WORKING WITH THE NCBI DATABASE

Sequencing and Bioinformatics Consortium (SBC)

LAND ACKNOWLEDGEMENT

We would like to acknowledge that:

- the University of British Columbia Vancouver Campus is located on the traditional, ancestral, and unceded territory of the xwməθkwəy̓əm (Musqueam) People, and
- the University of British Columbia Okanagan Campus is located on the traditional, ancestral, and unceded territory of the Syilx (Okanagan) Peoples.

HOUSEKEEPING

- We have muted everyone when entering. Please stay on mute unless you are asking a question.
- There will be a Q&A session at the end of the presentation. You can also ask questions using the Chat function, which we can address during the Q&A session.
- This webinar is being recorded.

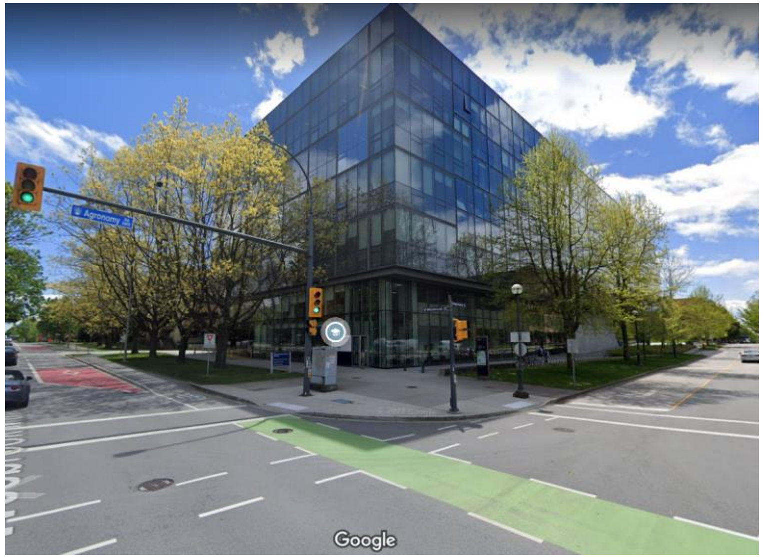
SEQUENCING AND BIOINFORMATICS CONSORTIUM

WHO ARE WE?

- Part of Vice President of Research and Innovation (VPRI)
  - Associate Director: Diane Miller
  - Research Manager: Sunita Sinha
  - Bioinformatics Specialist: Christine Yanta
  - Bioinformatics Team: Lenneth Li
  - Lab team: Shir Yi Toh and Alisha Goodbla

WHERE ARE WE?

- Pharmaceutical Sciences Building, UBC
  - Corner of Wesbrook Mall and Agronomy Road
  - Storefront on 3rd floor
  - Labs and sequencers on 6th floor



SEQUENCING AND BIOINFORMATICS CONSORTIUM

WHAT WE DO:

- Sample QC
    - Input material quality
    - Quantification
    - Purification
- Sequencing
    - Sanger
    - Next generation (Illumina, Pacbio)
    - Genotyping: microsatellite, AFLP, SNP
    - Cell line authentication
- Data analysis / Bioinformatics
- Storefront: Qiagen Biobar
- Customisation
    - Varied input materials
    - Custom protocols
- Troubleshooting
- User training

SEQUENCING AND BIOINFORMATICS CONSORTIUM

REACH OUT TO US:

- Website: https://sequencing.ubc.ca/about-consortium/contact-us
- Email: sequencing.centre@ubc.ca
- Consultation booking: https://sequencing.ubc.ca/virtual-office-hours

INTRODUCTION TO NCBI

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION (NCBI)

https://www.ncbi.nlm.nih.gov/

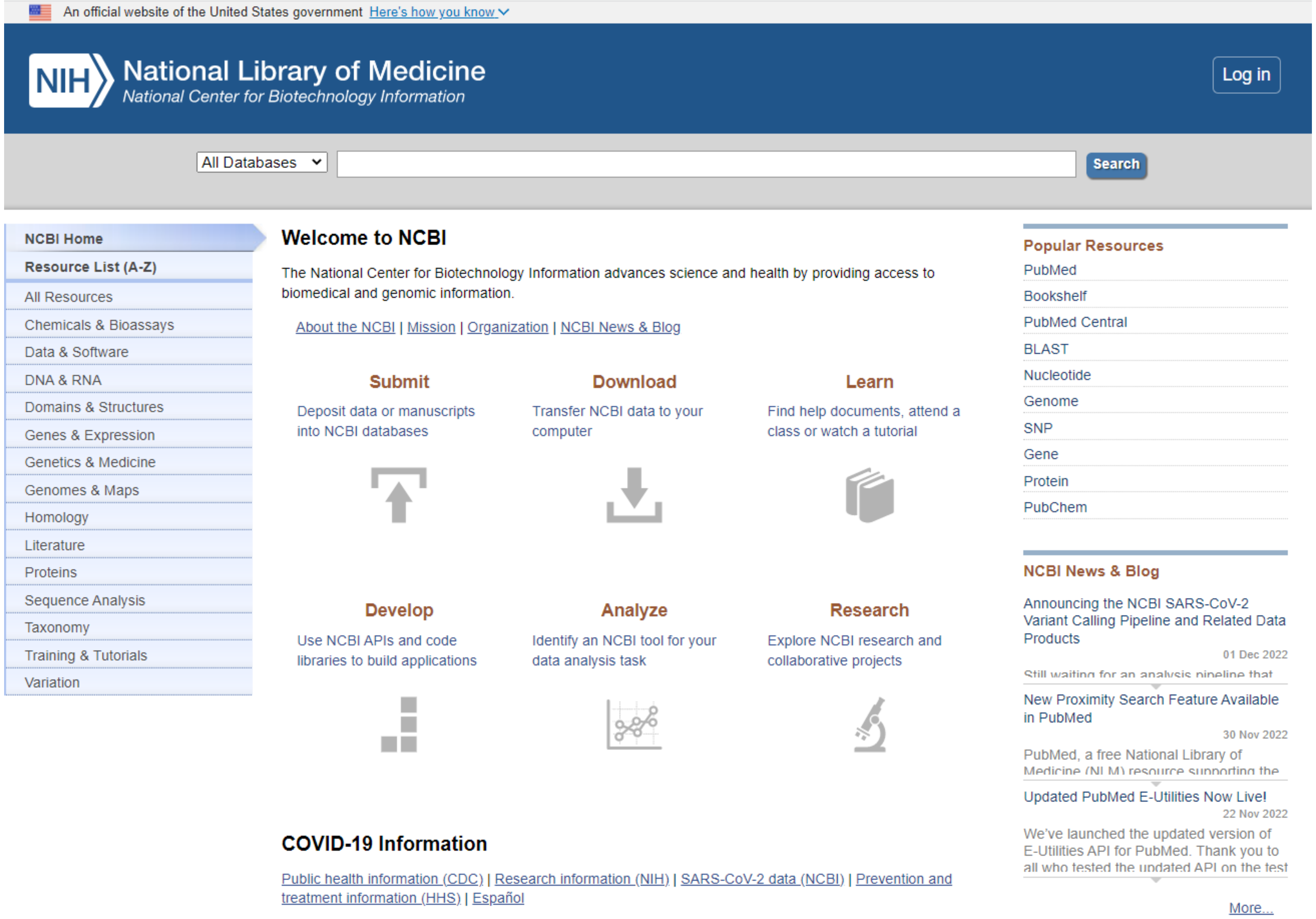
- Developed by the US government at the National Institutes of Health (NIH)

- Established for better informational processing methods for the vast amount of data being generated

- Comprehensive website for biologists:
  - ➢ Biology-related databases
  - ➢ Tools to view and analyze data

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION (NCBI)

https://www.ncbi.nlm.nih.gov/



An official website of the United States government Here's how you know

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

All Databases | Search

- **NCBI Home**
- **Resource List (A-Z)**
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation

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

Develop
Use NCBI APIs and code libraries to build applications

Analyze
Identify an NCBI tool for your data analysis task

Research
Explore NCBI research and collaborative projects

COVID-19 Information

Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI News & Blog

Announcing the NCBI SARS-CoV-2 Variant Calling Pipeline and Related Data Products
01 Dec 2022
Still waiting for an analysis pipeline that

New Proximity Search Feature Available in PubMed
30 Nov 2022
PubMed, a free National Library of Medicine (NLM) resource supporting the

Updated PubMed E-Utilities Now Live!
22 Nov 2022
We've launched the updated version of E-Utilities API for PubMed. Thank you to all who tested the updated API on the test

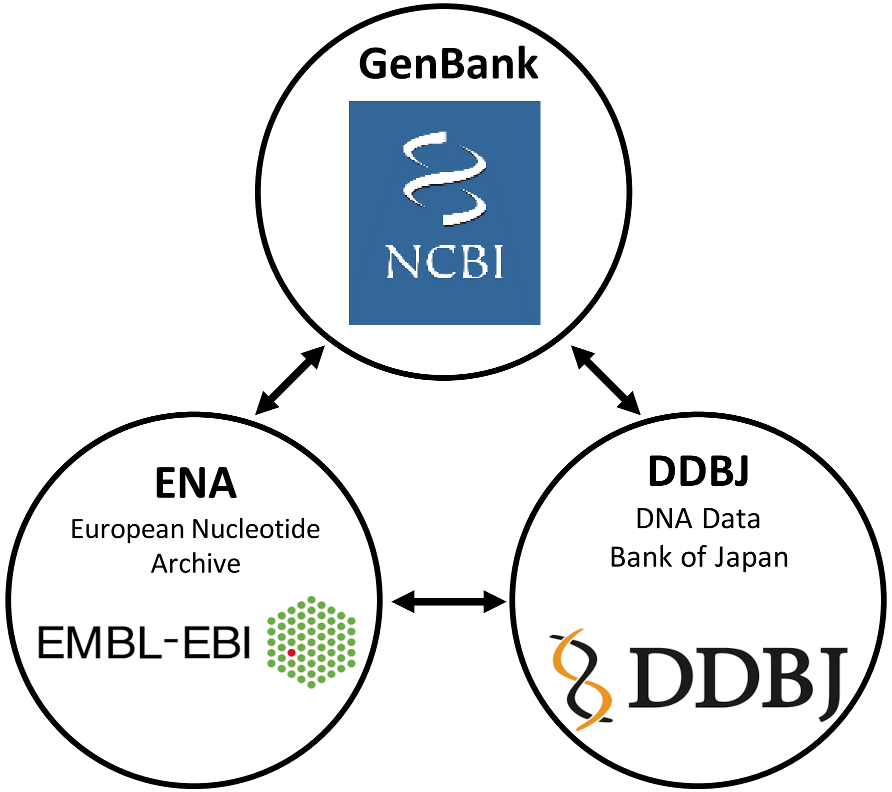
More...

UBC Sequencing and Bioinformatics Consortium

INTERNATIONAL NUCLEOTIDE SEQUENCE DATABASE COLLABORATION (INSDC)

The International Nucleotide Sequence Database Collaboration (INSDC) is a long-standing foundational initiative that operates between DDBJ, EMBL-EBI and NCBI.

INSDC covers the spectrum of data raw reads, through alignments and assemblies to functional annotation, enriched with contextual information relating to samples and experimental configurations.



GenBank

NCBI

ENA

European Nucleotide Archive

EMBL-EBI

DDBJ

DNA Data Bank of Japan

DDBJ

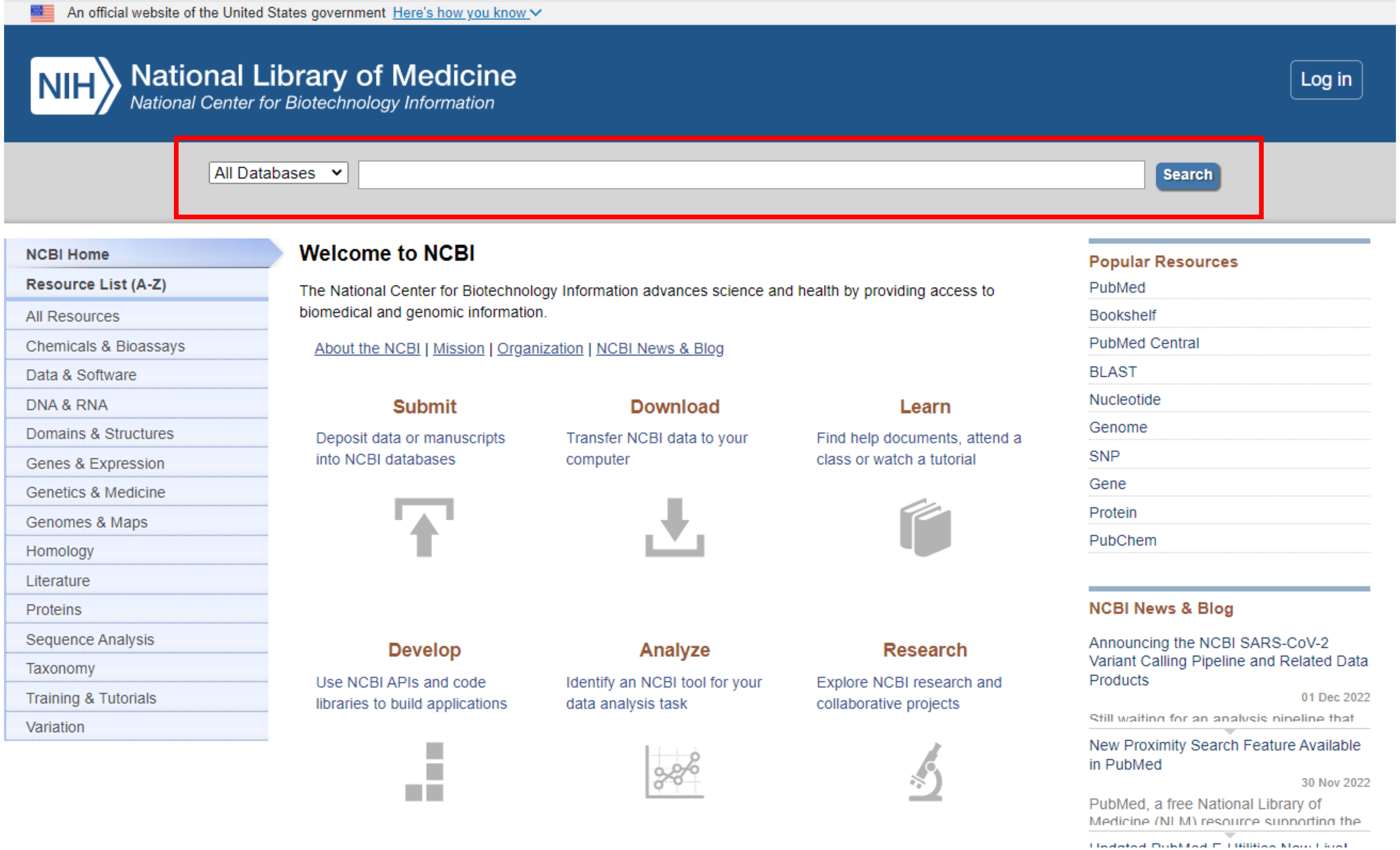
https://www.insdc.org/

NCBI RESOURCES - DATABASES

NCBI maintains over 40 databases, divided into 6 main categories:

1. **Literature**
    - PubMed, PubMed Central, NLM Catalog, Bookshelf, MeSH
2. **Genes**
    - GEO Profiles, Gene, GEO DataSets, PopSet, HomoloGene
3. **Proteins**
    - Protein, Identical Protein Groups, Protein Clusters, Structure, Protein Family Models, Conserved Domains
4. **Genomes**
    - Nucleotide, BioSample, SRA, Taxonomy, Assembly, BioProject, Genome, BioCollections
5. **Clinical**
    - dbSNP, dbVar, ClinVar, ClinicalTrials.gov, MedGen, GTR, dbGaP
6. **Chemicals**
    - PubChem Substance, PubChem Compound, PubChem BioAssay, BioSystems

NCBI RESOURCES



An official website of the United States government Here's how you know

National Library of Medicine
National Center for Biotechnology Information

- NCBI Home
- Resource List (A-Z)
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation

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

Develop
Use NCBI APIs and code libraries to build applications

Analyze
Identify an NCBI tool for your data analysis task

Research
Explore NCBI research and collaborative projects

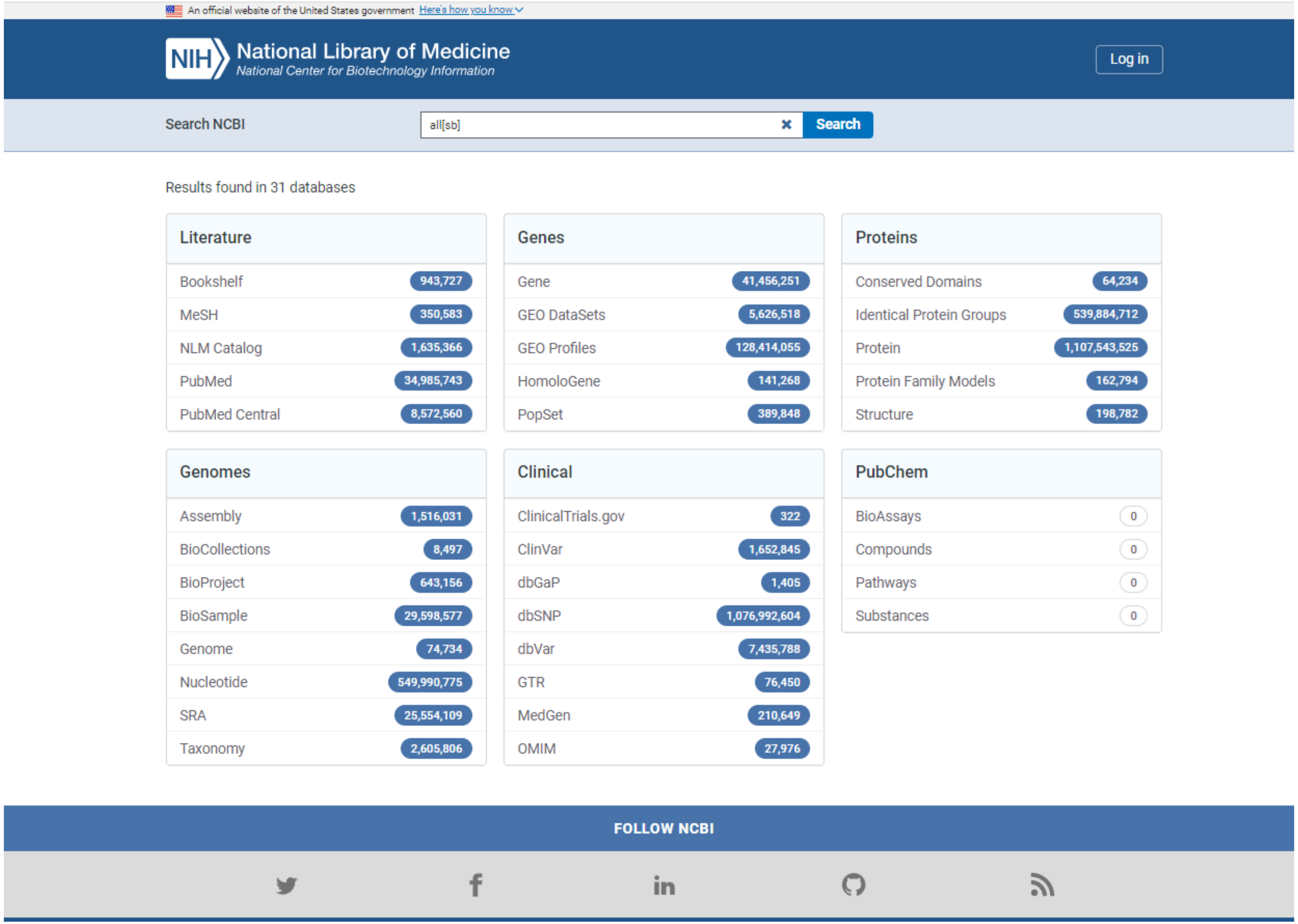
Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI News & Blog

Announcing the NCBI SARS-CoV-2 Variant Calling Pipeline and Related Data Products
01 Dec 2022
Still waiting for an analysis pipeline that

New Proximity Search Feature Available in PubMed
30 Nov 2022
PubMed, a free National Library of Medicine (NLM) resource supporting the



An official website of the United States government Here's how you know

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

Search NCBI all[sb] Search

Results found in 31 databases

Literature	
Bookshelf	943,727
MeSH	350,583
NLM Catalog	1,635,366
PubMed	34,985,743
PubMed Central	8,572,560

Genes	
Gene	41,456,251
GEO DataSets	5,626,518
GEO Profiles	128,414,055
HomoloGene	141,268
PopSet	389,848

Proteins	
Conserved Domains	64,234
Identical Protein Groups	539,884,712
Protein	1,107,543,525
Protein Family Models	162,794
Structure	198,782

