Browsing of Biological Databases and Identification of Bioinformatics Tools
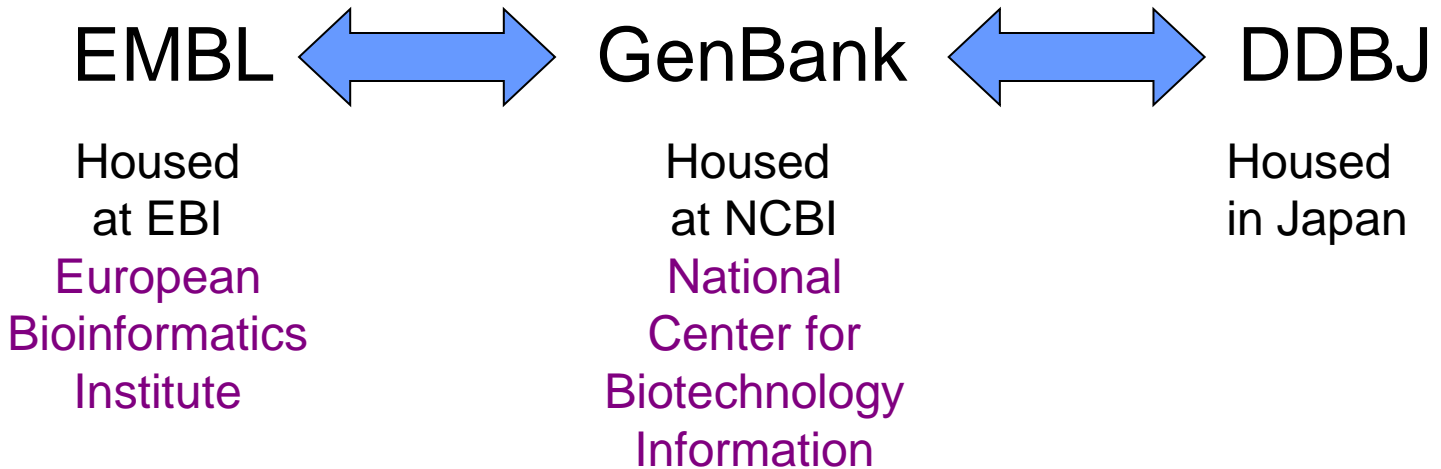
Databases

Primary (archival)

- GenBank /EMBL /DDBJ

- Uniprot

- PDB

- Medline (Pubmed)

Others are: Refseq, Taxon, OMIM etc

There are three major public DNA databases

EMBL ⟷ GenBank ⟷ DDBJ

Housed
at EBI

European
Bioinformatics
Institute

Housed
at NCBI

National
Center for
Biotechnology
Information

Housed
in Japan

>100,000 species are represented in GenBank

all species	128,941
viruses	6,137
bacteria	31,262
archaea	2,100
eukaryota	87,147

How can I retrieve Information Using PubMed at NCBI

PubMed

PubMed is...

- National Library of Medicine's search service
- links to participating online journals
- PubMed tutorial (via "Education" on side bar)

PMC
US National Library of Medicine
National Institutes of Health

PubMed ▼   [   ]   Search

Journal List   Advanced

PMC

PubMed Central® (PMC) is a free full-text archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health's National Library of Medicine (NIH/NLM).

PubReade

A whole new way to read scientific literature at PubMed Central

Get Started

PMC Overview

Users' Guide

Journal List

PMC FAQs

PMC Copyright Notice

Participate

Information for Publishers

How to Participate in PMC

Participation Agreements

File Submission Specifications

File Validation Tools

Keep Up to Date

New in PMC | RSS

PMC Announce Mail List

Utilities Announce Mail List

Tagging Guidelines Mail List

Other Resources

PMC International

Text Mining Collections

Developer Resources

PMC Citation Search

PMC Accessibility

5 MILLION Articles

are archived in PMC.

Content provided in part by:

2161
Full Participation
Journals

333
NIH Portfolio
Journals

4736
Selective Deposit
Journals

Public Access

Funders and PMC

How Papers Get Into PMC

NIH Manuscript Submission System

My Bibliography

PMCID/PMID/NIHMSID Converter

PubMed.gov
US National Library of Medicine
National Institutes of Health

| PubMed ▼ | Helicobacter pylori cagA | ⊗ | **Search** |

Create RSS   Create alert   Advanced

Help

Format: Summary ▾   Sort by: Most Recent ▾   Per page: 20 ▾

Send to ▾

Filters: Manage Filters

Article types
Clinical Trial
Review
Customize ...

