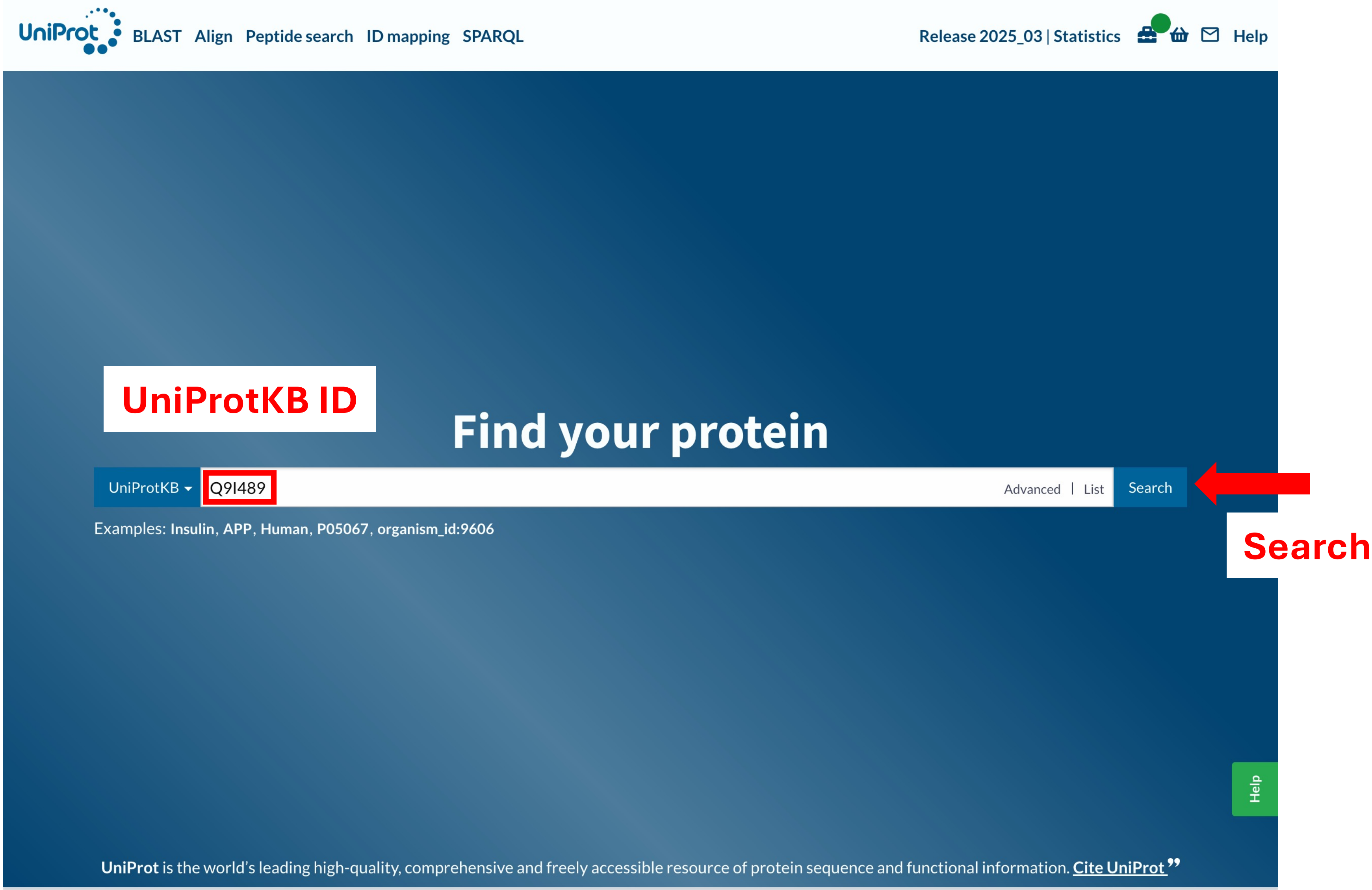
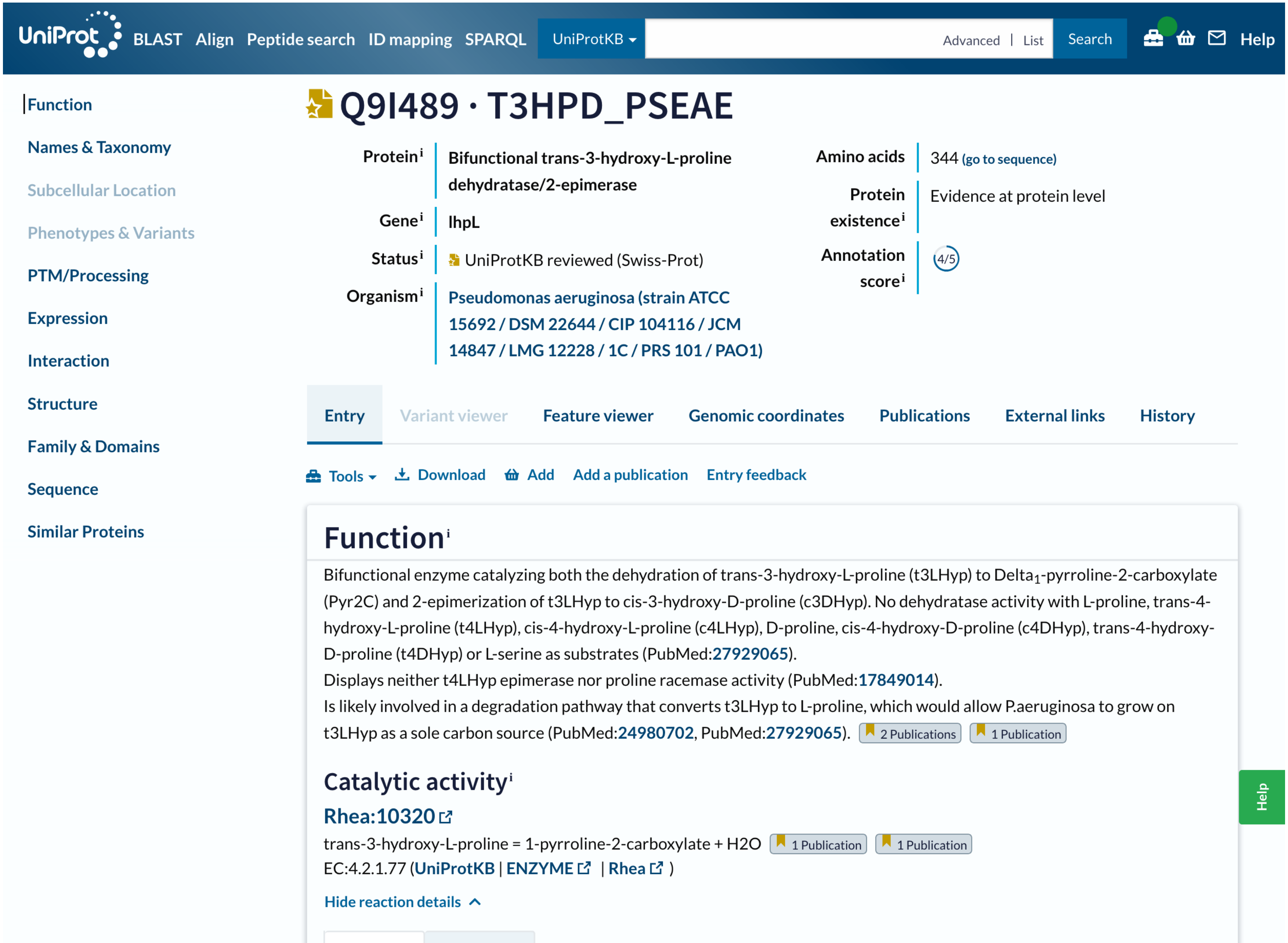
Accessing the protein and DNA sequences for an active UniProtKB ID



UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. Cite UniProt

https://www.uniprot.org/

UniProtKB annotation page



Q9I489 · T3HPD_PSEAE

Protein	Bifunctional trans-3-hydroxy-L-proline dehydratase/2-epimerase
Gene	lhpL
Status	UniProtKB reviewed (Swiss-Prot)
Organism	Pseudomonas aeruginosa (strain ATCC 15692 / DSM 22644 / CIP 104116 / JCM 14847 / LMG 12228 / 1C / PRS 101 / PAO1)
Amino acids	344 (go to sequence)
Protein existence	Evidence at protein level
Annotation score	4/5

Function

Bifunctional enzyme catalyzing both the dehydration of trans-3-hydroxy-L-proline (t3LHyp) to Delta$_1$-pyrroline-2-carboxylate (Pyr2C) and 2-epimerization of t3LHyp to cis-3-hydroxy-D-proline (c3DHyp). No dehydratase activity with L-proline, trans-4-hydroxy-L-proline (t4LHyp), cis-4-hydroxy-L-proline (c4LHyp), D-proline, cis-4-hydroxy-D-proline (c4DHyp), trans-4-hydroxy-D-proline (t4DHyp) or L-serine as substrates (PubMed:27929065).
Displays neither t4LHyp epimerase nor proline racemase activity (PubMed:17849014).
Is likely involved in a degradation pathway that converts t3LHyp to L-proline, which would allow P.aeruginosa to grow on t3LHyp as a sole carbon source (PubMed:24980702, PubMed:27929065). 2 Publications 1 Publication