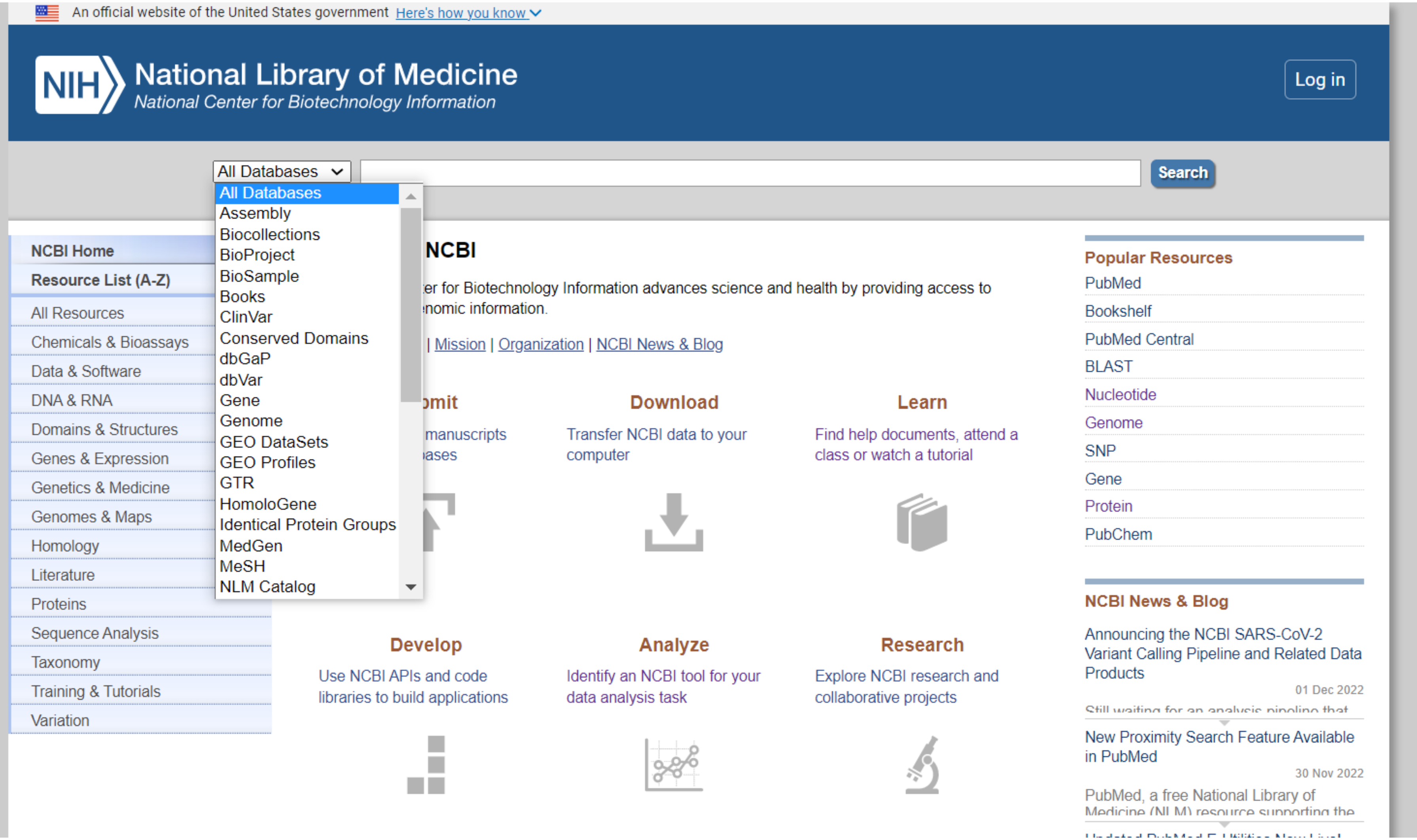
Genomes	
Assembly	1,516,031
BioCollections	8,497
BioProject	643,156
BioSample	29,598,577
Genome	74,734
Nucleotide	549,990,775
SRA	25,554,109
Taxonomy	2,605,806

Clinical	
ClinicalTrials.gov	322
ClinVar	1,652,845
dbGaP	1,405
dbSNP	1,076,992,604
dbVar	7,435,788
GTR	76,450
MedGen	210,649
OMIM	27,976

PubChem	
BioAssays	0
Compounds	0
Pathways	0
Substances	0

FOLLOW NCBI

NCBI RESOURCES



An official website of the United States government Here's how you know

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

All Databases
Search

- All Databases
- Assembly
- Biocollections
- BioProject
- BioSample
- Books
- ClinVar
- Conserved Domains
- dbGaP
- dbVar
- Gene
- Genome
- GEO DataSets
- GEO Profiles
- GTR
- HomoloGene
- Identical Protein Groups
- MedGen
- MeSH
- NLM Catalog

NCBI Home

Resource List (A-Z)

- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation

NCBI

er for Biotechnology Information advances science and health by providing access to nomic information.

| Mission | Organization | NCBI News & Blog

mit

manuscripts ases

Download

Transfer NCBI data to your computer

Learn

Find help documents, attend a class or watch a tutorial

Develop

Use NCBI APIs and code libraries to build applications

Analyze

Identify an NCBI tool for your data analysis task

Research

Explore NCBI research and collaborative projects

Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI News & Blog

Announcing the NCBI SARS-CoV-2 Variant Calling Pipeline and Related Data Products

01 Dec 2022

Still waiting for an analysis pipeline that

New Proximity Search Feature Available in PubMed

30 Nov 2022

PubMed, a free National Library of Medicine (NLM) resource supporting the

NCBI RESOURCES - TOOLS

NCBI hosts a variety of tools to analyze the datasets:

1. **Literature**
   - LinkOut, PubMed Clinical Queries
2. **Genes**
   - BLAST, Digital Differential Display, ORF Finder, Primer-Blast, Sequence Viewer, etc.
3. **Proteins**
   - Amino Acid Explorer, BLAST, CDTree, Cn3D, COBALT, Related Structures, etc.
4. **Genomes**
   - Genome Data Viewer, Assembly Archive, BLAST, e-PCR, Genome Workbench, ProSplign, VecScreen, Viral Genotyping Tool, etc.
5. **Clinical**
   - PheGenI
6. **Chemicals**
   - BioAssay Services, PubChem Standardization Service, PubChem Structure Search

NCBI RESOURCES

- NCBI maintains ~40 databases, divided into 6 main categories
- NCBI hosts a variety of tools to analyze the datasets:

1. **Literature**
2. **Genes**
3. **Proteins**
4. **Genomes**
5. **Clinical**
6. **Chemicals**

DATABASES

NCBI DATABASES

Database

- organized collection of structured information

Two types of databases on NCBI

1. Archival
    - ➢ Redundant (may have many sequence records per gene/protein/genome)
    - ➢ Submitter maintains record
        - ➢ No controlled vocabulary, variation in annotation of biological features

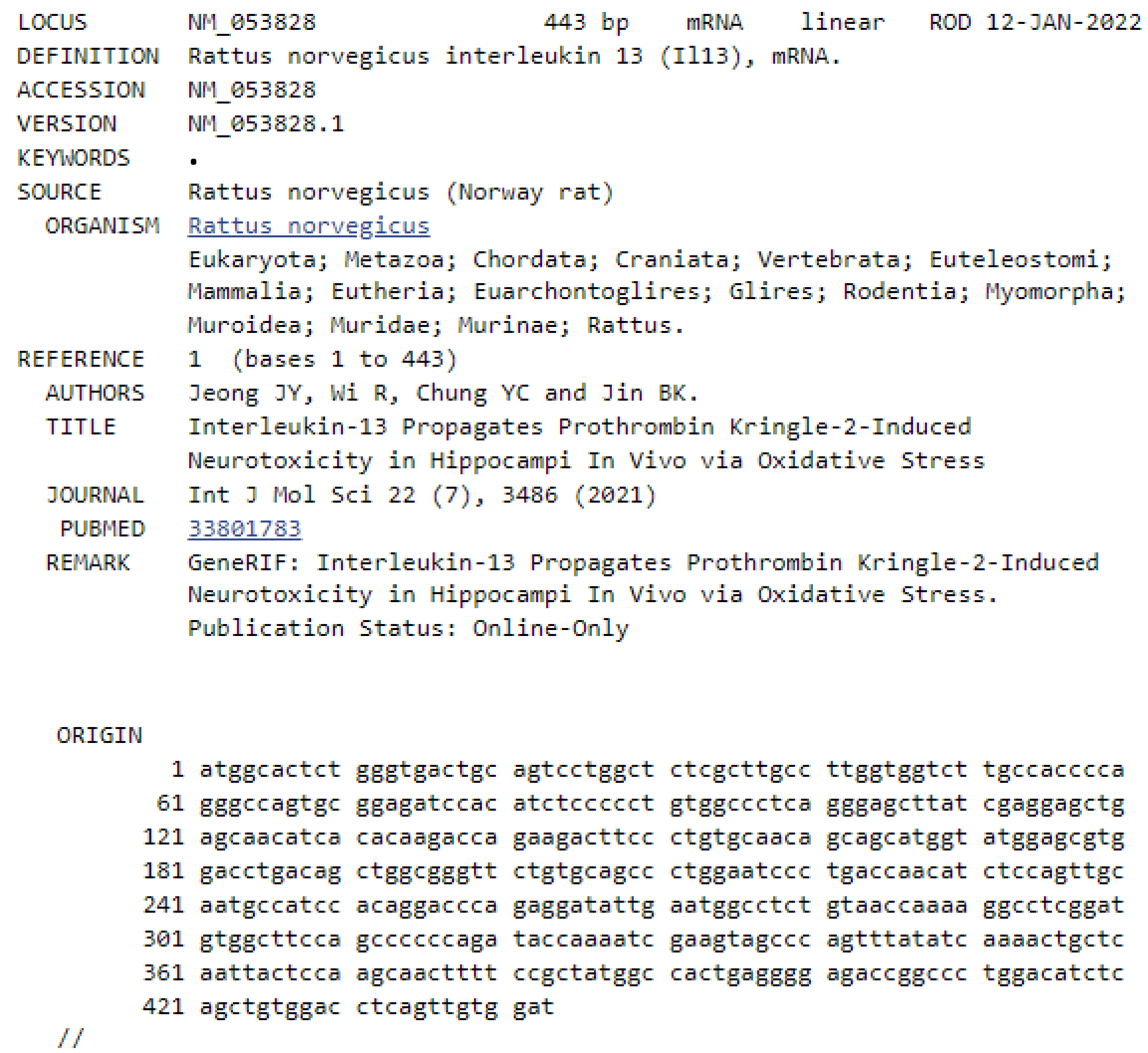
NCBI'S ARCHIVAL DATABASE - GENBANK

https://www.ncbi.nlm.nih.gov/genbank/

GenBank

- Genetic sequence database, provide access to most-up-to-date and comprehensive annotated collection of all publicly available DNA sequences
- No restrictions on the user

GENBANK FLAT FILE FORMAT



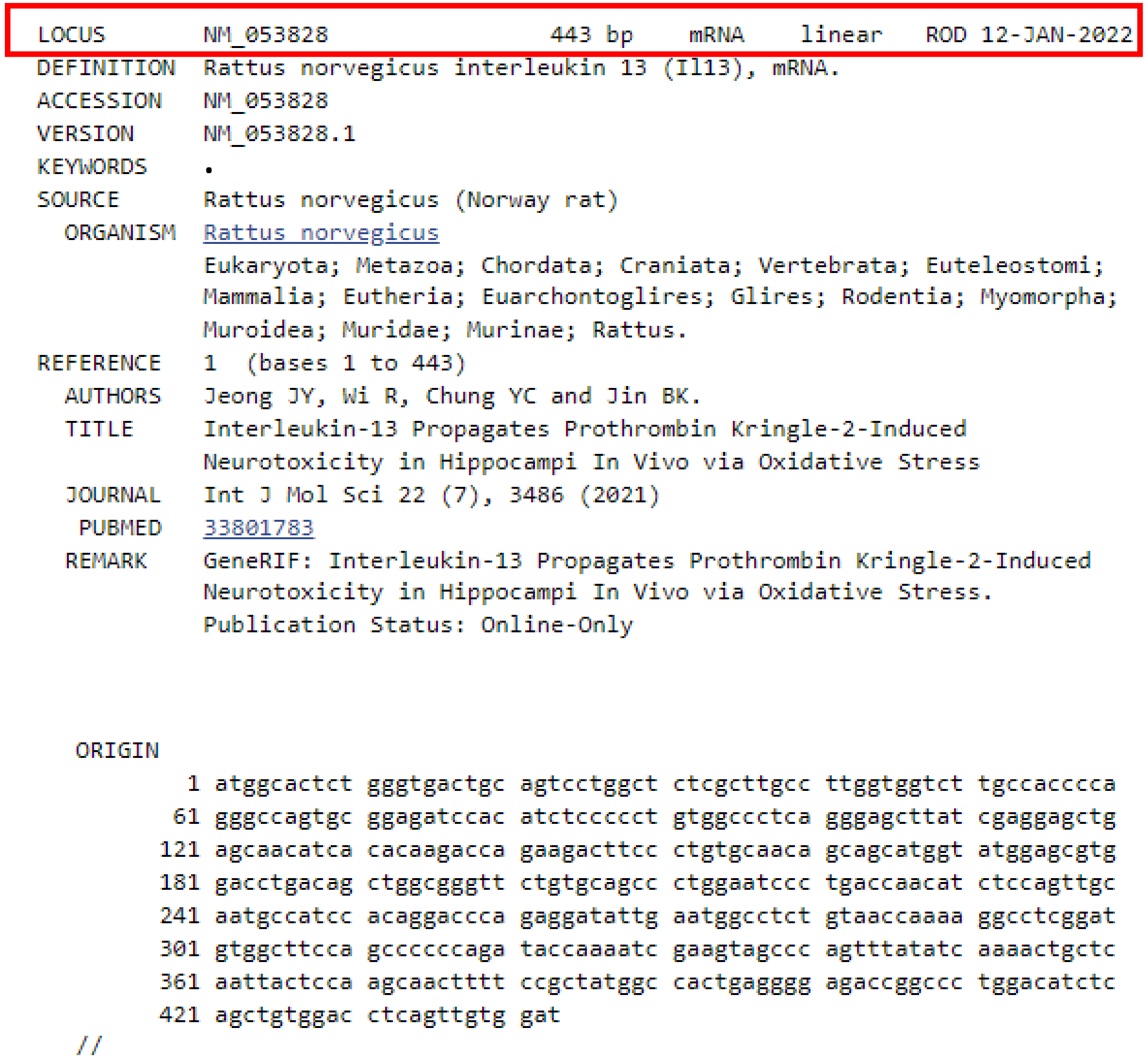
```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctcccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

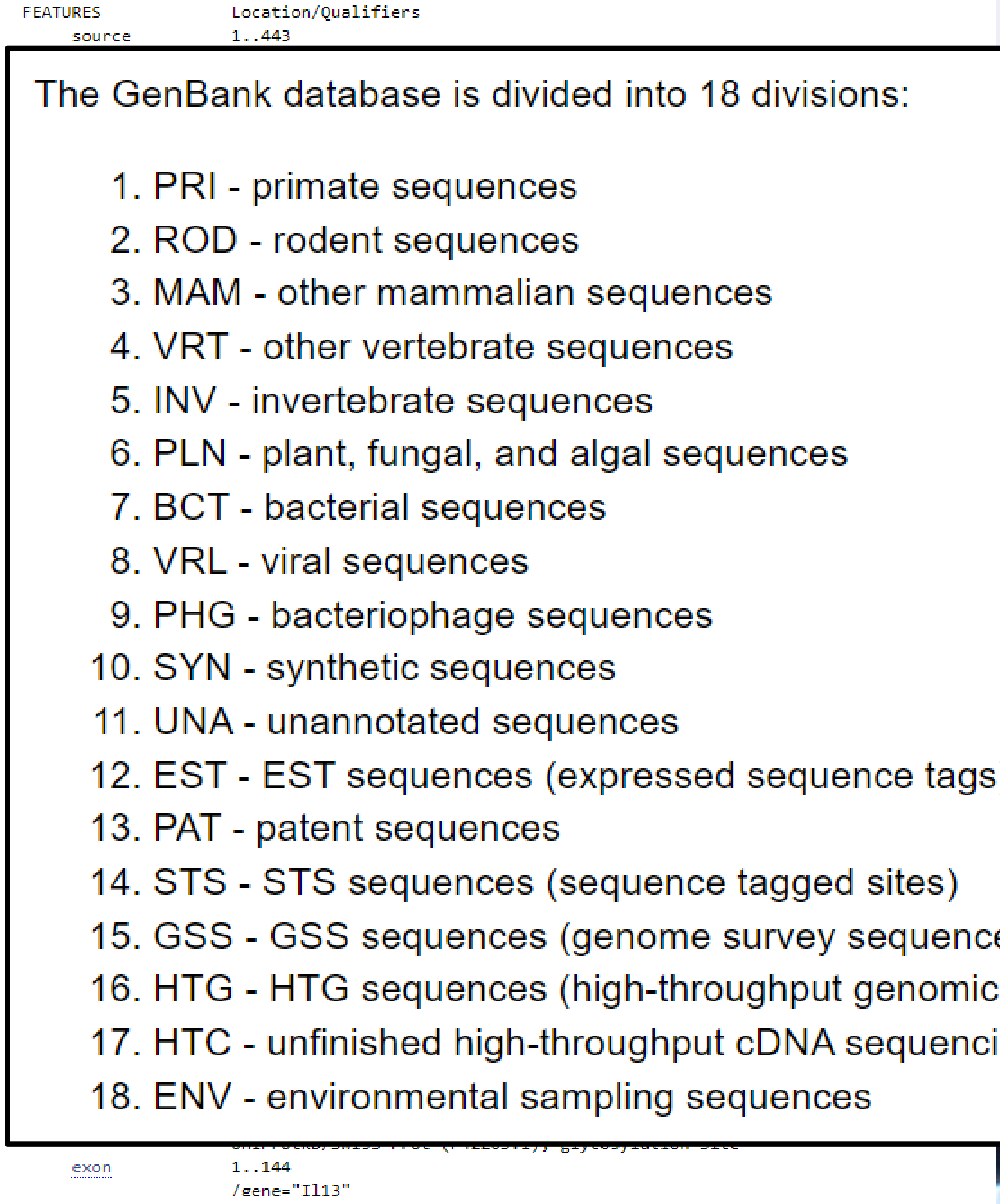
https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```



```
FEATURES             Location/Qualifiers
     source          1..443
```

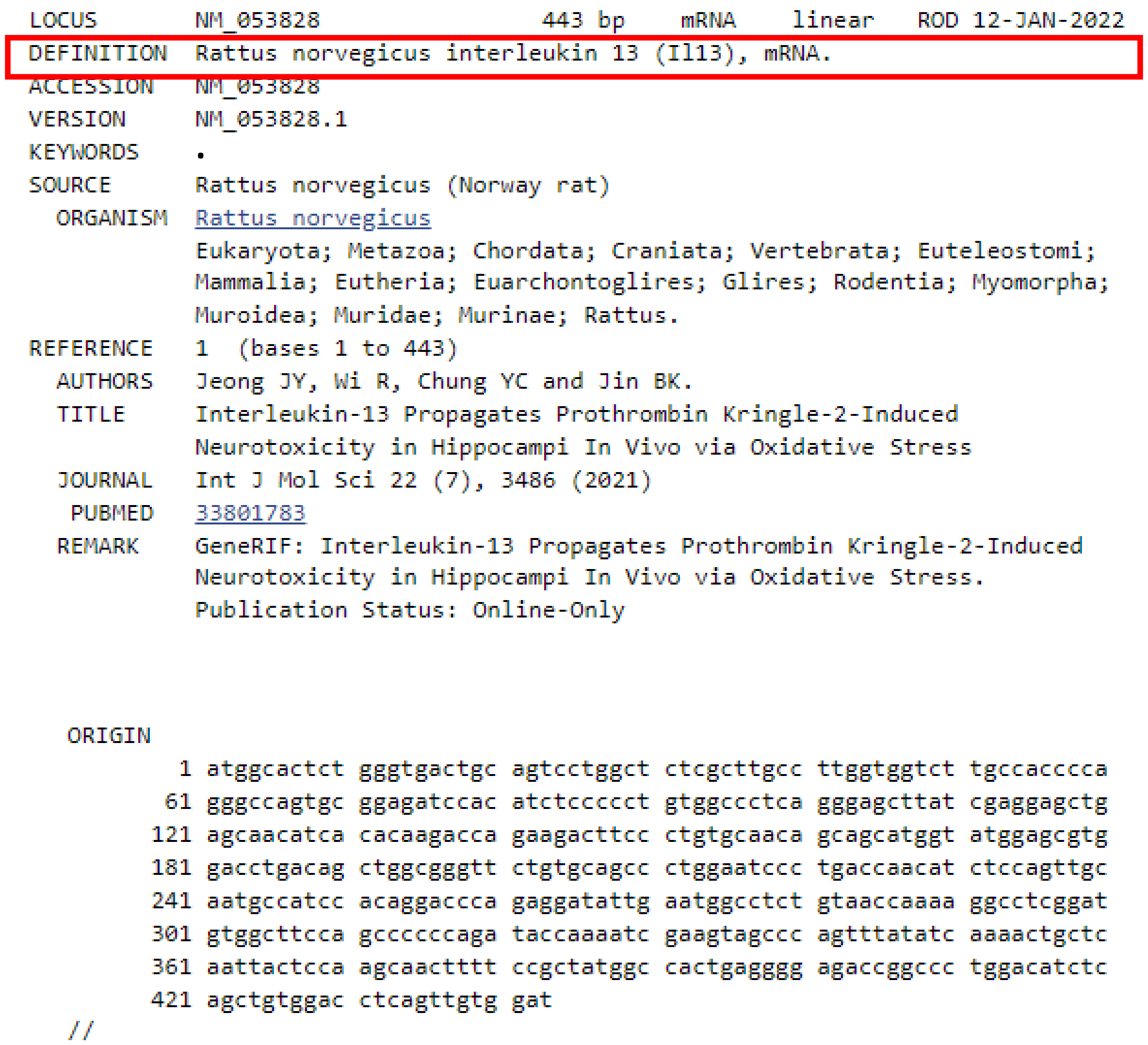
The GenBank database is divided into 18 divisions:

1. PRI - primate sequences
2. ROD - rodent sequences
3. MAM - other mammalian sequences
4. VRT - other vertebrate sequences
5. INV - invertebrate sequences
6. PLN - plant, fungal, and algal sequences
7. BCT - bacterial sequences
8. VRL - viral sequences
9. PHG - bacteriophage sequences
10. SYN - synthetic sequences
11. UNA - unannotated sequences
12. EST - EST sequences (expressed sequence tags)
13. PAT - patent sequences
14. STS - STS sequences (sequence tagged sites)
15. GSS - GSS sequences (genome survey sequences)
16. HTG - HTG sequences (high-throughput genomic sequences)
17. HTC - unfinished high-throughput cDNA sequencing
18. ENV - environmental sampling sequences