Text availability   clear
Abstract
✓ Free full text
Full text

Publication dates   clear
✓ 5 years
10 years
Custom range...

Species
Humans
Other Animals

Clear all

Show additional filters

Best matches for Helicobacter pylori cagA:

Clinical relevance of Helicobacter pylori vacA and cagA genotypes in gastric carcinoma.
Ferreira RM et al. Best Pract Res Clin Gastroenterol. (2014)

Medicinal plant activity on Helicobacter pylori related diseases.
Wang YC et al. World J Gastroenterol. (2014)

Five-year monitoring of considerable changes in tyrosine phosphorylation motifs of the
Helicobacter pylori cagA gene in Iran.
Kargar M et al. J Appl Genet. (2014)

[**Switch to our new best match sort order**]

Search results

Items: 1 to 20 of 487

<< First   < Prev   Page [1] of 25   Next >   Last >>

ⓘ Filters activated: Free full text, published in the last 5 years. Clear all to show 3318 items.

☐ HELICOBACTER PYLORI INFECTION AND IMMUNE PROFILE OF PATIENTS WITH
1. DIFFERENT GASTRODUODENAL DISEASES.
Vinagre RMDF, Vinagre IDF, Vilar-E-Silva A, Fecury AA, Martins LC.
Arq Gastroenterol. 2018 Apr-Jun;55(2):122-127. doi: 10.1590/S0004-2803.201800000-21.
PMID: 30043859   Free Article
Similar articles

☐ Prevalence of CagA and antimicrobial sensitivity of H. pylori isolates of patients with gastric cancer
2. in Egypt.
Al-Eraky DM, Helmy OM, Ragab YM, Abdul-Khalek Z, El-Seidi EA, Ramadan MA.
Infect Agent Cancer. 2018 Jul 16;13:24. doi: 10.1186/s13027-018-0198-1. eCollection 2018.
PMID: 30026792   Free PMC Article
Similar articles

☐ The interplay between mutations in cagA, 23S rRNA, gyrA and drug resistance in Helicobacter
3. pylori.
Vianna JS, Ramis IB, Ramos DF, Gastal OL, Silva RAD, Gonçalves CV, Silva PEAD.
Rev Inst Med Trop Sao Paulo. 2018;60:e25. doi: 10.1590/s1678-9946201860025. Epub 2018 Jun 28.

Sort by:

| Best match | **Most recent** |

Results by year

Download CSV

Titles with your search terms

Diminished circulating concentration of
interleukin-35 in Helicobacte [Helicobacter. 2018]

CagA and vacA allelic combination of
Helicobacter pylori in gas [Microb Pathog. 2018]

Intramolecular energies of the cytotoxic protein
CagA of Helicobacter [Comput Biol Chem. 2018]

See more...

Find related data

Database: Select ▼

[Find items]

Search details

(("helicobacter pylori"[MeSH Terms]
OR ("helicobacter"[All Fields] AND
"pylori"[All Fields]) OR
"helicobacter pylori"[All Fields])

OMIM

OMIM is…

- Online Mendelian Inheritance in Man
- catalog of human genes and genetic disorders
- Information about particular diseases

Books

Books is…

• searchable resource of on-line books

Taxanomy

TaxBrowser is…

• browser for the major divisions of living organisms   (archaea, bacteria, eukaryota, viruses)

• taxonomy information such as nucleotides and Proteins

Structure

Structure site includes…

- Molecular Modelling Database (MMDB)
- biopolymer structures obtained from the Protein Data Bank (PDB)
- 3D-structure viewer

SEQUENCE RETRIEVAL

Sequence is the nucleotides in a gene or the amino acids in a protein in their right order.

How to locate sequences:

Go to databases (I. e NCBI, iProClass, UNIPROT, etc).
Search for protein(s) or nucleotides sequence(s).
Retrieve sequences in FASTA format.

Nucleotide Sequence Retrieval

1   Go to the link for NCBI :
    http://www.ncbi.nlm.nih.gov/

2   In the search window, scroll and select
    nucleotide

3   Lower search window type the name of
    your gene or protein

4   Click search button

Output

- 1: Look for your query gene or protein with complete mRNA cds

- 2: Below the query gene found; look for FASTA. Click the FASTA to retrieve your nucleotide sequences in FASTA format.

- 3: Copy and save your sequence as a text file using notepad to the relevant folder.