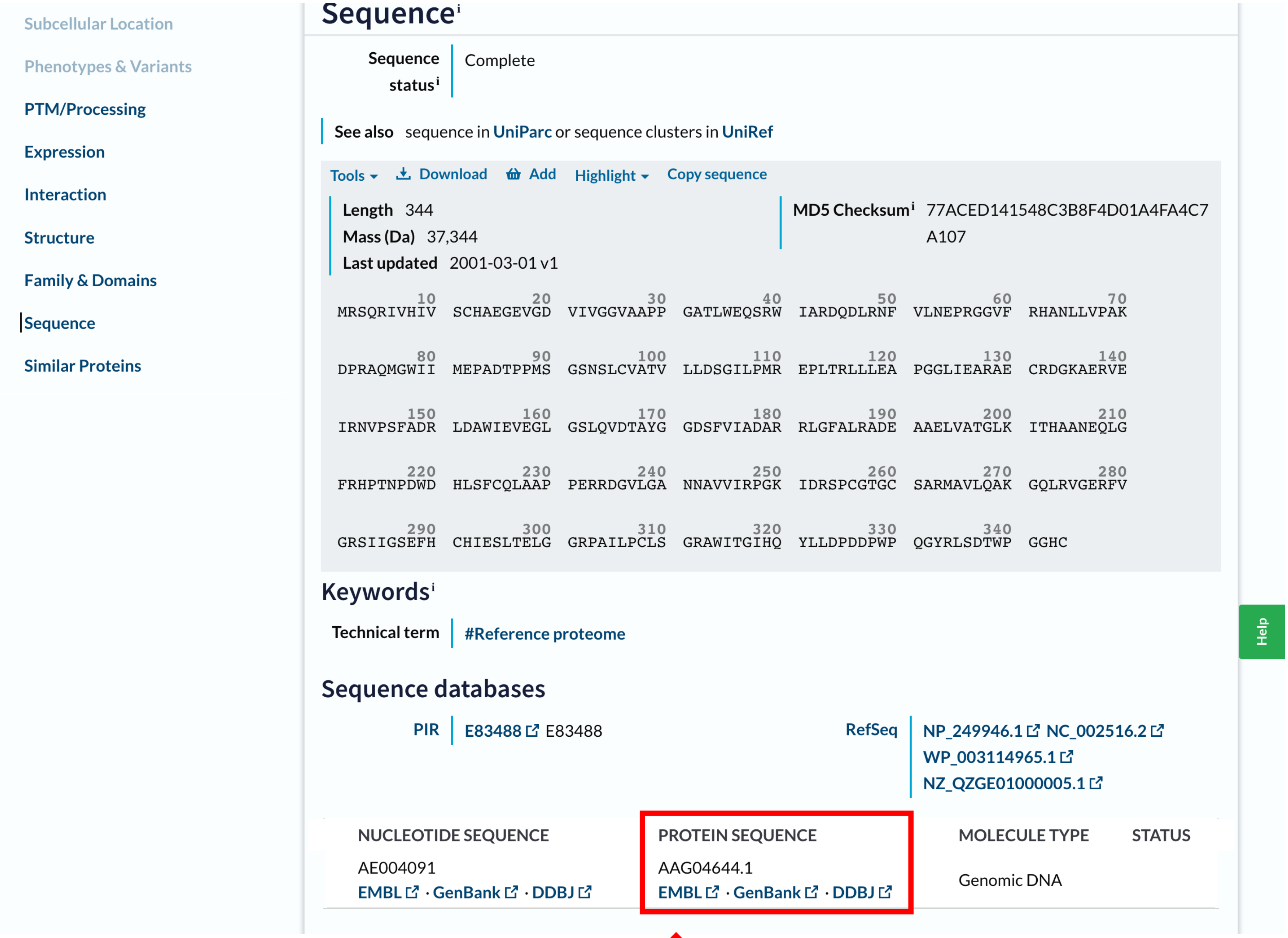
Catalytic activity

Rhea:10320

trans-3-hydroxy-L-proline = 1-pyrroline-2-carboxylate + H2O 1 Publication 1 Publication
EC:4.2.1.77 (UniProtKB | ENZYME | Rhea)

Hide reaction details

Scroll down; locate the PROTEIN SEQUENCE links



Sequence[i]

Sequence status[i]	Complete

See also sequence in UniParc or sequence clusters in UniRef

Length 344
Mass (Da) 37,344
Last updated 2001-03-01 v1

MD5 Checksum[i] 77ACED141548C3B8F4D01A4FA4C7A107

```
        10         20         30         40         50         60         70
MRSQRIVHIV SCHAEGEVGD VIVGGVAAPP GATLWEQSRW IARDQDLRNF VLNEPRGGVF RHANLLVPAK

        80         90        100        110        120        130        140
DPRAQMGWII MEPADTPPMS GSNSLCVATV LLDSGILPMR EPLTRLLLEA PGGLIEARAE CRDGKAERVE

       150        160        170        180        190        200        210
IRNVPSFADR LDAWIEVEGL GSLQVDTAYG GDSFVIADAR RLGFALRADE AAELVATGLK ITHAANEQLG

       220        230        240        250        260        270        280
FRHPTNPDWD HLSFCQLAAP PERRDGVLGA NNAVVIRPGK IDRSPCGTGC SARMAVLQAK GQLRVGERFV

       290        300        310        320        330        340
GRSIIGSEFH CHIESLTELG GRPAILPCLS GRAWITGIHQ YLLDPDDPWP QGYRLSDTWP GGHC
```

Keywords[i]

Technical term	#Reference proteome

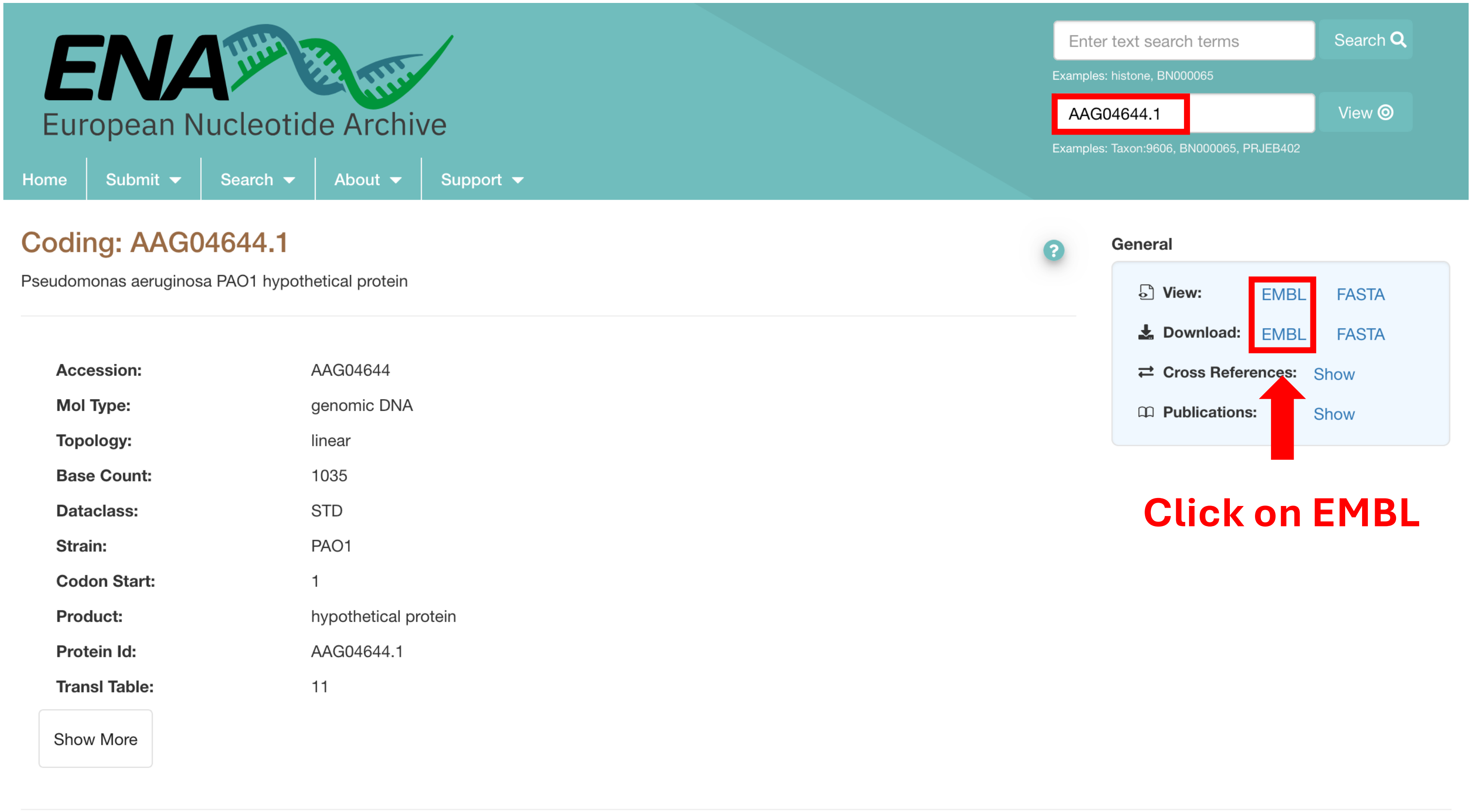
Sequence databases

PIR	E83488 E83488	RefSeq	NP_249946.1 NC_002516.2 WP_003114965.1 NZ_QZGE01000005.1

NUCLEOTIDE SEQUENCE	PROTEIN SEQUENCE	MOLECULE TYPE	STATUS
AE004091 EMBL · GenBank · DDBJ	AAG04644.1 EMBL · GenBank · DDBJ	Genomic DNA	