```
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



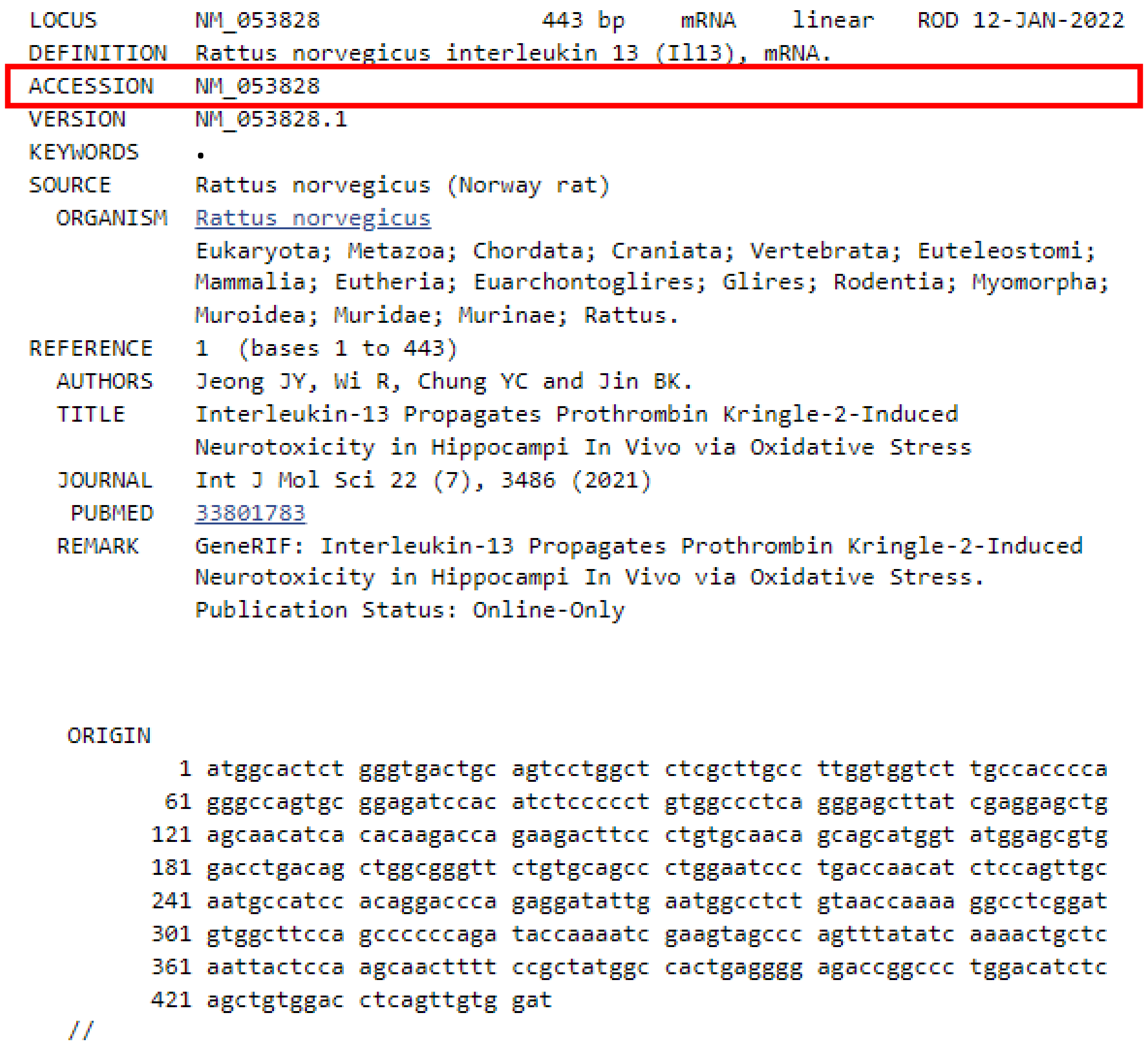
```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctcccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



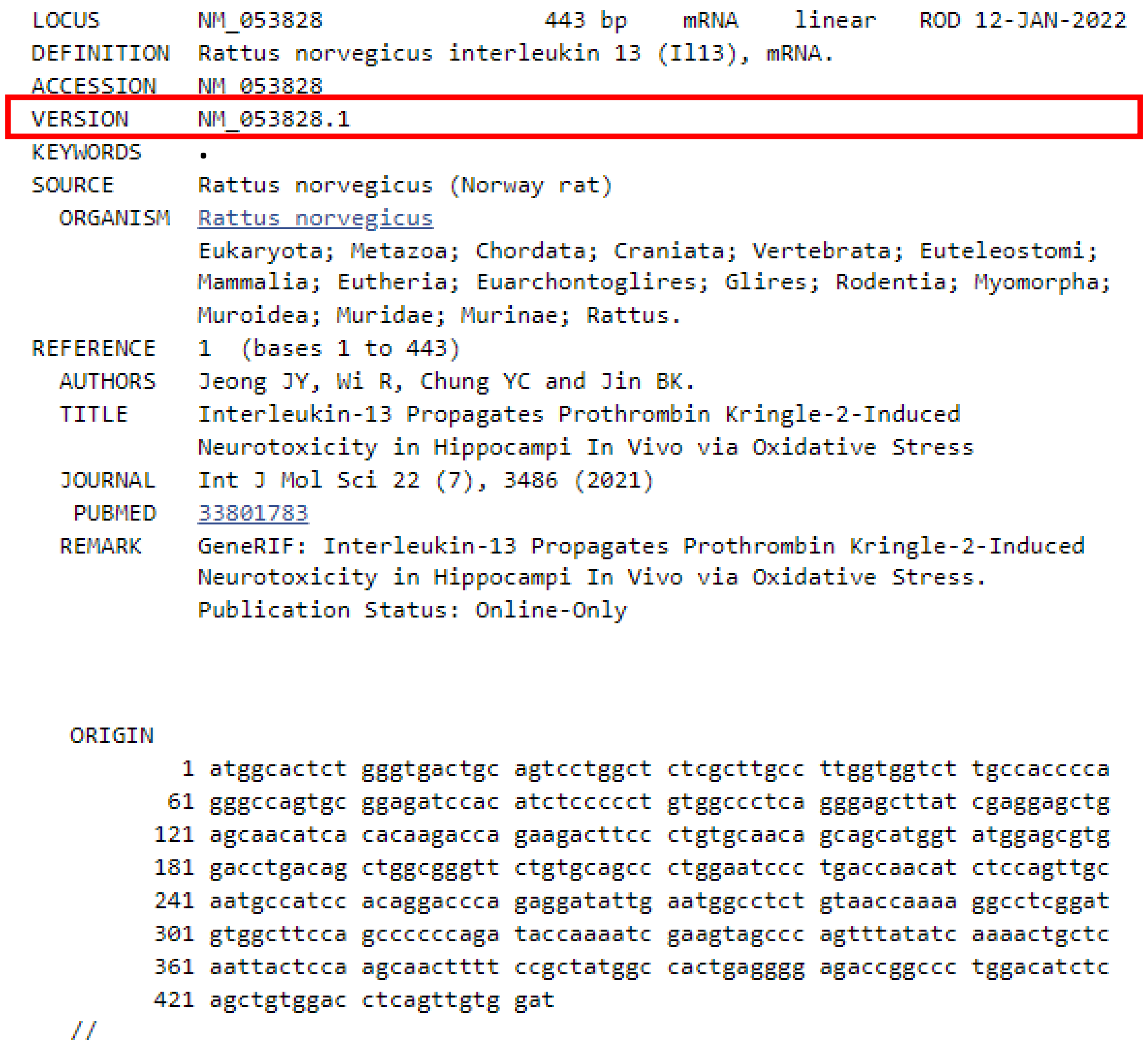
```
LOCUS       NM_053828               443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctcccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gcccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



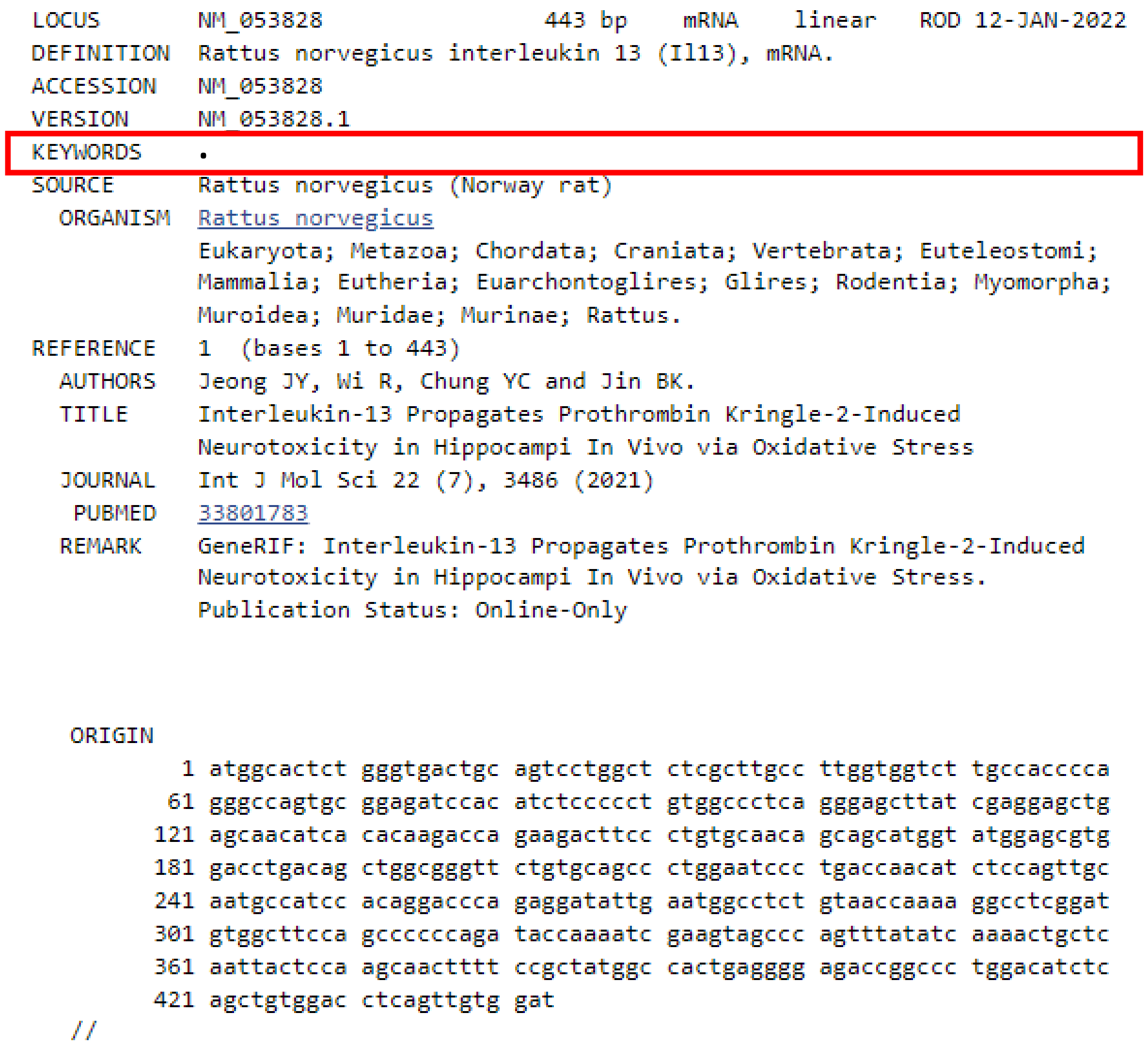
```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gcccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#Definition

GENBANK FLAT FILE FORMAT



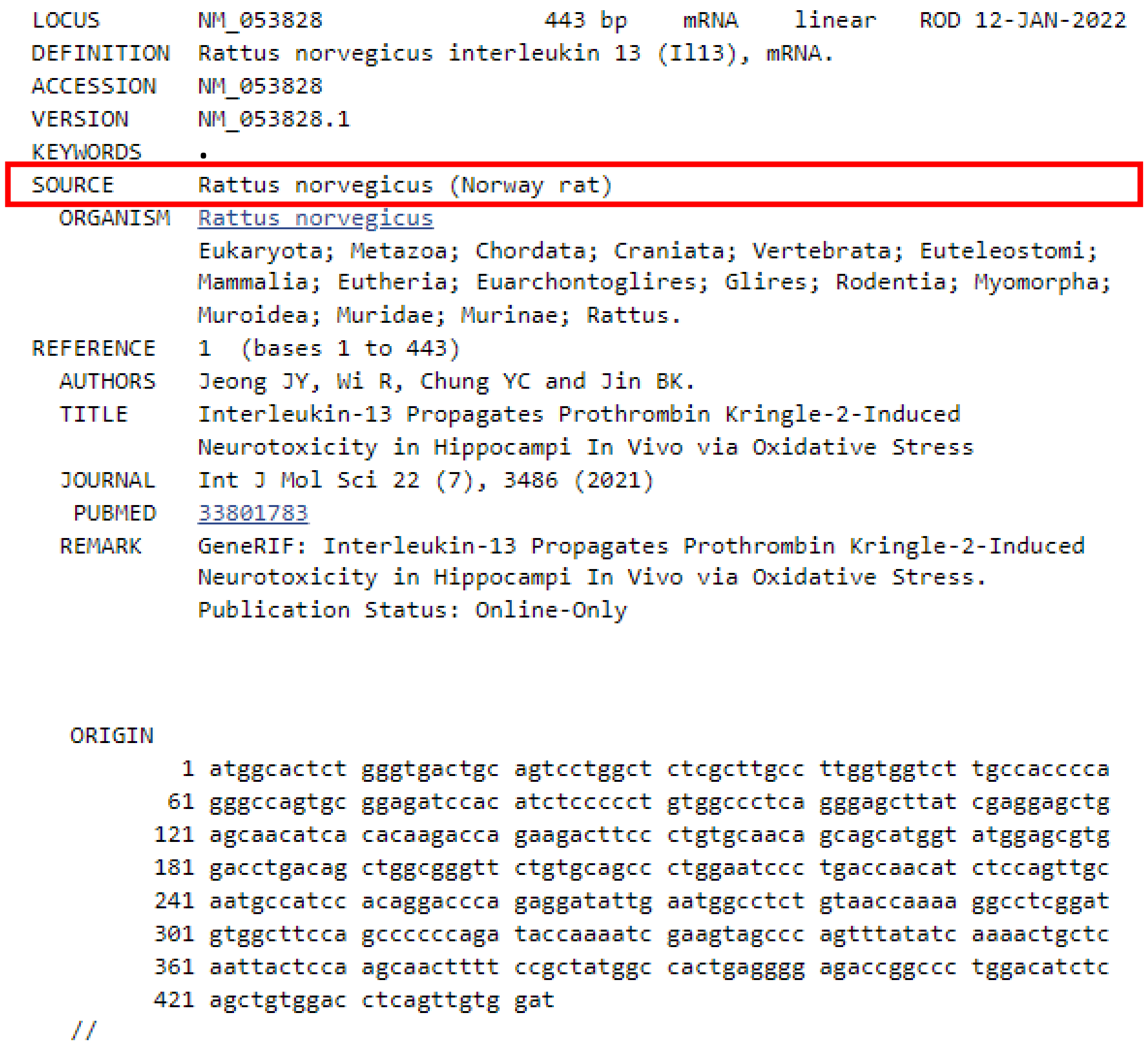
```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctcccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gcccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



