

# **Veritabanları & veri formatları**

- NCBI Home
- Resource List (A-Z)
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology

## Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

### Submit

Deposit data or manuscripts  
into NCBI databases



### Download

Transfer NCBI data to your  
computer



### Learn

Find help documents, attend a  
class or watch a tutorial



## Popular Resources

[PubMed](#)

[Bookshelf](#)

[PubMed Central](#)

[BLAST](#)

[Nucleotide](#)

[Genome](#)

[SNP](#)

[Gene](#)

[Protein](#)

[PubChem](#)

# Genbank

```
LOCUS       JQ907522                847 bp    DNA    linear    BCT 18-APR-2014
DEFINITION  Escherichia coli strain 6 perosamine synthetase (rfbE) gene,
            partial cds.
ACCESSION   JQ907522
VERSION     JQ907522.1  GI:392859036
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1 (bases 1 to 847)
  AUTHORS   Wang,C., Yang,X. and Cheng,H.
  TITLE     Escherichia coli
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 847)
  AUTHORS   Wang,C., Yang,X. and Cheng,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-APR-2012) College of Husbandry and Veterinary, Henan
            Agriculture University, Wenhua Road 95, ZhengZhou, HeNan 450002,
            China
UNIMARK     JQ907522 features
            JQ907522
```

```

FEATURES             Location/Qualifiers
     source            1..847
                        /organism="Escherichia coli"
                        /mol_type="genomic DNA"
                        /strain=6
                        /serotype=0157
                        /isolation_source="feces"
                        /db_xref="taxon:562"
                        /country="China"
                        /collected_by="H.Cheng"
                        /identified_by="H.Cheng"
                        /PCR_primers="fwd_name: rfbe-f, fwd_seq:
                        gtaatggaacggttgctc, rev_name: rfbe-r, rev_seq:
                        attccacgccaaccaag"
     gene              1..847
                        /gene="rfbE"
     CDS                1..847
                        /gene="rfbE"
                        /codon_start=2
                        /transl_table=11
                        /product="perosamine synthetase"
                        /protein_id="AFM85304.1"
                        /db_xref="GI:392859037"
                        /translation="NEKVALHLALLALGISEGDEVIVPTLTYIASVNAIKYTGATPIF
                        VDSNETWQMSVSDIEQKITNKTKAIMCVHLYGHPCDMEQIVELAKSRNLFVIEDCAE
                        AFGSKYKYGKVVGTFGDISTFSFFGNKTITTTGEGGMVVTNDKTLYDRCLHFKGQGLAVH
                        RQYWHDVIGYNYRMTNICA AIGLAQLEQADD FISRKREIADIYKKNINSLVQVHKESK
                        DVFHTYWMVSILTRTAEEREELRNHLADKLIETRPVFPVHTMPMYSEKYQKHPIAED
                        LGGGCK"
     primer_bind       386..407
                        /name="rfbef"
                        /sequence="rfbef: 5'-ACTATTACTACAGGTGAAGGTG-3'"
                        /ugene_group="PCR"

```

```
misc_feature      386..702
                  /length=317
                  /ugene_name="product"
                  /ugene_group="PCR"
primer_bind       complement(682..702)
                  /name="rfber"
                  /sequence="rfber: 5'-GCGGTCCTAGTTAGAATTGAG-3'"
                  /ugene_group="PCR"
```