Click on EMBL

ENA page for Coding information opens



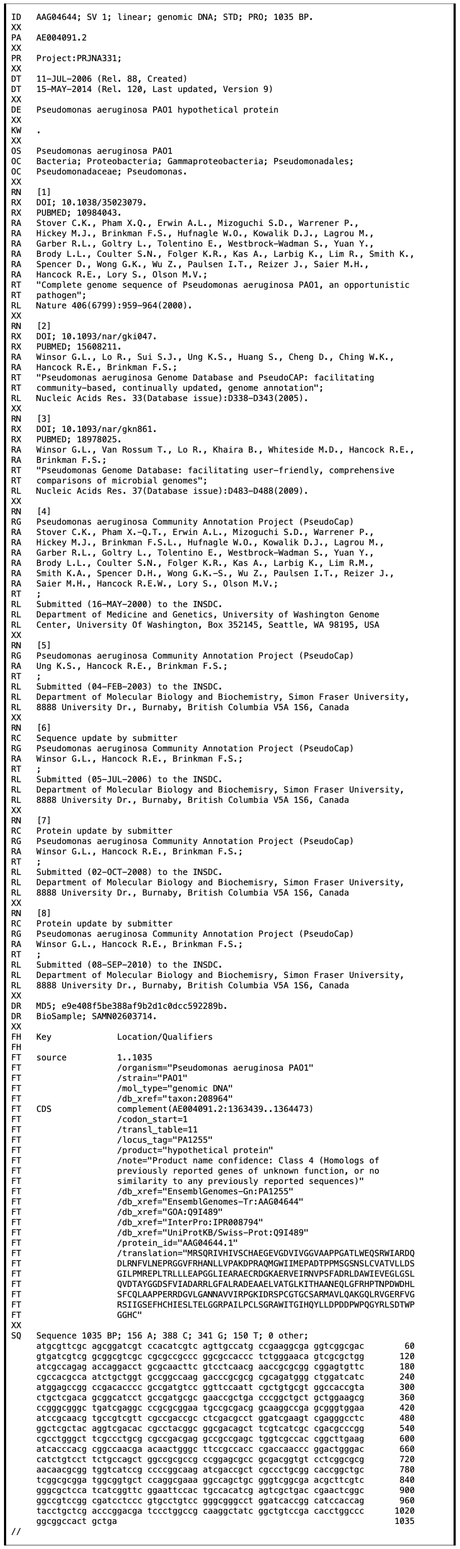
Coding: AAG04644.1

Pseudomonas aeruginosa PAO1 hypothetical protein

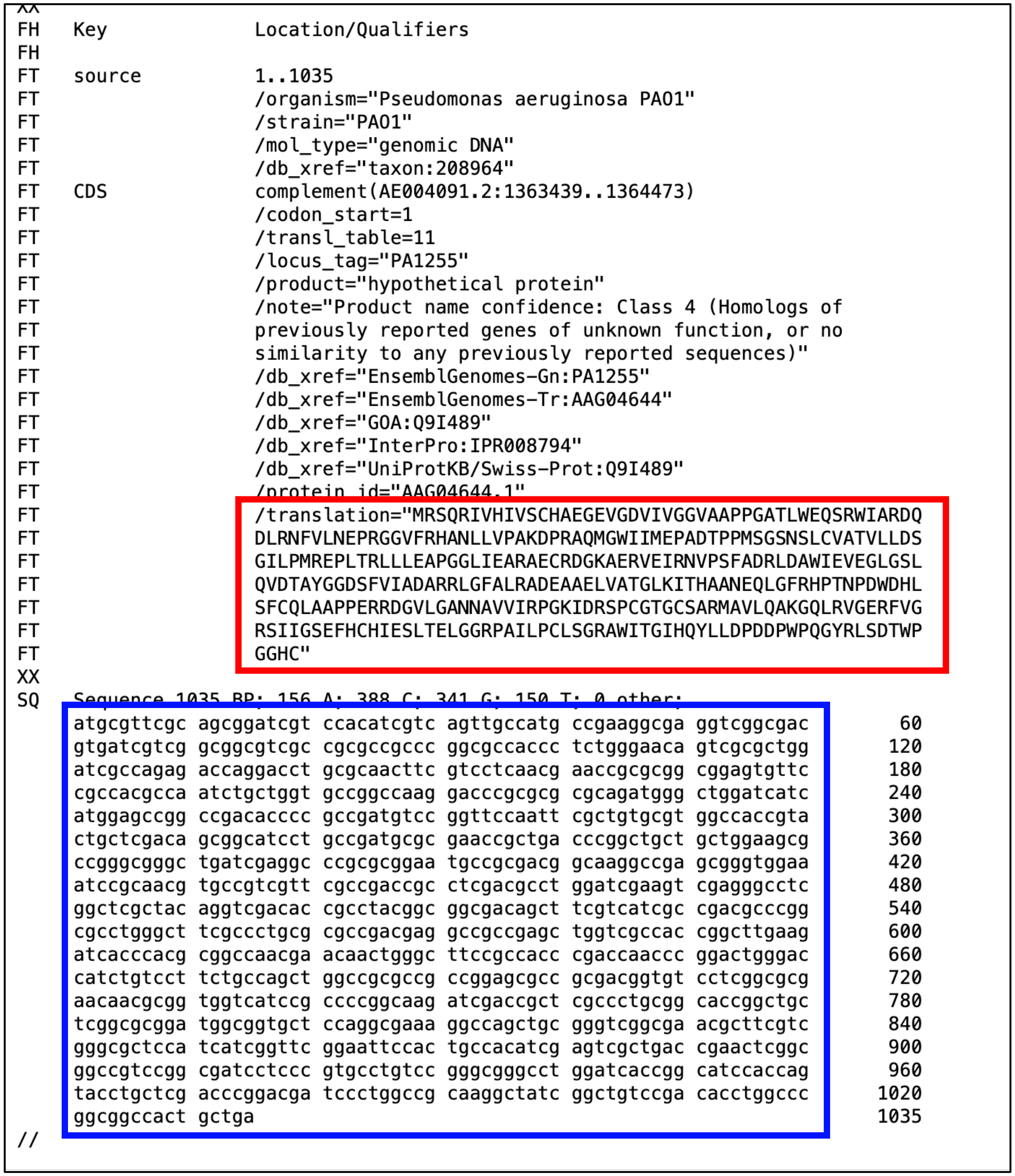
Accession:	AAG04644
Mol Type:	genomic DNA
Topology:	linear
Base Count:	1035
Dataclass:	STD
Strain:	PAO1
Codon Start:	1
Product:	hypothetical protein
Protein Id:	AAG04644.1
Transl Table:	11