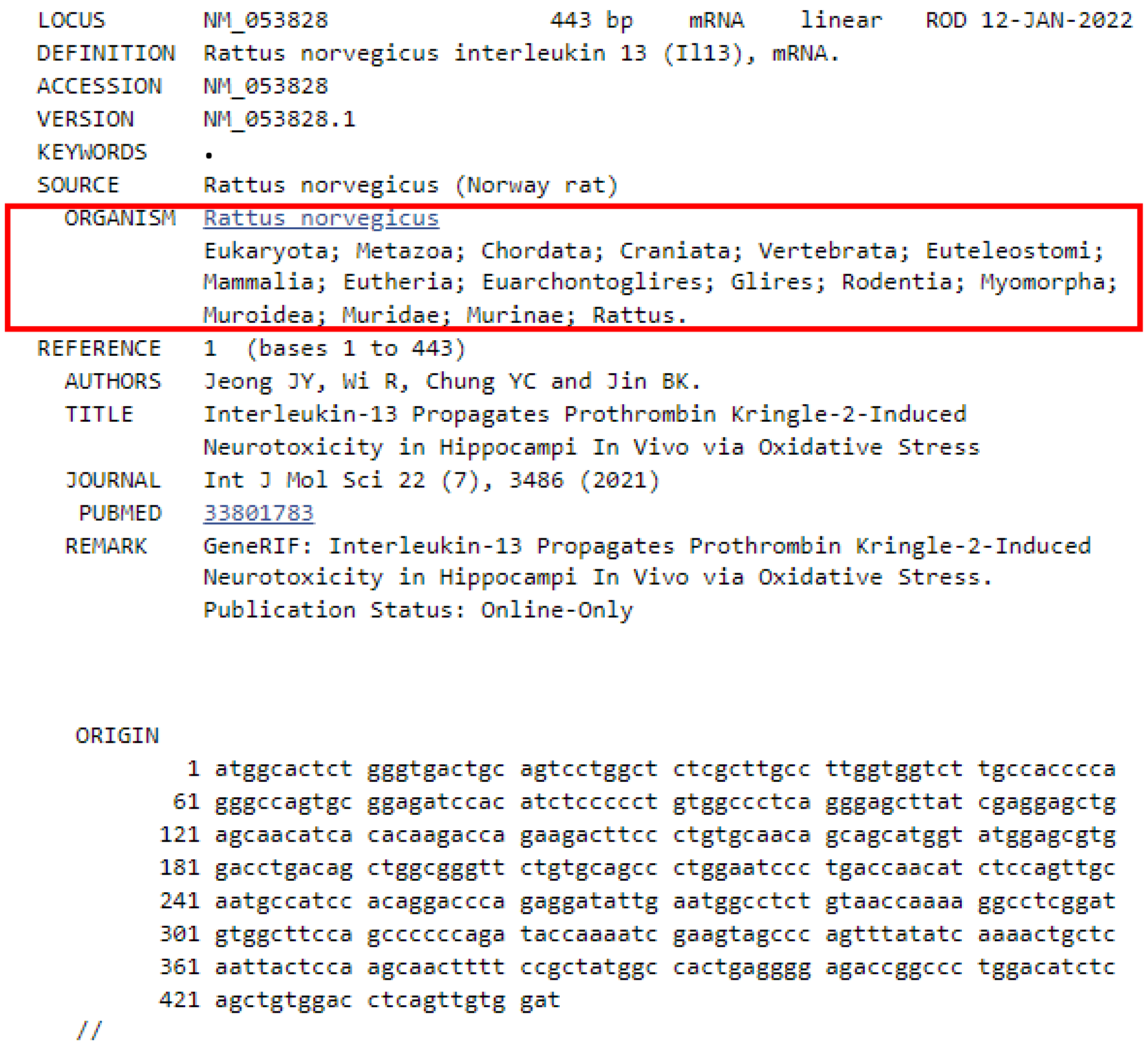
```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gcccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



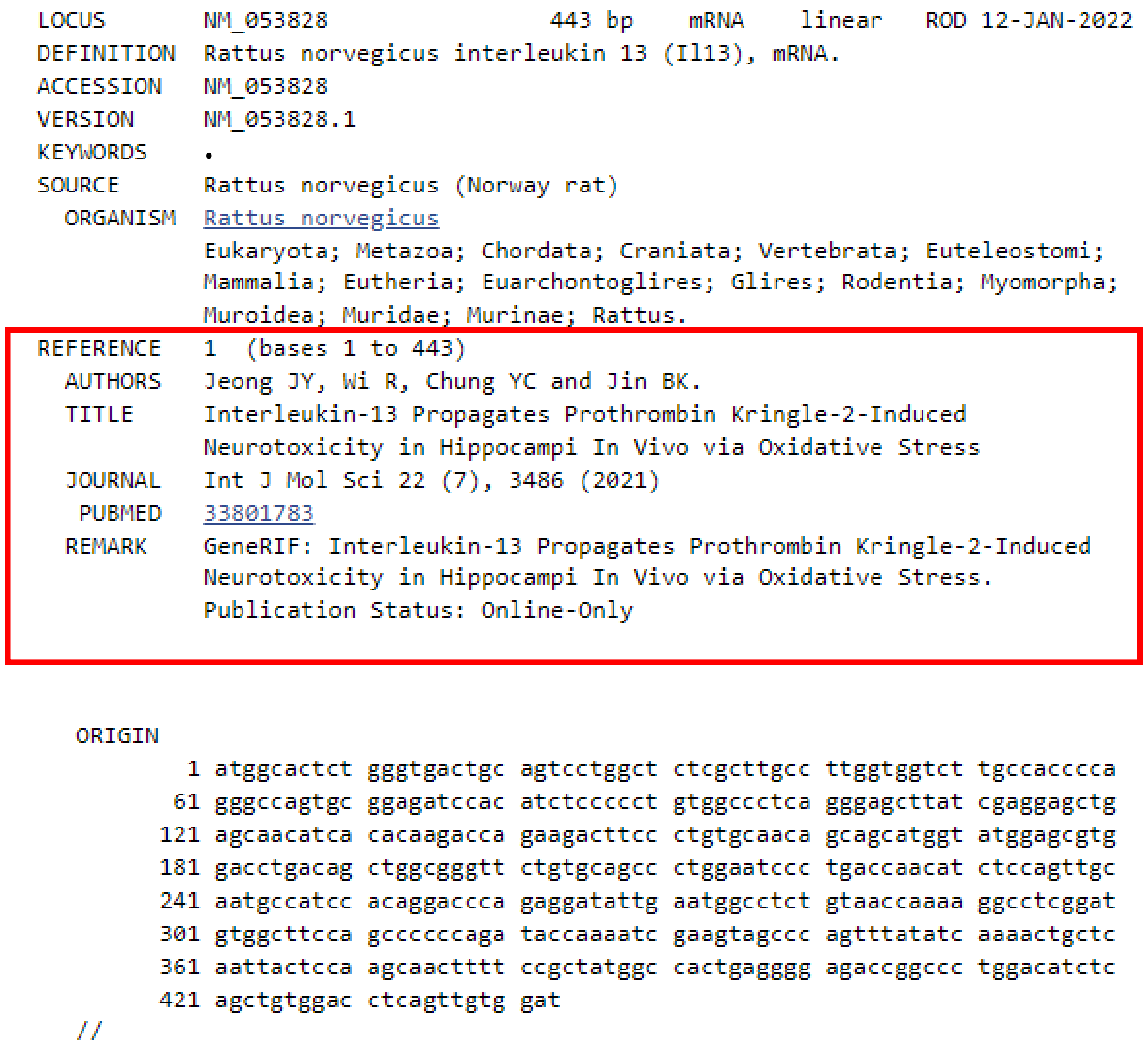
```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



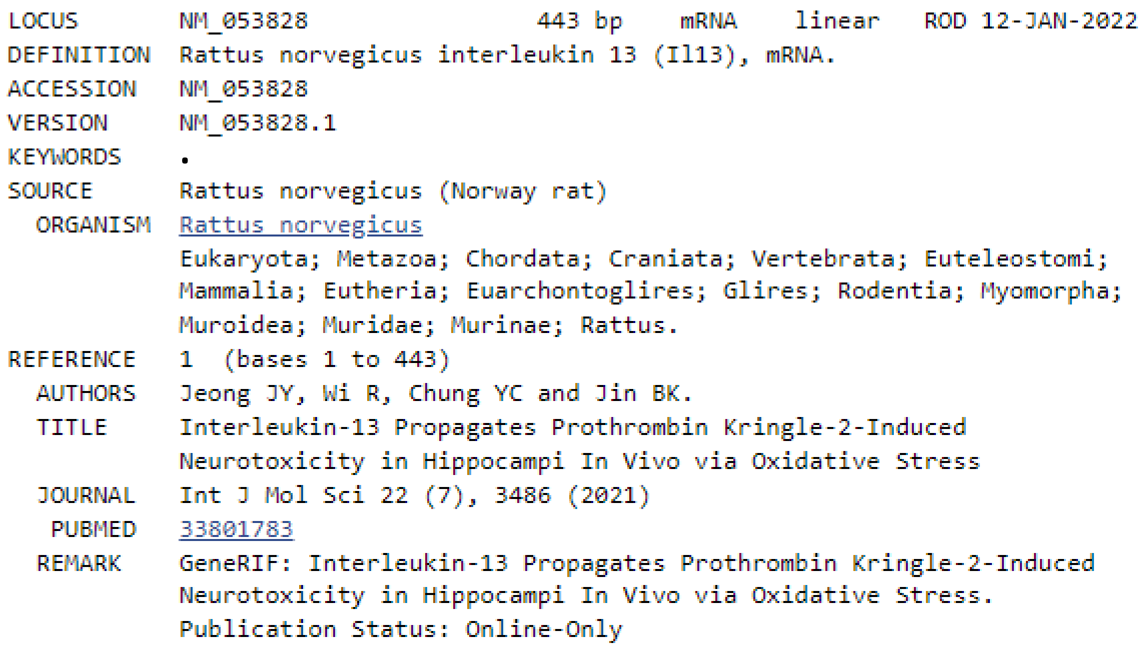
```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only

ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

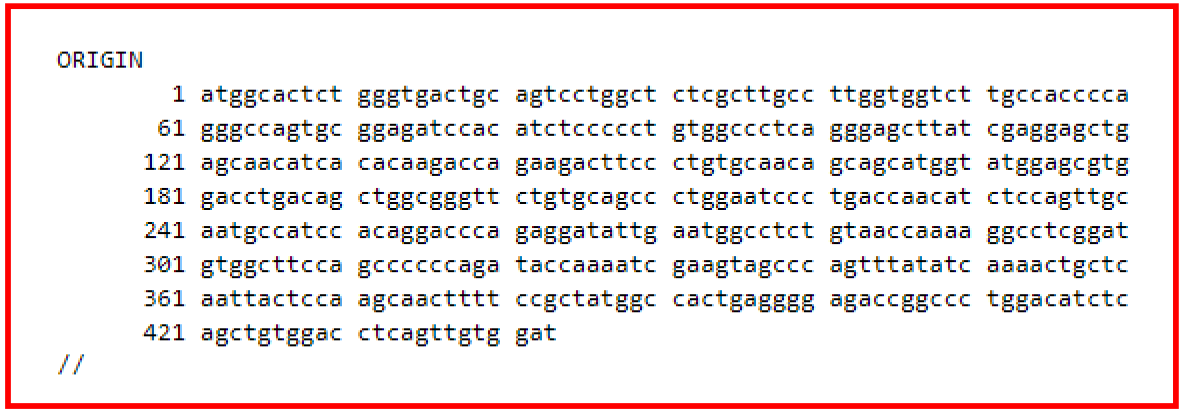
```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only
```

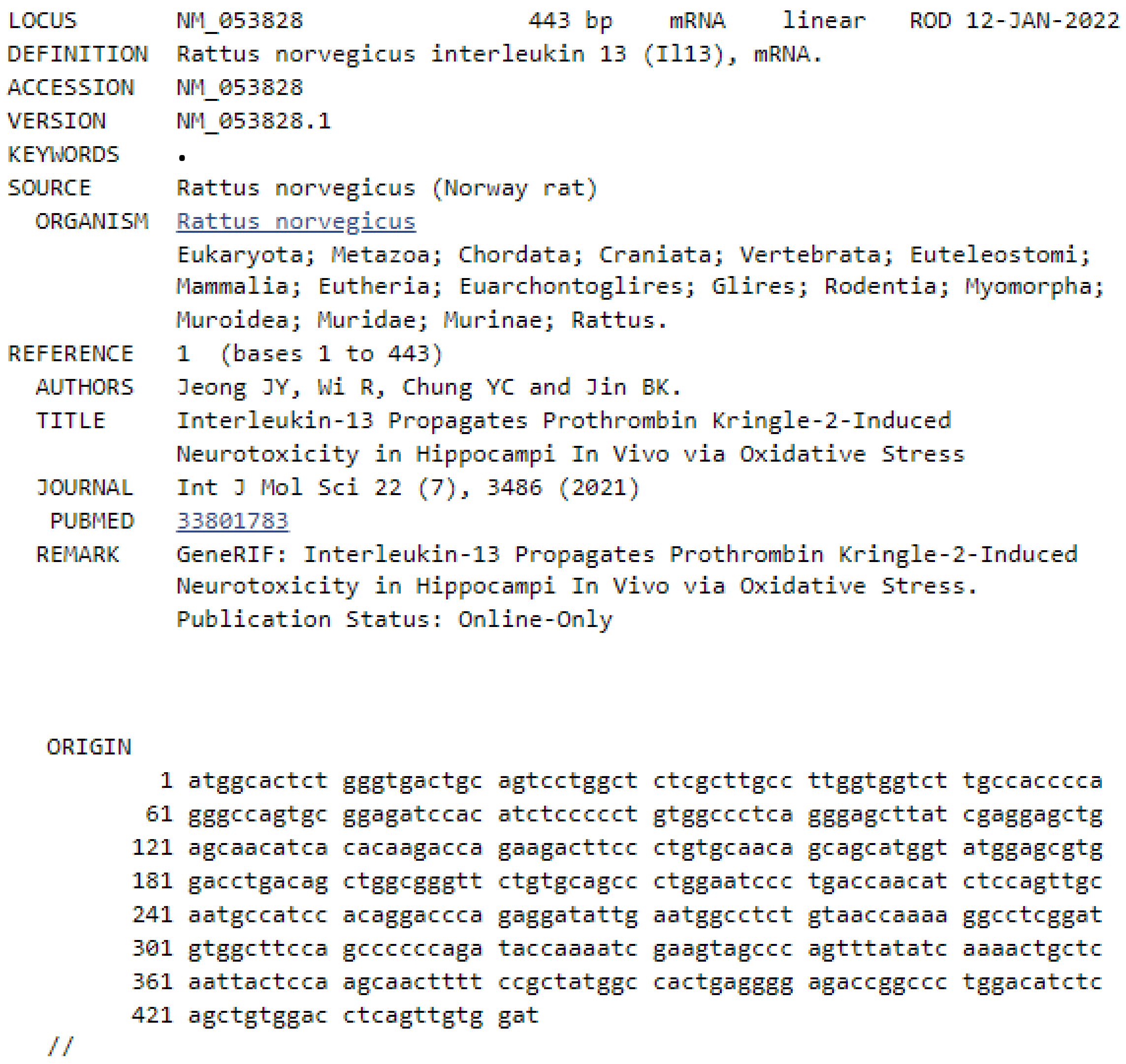


```
ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
//
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

GENBANK FLAT FILE FORMAT



```
LOCUS       NM_053828                443 bp    mRNA    linear   ROD 12-JAN-2022
DEFINITION  Rattus norvegicus interleukin 13 (Il13), mRNA.
ACCESSION   NM_053828
VERSION     NM_053828.1
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha;
            Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Jeong JY, Wi R, Chung YC and Jin BK.
  TITLE     Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress
  JOURNAL   Int J Mol Sci 22 (7), 3486 (2021)
   PUBMED   33801783
  REMARK    GeneRIF: Interleukin-13 Propagates Prothrombin Kringle-2-Induced
            Neurotoxicity in Hippocampi In Vivo via Oxidative Stress.
            Publication Status: Online-Only


   ORIGIN
        1 atggcactct gggtgactgc agtcctggct ctcgcttgcc ttggtggtct tgccacccca
       61 gggccagtgc ggagatccac atctccccct gtggccctca gggagcttat cgaggagctg
      121 agcaacatca cacaagacca gaagacttcc ctgtgcaaca gcagcatggt atggagcgtg
      181 gacctgacag ctggcgggtt ctgtgcagcc ctggaatccc tgaccaacat ctccagttgc
      241 aatgccatcc acaggaccca gaggatattg aatggcctct gtaaccaaaa ggcctcggat
      301 gtggcttcca gcccccccaga taccaaaatc gaagtagccc agtttatatc aaaactgctc
      361 aattactcca agcaactttt ccgctatggc cactgagggg agaccggccc tggacatctc
      421 agctgtggac ctcagttgtg gat
   //
```

```
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q22"
     gene            1..443
                     /gene="Il13"
                     /note="interleukin 13"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
     CDS             1..396
                     /gene="Il13"
                     /note="IL-13; T-cell activation protein P600"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="NP_446280.1"
                     /db_xref="GeneID:116553"
                     /db_xref="RGD:68949"
                     /translation="MALWVTAVLALACLGGLATPGPVRRSTSPPVALRELIEELSNIT
                     QDQKTSLCNSSMVWSVDLTAGGFCAALESLTNISSCNAIHRTQRILNGLCNQKASDVA
                     SSPPDTKIEVAQFISKLLNYSKQLFRYGH"
     sig_peptide     1..21
                     /gene="Il13"
     mat_peptide     22..393
                     /gene="Il13"
                     /product="interleukin-13"
     misc_feature    124..126
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    157..159
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    226..228
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     misc_feature    361..363
                     /gene="Il13"
                     /note="N-linked (GlcNAc...) asparagine.
                     /evidence=ECO:0000255; propagated from
                     UniProtKB/Swiss-Prot (P42203.1); glycosylation site"
     exon            1..144
                     /gene="Il13"
```

https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#DefinitionB

NCBI DATABASES

Database

- organized collection of structured information

Two types of databases on NCBI

1. Archival
    - ➢ Redundant (may have many sequence records per gene/protein/genome)
    - ➢ Submitter maintains record
        - ➢ No controlled vocabulary, variation in annotation of biological features
2. **Curated**
    - ➢ **Non-redundant (only one record for each gene/protein/genome)**
        - ➢ **Encapsulates current understanding**
    - ➢ **Maintained by expert(s)**

REFSEQ

Uses of Refseq

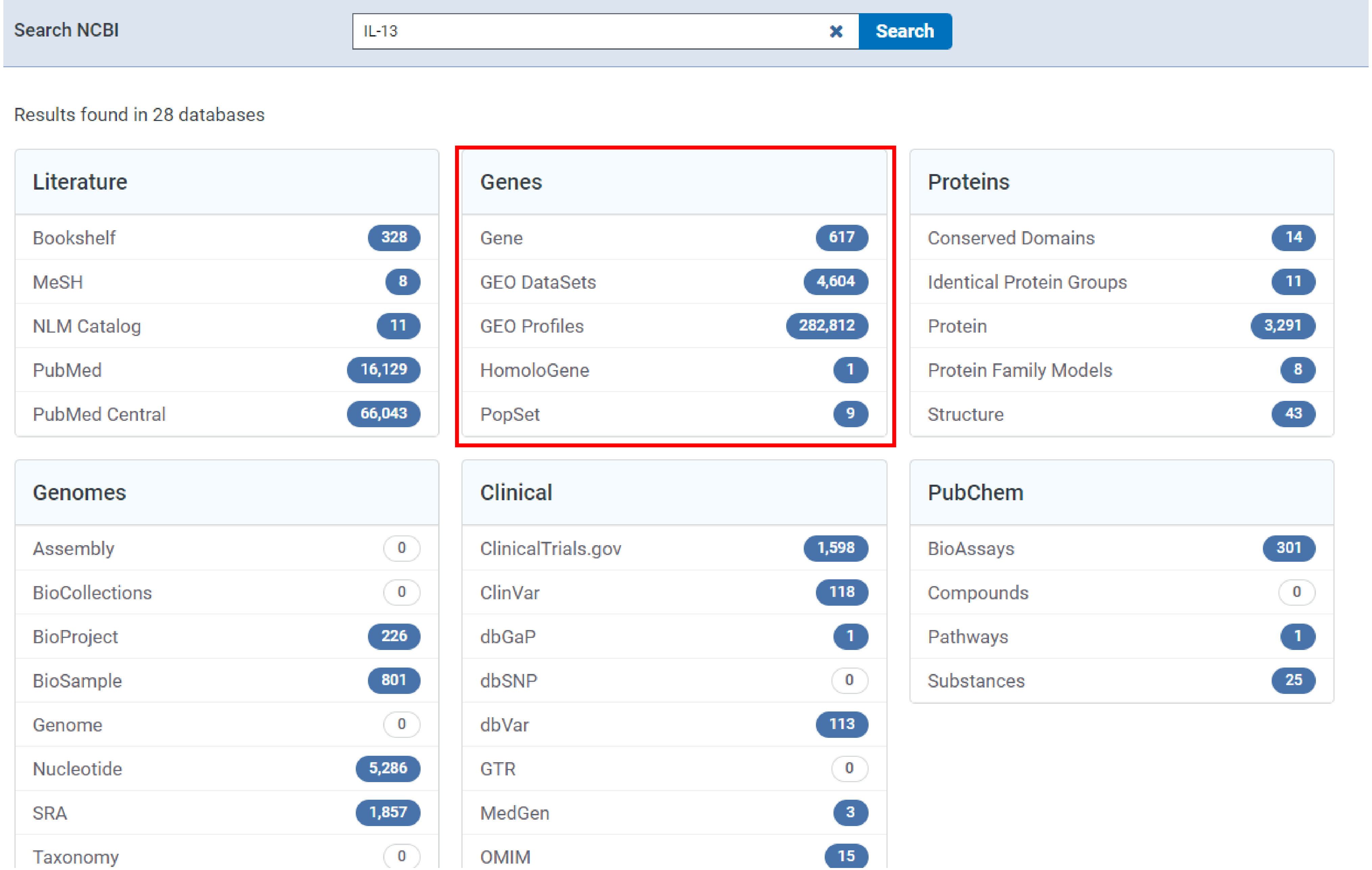
- ➢ Serves as the basis for medical, functional, and diversity studies
- ➢ It provides a stable reference for gene identification and characterization.
- ➢ Mutation analysis, expression studies, polymorphism discovery, and comparative analyses.