Amino acid Sequence Retrieval

1 Go to the link for UNIPROT
   http://www.uniprot.org/

2 In the query window, type your gene
   name

3 Click the search button

4 Wait for your result to appear. Then click
   the gene id number.

• This takes you to the query gene/protein
  information page

BLAST

BLAST is…

- Basic Local Alignment Search Tool
- NCBI's sequence similarity search tool
- supports analysis of DNA and protein databases

BLAST

BLAST (Basic Alignment Search Tool allows rapid sequence comparison of a query sequence against a database. The BLAST algorithm is fast, accurate and Web-accessible .

BLAST

BLAST searching is fundamental to understanding
the relatedness of any favorite query sequence
to other known proteins or DNA sequences.

Applications include
• discovering new genes or proteins
• discovering variants of genes or proteins
• exploring protein structure and function

BLAST

NCBI → Resources → Sequence Analysis → BLAST

BLAST

✓Select the BLAST program

✓ Enter query sequence in FASTA format

✓ Choose optional parameter i. e Highly similar srquences

✓ Click Blast

BLAST

• The score of a pairwise alignment includes positive values for exact matches, and other scores for mismatches and gaps.

• Score reflects degree of similarity

Protein Sequence Analysis

- **Click or Paste the web link to your browser to enter the ProtParam tool.**

- **The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) for protein sequences**

- **http://web.expasy.org/protparam/**

- **your sequence in the Protein query box in FASTA format or SwissProt accession number**

- **Click compute parameters.**

Translation of Protein to nucleotide

- Emboss Transeq ⟶ Sequence Translation

⬇

Launch back transeq (Protein sequence back translation)