Click on EMBL

File with protein and DNA sequences opens

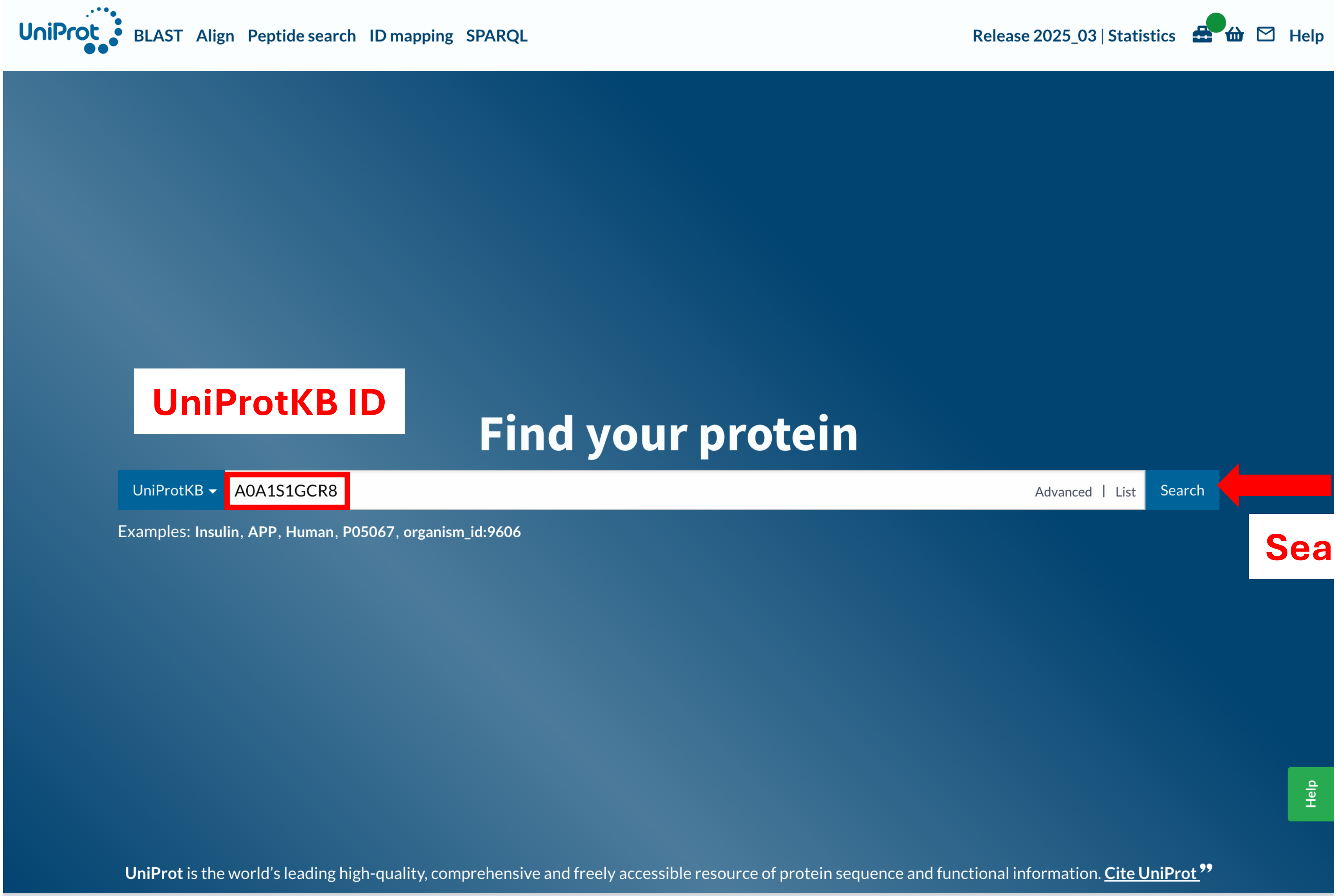


```
ID   AAG04644; SV 1; linear; genomic DNA; STD; PRO; 1035 BP.
XX
PA   AE004091.2
XX
PR   Project:PRJNA331;
XX
DT   11-JUL-2006 (Rel. 88, Created)
DT   15-MAY-2014 (Rel. 120, Last updated, Version 9)
XX
DE   Pseudomonas aeruginosa PAO1 hypothetical protein
XX
KW   .
XX
OS   Pseudomonas aeruginosa PAO1
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
XX
RN   [1]
RX   DOI; 10.1038/35023079.
RX   PUBMED; 10984043.
RA   Stover C.K., Pham X.Q., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R., Smith K.,
RA   Spencer D., Wong G.K., Wu Z., Paulsen I.T., Reizer J., Saier M.H.,
RA   Hancock R.E., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic
RT   pathogen";
RL   Nature 406(6799):959-964(2000).
XX
RN   [2]
RX   DOI; 10.1093/nar/gki047.
RX   PUBMED; 15608211.
RA   Winsor G.L., Lo R., Sui S.J., Ung K.S., Huang S., Cheng D., Ching W.K.,
RA   Hancock R.E., Brinkman F.S.;
RT   "Pseudomonas aeruginosa Genome Database and PseudoCAP: facilitating
RT   community-based, continually updated, genome annotation";
RL   Nucleic Acids Res. 33(Database issue):D338-D343(2005).
XX
RN   [3]
RX   DOI; 10.1093/nar/gkn861.
RX   PUBMED; 18978025.
RA   Winsor G.L., Van Rossum T., Lo R., Khaira B., Whiteside M.D., Hancock R.E.,
RA   Brinkman F.S.;
RT   "Pseudomonas Genome Database: facilitating user-friendly, comprehensive
RT   comparisons of microbial genomes";
RL   Nucleic Acids Res. 37(Database issue):D483-D488(2009).
XX
RN   [4]
RG   Pseudomonas aeruginosa Community Annotation Project (PseudoCap)
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J.,
RA   Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   ;
RL   Submitted (16-MAY-2000) to the INSDC.
RL   Department of Medicine and Genetics, University of Washington Genome
RL   Center, University Of Washington, Box 352145, Seattle, WA 98195, USA
XX
RN   [5]
RG   Pseudomonas aeruginosa Community Annotation Project (PseudoCap)
RA   Ung K.S., Hancock R.E., Brinkman F.S.;
RT   ;
RL   Submitted (04-FEB-2003) to the INSDC.
RL   Department of Molecular Biology and Biochemistry, Simon Fraser University,
RL   8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada
XX
RN   [6]
RC   Sequence update by submitter
RG   Pseudomonas aeruginosa Community Annotation Project (PseudoCap)
RA   Winsor G.L., Hancock R.E., Brinkman F.S.;
RT   ;
RL   Submitted (05-JUL-2006) to the INSDC.
RL   Department of Molecular Biology and Biochemisry, Simon Fraser University,
RL   8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada
XX
RN   [7]
RC   Protein update by submitter
RG   Pseudomonas aeruginosa Community Annotation Project (PseudoCap)
RA   Winsor G.L., Hancock R.E., Brinkman F.S.;
RT   ;
RL   Submitted (02-OCT-2008) to the INSDC.
RL   Department of Molecular Biology and Biochemisry, Simon Fraser University,
RL   8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada
XX
RN   [8]
RC   Protein update by submitter
RG   Pseudomonas aeruginosa Community Annotation Project (PseudoCap)
RA   Winsor G.L., Hancock R.E., Brinkman F.S.;
RT   ;
RL   Submitted (08-SEP-2010) to the INSDC.
RL   Department of Molecular Biology and Biochemisry, Simon Fraser University,
RL   8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada
XX
DR   MD5; e9e408f5be388af9b2d1c0dcc592289b.
DR   BioSample; SAMN02603714.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..1035
FT                   /organism="Pseudomonas aeruginosa PAO1"
FT                   /strain="PAO1"
FT                   /mol_type="genomic DNA"
FT                   /db_xref="taxon:208964"
FT   CDS             complement(AE004091.2:1363439..1364473)
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /locus_tag="PA1255"
FT                   /product="hypothetical protein"
FT                   /note="Product name confidence: Class 4 (Homologs of
FT                   previously reported genes of unknown function, or no
FT                   similarity to any previously reported sequences)"
FT                   /db_xref="EnsemblGenomes-Gn:PA1255"
FT                   /db_xref="EnsemblGenomes-Tr:AAG04644"
FT                   /db_xref="GOA:Q9I489"
FT                   /db_xref="InterPro:IPR008794"
FT                   /db_xref="UniProtKB/Swiss-Prot:Q9I489"
FT                   /protein_id="AAG04644.1"
FT                   /translation="MRSQRIVHIVSCHAEGEVGDVIVGGVAAPPGATLWEQSRWIARDQ
FT                   DLRNFVLNEPRGGVFRHANLLVPAKDPRAQMGWIIMEPADTPPMSGSNSLCVATVLLDS
FT                   GILPMREPLTRLLLEAPGGLIEARAECRDGKAERVEIRNVPSFADRLDAWIEVEGLGSL
FT                   QVDTAYGGDSFVIADARRLGFALRADEAAELVATGLKITHAANEQLGFRHPTNPDWDHL
FT                   SFCQLAAPPERRDGVLGANNAVVIRPGKIDRSPCGTGCSARMAVLQAKGQLRVGERFVG
FT                   RSIIGSEFHCHIESLTELGGRPAILPCLSGRAWITGIHQYLLDPDDPWPQGYRLSDTWP
FT                   GGHC"
XX
SQ   Sequence 1035 BP; 156 A; 388 C; 341 G; 150 T; 0 other;
     atgcgttcgc agcggatcgt ccacatcgtc agttgccatg ccgaaggcga ggtcggcgac        60
     gtgatcgtcg gcggcgtcgc cgcgccgccc ggcgccaccc tctgggaaca gtcgcgctgg       120
     atcgccagag accaggacct gcgcaacttc gtcctcaacg aaccgcgcgg cggagtgttc       180
     cgccacgcca atctgctggt gccggccaag gacccgcgcg cgcagatggg ctggatcatc       240
     atggagccgg ccgacacccc gccgatgtcc ggttccaatt cgctgtgcgt ggccaccgta       300
     ctgctcgaca gcggcatcct gccgatgcgc gaaccgctga cccggctgct gctggaagcg       360
     ccgggcgggc tgatcgaggc ccgcgcggaa tgccgcgacg gcaaggccga gcgggtggaa       420
     atccgcaacg tgccgtcgtt cgccgaccgc ctcgacgcct ggatcgaagt cgagggcctc       480
     ggctcgctac aggtcgacac cgcctacggc ggcgacagct tcgtcatcgc cgacgcccgg       540
     cgcctgggct tcgccctgcg cgccgacgag gccgccgagc tggtcgccac cggcttgaag       600
     atcacccacg cggccaacga acaactgggc ttccgccacc cgaccaaccc ggactgggac       660
     catctgtcct tctgccagct ggccgcgccg ccggagcgcc gcgacggtgt cctcggcgcg       720
     aacaacgcgg tggtcatccg ccccggcaag atcgaccgct cgccctgcgg caccggctgc       780
     tcggcgcgga tggcggtgct ccaggcgaaa ggccagctgc gggtcggcga acgcttcgtc       840
     gggcgctcca tcatcggttc ggaattccac tgccacatcg agtcgctgac cgaactcggc       900
     ggccgtccgg cgatcctccc gtgcctgtcc gggcgggcct ggatcaccgg catccaccag       960
     tacctgctcg acccggacga tccctggccg caaggctatc ggctgtccga cacctggccc      1020
     ggcggccact gctga                                                      1035
//
```



```
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..1035
FT                   /organism="Pseudomonas aeruginosa PAO1"
FT                   /strain="PAO1"
FT                   /mol_type="genomic DNA"
FT                   /db_xref="taxon:208964"
FT   CDS             complement(AE004091.2:1363439..1364473)
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /locus_tag="PA1255"
FT                   /product="hypothetical protein"
FT                   /note="Product name confidence: Class 4 (Homologs of
FT                   previously reported genes of unknown function, or no
FT                   similarity to any previously reported sequences)"
FT                   /db_xref="EnsemblGenomes-Gn:PA1255"
FT                   /db_xref="EnsemblGenomes-Tr:AAG04644"
FT                   /db_xref="GOA:Q9I489"
FT                   /db_xref="InterPro:IPR008794"
FT                   /db_xref="UniProtKB/Swiss-Prot:Q9I489"
FT                   /protein_id="AAG04644.1"
FT                   /translation="MRSQRIVHIVSCHAEGEVGDVIVGGVAAPPGATLWEQSRWIARDQ
FT                   DLRNFVLNEPRGGVFRHANLLVPAKDPRAQMGWIIMEPADTPPMSGSNSLCVATVLLDS
FT                   GILPMREPLTRLLLEAPGGLIEARAECRDGKAERVEIRNVPSFADRLDAWIEVEGLGSL
FT                   QVDTAYGGDSFVIADARRLGFALRADEAAELVATGLKITHAANEQLGFRHPTNPDWDHL
FT                   SFCQLAAPPERRDGVLGANNAVVIRPGKIDRSPCGTGCSARMAVLQAKGQLRVGERFVG
FT                   RSIIGSEFHCHIESLTELGGRPAILPCLSGRAWITGIHQYLLDPDDPWPQGYRLSDTWP
FT                   GGHC"
XX
SQ   Sequence 1035 BP; 156 A; 388 C; 341 G; 150 T; 0 other;
     atgcgttcgc agcggatcgt ccacatcgtc agttgccatg ccgaaggcga ggtcggcgac        60
     gtgatcgtcg gcggcgtcgc cgcgccgccc ggcgccaccc tctgggaaca gtcgcgctgg       120
     atcgccagag accaggacct gcgcaacttc gtcctcaacg aaccgcgcgg cggagtgttc       180
     cgccacgcca atctgctggt gccggccaag gacccgcgcg cgcagatggg ctggatcatc       240
     atggagccgg ccgacacccc gccgatgtcc ggttccaatt cgctgtgcgt ggccaccgta       300
     ctgctcgaca gcggcatcct gccgatgcgc gaaccgctga cccggctgct gctggaagcg       360
     ccgggcgggc tgatcgaggc ccgcgcggaa tgccgcgacg gcaaggccga gcgggtggaa       420
     atccgcaacg tgccgtcgtt cgccgaccgc ctcgacgcct ggatcgaagt cgagggcctc       480
     ggctcgctac aggtcgacac cgcctacggc ggcgacagct tcgtcatcgc cgacgcccgg       540
     cgcctgggct tcgccctgcg cgccgacgag gccgccgagc tggtcgccac cggcttgaag       600
     atcacccacg cggccaacga acaactgggc ttccgccacc cgaccaaccc ggactgggac       660
     catctgtcct tctgccagct ggccgcgccg ccggagcgcc gcgacggtgt cctcggcgcg       720
     aacaacgcgg tggtcatccg ccccggcaag atcgaccgct cgccctgcgg caccggctgc       780
     tcggcgcgga tggcggtgct ccaggcgaaa ggccagctgc gggtcggcga acgcttcgtc       840
     gggcgctcca tcatcggttc ggaattccac tgccacatcg agtcgctgac cgaactcggc       900
     ggccgtccgg cgatcctccc gtgcctgtcc gggcgggcct ggatcaccgg catccaccag       960
     tacctgctcg acccggacga tccctggccg caaggctatc ggctgtccga cacctggccc      1020
     ggcggccact gctga                                                      1035
//
```