- ➢ RefSeq Accession numbers:
  - Genomic DNA
    - NC_123456 - complete genome, complete chromosome, complete plasmid
    - NG_123456 - genomic region
    - NT_123456 - genomic contig
  - mRNA - NM_123456
  - Protein - NP_123456
  - Gene and protein models from genome annotation projects:
    - XM_123456 - mRNA
    - XR_123456 - RNA (non-coding transcripts)
    - XP_123456 - protein

https://www.ncbi.nlm.nih.gov/refseq/

NCBI DATABASES

1. **Gene:**
   - **Gene sequences and annotations used as references for the study of orthologs structure, expression, and evolution**
2. Protein:
   - Protein sequences, 3-D structures, and tools for the study of functional protein domains and active sites
3. Genome:
   - Genome sequence assemblies, large-scale functional genomics data, and source biological samples
4. Clinical:
   - Heritable DNA variations, associations with human pathologies, and clinical diagnostics and treatments



Search NCBI: IL-13

Results found in 28 databases

Literature	
Bookshelf	328
MeSH	8
NLM Catalog	11
PubMed	16,129
PubMed Central	66,043

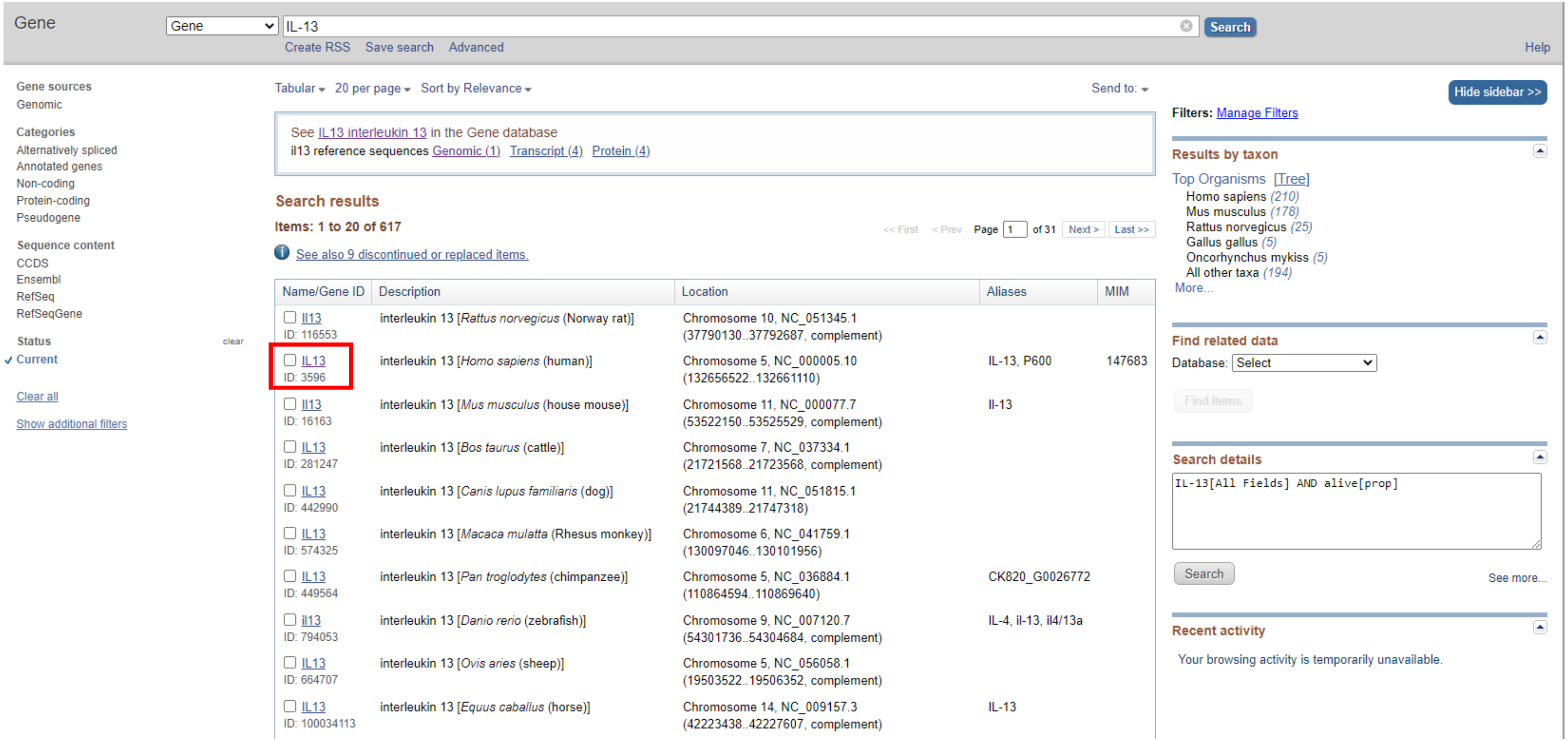
Genes	
Gene	617
GEO DataSets	4,604
GEO Profiles	282,812
HomoloGene	1
PopSet	9

Proteins	
Conserved Domains	14
Identical Protein Groups	11
Protein	3,291
Protein Family Models	8
Structure	43

Genomes	
Assembly	0
BioCollections	0
BioProject	226
BioSample	801
Genome	0
Nucleotide	5,286
SRA	1,857
Taxonomy	0

Clinical	
ClinicalTrials.gov	1,598
ClinVar	118
dbGaP	1
dbSNP	0
dbVar	113
GTR	0
MedGen	3
OMIM	15

PubChem	
BioAssays	301
Compounds	0
Pathways	1
Substances	25



IL-13

Gene sources
Genomic

Categories
Alternatively spliced
Annotated genes
Non-coding
Protein-coding
Pseudogene

Sequence content
CCDS
Ensembl
RefSeq
RefSeqGene

Status clear
✓ Current

Clear all

Show additional filters

See IL13 interleukin 13 in the Gene database
il13 reference sequences Genomic (1) Transcript (4) Protein (4)

Search results

Items: 1 to 20 of 617

Page 1 of 31

See also 9 discontinued or replaced items.

Name/Gene ID	Description	Location	Aliases	MIM
☐ Il13 ID: 116553	interleukin 13 [*Rattus norvegicus* (Norway rat)]	Chromosome 10, NC_051345.1 (37790130..37792687, complement)		
☐ IL13 ID: 3596	interleukin 13 [*Homo sapiens* (human)]	Chromosome 5, NC_000005.10 (132656522..132661110)	IL-13, P600	147683
☐ Il13 ID: 16163	interleukin 13 [*Mus musculus* (house mouse)]	Chromosome 11, NC_000077.7 (53522150..53525529, complement)	Il-13	
☐ IL13 ID: 281247	interleukin 13 [*Bos taurus* (cattle)]	Chromosome 7, NC_037334.1 (21721568..21723568, complement)		
☐ IL13 ID: 442990	interleukin 13 [*Canis lupus familiaris* (dog)]	Chromosome 11, NC_051815.1 (21744389..21747318)		
☐ IL13 ID: 574325	interleukin 13 [*Macaca mulatta* (Rhesus monkey)]	Chromosome 6, NC_041759.1 (130097046..130101956)		
☐ IL13 ID: 449564	interleukin 13 [*Pan troglodytes* (chimpanzee)]	Chromosome 5, NC_036884.1 (110864594..110869640)	CK820_G0026772	
☐ il13 ID: 794053	interleukin 13 [*Danio rerio* (zebrafish)]	Chromosome 9, NC_007120.7 (54301736..54304684, complement)	IL-4, il-13, il4/13a	
☐ IL13 ID: 664707	interleukin 13 [*Ovis aries* (sheep)]	Chromosome 5, NC_056058.1 (19503522..19506352, complement)		
☐ IL13 ID: 100034113	interleukin 13 [*Equus caballus* (horse)]	Chromosome 14, NC_009157.3 (42223438..42227607, complement)	IL-13	

Filters: Manage Filters

Results by taxon

Top Organisms [Tree]
- Homo sapiens *(210)*
- Mus musculus *(178)*
- Rattus norvegicus *(25)*
- Gallus gallus *(5)*
- Oncorhynchus mykiss *(5)*
- All other taxa *(194)*

More...

Find related data

Database: Select

Find items

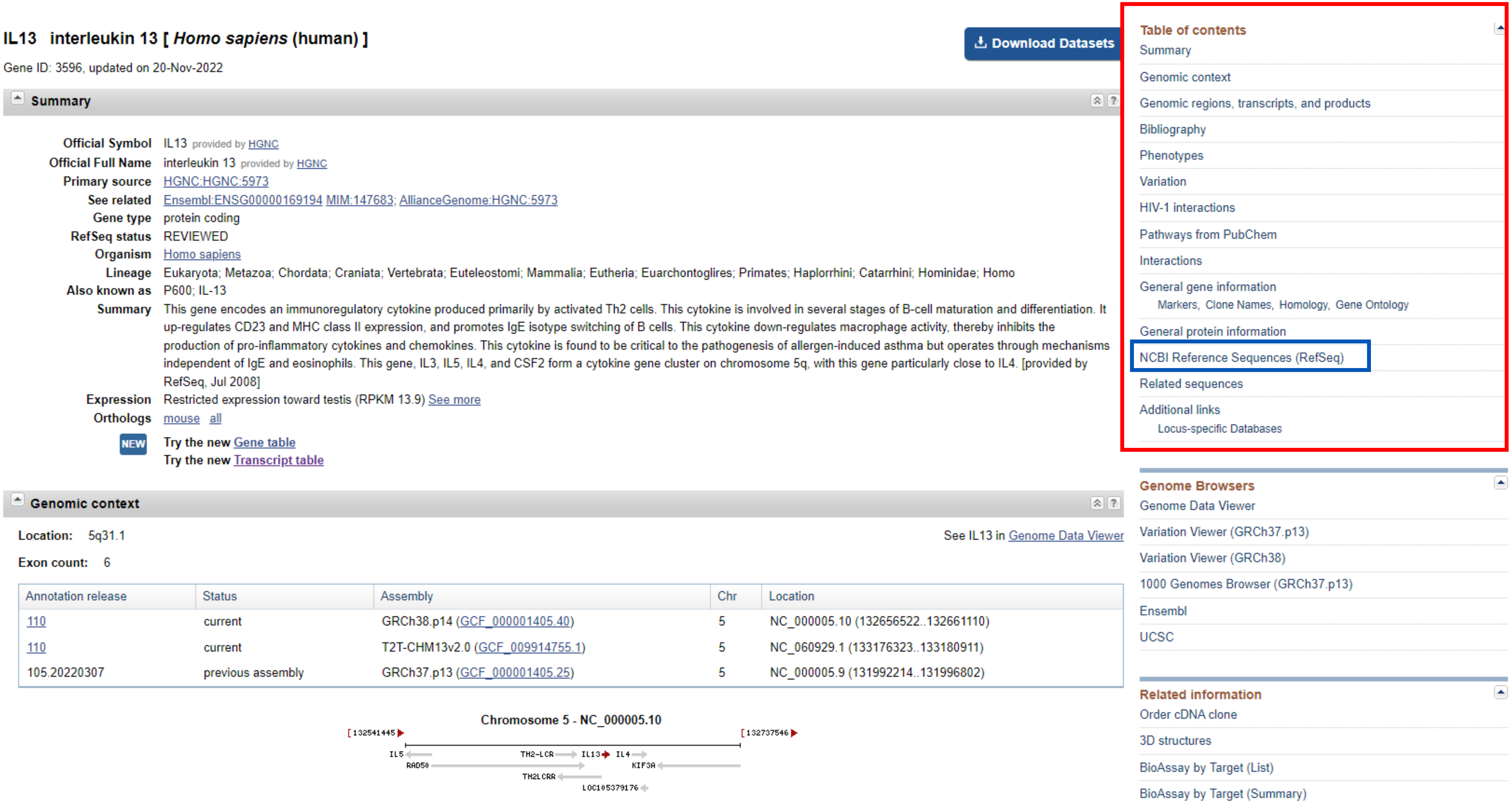
Search details

```
IL-13[All Fields] AND alive[prop]
```

Search See more...

Recent activity

Your browsing activity is temporarily unavailable.



IL13 interleukin 13 [*Homo sapiens* (human)]

Gene ID: 3596, updated on 20-Nov-2022

Summary

Official Symbol	IL13 provided by HGNC
Official Full Name	interleukin 13 provided by HGNC
Primary source	HGNC:HGNC:5973
See related	Ensembl:ENSG00000169194 MIM:147683; AllianceGenome:HGNC:5973
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as	P600; IL-13
Summary	This gene encodes an immunoregulatory cytokine produced primarily by activated Th2 cells. This cytokine is involved in several stages of B-cell maturation and differentiation. It up-regulates CD23 and MHC class II expression, and promotes IgE isotype switching of B cells. This cytokine down-regulates macrophage activity, thereby inhibits the production of pro-inflammatory cytokines and chemokines. This cytokine is found to be critical to the pathogenesis of allergen-induced asthma but operates through mechanisms independent of IgE and eosinophils. This gene, IL3, IL5, IL4, and CSF2 form a cytokine gene cluster on chromosome 5q, with this gene particularly close to IL4. [provided by RefSeq, Jul 2008]
Expression	Restricted expression toward testis (RPKM 13.9) See more
Orthologs	mouse all

NEW **Try the new** Gene table
Try the new Transcript table

Genomic context

Location: 5q31.1

See IL13 in Genome Data Viewer

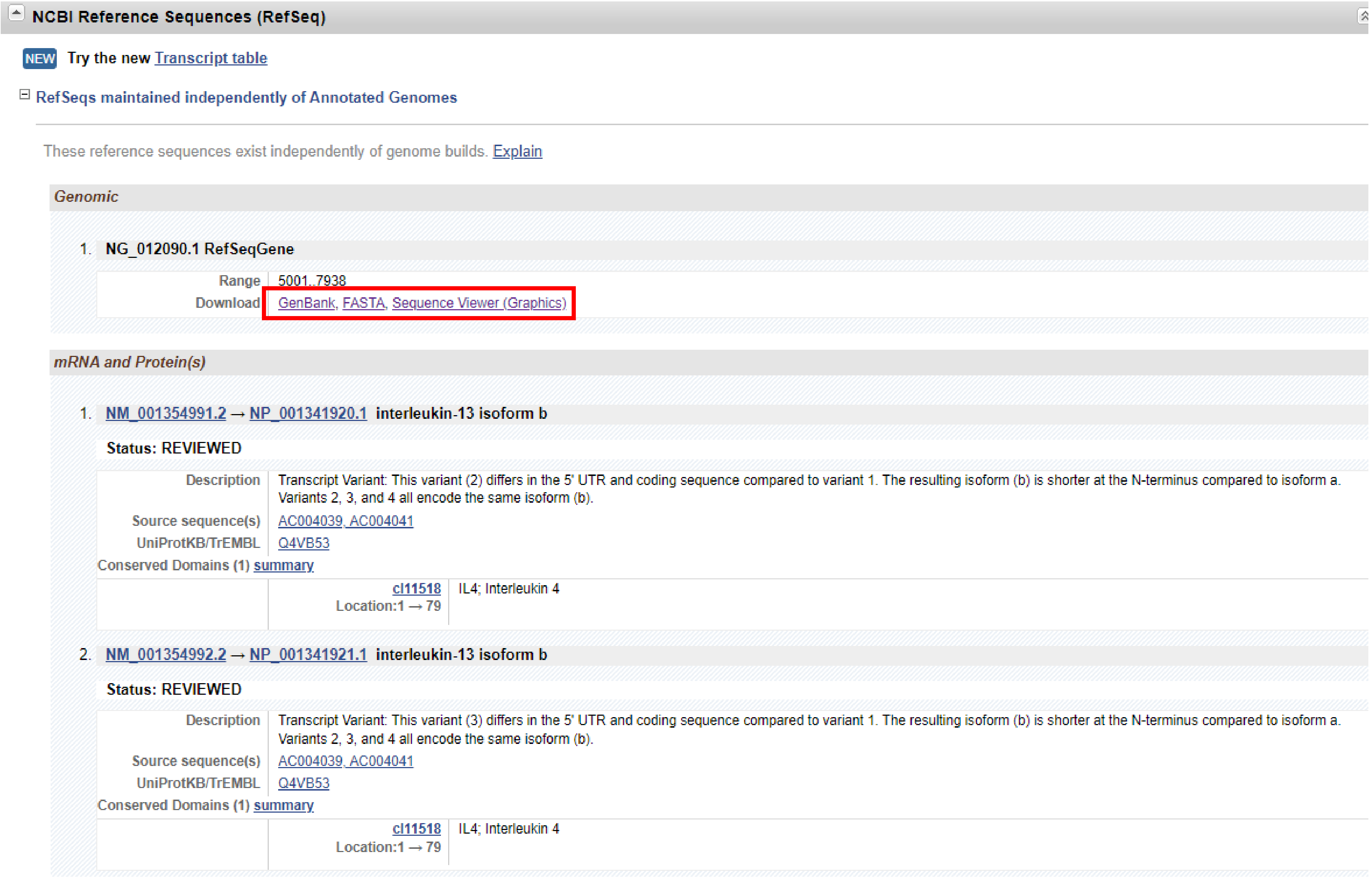
Exon count: 6

Annotation release	Status	Assembly	Chr	Location
110	current	GRCh38.p14 (GCF_000001405.40)	5	NC_000005.10 (132656522..132661110)
110	current	T2T-CHM13v2.0 (GCF_009914755.1)	5	NC_060929.1 (133176323..133180911)
105.20220307	previous assembly	GRCh37.p13 (GCF_000001405.25)	5	NC_000005.9 (131992214..131996802)

Table of contents
- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Bibliography
- Phenotypes
- Variation
- HIV-1 interactions
- Pathways from PubChem
- Interactions
- General gene information
  - Markers, Clone Names, Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links
  - Locus-specific Databases

Genome Browsers
- Genome Data Viewer
- Variation Viewer (GRCh37.p13)
- Variation Viewer (GRCh38)
- 1000 Genomes Browser (GRCh37.p13)
- Ensembl
- UCSC

Related information
- Order cDNA clone
- 3D structures
- BioAssay by Target (List)
- BioAssay by Target (Summary)



Try the new Transcript table

RefSeqs maintained independently of Annotated Genomes

These reference sequences exist independently of genome builds. Explain

Genomic

1. NG_012090.1 RefSeqGene

Range	5001..7938
Download	GenBank, FASTA, Sequence Viewer (Graphics)

mRNA and Protein(s)

1. NM_001354991.2 → NP_001341920.1 interleukin-13 isoform b

Status: REVIEWED

Description	Transcript Variant: This variant (2) differs in the 5' UTR and coding sequence compared to variant 1. The resulting isoform (b) is shorter at the N-terminus compared to isoform a. Variants 2, 3, and 4 all encode the same isoform (b).
Source sequence(s)	AC004039, AC004041
UniProtKB/TrEMBL	Q4VB53

Conserved Domains (1) summary

cl11518 Location:1 → 79	IL4; Interleukin 4

2. NM_001354992.2 → NP_001341921.1 interleukin-13 isoform b

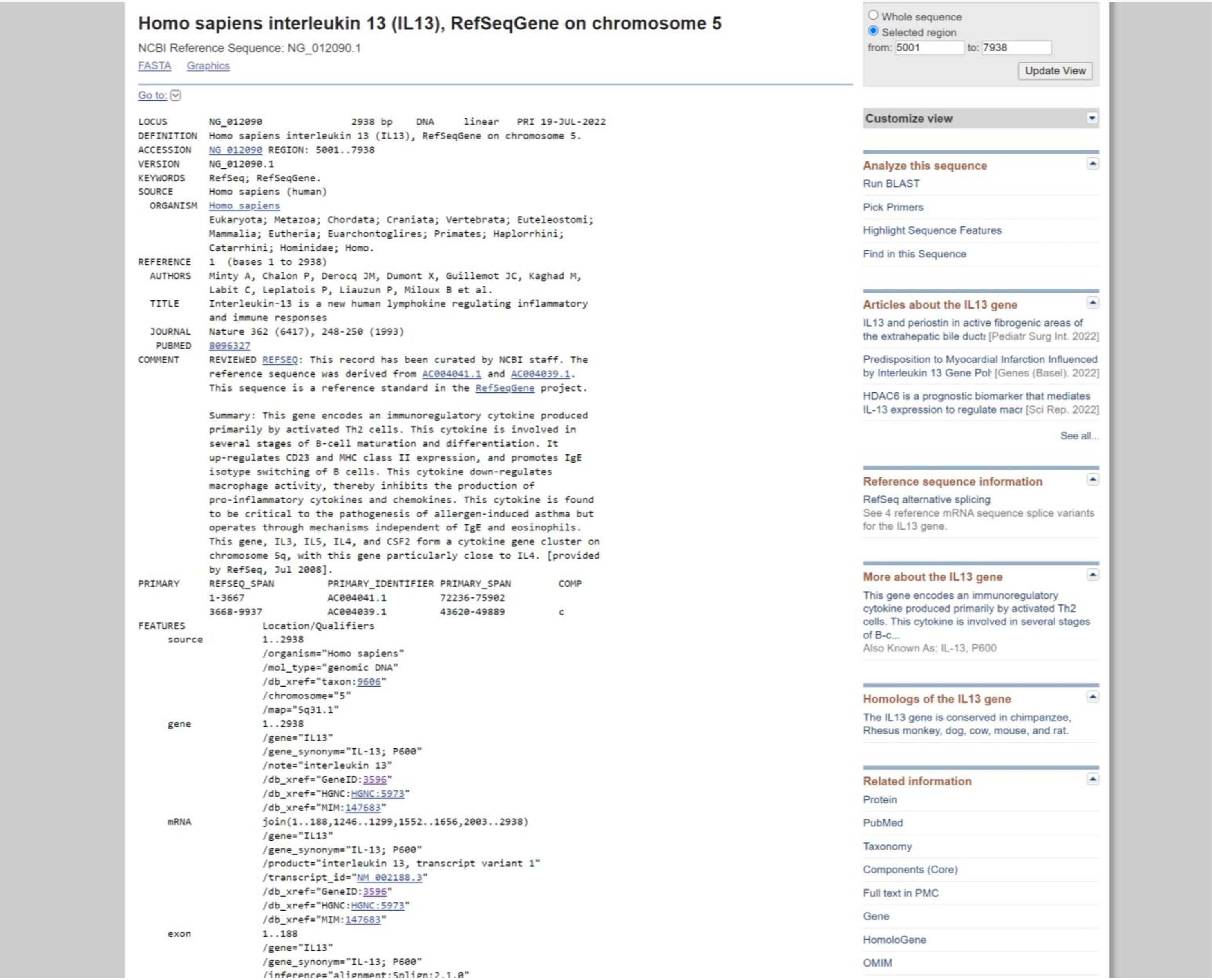
Status: REVIEWED

Description	Transcript Variant: This variant (3) differs in the 5' UTR and coding sequence compared to variant 1. The resulting isoform (b) is shorter at the N-terminus compared to isoform a. Variants 2, 3, and 4 all encode the same isoform (b).
Source sequence(s)	AC004039, AC004041
UniProtKB/TrEMBL	Q4VB53

