web portal supports analysis of genomes, genome assemblies, metagenomic contigs, or proteomes. The command line tool additionally supports analysis of metagenomic reads and k-mer prediction of pathogen-of-origin for AMR genes.

**Web portal - RGI 6.0.5, CARD 4.0.1**: Open Reading Frame (ORF) prediction using Prodigal, homolog detection using DIAMOND, and Strict significance based on CARD curated bitscore cut-offs. Options included for percent identity filtering, low quality/coverage assemblies, merged metagenomic reads, small plasmids or assembly contigs (<20,000 bp).

Online RGI results cached for 7 days. As the CARD curation evolves, the results of the RGI evolve. RGI targets, reference sequences, and significance cut-offs are under constant curation. Full documentation for the RGI can be found at GitHub.

Use RGI:

**Enter a GenBank accession(s):** AY528425.1

Nucleotide sequences will undergo ORF calling to generate predicted protein sequences. Examples: JN420336.1, AY123251.1, HQ451074.1, AL123456

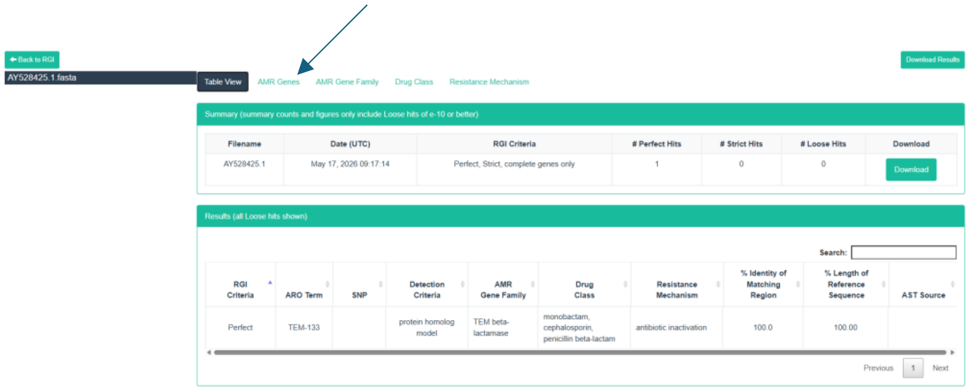
**Upload FASTA sequence file(s):** Nobena datoteka ni izbrana

Upload a **plain text file** containing DNA or protein sequence(s) in FASTA format (20 Mb limit). The file can contain more than one FASTA formatted sequence, such as assembly contigs or multiple proteins. Each file will be treated as a single sample.

**Select Data Type:**
● DNA sequence
○ Protein sequence

**Select Criteria:**
● Perfect and Strict hits only
○ Perfect, Strict and Loose hits

**Nudge ≥95% identity Loose hits to Strict:**
● Exclude nudge



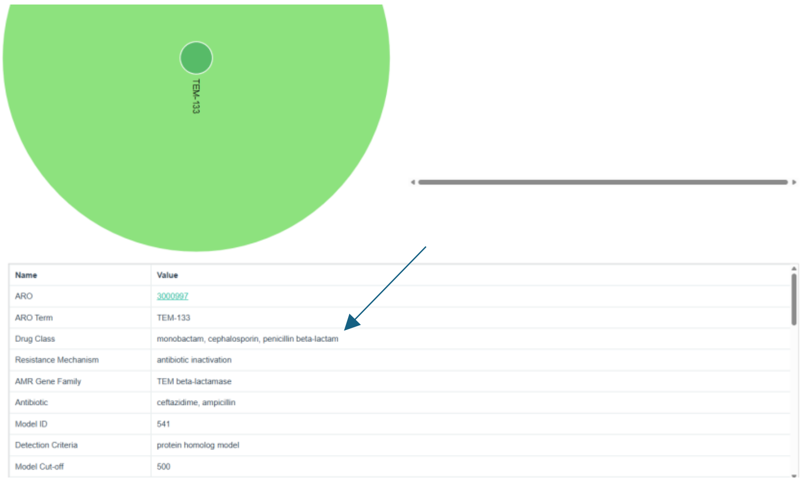
Summary (summary counts and figures only include Loose hits of e-10 or better)

| Filename | Date (UTC) | RGI Criteria | # Perfect Hits | # Strict Hits | # Loose Hits | Download |
|---|---|---|---|---|---|---|
| AY528425.1 | May 17, 2026 09:17:14 | Perfect, Strict, complete genes only | 1 | 0 | 0 | Download |

Results (all Loose hits shown)

Search:

| RGI Criteria | ARO Term | SNP | Detection Criteria | AMR Gene Family | Drug Class | Resistance Mechanism | % Identity of Matching Region | % Length of Reference Sequence | AST Source |
|---|---|---|---|---|---|---|---|---|---|
| Perfect | TEM-133 |  | protein homolog model | TEM beta-lactamase | monobactam, cephalosporin, penicillin beta-lactam | antibiotic inactivation | 100.0 | 100.00 |  |



| Name | Value |
|---|---|
| ARO | 3000997 |
| ARO Term | TEM-133 |
| Drug Class | monobactam, cephalosporin, penicillin beta-lactam |
| Resistance Mechanism | antibiotic inactivation |
| AMR Gene Family | TEM beta-lactamase |
| Antibiotic | ceftazidime, ampicillin |
| Model ID | 541 |
| Detection Criteria | protein homolog model |
| Model Cut-off | 500 |

Odgovori:

1. Monobactam, cephalosporin, penicillin beta-lactam.
2. Perfect, kar je najvišja stopnja zanesljivosti.
3. Da, pričakujem podobne rezultate, ker sta obe bakteriji gramnegativni enterobakteriji iz iste družine Enterobacteriaceae. Obe prav tako nosita gen za beta-laktamazo.