Accessing protein and DNA sequences for archived UniProtKB ID



A0A1S1GCR8

Examples: **Insulin**, **APP**, **Human**, **P05067**, **organism_id:9606**

UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. Cite UniProt

UniProtKB ID

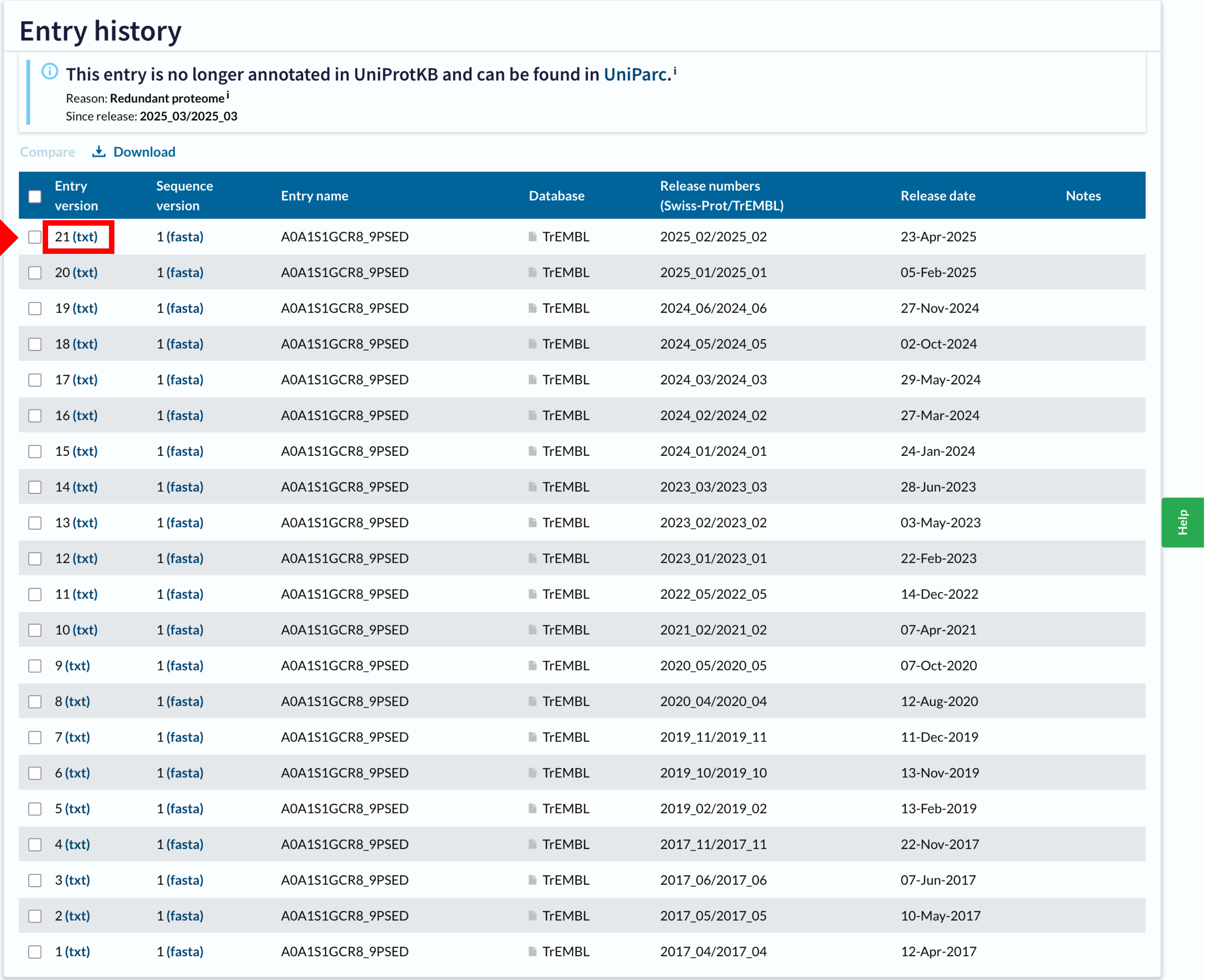
Search

https://www.uniprot.org/

A0A1S1GCR8

File with Entry history opens; click on most recent version text link

Click



Entry history

This entry is no longer annotated in UniProtKB and can be found in UniParc.