Conserved Domains (1) summary

cl11518 Location:1 → 79	IL4; Interleukin 4

GENBANK RECORD



Homo sapiens interleukin 13 (IL13), RefSeqGene on chromosome 5

NCBI Reference Sequence: NG_012090.1

```
LOCUS       NG_012090               2938 bp    DNA     linear   PRI 19-JUL-2022
DEFINITION  Homo sapiens interleukin 13 (IL13), RefSeqGene on chromosome 5.
ACCESSION   NG_012090 REGION: 5001..7938
VERSION     NG_012090.1
KEYWORDS    RefSeq; RefSeqGene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2938)
  AUTHORS   Minty A, Chalon P, Derocq JM, Dumont X, Guillemot JC, Kaghad M,
            Labit C, Leplatois P, Liauzun P, Miloux B et al.
  TITLE     Interleukin-13 is a new human lymphokine regulating inflammatory
            and immune responses
  JOURNAL   Nature 362 (6417), 248-250 (1993)
   PUBMED   8096327
COMMENT     REVIEWED REFSEQ: This record has been curated by NCBI staff. The
            reference sequence was derived from AC004041.1 and AC004039.1.
            This sequence is a reference standard in the RefSeqGene project.

            Summary: This gene encodes an immunoregulatory cytokine produced
            primarily by activated Th2 cells. This cytokine is involved in
            several stages of B-cell maturation and differentiation. It
            up-regulates CD23 and MHC class II expression, and promotes IgE
            isotype switching of B cells. This cytokine down-regulates
            macrophage activity, thereby inhibits the production of
            pro-inflammatory cytokines and chemokines. This cytokine is found
            to be critical to the pathogenesis of allergen-induced asthma but
            operates through mechanisms independent of IgE and eosinophils.
            This gene, IL3, IL5, IL4, and CSF2 form a cytokine gene cluster on
            chromosome 5q, with this gene particularly close to IL4. [provided
            by RefSeq, Jul 2008].
PRIMARY     REFSEQ_SPAN         PRIMARY_IDENTIFIER PRIMARY_SPAN        COMP
            1-3667              AC004041.1         72236-75902
            3668-9937           AC004039.1         43620-49889         c
FEATURES             Location/Qualifiers
     source          1..2938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q31.1"
     gene            1..2938
                     /gene="IL13"
                     /gene_synonym="IL-13; P600"
                     /note="interleukin 13"
                     /db_xref="GeneID:3596"
                     /db_xref="HGNC:HGNC:5973"
                     /db_xref="MIM:147683"
     mRNA            join(1..188,1246..1299,1552..1656,2003..2938)
                     /gene="IL13"
                     /gene_synonym="IL-13; P600"
                     /product="interleukin 13, transcript variant 1"
                     /transcript_id="NM_002188.3"
                     /db_xref="GeneID:3596"
                     /db_xref="HGNC:HGNC:5973"
                     /db_xref="MIM:147683"
     exon            1..188
                     /gene="IL13"
                     /gene_synonym="IL-13; P600"
                     /inference="alignment:Splign:2.1.0"
```

○ Whole sequence
● Selected region
from: 5001 to: 7938

Analyze this sequence

Articles about the IL13 gene

IL13 and periostin in active fibrogenic areas of the extrahepatic bile ducts [Pediatr Surg Int. 2022]

Predisposition to Myocardial Infarction Influenced by Interleukin 13 Gene Pol [Genes (Basel). 2022]

HDAC6 is a prognostic biomarker that mediates IL-13 expression to regulate macr [Sci Rep. 2022]

Reference sequence information

RefSeq alternative splicing
See 4 reference mRNA sequence splice variants for the IL13 gene.

More about the IL13 gene

This gene encodes an immunoregulatory cytokine produced primarily by activated Th2 cells. This cytokine is involved in several stages of B-c...
Also Known As: IL-13, P600

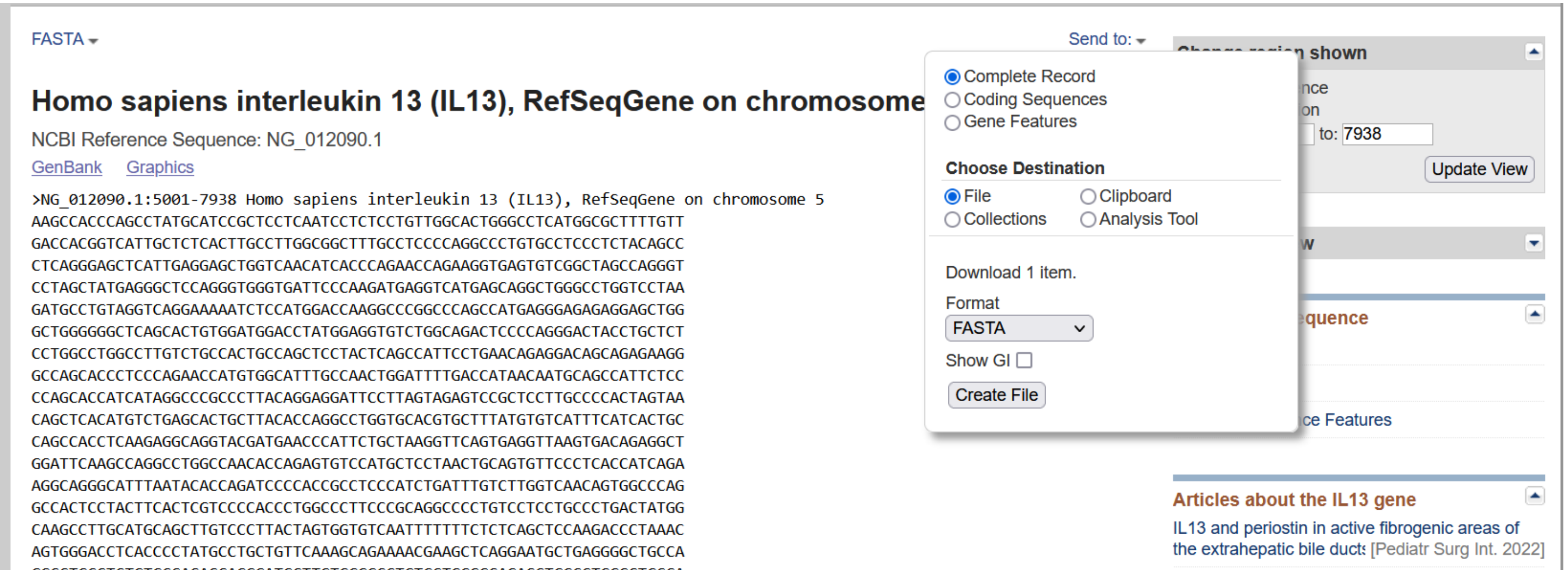
Homologs of the IL13 gene

The IL13 gene is conserved in chimpanzee, Rhesus monkey, dog, cow, mouse, and rat.

Related information

FASTA FILE

- Used for storing sequence reads only
- Mostly used for reference data



Homo sapiens interleukin 13 (IL13), RefSeqGene on chromosome

NCBI Reference Sequence: NG_012090.1

```
>NG_012090.1:5001-7938 Homo sapiens interleukin 13 (IL13), RefSeqGene on chromosome 5
AAGCCACCCAGCCTATGCATCCGCTCCTCAATCCTCTCCTGTTGGCACTGGGCCTCATGGCGCTTTTGTT
GACCACGGTCATTGCTCTCACTTGCCTTGGCGGCTTTGCCTCCCCAGGCCCTGTGCCTCCCTCTACAGCC
CTCAGGGAGCTCATTGAGGAGCTGGTCAACATCACCCAGAACCAGAAGGTGAGTGTCGGCTAGCCAGGGT
CCTAGCTATGAGGGCTCCAGGGTGGGTGATTCCCAAGATGAGGTCATGAGCAGGCTGGGCCTGGTCCTAA
GATGCCTGTAGGTCAGGAAAAATCTCCATGGACCAAGGCCCGGCCCAGCCATGAGGGAGAGAGGAGCTGG
GCTGGGGGGCTCAGCACTGTGGATGGACCTATGGAGGTGTCTGGCAGACTCCCCAGGGACTACCTGCTCT
CCTGGCCTGGCCTTGTCTGCCACTGCCAGCTCCTACTCAGCCATTCCTGAACAGAGGACAGCAGAGAAGG
GCCAGCACCCTCCCAGAACCATGTGGCATTTGCCAACTGGATTTTGACCATAACAATGCAGCCATTCTCC
CCAGCACCATCATAGGCCCGCCCTTACAGGAGGATTCCTTAGTAGAGTCCGCTCCTTGCCCCACTAGTAA
CAGCTCACATGTCTGAGCACTGCTTACACCAGGCCTGGTGCACGTGCTTTATGTGTCATTTCATCACTGC
CAGCCACCTCAAGAGGCAGGTACGATGAACCCATTCTGCTAAGGTTCAGTGAGGTTAAGTGACAGAGGCT
GGATTCAAGCCAGGCCTGGCCAACACCAGAGTGTCCATGCTCCTAACTGCAGTGTTCCCTCACCATCAGA
AGGCAGGGCATTTAATACACCAGATCCCCACCGCCTCCCATCTGATTTGTCTTGGTCAACAGTGGCCCAG
GCCACTCCTACTTCACTCGTCCCCACCCTGGCCCTTCCCGCAGGCCCCTGTCCTCCTGCCCTGACTATGG
CAAGCCTTGCATGCAGCTTGTCCCTTACTAGTGGTGTCAATTTTTTTCTCTCAGCTCCAAGACCCTAAAC
AGTGGGACCTCACCCCTATGCCTGCTGTTCAAAGCAGAAAACGAAGCTCAGGAATGCTGAGGGGCTGCCA
```

● Complete Record
○ Coding Sequences
○ Gene Features

Choose Destination

● File ○ Clipboard
○ Collections ○ Analysis Tool

Download 1 item.

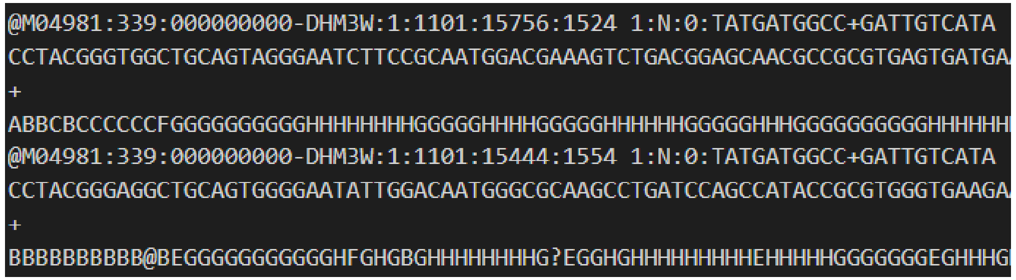
Format: FASTA

Show GI ☐

Articles about the IL13 gene

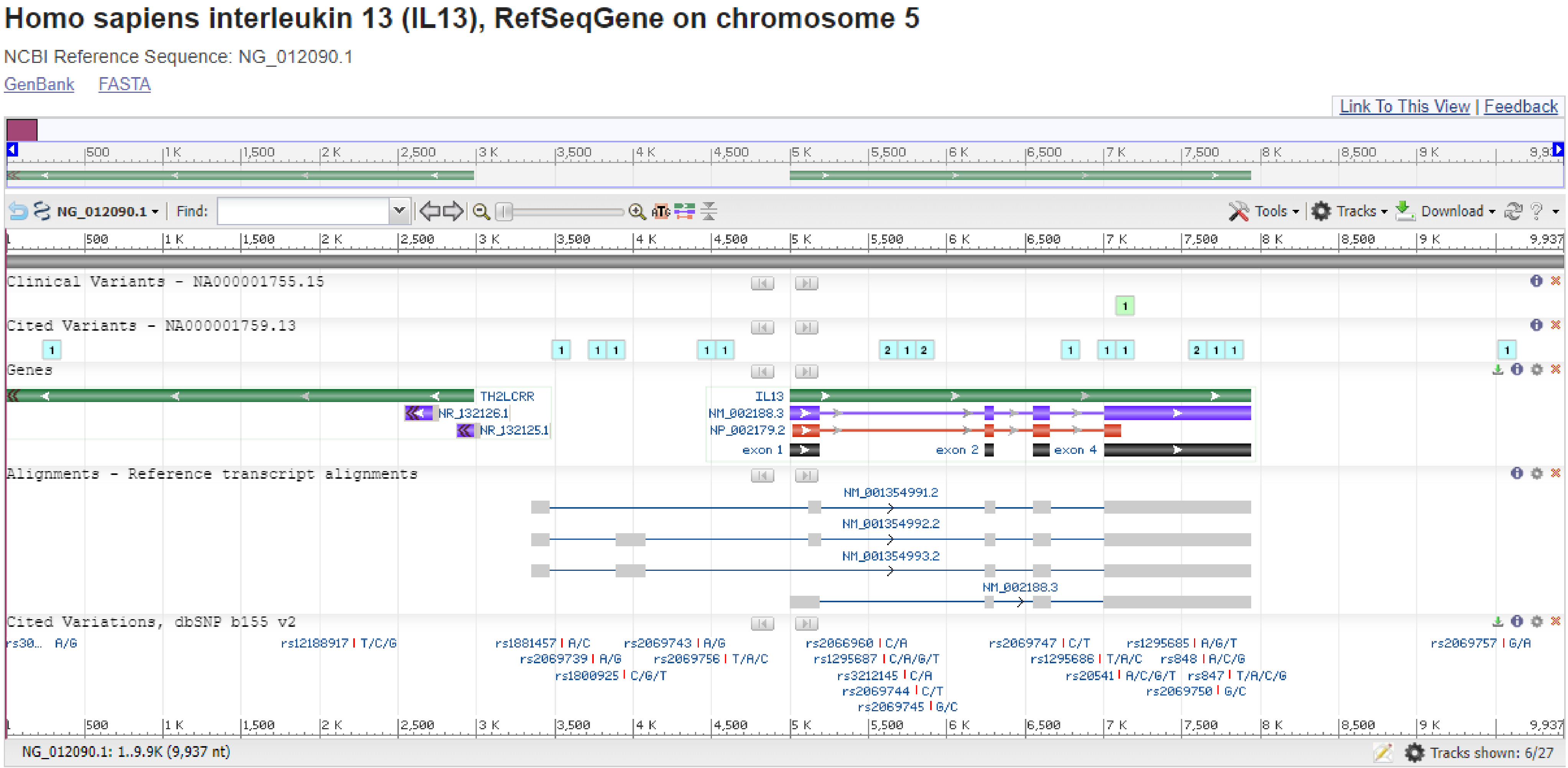
IL13 and periostin in active fibrogenic areas of the extrahepatic bile ducts [Pediatr Surg Int. 2022]

- Different than FastQ files (detailed explanation in previous workshop)
  - Contains read quality



```
@M04981:339:000000000-DHM3W:1:1101:15756:1524 1:N:0:TATGATGGCC+GATTGTCATA
CCTACGGGTGGCTGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGA
+
ABBCBCCCCCCFGGGGGGGGGGHHHHHHHHGGGGGHHHHGGGGGHHHHHHGGGGGHHHGGGGGGGGGGHHHHHH
@M04981:339:000000000-DHM3W:1:1101:15444:1554 1:N:0:TATGATGGCC+GATTGTCATA
CCTACGGGAGGCTGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATACCGCGTGGGTGAAGA
+
BBBBBBBBBB@BEGGGGGGGGGGGHFGHGBGHHHHHHHHG?EGGHGHHHHHHHHHEHHHHHGGGGGGEGHHHG
```

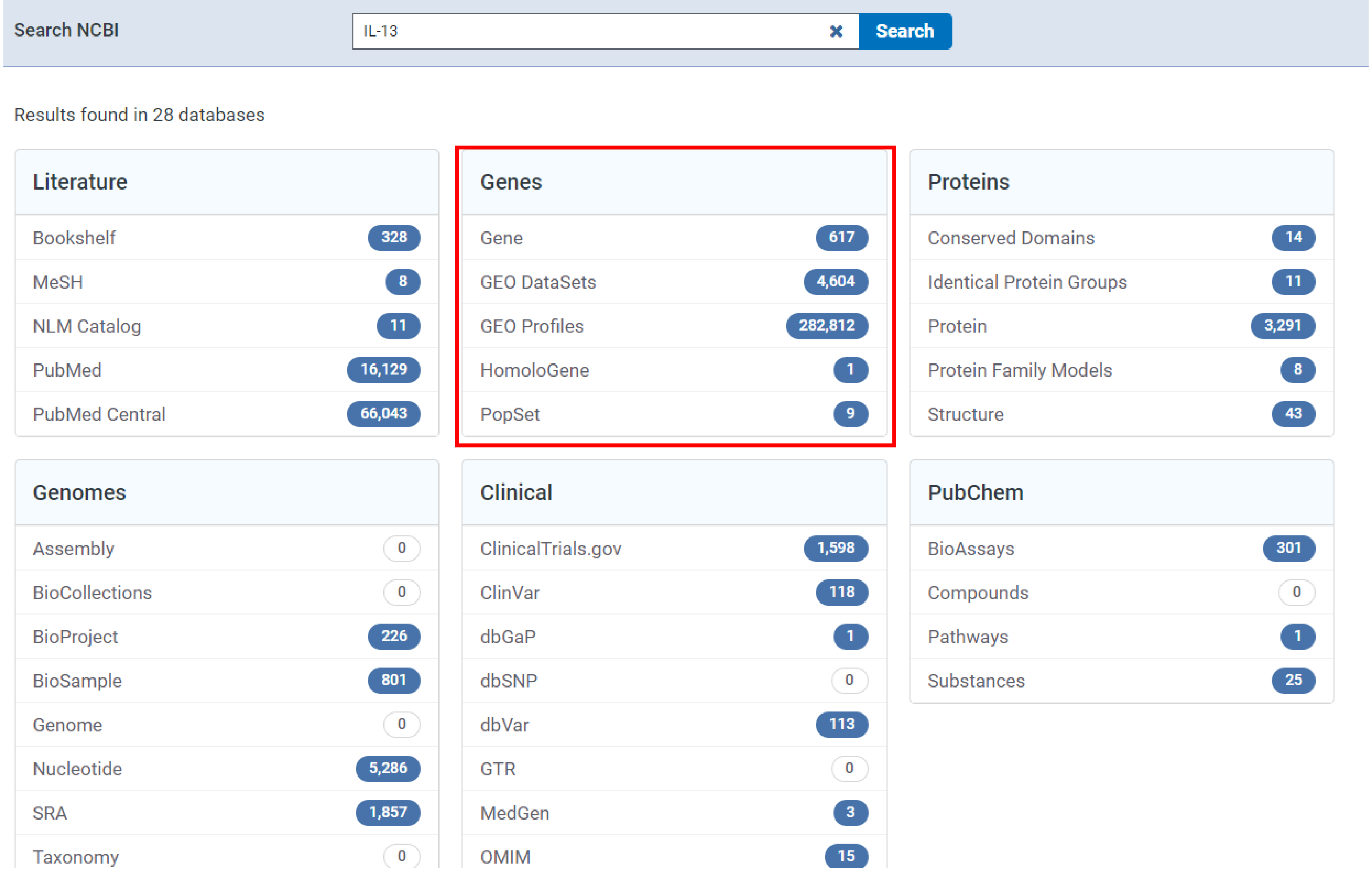
SEQUENCE VIEWER



Homo sapiens interleukin 13 (IL13), RefSeqGene on chromosome 5

NCBI Reference Sequence: NG_012090.1

GenBank FASTA



Results found in 28 databases

Literature	
Bookshelf	328
MeSH	8
NLM Catalog	11
PubMed	16,129
PubMed Central	66,043

Genes	
Gene	617
GEO DataSets	4,604
GEO Profiles	282,812
HomoloGene	1
PopSet	9

Proteins	
Conserved Domains	14
Identical Protein Groups	11
Protein	3,291
Protein Family Models	8
Structure	43