⬇
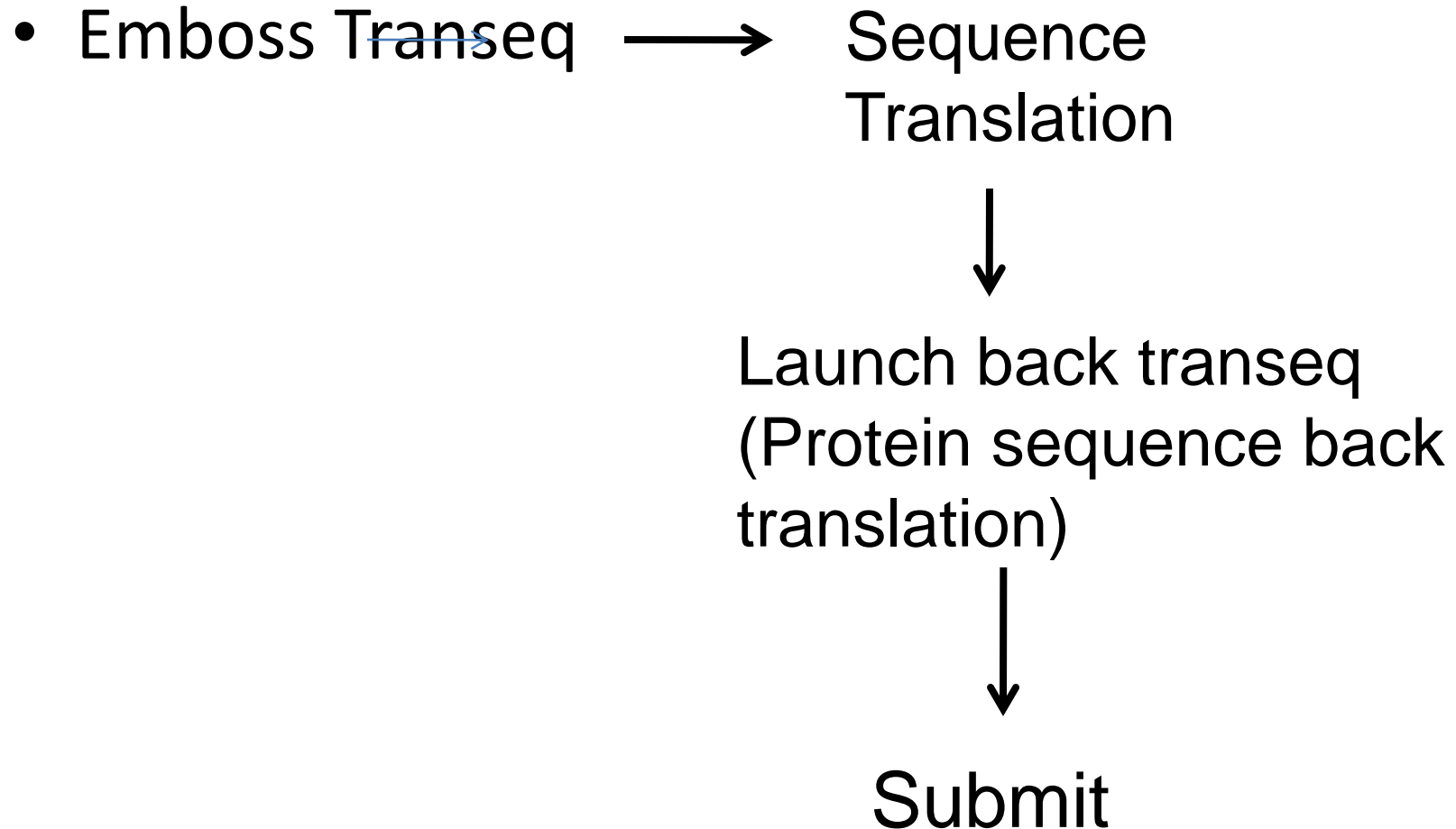
Submit

Translation of nucleotides to proteins

- Emboss Transeq $\longrightarrow$ Enter or paste DNA sequence

Submit

EMBOSS Transeq

EMBOSS Transeq translates nucleic acid sequences to their corresponding peptide sequences. It can translate to the three forward and three reverse frames, and output multiple frame translations at once.

STEP 1 - Enter your input sequence

Enter or paste a [DNA/RNA ▾] sequence in any supported format:

```
GTTTTATTACCAATTGATGGTTCAGAACATTCAAAACGTGCTGTTAATTGGTATTTAACA
GAATTTTGTAAACCAGATGATCATACATATTTTTTACATGTTGTTGAATCACATTATTCA
AAAACAACAGCTATTGAATCACATGATCATGCTAAAGAATTATCATCAAATTTAAATAAA
AATATTAAATCAAATGCTCAATTAGGTAAATTATTAGGTGATAAATTACATGATGATTTA
GAAAAATCACATATTCAAATGGAATATATTATGCAAATTGGTAATAAACCAGGTGAATTA
ATTGTTGATTTAATTAAAAAAATTATCAGTTGATGTTGTTTTAATTGGTAATCGTGGTTTA
GGTGCTTTACGTCGTACATTTTTAGGTTCAGTTTCAGAATATGTTTTACATCATTGTAAT
GTTCCATTTATTATTATTCCACCACCATTATGTTCA
```

Or, upload a file: [Choose file] No file chosen   See example inputs

STEP 2 - Select Parameters

FRAME   CODON TABLE

[1   ▾]   [Standard Code   ▾]

The default settings will fulfill the needs of most users.

[More options...] *(Click here, if you want to view or change the default settings.)*

STEP 3 - Submit your job

☐ Be notified by email *(Tick this box if you want to be notified by email when the results are available)*

[Submit]

Multiple sequence alignment

- A **multiple sequence alignment (MSA)** is a sequence alignment of three or more biological sequences, generally protein, DNA, or RNA.

- In many cases, the input set of query sequences are assumed to have an evolutionary relationship by which they share a lineage and are descended from a common ancestor.

- From the resulting MSA, sequence homology can be inferred and phylogenetic analysis can be conducted to assess the sequences' shared evolutionary origins.

- Multiple sequence alignment is often used to assess sequence conservation of protein domains, tertiary and secondary structures, and even individual amino acids or nucleotides

MSA- Clustal Omega

✓ Go to :

https://www.ebi.ac.uk/Tools/msa/clustalo/

✓ Input DNA Sequences

✓ Submit

Phylogenetic Analysis I

- This is a known as molecular phylogenetics, which is the analysis of hereditary molecular differences, mainly in DNA sequences, to gain information on an organism's evolutionary relationships

- It finds evolutionary ties between organisms

- It provides relationships between an ancestral sequence and its descendants

- The result of a molecular phylogenetic analysis is expressed or can be drawn in a hierarchical diagram called a cladogram or phylogram (phylogenetic tree).

- Molecular phylogenetics is one aspect of **molecular systematics**, a broader term that also includes the use of molecular data in taxonomy and biogeography.

TreeTop - Phylogenetic Tree Prediction

- Go the TreeTop web link at http://www.genebee.msu.su/services/phtree_reduced.html

- In the alignment window, copy and paste your ClustalW save MSA.

- It is preferable to activate Bootstrap.

- Include your email to receive result or you can wait for the analysis to be completed.

- Click submit Query button

Phylogenetic Tree Prediction using Clustal Omega

✓ Go to :
https://www.ebi.ac.uk/Tools/msa/clustalo/

✓ Input DNA Sequences

✓ Submit

✓ Click on Phylogenetic tree