Reason: **Redundant proteome**

Since release: **2025_03/2025_03**

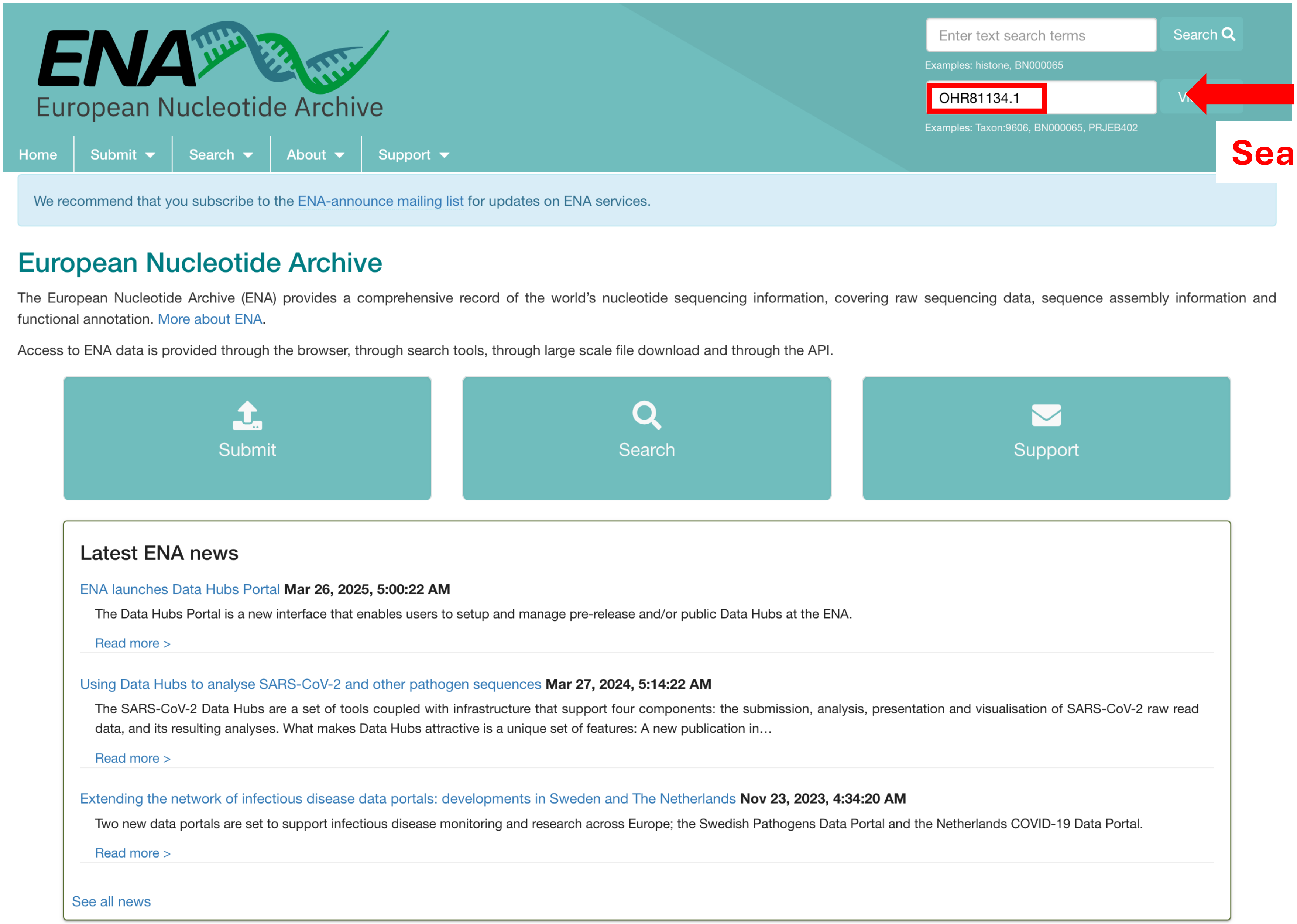
Compare Download

☐	Entry version	Sequence version	Entry name	Database	Release numbers (Swiss-Prot/TrEMBL)	Release date	Notes
☐	21 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2025_02/2025_02	23-Apr-2025	
☐	20 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2025_01/2025_01	05-Feb-2025	
☐	19 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2024_06/2024_06	27-Nov-2024	
☐	18 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2024_05/2024_05	02-Oct-2024	
☐	17 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2024_03/2024_03	29-May-2024	
☐	16 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2024_02/2024_02	27-Mar-2024	
☐	15 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2024_01/2024_01	24-Jan-2024	
☐	14 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2023_03/2023_03	28-Jun-2023	
☐	13 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2023_02/2023_02	03-May-2023	
☐	12 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2023_01/2023_01	22-Feb-2023	
☐	11 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2022_05/2022_05	14-Dec-2022	
☐	10 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2021_02/2021_02	07-Apr-2021	
☐	9 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2020_05/2020_05	07-Oct-2020	
☐	8 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2020_04/2020_04	12-Aug-2020	
☐	7 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2019_11/2019_11	11-Dec-2019	
☐	6 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2019_10/2019_10	13-Nov-2019	
☐	5 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2019_02/2019_02	13-Feb-2019	
☐	4 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2017_11/2017_11	22-Nov-2017	
☐	3 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2017_06/2017_06	07-Jun-2017	
☐	2 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2017_05/2017_05	10-May-2017	
☐	1 (txt)	1 (fasta)	A0A1S1GCR8_9PSED	TrEMBL	2017_04/2017_04	12-Apr-2017	

Entry version opens , with EMBL Protein ID and sequence

```
ID   A0A1S1GCR8_9PSED        Unreviewed;       308 AA.
AC   A0A1S1GCR8;
DT   12-APR-2017, integrated into UniProtKB/TrEMBL.
DT   12-APR-2017, sequence version 1.
DT   02-APR-2025, entry version 21.
DE   SubName: Full=Hydroxyproline-2-epimerase {ECO:0000313|EMBL:OHR81134.1}
GN   ORFNames=HMPREF3289_28740 {ECO:0000313|EMBL:OHR81134.1};
OS   Pseudomonas sp. HMSC75E02.
OC   Bacteria; Pseudomonadati; Pseudomonadota; Gammaproteobacteria;
OC   Pseudomonadales; Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=1608908 {ECO:0000313|EMBL:OHR81134.1, ECO:0000313|Proteomes:UP000180041};
RN   [1] {ECO:0000313|EMBL:OHR81134.1, ECO:0000313|Proteomes:UP000180041}
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=HMSC75E02 {ECO:0000313|EMBL:OHR81134.1,
RC   ECO:0000313|Proteomes:UP000180041};
RA   Wen L., He K., Yang H.;
RL   Submitted (FEB-2016) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the proline racemase family.
CC       {ECO:0000256|ARBA:ARBA00007529}.
CC   -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ
CC       whole genome shotgun (WGS) entry which is preliminary data.
CC       {ECO:0000313|EMBL:OHR81134.1}.
CC   ---------------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see https://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution (CC BY 4.0) License
CC   ---------------------------------------------------------------------------
DR   EMBL; LTNB01000780; OHR81134.1; -; Genomic_DNA.
DR   RefSeq; WP_009619554.1; NZ_KV838524.1.
DR   AlphaFoldDB; A0A1S1GCR8; -.
DR   Proteomes; UP000180041; Unassembled WGS sequence.
DR   GO; GO:0047580; F:4-hydroxyproline epimerase activity; IEA:UniProtKB-ARBA.
DR   FunFam; 3.10.310.10:FF:000012; 4-hydroxyproline 2-epimerase; 1.
DR   Gene3D; 3.10.310.10; Diaminopimelate Epimerase, Chain A, domain 1; 2.
DR   InterPro; IPR008794; Pro_racemase_fam.
DR   PANTHER; PTHR33442; TRANS-3-HYDROXY-L-PROLINE DEHYDRATASE; 1.
DR   PANTHER; PTHR33442:SF1; TRANS-3-HYDROXY-L-PROLINE DEHYDRATASE; 1.
DR   Pfam; PF05544; Pro_racemase; 1.
DR   PIRSF; PIRSF029792; Pro_racemase; 1.
DR   SFLD; SFLDS00028; Proline_Racemase; 1.
DR   SUPFAM; SSF54506; Diaminopimelate epimerase-like; 1.
PE   3: Inferred from homology;
SQ   SEQUENCE   308 AA;  33287 MW;  CE4ED53CC5596C3A CRC64;
     MKRVHIIDSH TGGEPTRLVM KGFPELAGAS LAEKRNALRE QHDHWRRACL LEPRGNDVLV
     GALYCEPVSA GATCGVIFFN NAGYLNMCGH GTIGLVASLQ HLGLIGPGVH KIDTPVGPVS
     ATLHEDGAVT IGNVPAYRYR RQVPVEVPGH GRVYGDIAWG GNWFFLVSEH GQRLALDNVE
     ALTDFTWALL QALEAQGIHG EGGAPIDHVE LFADDEEADS RNFVMCPGKA YDRSPCGTGT
     SAKLACLAAD GKLAPGQAWV QAGITGSRFE GRYEWEGERI RPFITGRAYM TADSTLLIDE
     QDPFAWGI
//
```

Accessing the protein and DNA sequences for an EMBL Protein ID



Search

We recommend that you subscribe to the ENA-announce mailing list for updates on ENA services.

European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation. More about ENA.

Access to ENA data is provided through the browser, through search tools, through large scale file download and through the API.

Latest ENA news

ENA launches Data Hubs Portal **Mar 26, 2025, 5:00:22 AM**

The Data Hubs Portal is a new interface that enables users to setup and manage pre-release and/or public Data Hubs at the ENA.

Read more >

Using Data Hubs to analyse SARS-CoV-2 and other pathogen sequences **Mar 27, 2024, 5:14:22 AM**

The SARS-CoV-2 Data Hubs are a set of tools coupled with infrastructure that support four components: the submission, analysis, presentation and visualisation of SARS-CoV-2 raw read data, and its resulting analyses. What makes Data Hubs attractive is a unique set of features: A new publication in…

Read more >

Extending the network of infectious disease data portals: developments in Sweden and The Netherlands **Nov 23, 2023, 4:34:20 AM**