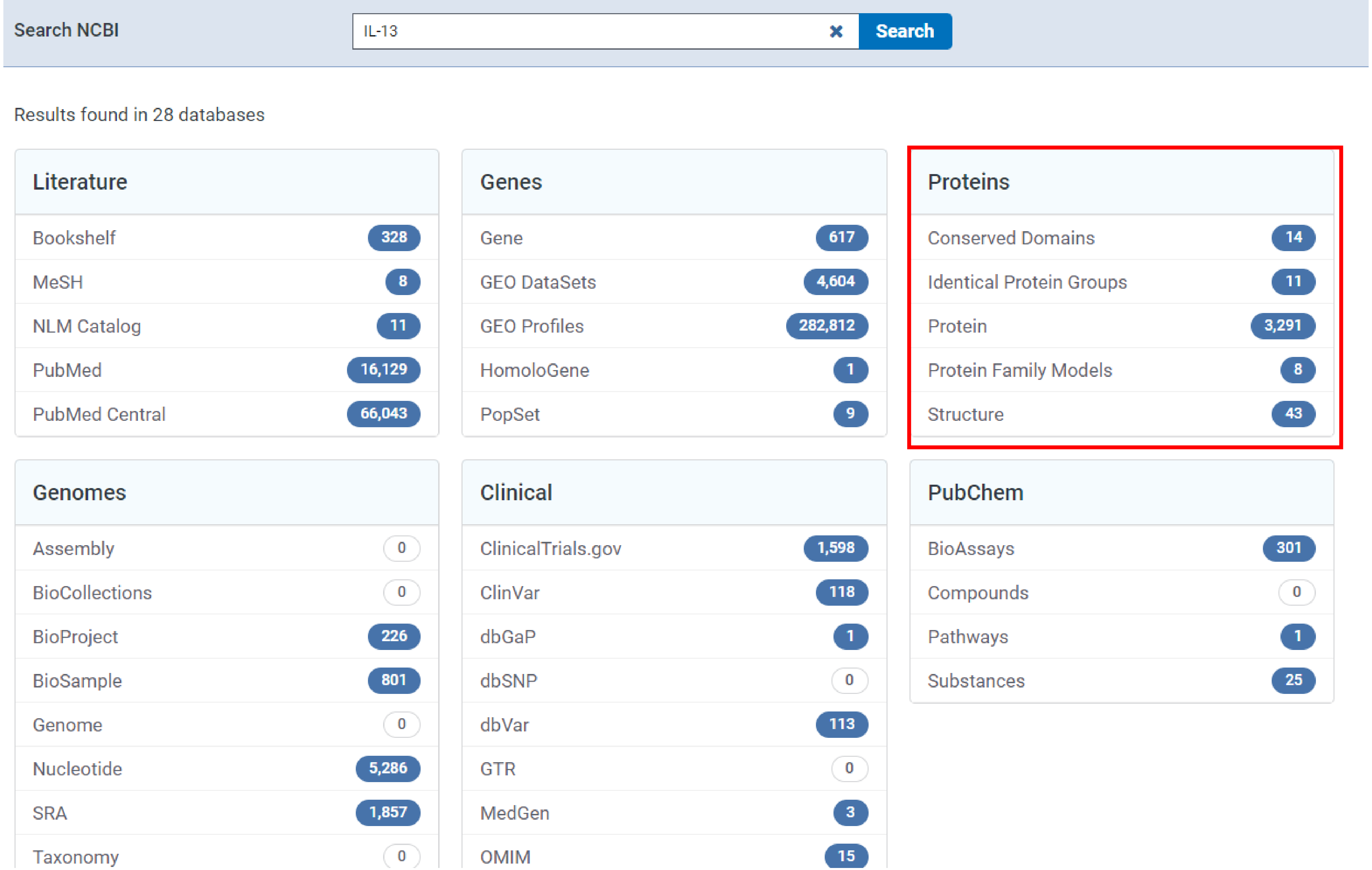
Genomes	
Assembly	0
BioCollections	0
BioProject	226
BioSample	801
Genome	0
Nucleotide	5,286
SRA	1,857
Taxonomy	0

Clinical	
ClinicalTrials.gov	1,598
ClinVar	118
dbGaP	1
dbSNP	0
dbVar	113
GTR	0
MedGen	3
OMIM	15

PubChem	
BioAssays	301
Compounds	0
Pathways	1
Substances	25

NCBI DATABASES

1. Gene:
    - Gene sequences and annotations used as references for the study of orthologs structure, expression, and evolution
2. **Protein:**
    - **Protein sequences, 3-D structures, and tools for the study of functional protein domains and active sites**
3. Genome:
    - Genome sequence assemblies, large-scale functional genomics data, and source biological samples
4. Clinical:
    - Heritable DNA variations, associations with human pathologies, and clinical diagnostics and treatments



Results found in 28 databases

Literature	
Bookshelf	328
MeSH	8
NLM Catalog	11
PubMed	16,129
PubMed Central	66,043

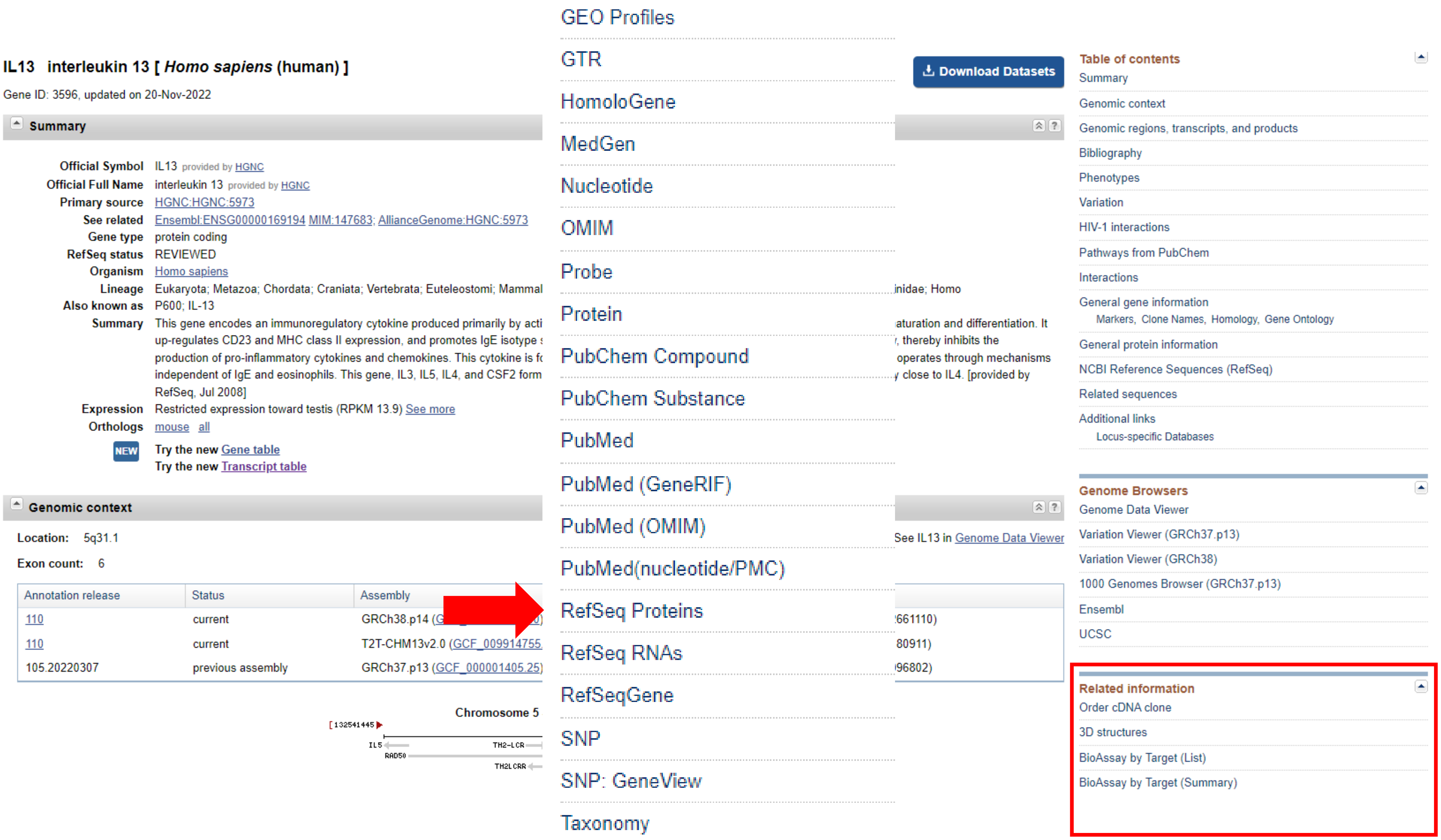
Genes	
Gene	617
GEO DataSets	4,604
GEO Profiles	282,812
HomoloGene	1
PopSet	9

Proteins	
Conserved Domains	14
Identical Protein Groups	11
Protein	3,291
Protein Family Models	8
Structure	43

Genomes	
Assembly	0
BioCollections	0
BioProject	226
BioSample	801
Genome	0
Nucleotide	5,286
SRA	1,857
Taxonomy	0

Clinical	
ClinicalTrials.gov	1,598
ClinVar	118
dbGaP	1
dbSNP	0
dbVar	113
GTR	0
MedGen	3
OMIM	15

PubChem	
BioAssays	301
Compounds	0
Pathways	1
Substances	25



IL13 interleukin 13 [*Homo sapiens* (human)]

Gene ID: 3596, updated on 20-Nov-2022

Summary

Official Symbol	IL13 provided by HGNC
Official Full Name	interleukin 13 provided by HGNC
Primary source	HGNC:HGNC:5973
See related	Ensembl:ENSG00000169194 MIM:147683; AllianceGenome:HGNC:5973
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammal ...inidae; Homo
Also known as	P600; IL-13
Summary	This gene encodes an immunoregulatory cytokine produced primarily by acti ...aturation and differentiation. It up-regulates CD23 and MHC class II expression, and promotes IgE isotype s ..., thereby inhibits the production of pro-inflammatory cytokines and chemokines. This cytokine is fo ... operates through mechanisms independent of IgE and eosinophils. This gene, IL3, IL5, IL4, and CSF2 form ...y close to IL4. [provided by RefSeq, Jul 2008]
Expression	Restricted expression toward testis (RPKM 13.9) See more
Orthologs	mouse all

NEW Try the new Gene table
Try the new Transcript table

Genomic context

Location: 5q31.1

Exon count: 6

See IL13 in Genome Data Viewer

Annotation release	Status	Assembly	
110	current	GRCh38.p14 (G	661110)
110	current	T2T-CHM13v2.0 (GCF_009914755	80911)
105.20220307	previous assembly	GRCh37.p13 (GCF_000001405.25)	96802)

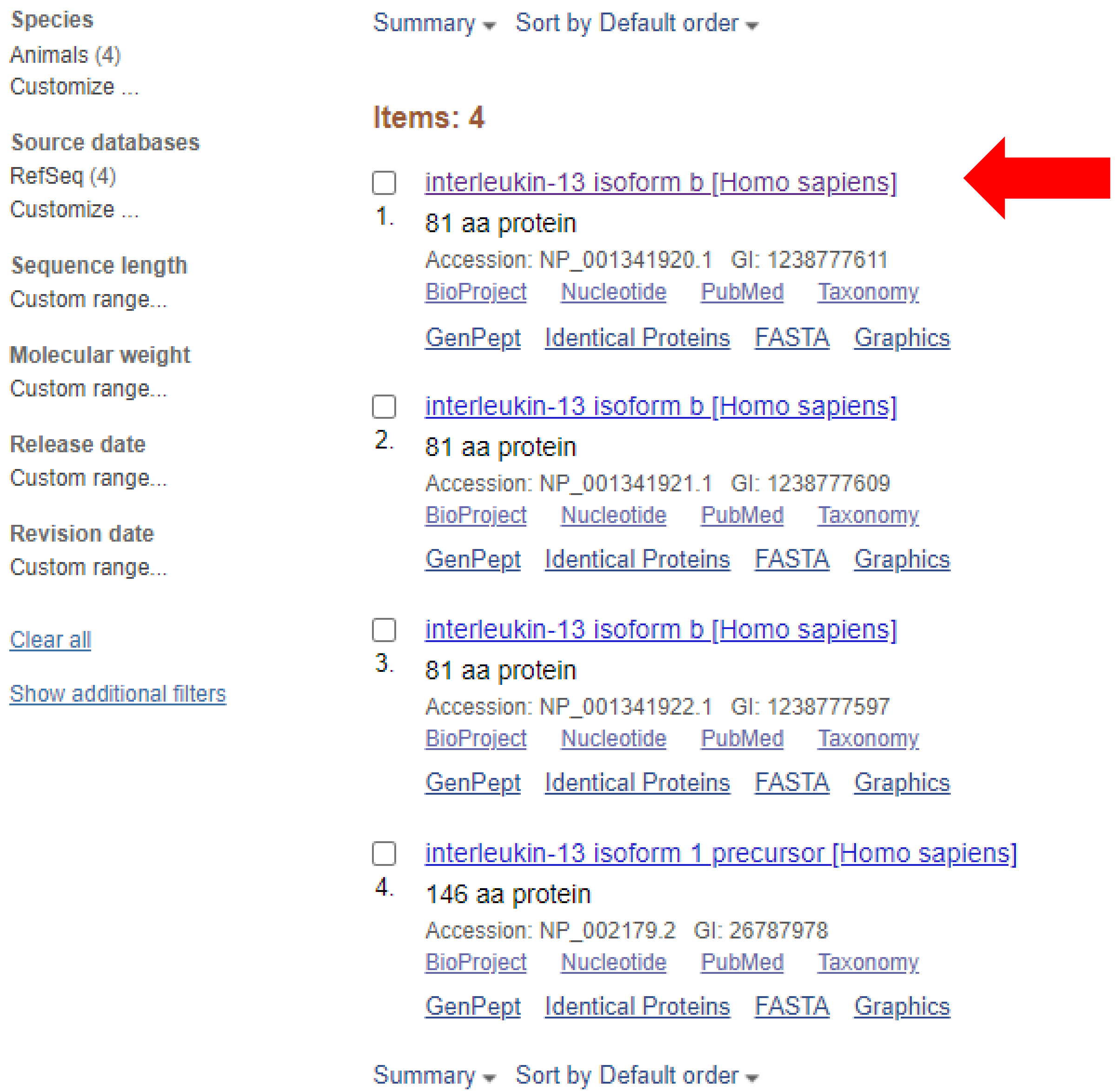
Chromosome 5

GEO Profiles
GTR
HomoloGene
MedGen
Nucleotide
OMIM
Probe
Protein
PubChem Compound
PubChem Substance
PubMed
PubMed (GeneRIF)
PubMed (OMIM)
PubMed(nucleotide/PMC)
RefSeq Proteins
RefSeq RNAs
RefSeqGene
SNP
SNP: GeneView
Taxonomy

Table of contents
- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Bibliography
- Phenotypes
- Variation
- HIV-1 interactions
- Pathways from PubChem
- Interactions
- General gene information
  - Markers, Clone Names, Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links
  - Locus-specific Databases

Genome Browsers
- Genome Data Viewer
- Variation Viewer (GRCh37.p13)
- Variation Viewer (GRCh38)
- 1000 Genomes Browser (GRCh37.p13)
- Ensembl
- UCSC

Related information
- Order cDNA clone
- 3D structures
- BioAssay by Target (List)
- BioAssay by Target (Summary)



Species
Animals (4)
Customize ...

Source databases
RefSeq (4)
Customize ...

Sequence length
Custom range...

Molecular weight
Custom range...

Release date
Custom range...

Revision date
Custom range...

Clear all

Show additional filters

Summary ▾ Sort by Default order ▾

Items: 4

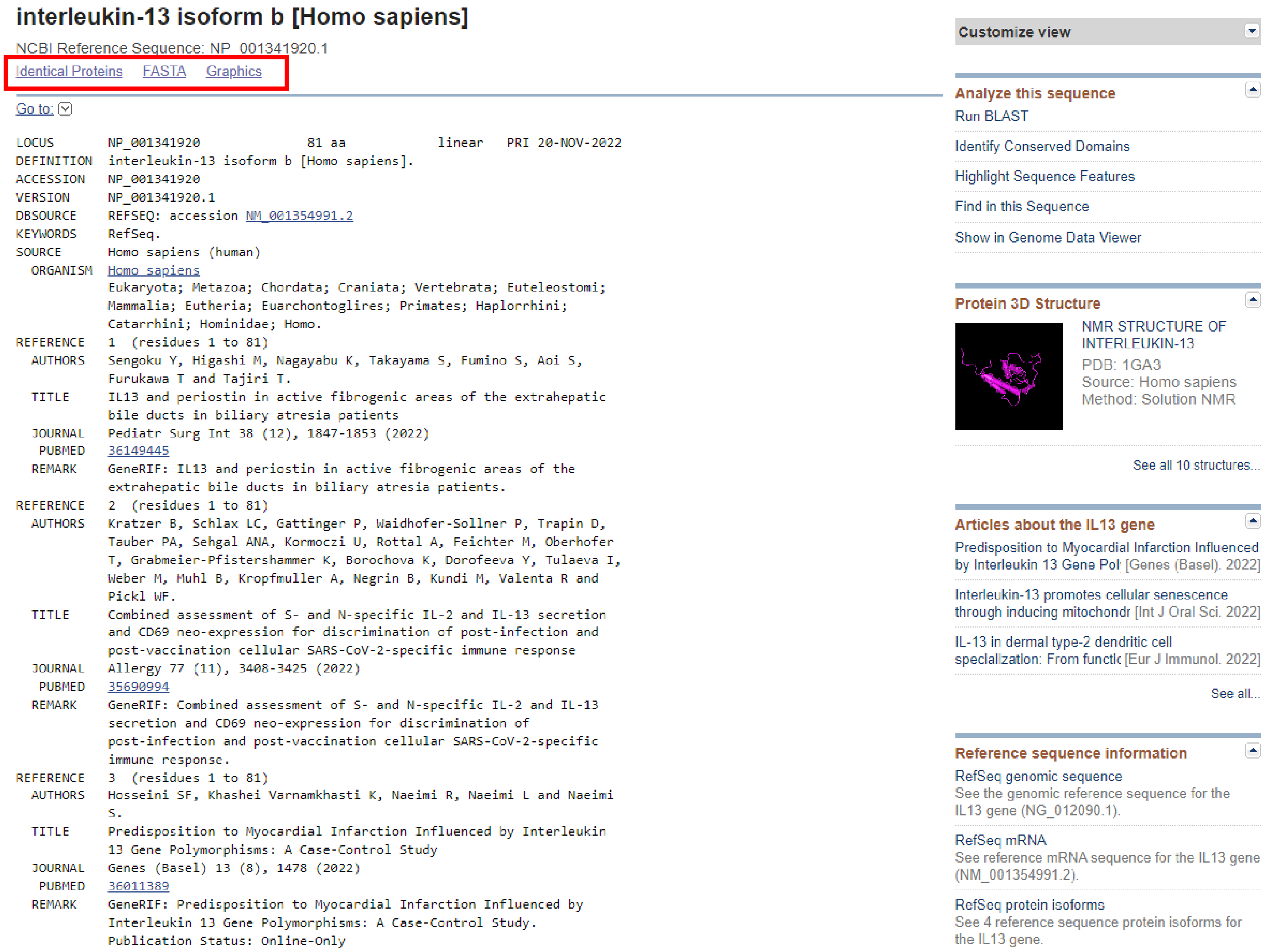
☐ 1. interleukin-13 isoform b [Homo sapiens]
81 aa protein
Accession: NP_001341920.1 GI: 1238777611
BioProject Nucleotide PubMed Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 2. interleukin-13 isoform b [Homo sapiens]
81 aa protein
Accession: NP_001341921.1 GI: 1238777609
BioProject Nucleotide PubMed Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 3. interleukin-13 isoform b [Homo sapiens]
81 aa protein
Accession: NP_001341922.1 GI: 1238777597
BioProject Nucleotide PubMed Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 4. interleukin-13 isoform 1 precursor [Homo sapiens]
146 aa protein
Accession: NP_002179.2 GI: 26787978
BioProject Nucleotide PubMed Taxonomy
GenPept Identical Proteins FASTA Graphics

Summary ▾ Sort by Default order ▾



interleukin-13 isoform b [Homo sapiens]

NCBI Reference Sequence: NP_001341920.1

```
LOCUS       NP_001341920              81 aa            linear   PRI 20-NOV-2022
DEFINITION  interleukin-13 isoform b [Homo sapiens].
ACCESSION   NP_001341920
VERSION     NP_001341920.1
DBSOURCE    REFSEQ: accession NM_001354991.2
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 81)
  AUTHORS   Sengoku Y, Higashi M, Nagayabu K, Takayama S, Fumino S, Aoi S,
            Furukawa T and Tajiri T.
  TITLE     IL13 and periostin in active fibrogenic areas of the extrahepatic
            bile ducts in biliary atresia patients
  JOURNAL   Pediatr Surg Int 38 (12), 1847-1853 (2022)
   PUBMED   36149445
  REMARK    GeneRIF: IL13 and periostin in active fibrogenic areas of the
            extrahepatic bile ducts in biliary atresia patients.
REFERENCE   2  (residues 1 to 81)
  AUTHORS   Kratzer B, Schlax LC, Gattinger P, Waidhofer-Sollner P, Trapin D,
            Tauber PA, Sehgal ANA, Kormoczi U, Rottal A, Feichter M, Oberhofer
            T, Grabmeier-Pfistershammer K, Borochova K, Dorofeeva Y, Tulaeva I,
            Weber M, Muhl B, Kropfmuller A, Negrin B, Kundi M, Valenta R and
            Pickl WF.
  TITLE     Combined assessment of S- and N-specific IL-2 and IL-13 secretion
            and CD69 neo-expression for discrimination of post-infection and
            post-vaccination cellular SARS-CoV-2-specific immune response
  JOURNAL   Allergy 77 (11), 3408-3425 (2022)
   PUBMED   35690994
  REMARK    GeneRIF: Combined assessment of S- and N-specific IL-2 and IL-13
            secretion and CD69 neo-expression for discrimination of
            post-infection and post-vaccination cellular SARS-CoV-2-specific
            immune response.
REFERENCE   3  (residues 1 to 81)
  AUTHORS   Hosseini SF, Khashei Varnamkhasti K, Naeimi R, Naeimi L and Naeimi
            S.
  TITLE     Predisposition to Myocardial Infarction Influenced by Interleukin
            13 Gene Polymorphisms: A Case-Control Study
  JOURNAL   Genes (Basel) 13 (8), 1478 (2022)
   PUBMED   36011389
  REMARK    GeneRIF: Predisposition to Myocardial Infarction Influenced by
            Interleukin 13 Gene Polymorphisms: A Case-Control Study.
            Publication Status: Online-Only
```

Analyze this sequence

- Run BLAST
- Identify Conserved Domains
- Highlight Sequence Features
- Find in this Sequence
- Show in Genome Data Viewer

Protein 3D Structure

NMR STRUCTURE OF INTERLEUKIN-13
PDB: 1GA3
Source: Homo sapiens
Method: Solution NMR

See all 10 structures...