#### ORIGIN

```
   1 TAATGAAAAG GTTGCTCTTC ATTTAGCTTT GTTAGCGTTA GGTATATCGG AAGGAGATGA
  61 AGTTATTGTT CCAACACTGA CATATATAGC ATCAGTTAAT GCTATAAAAT ACACAGGAGC
 121 CACCCCCATT TTCGTTGATT CAGATAATGA AACTTGGCAA ATGTCTGTTA GTGACATAGA
 181 ACAAAAATC  ACTAATAAAA CTAAAGCTAT TATGTGTGTC CATTTATACG GACATCCATG
 241 TGATATGGAA CAAATTGTAG  AACTGGCCAA AAGTAGAAAT TTGTTTGTA  TTGAAGATTG
 301 CGCTGAAGCC TTTGGTTCTA  AATATAAAGG TAAATATGTG GGAACATTTG  GAGATATTTT
 361 TACTTTTAGC TTTTTTGGAA  ATAAAACAT  TACTACAGGT  GAAGGTGGAA  TGGTTGTCAC
 421 GAATGACAAA AACTTTTATG  ACCGTTGTTT ACATTTTAAA  GGCCAAGGAT  TAGCTGTACA
 481 TAGGCAATAT TGGCATGACG  TTATAGGCTA CAATTATAGG  ATGACAAATA  TCTGCGCTGC
 541 TATAGGATTA GCCCAGTTAG  AACAAGCTGA TGATTTTATA  TCACGAAAAC  GTGAAATTGC
 601 TGATATTTAT AAAAAAATA  TCAACAGTCT TGTACAAGTC  CACAAGGAAA  GTAAAGATGT
 661 TTTTCACACT TATTGGATGG  TCTCAATTCT  AACTAGGACC  GCAGAGGAAA  GAGAGGAATT
 721 AAGGAATCAC CTTGCAGATA  AACTCATCGA  AACAAGGCCA  GTTTTTTACC  CTGTCCACAC
 781 GATGCCAATG TACTCGGAAA  AATATCAAAA  GCACCCTATA  GCTGAGGATC  TTGGTGGGGG
 841 TGGAAAA
```

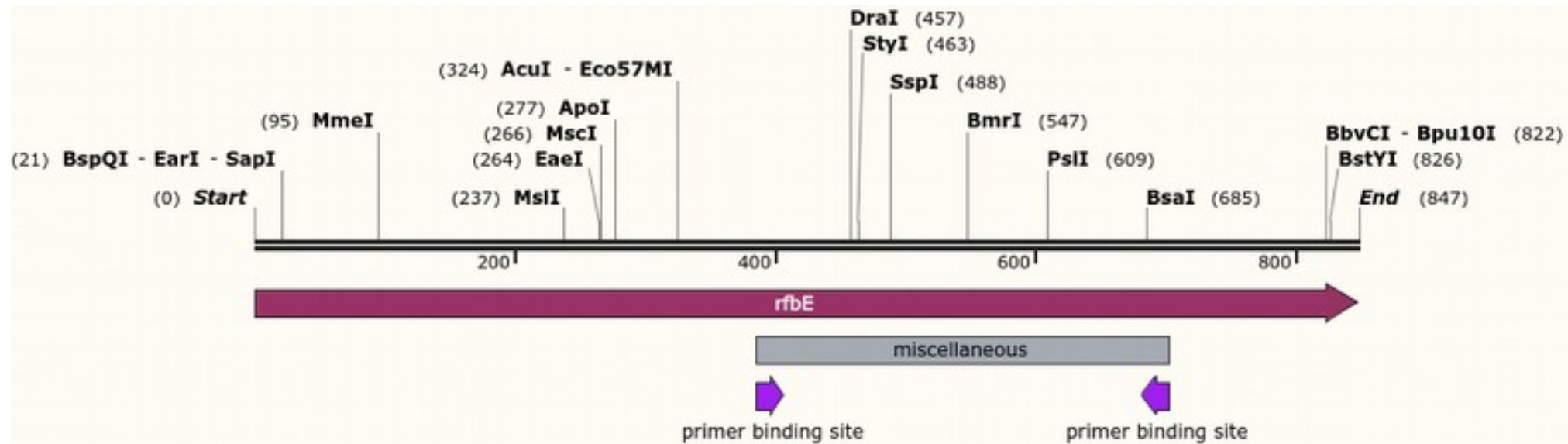
//

```
from Bio import SeqIO
from Bio.SeqRecord import SeqRecord
from Bio.SeqFeature import FeatureLocation, CompoundLocation
```

```
if __name__ == '__main__':
```

```
    r = SeqIO.read("gck.gb", "genbank")
    # @type r : SeqRecord
    # string type of annotations:
    print("id ", r.id)
    print("name ", r.name)
    print("description ", r.description)
    print("seq ", r.seq)
    print("letter annotations ", r.letter_annotations)
```

```
for rec in SeqIO.parse(filename, 'embl'):
    rid = rec.id # accession
    desc = rec.description # dscription
    organism = None
    print(rid)
    #print(desc)
    for a in rec.annotations:
        #print('{0} -> {1}'.format(a, rec.annotations[a]))
        if 'organism' in a:
            organism = rec.annotations[a]
    seq = str(rec.seq)
    dbBlock.append((rid, desc, organism, seq))
    for feature in rec.features:
        featBlock.append((featId,
                           feature.type,
                           feature.location.start,
                           feature.location.end,
                           rid))
        for qualifier in feature.qualifiers:
            qualBlock.append((featId,
                               qualifier,
                               feature.qualifiers[qualifier][0]))
    featId = featId + 1
```