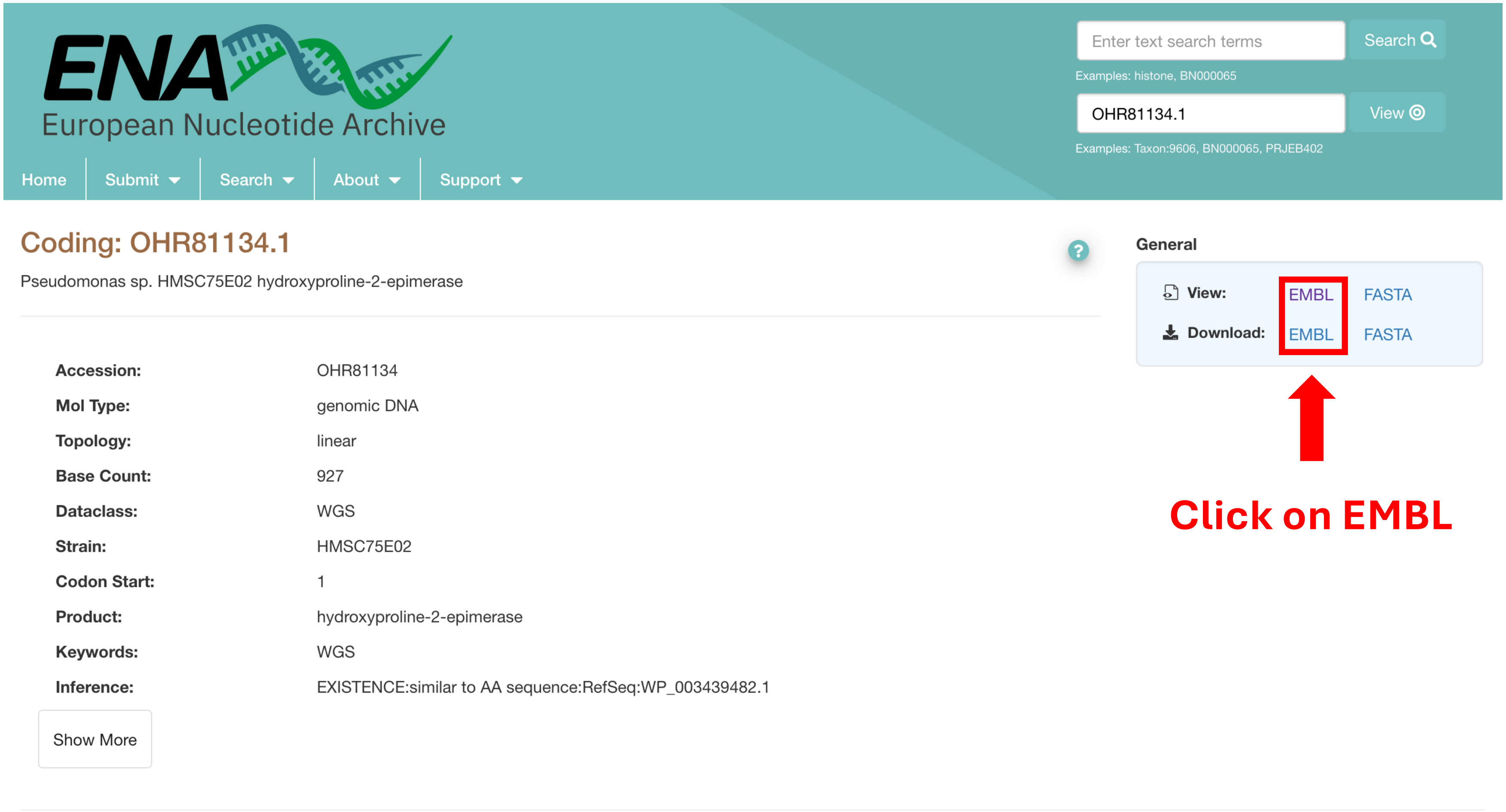
Two new data portals are set to support infectious disease monitoring and research across Europe; the Swedish Pathogens Data Portal and the Netherlands COVID-19 Data Portal.

Read more >

See all news

https://www.ebi.ac.uk/ena/browser/home

ENA page for Coding information opens



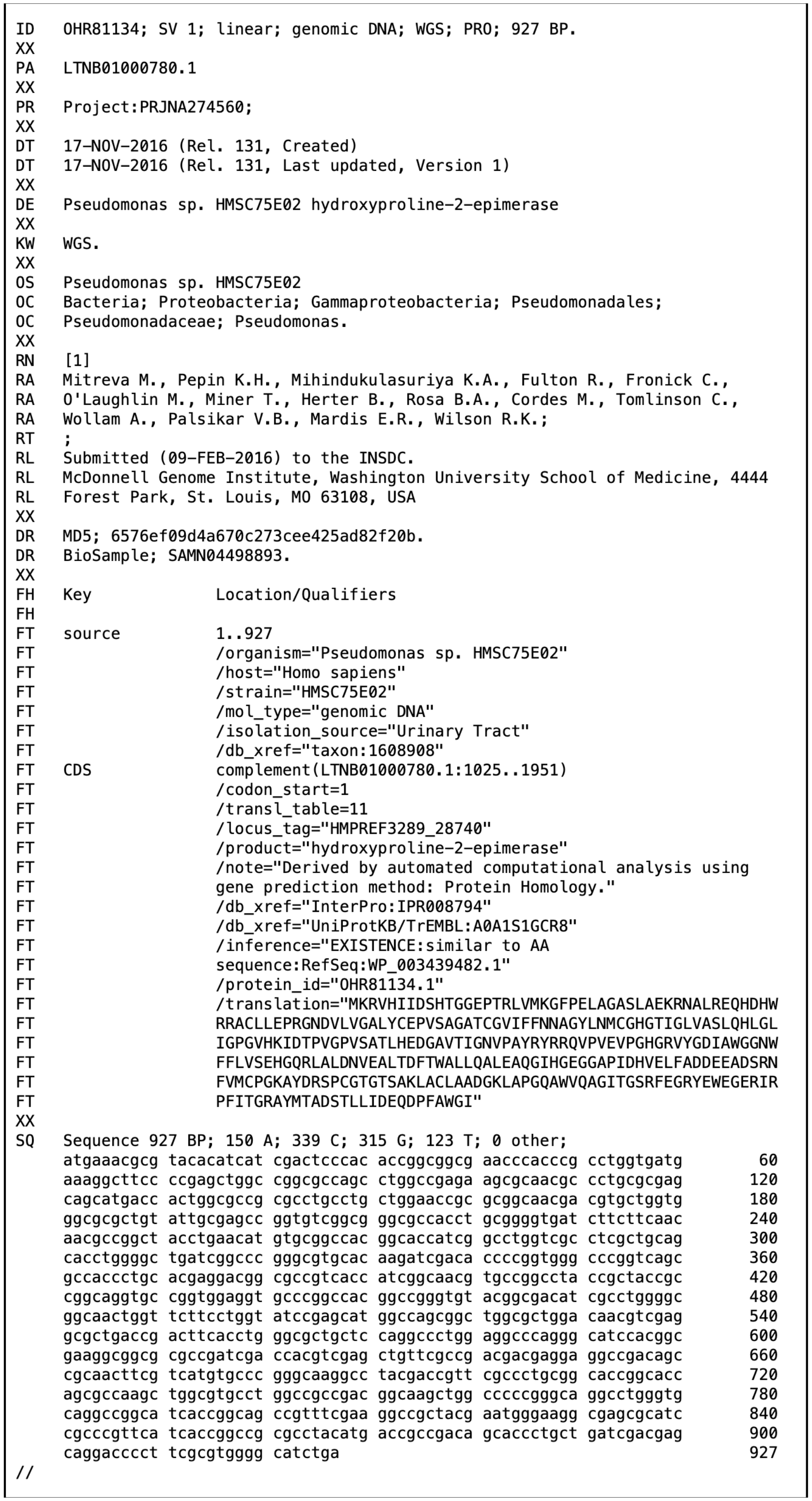
Coding: OHR81134.1

Pseudomonas sp. HMSC75E02 hydroxyproline-2-epimerase

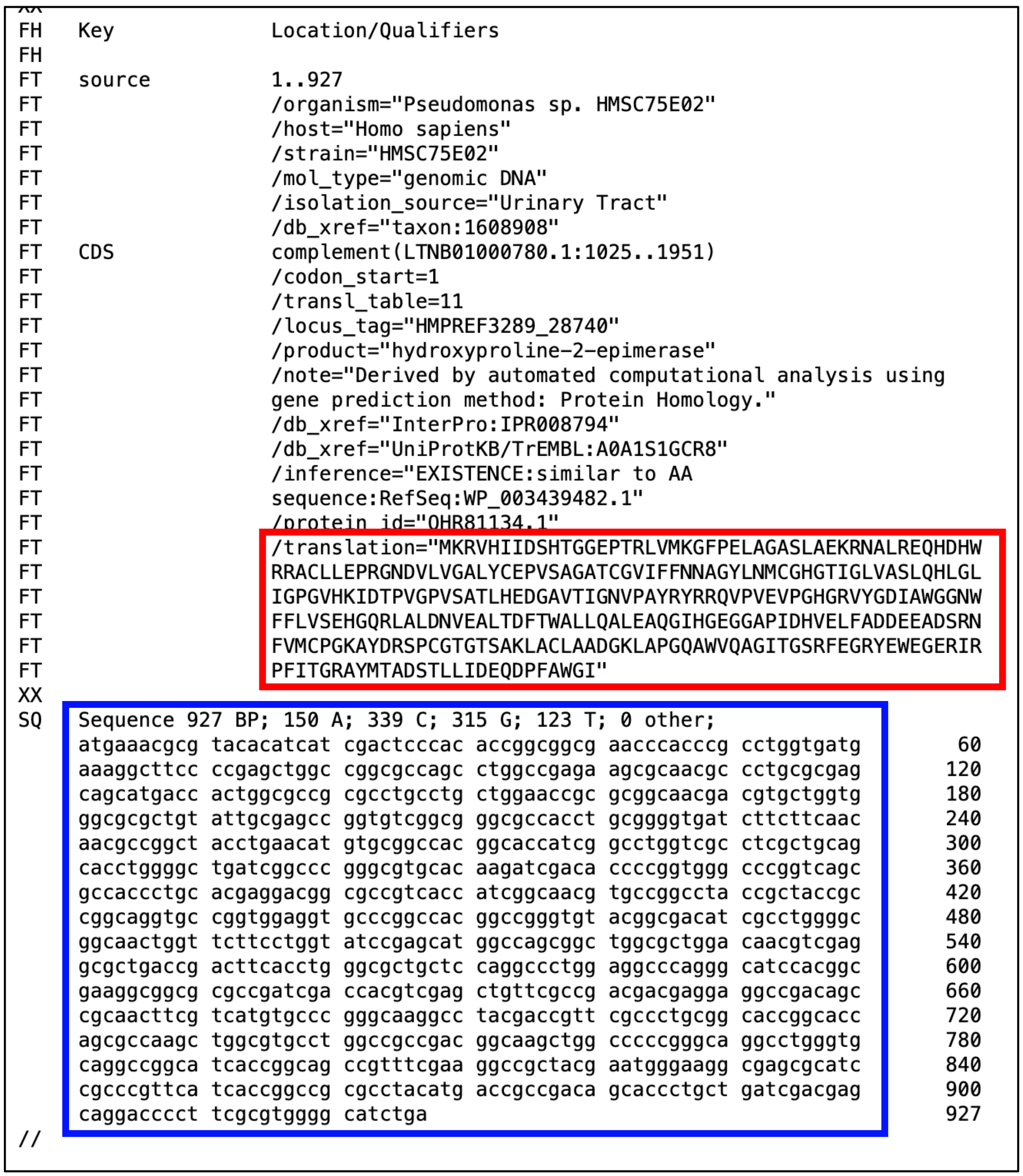
Accession:	OHR81134
Mol Type:	genomic DNA
Topology:	linear
Base Count:	927
Dataclass:	WGS
Strain:	HMSC75E02
Codon Start:	1
Product:	hydroxyproline-2-epimerase
Keywords:	WGS
Inference:	EXISTENCE:similar to AA sequence:RefSeq:WP_003439482.1