# UniProt

```
ID   GDA4_WHEAT           Reviewed;           297 AA.
AC   P04724;
DT   13-AUG-1987, integrated into UniProtKB/Swiss-Prot.
DT   13-AUG-1987, sequence version 1.
DT   23-MAY-2018, entry version 84.
DE   RecName: Full=Alpha/beta-gliadin A-IV;
DE   AltName: Full=Prolamin;
DE   Flags: Precursor;
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BOP clade;
OC   Pooideae; Triticoideae; Triticeae; Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RX   PubMed=2989281;
RA   Okita T.W., Cheesbrough V., Reeves C.D.;
RT   "Evolution and heterogeneity of the alpha-/beta-type and gamma-type
RT   gliadin DNA sequences.";
RL   J. Biol. Chem. 260:8203-8213(1985).
```



DR EMBL; M11075; AAA34282.1; -; mRNA.  
DR PIR; D22364; D22364.  
DR PIR; T06500; T06500.  
DR UniGene; Ta.90694; -.  
DR ProteinModelPortal; P04724; -.  
DR PRIDE; P04724; -.  
DR Proteomes; UP000019116; Unplaced.  
DR ExpressionAtlas; P04724; baseline.  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA:UniProtKB-KW.  
DR InterPro; IPR036312; Bifun\_inhib/LTP/seed\_sf.  
DR InterPro; IPR016140; Bifunc\_inhib/LTP/seed\_store.  
DR InterPro; IPR001954; Glia\_glutenin.  
DR Pfam; PF13016; Gliadin; 1.  
DR PRINTS; PR00208; GLIADGLUTEN.  
DR SMART; SM00499; AAI; 1.  
DR SUPFAM; SSF47699; SSF47699; 2.  
PE 2: Evidence at transcript level;  
KW Allergen; Complete proteome; Reference proteome; Repeat;  
KW Seed storage protein; Signal; Storage protein.  
FT SIGNAL 1 20  
FT CHAIN 21 297 Alpha/beta-gliadin A-IV.  
FT /FTid=PRO\_0000032271.  
SQ SEQUENCE 297 AA; 34239 MW; 0025ED289AE9588B CRC64;  
MKTFLILALR AIVATTATIA VRVPVQLQP QNPSQQQPQK QVPLVQQQF PGQQQPFPPQ  
QPYPQQQFP SQQPYMLQP FPQPQLPYPQ PQLPYPQPQ FRPQQSYPQP QPQYSQPQP  
ISQQQQQQQQ QQQQQQILQ QILQQQLIPC RDVVLQQHSI AHGSSQVLQQ STYQLVQQFC  
CQQLWQIPEQ SRCQAIHNVV HAILHQQQQ QQQQQQQQQ QPLSQVCFQQ SQQQYPSGQG  
SFQPSQNPQ AQGSVQPQL PQFEEIRNLA LETLPAMCNV YIPPYCTIAP VGIFGTN

//