Click on EMBL

File with protein and DNA sequences opens



```
ID   OHR81134; SV 1; linear; genomic DNA; WGS; PRO; 927 BP.
XX
PA   LTNB01000780.1
XX
PR   Project:PRJNA274560;
XX
DT   17-NOV-2016 (Rel. 131, Created)
DT   17-NOV-2016 (Rel. 131, Last updated, Version 1)
XX
DE   Pseudomonas sp. HMSC75E02 hydroxyproline-2-epimerase
XX
KW   WGS.
XX
OS   Pseudomonas sp. HMSC75E02
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
XX
RN   [1]
RA   Mitreva M., Pepin K.H., Mihindukulasuriya K.A., Fulton R., Fronick C.,
RA   O'Laughlin M., Miner T., Herter B., Rosa B.A., Cordes M., Tomlinson C.,
RA   Wollam A., Palsikar V.B., Mardis E.R., Wilson R.K.;
RT   ;
RL   Submitted (09-FEB-2016) to the INSDC.
RL   McDonnell Genome Institute, Washington University School of Medicine, 4444
RL   Forest Park, St. Louis, MO 63108, USA
XX
DR   MD5; 6576ef09d4a670c273cee425ad82f20b.
DR   BioSample; SAMN04498893.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..927
FT                   /organism="Pseudomonas sp. HMSC75E02"
FT                   /host="Homo sapiens"
FT                   /strain="HMSC75E02"
FT                   /mol_type="genomic DNA"
FT                   /isolation_source="Urinary Tract"
FT                   /db_xref="taxon:1608908"
FT   CDS             complement(LTNB01000780.1:1025..1951)
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /locus_tag="HMPREF3289_28740"
FT                   /product="hydroxyproline-2-epimerase"
FT                   /note="Derived by automated computational analysis using
FT                   gene prediction method: Protein Homology."
FT                   /db_xref="InterPro:IPR008794"
FT                   /db_xref="UniProtKB/TrEMBL:A0A1S1GCR8"
FT                   /inference="EXISTENCE:similar to AA
FT                   sequence:RefSeq:WP_003439482.1"
FT                   /protein_id="OHR81134.1"
FT                   /translation="MKRVHIIDSHTGGEPTRLVMKGFPELAGASLAEKRNALREQHDHW
FT                   RRACLLEPRGNDVLVGALYCEPVSAGATCGVIFFNNAGYLNMCGHGTIGLVASLQHLGL
FT                   IGPGVHKIDTPVGPVSATLHEDGAVTIGNVPAYRYRRQVPVEVPGHGRVYGDIAWGGNW
FT                   FFLVSEHGQRLALDNVEALTDFTWALLQALEAQGIHGEGGAPIDHVELFADDEEADSRN
FT                   FVMCPGKAYDRSPCGTGTSAKLACLAADGKLAPGQAWVQAGITGSRFEGRYEWEGERIR
FT                   PFITGRAYMTADSTLLIDEQDPFAWGI"
XX
SQ   Sequence 927 BP; 150 A; 339 C; 315 G; 123 T; 0 other;
     atgaaacgcg tacacatcat cgactcccac accggcggcg aacccacccg cctggtgatg        60
     aaaggcttcc ccgagctggc cggcgccagc ctggccgaga agcgcaacgc cctgcgcgag       120
     cagcatgacc actggcgccg cgcctgcctg ctggaaccgc gcggcaacga cgtgctggtg       180
     ggcgcgctgt attgcgagcc ggtgtcggcg ggcgccacct gcggggtgat cttcttcaac       240
     aacgccggct acctgaacat gtgcggccac ggcaccatcg gcctggtcgc ctcgctgcag       300
     cacctggggc tgatcggccc gggcgtgcac aagatcgaca ccccggtggg cccggtcagc       360
     gccaccctgc acgaggacgg cgccgtcacc atcggcaacg tgccggccta ccgctaccgc       420
     cggcaggtgc cggtggaggt gcccggccac ggccgggtgt acggcgacat cgcctggggc       480
     ggcaactggt tcttcctggt atccgagcat ggccagcggc tggcgctgga caacgtcgag       540
     gcgctgaccg acttcacctg ggcgctgctc caggccctgg aggcccaggg catccacggc       600
     gaaggcggcg cgccgatcga ccacgtcgag ctgttcgccg acgacgagga ggccgacagc       660
     cgcaacttcg tcatgtgccc gggcaaggcc tacgaccgtt cgccctgcgg caccggcacc       720
     agcgccaagc tggcgtgcct ggccgccgac ggcaagctgg cccccgggca ggcctgggtg       780
     caggccggca tcaccggcag ccgtttcgaa ggccgctacg aatgggaagg cgagcgcatc       840
     cgcccgttca tcaccggccg cgcctacatg accgccgaca gcaccctgct gatcgacgag       900
     caggacccct tcgcgtgggg catctga                                           927
//
```



```
FH   Key             Location/Qualifiers
FH
FT   source          1..927
FT                   /organism="Pseudomonas sp. HMSC75E02"
FT                   /host="Homo sapiens"
FT                   /strain="HMSC75E02"
FT                   /mol_type="genomic DNA"
FT                   /isolation_source="Urinary Tract"
FT                   /db_xref="taxon:1608908"
FT   CDS             complement(LTNB01000780.1:1025..1951)
FT                   /codon_start=1
FT                   /transl_table=11
FT                   /locus_tag="HMPREF3289_28740"
FT                   /product="hydroxyproline-2-epimerase"
FT                   /note="Derived by automated computational analysis using
FT                   gene prediction method: Protein Homology."
FT                   /db_xref="InterPro:IPR008794"
FT                   /db_xref="UniProtKB/TrEMBL:A0A1S1GCR8"
FT                   /inference="EXISTENCE:similar to AA
FT                   sequence:RefSeq:WP_003439482.1"
FT                   /protein_id="OHR81134.1"
FT                   /translation="MKRVHIIDSHTGGEPTRLVMKGFPELAGASLAEKRNALREQHDHW
FT                   RRACLLEPRGNDVLVGALYCEPVSAGATCGVIFFNNAGYLNMCGHGTIGLVASLQHLGL
FT                   IGPGVHKIDTPVGPVSATLHEDGAVTIGNVPAYRYRRQVPVEVPGHGRVYGDIAWGGNW
FT                   FFLVSEHGQRLALDNVEALTDFTWALLQALEAQGIHGEGGAPIDHVELFADDEEADSRN
FT                   FVMCPGKAYDRSPCGTGTSAKLACLAADGKLAPGQAWVQAGITGSRFEGRYEWEGERIR
FT                   PFITGRAYMTADSTLLIDEQDPFAWGI"
XX
SQ   Sequence 927 BP; 150 A; 339 C; 315 G; 123 T; 0 other;
     atgaaacgcg tacacatcat cgactcccac accggcggcg aacccacccg cctggtgatg        60
     aaaggcttcc ccgagctggc cggcgccagc ctggccgaga agcgcaacgc cctgcgcgag       120
     cagcatgacc actggcgccg cgcctgcctg ctggaaccgc gcggcaacga cgtgctggtg       180
     ggcgcgctgt attgcgagcc ggtgtcggcg ggcgccacct gcggggtgat cttcttcaac       240
     aacgccggct acctgaacat gtgcggccac ggcaccatcg gcctggtcgc ctcgctgcag       300
     cacctggggc tgatcggccc gggcgtgcac aagatcgaca ccccggtggg cccggtcagc       360
     gccaccctgc acgaggacgg cgccgtcacc atcggcaacg tgccggccta ccgctaccgc       420
     cggcaggtgc cggtggaggt gcccggccac ggccgggtgt acggcgacat cgcctggggc       480
     ggcaactggt tcttcctggt atccgagcat ggccagcggc tggcgctgga caacgtcgag       540
     gcgctgaccg acttcacctg ggcgctgctc caggccctgg aggcccaggg catccacggc       600
     gaaggcggcg cgccgatcga ccacgtcgag ctgttcgccg acgacgagga ggccgacagc       660
     cgcaacttcg tcatgtgccc gggcaaggcc tacgaccgtt cgccctgcgg caccggcacc       720
     agcgccaagc tggcgtgcct ggccgccgac ggcaagctgg cccccgggca ggcctgggtg       780
     caggccggca tcaccggcag ccgtttcgaa ggccgctacg aatgggaagg cgagcgcatc       840
     cgcccgttca tcaccggccg cgcctacatg accgccgaca gcaccctgct gatcgacgag       900
     caggacccct tcgcgtgggg catctga                                           927
//
```