Articles about the IL13 gene

Predisposition to Myocardial Infarction Influenced by Interleukin 13 Gene Pol [Genes (Basel). 2022]

Interleukin-13 promotes cellular senescence through inducing mitochondr [Int J Oral Sci. 2022]

IL-13 in dermal type-2 dendritic cell specialization: From functic [Eur J Immunol. 2022]

See all...

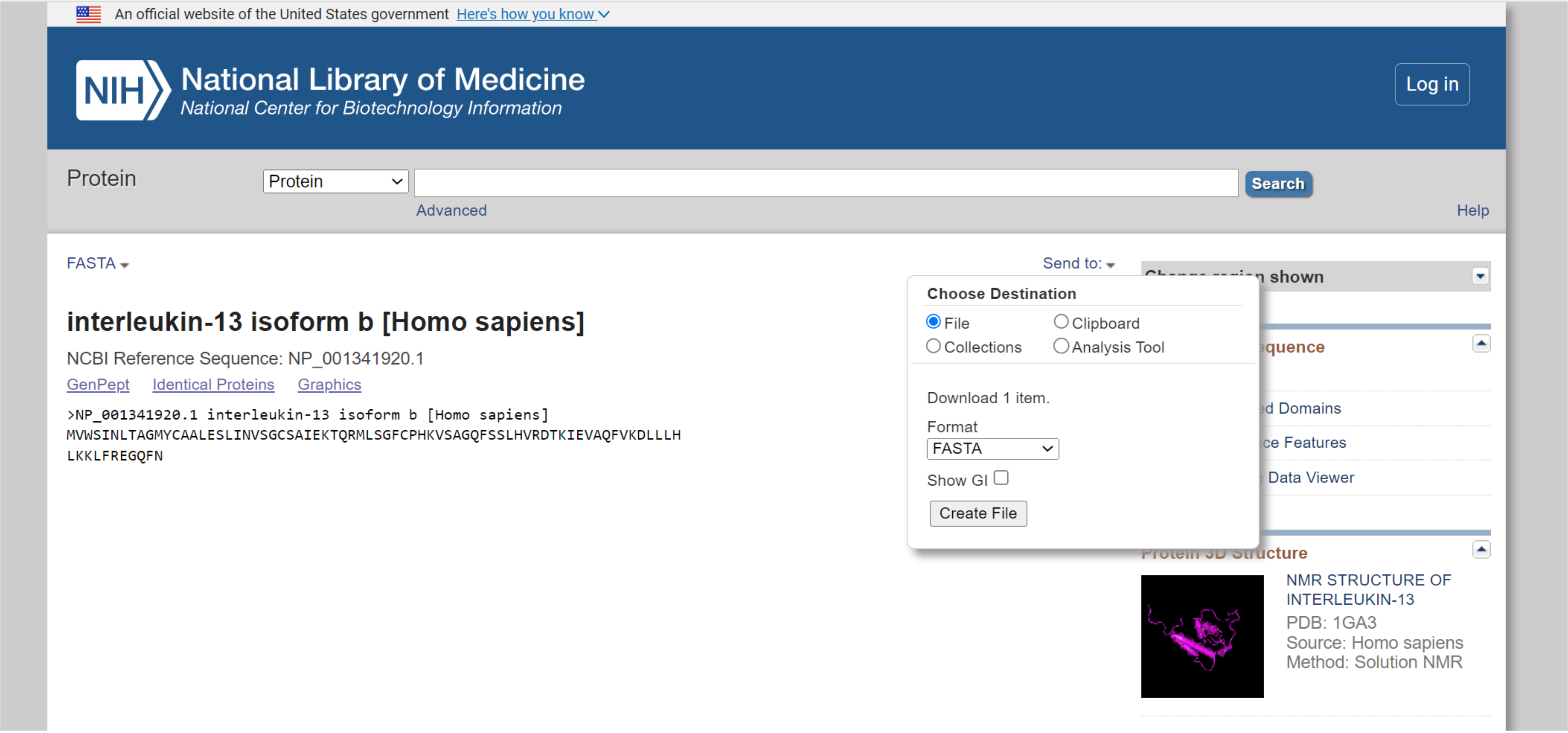
Reference sequence information

RefSeq genomic sequence
See the genomic reference sequence for the IL13 gene (NG_012090.1).

RefSeq mRNA
See reference mRNA sequence for the IL13 gene (NM_001354991.2).

RefSeq protein isoforms
See 4 reference sequence protein isoforms for the IL13 gene.

PROTEIN FASTA FILE



An official website of the United States government Here's how you know

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

Protein

Protein ▾ Search

Advanced

Help

FASTA ▾

Send to: ▾

interleukin-13 isoform b [Homo sapiens]

NCBI Reference Sequence: NP_001341920.1

GenPept Identical Proteins Graphics

```
>NP_001341920.1 interleukin-13 isoform b [Homo sapiens]
MVWSINLTAGMYCAALESLINVSGCSAIEKTQRMLSGFCPHKVSAGQFSSLHVRDTKIEVAQFVKDLLLH
LKKLFREGQFN
```

Choose Destination

● File ○ Clipboard
○ Collections ○ Analysis Tool

Download 1 item.

Format
FASTA ▾

Show GI ☐

Create File

Change region shown

quence

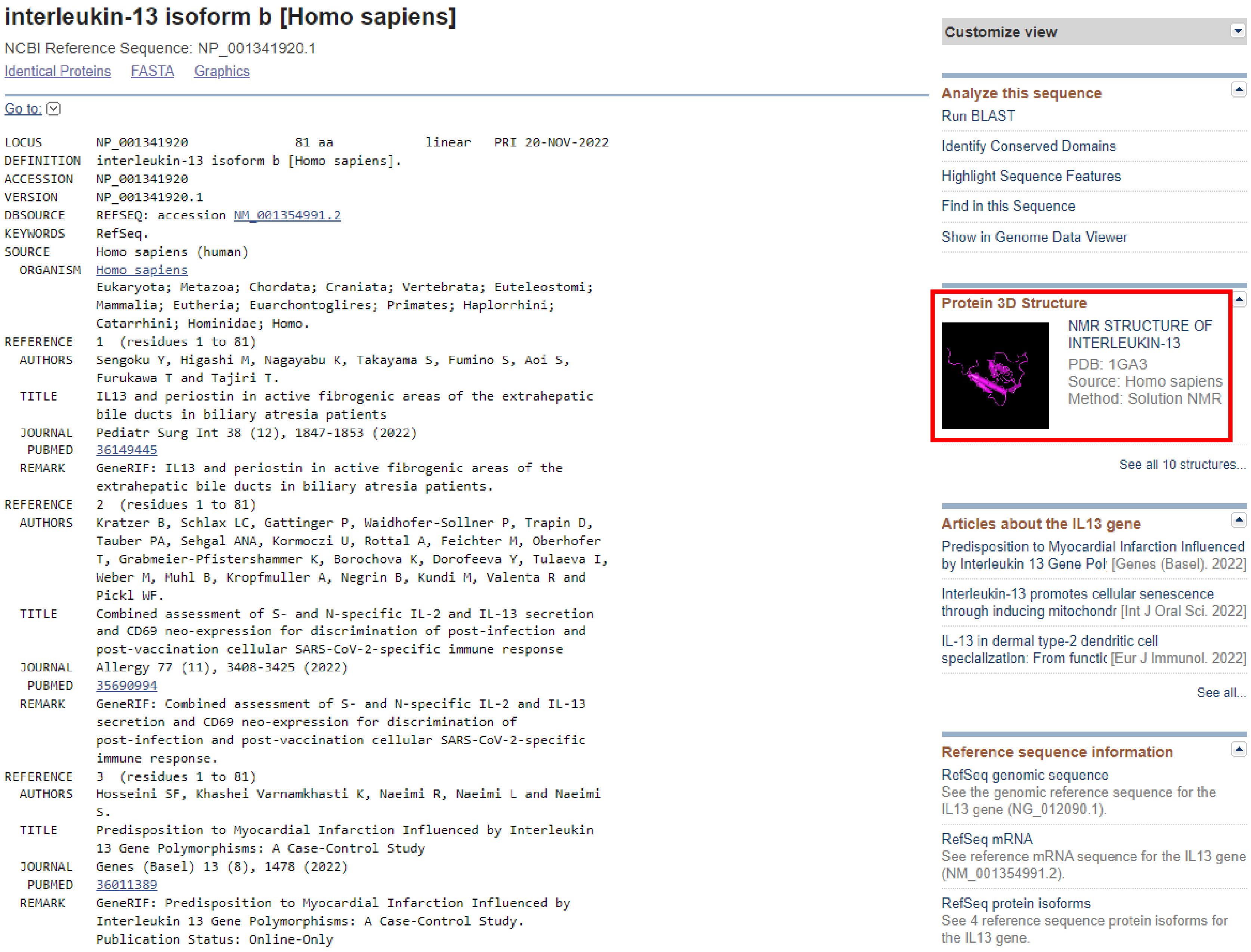
d Domains

ce Features

Data Viewer

Protein 3D Structure

NMR STRUCTURE OF INTERLEUKIN-13
PDB: 1GA3
Source: Homo sapiens
Method: Solution NMR



interleukin-13 isoform b [Homo sapiens]

NCBI Reference Sequence: NP_001341920.1

```
LOCUS       NP_001341920              81 aa            linear   PRI 20-NOV-2022
DEFINITION  interleukin-13 isoform b [Homo sapiens].
ACCESSION   NP_001341920
VERSION     NP_001341920.1
DBSOURCE    REFSEQ: accession NM_001354991.2
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 81)
  AUTHORS   Sengoku Y, Higashi M, Nagayabu K, Takayama S, Fumino S, Aoi S,
            Furukawa T and Tajiri T.
  TITLE     IL13 and periostin in active fibrogenic areas of the extrahepatic
            bile ducts in biliary atresia patients
  JOURNAL   Pediatr Surg Int 38 (12), 1847-1853 (2022)
   PUBMED   36149445
  REMARK    GeneRIF: IL13 and periostin in active fibrogenic areas of the
            extrahepatic bile ducts in biliary atresia patients.
REFERENCE   2  (residues 1 to 81)
  AUTHORS   Kratzer B, Schlax LC, Gattinger P, Waidhofer-Sollner P, Trapin D,
            Tauber PA, Sehgal ANA, Kormoczi U, Rottal A, Feichter M, Oberhofer
            T, Grabmeier-Pfistershammer K, Borochova K, Dorofeeva Y, Tulaeva I,
            Weber M, Muhl B, Kropfmuller A, Negrin B, Kundi M, Valenta R and
            Pickl WF.
  TITLE     Combined assessment of S- and N-specific IL-2 and IL-13 secretion
            and CD69 neo-expression for discrimination of post-infection and
            post-vaccination cellular SARS-CoV-2-specific immune response
  JOURNAL   Allergy 77 (11), 3408-3425 (2022)
   PUBMED   35690994
  REMARK    GeneRIF: Combined assessment of S- and N-specific IL-2 and IL-13
            secretion and CD69 neo-expression for discrimination of
            post-infection and post-vaccination cellular SARS-CoV-2-specific
            immune response.
REFERENCE   3  (residues 1 to 81)
  AUTHORS   Hosseini SF, Khashei Varnamkhasti K, Naeimi R, Naeimi L and Naeimi
            S.
  TITLE     Predisposition to Myocardial Infarction Influenced by Interleukin
            13 Gene Polymorphisms: A Case-Control Study
  JOURNAL   Genes (Basel) 13 (8), 1478 (2022)
   PUBMED   36011389
  REMARK    GeneRIF: Predisposition to Myocardial Infarction Influenced by
            Interleukin 13 Gene Polymorphisms: A Case-Control Study.
            Publication Status: Online-Only
```

Analyze this sequence

- Run BLAST
- Identify Conserved Domains
- Highlight Sequence Features
- Find in this Sequence
- Show in Genome Data Viewer

Protein 3D Structure

NMR STRUCTURE OF INTERLEUKIN-13
PDB: 1GA3
Source: Homo sapiens
Method: Solution NMR

See all 10 structures...

Articles about the IL13 gene

- Predisposition to Myocardial Infarction Influenced by Interleukin 13 Gene Pol [Genes (Basel). 2022]
- Interleukin-13 promotes cellular senescence through inducing mitochondr [Int J Oral Sci. 2022]
- IL-13 in dermal type-2 dendritic cell specialization: From functic [Eur J Immunol. 2022]

See all...

Reference sequence information

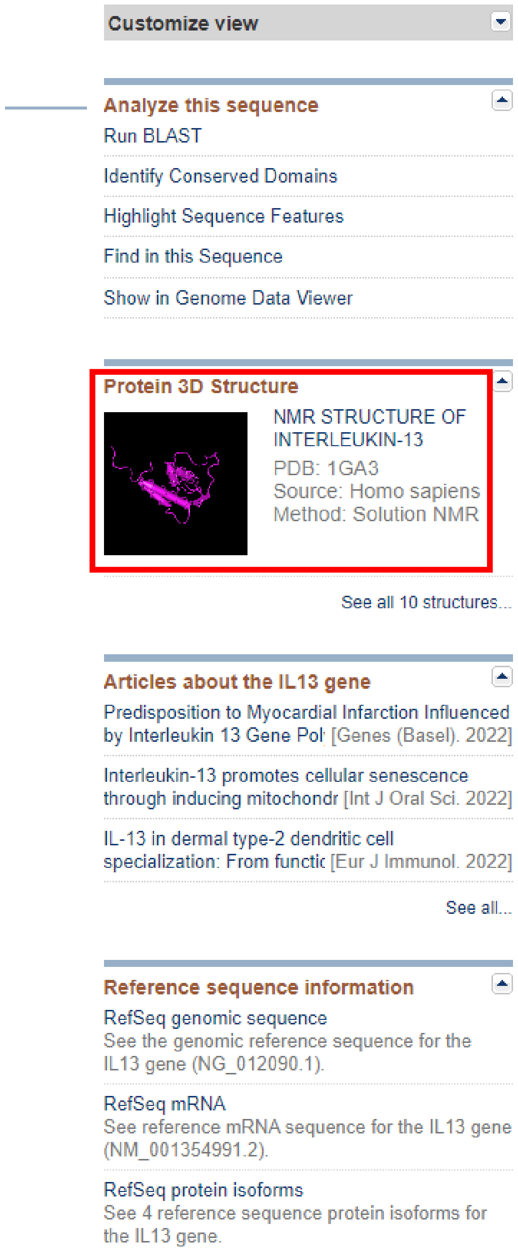
RefSeq genomic sequence
See the genomic reference sequence for the IL13 gene (NG_012090.1).

RefSeq mRNA
See reference mRNA sequence for the IL13 gene (NM_001354991.2).

RefSeq protein isoforms
See 4 reference sequence protein isoforms for the IL13 gene.

ICN3D

- Web-based protein structure viewer that is interactive



Customize view

Analyze this sequence

Run BLAST

Identify Conserved Domains

Highlight Sequence Features

Find in this Sequence

Show in Genome Data Viewer

Protein 3D Structure

NMR STRUCTURE OF INTERLEUKIN-13

PDB: 1GA3

Source: Homo sapiens

Method: Solution NMR

See all 10 structures...

Articles about the IL13 gene

Predisposition to Myocardial Infarction Influenced by Interleukin 13 Gene Pol [Genes (Basel). 2022]

Interleukin-13 promotes cellular senescence through inducing mitochondr [Int J Oral Sci. 2022]

IL-13 in dermal type-2 dendritic cell specialization: From functic [Eur J Immunol. 2022]

See all...

Reference sequence information

RefSeq genomic sequence

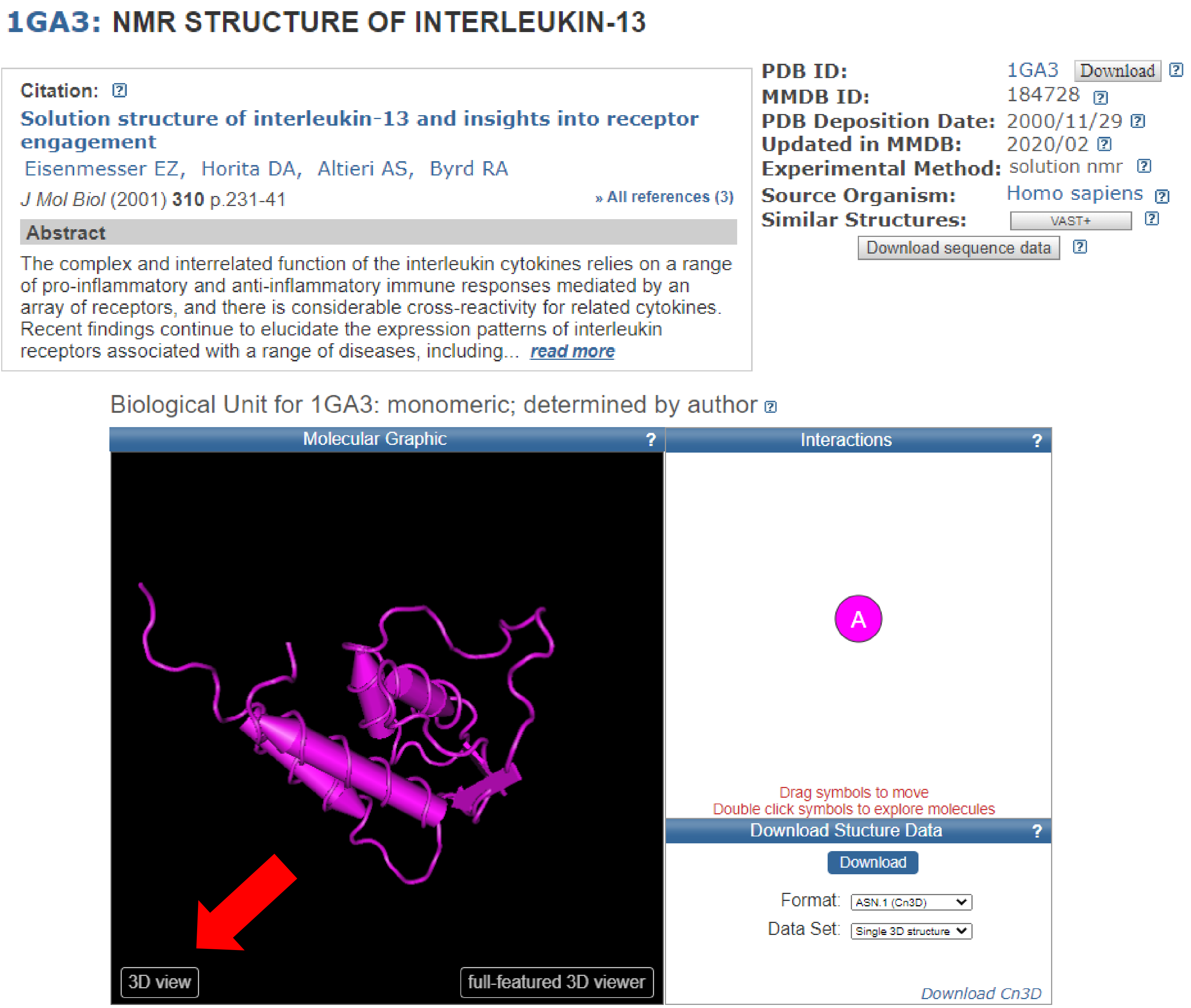
See the genomic reference sequence for the IL13 gene (NG_012090.1).

RefSeq mRNA

See reference mRNA sequence for the IL13 gene (NM_001354991.2).

RefSeq protein isoforms

See 4 reference sequence protein isoforms for the IL13 gene.



1GA3: NMR STRUCTURE OF INTERLEUKIN-13

Citation:

Solution structure of interleukin-13 and insights into receptor engagement

Eisenmesser EZ, Horita DA, Altieri AS, Byrd RA

J Mol Biol (2001) **310** p.231-41

» All references (3)

Abstract

The complex and interrelated function of the interleukin cytokines relies on a range of pro-inflammatory and anti-inflammatory immune responses mediated by an array of receptors, and there is considerable cross-reactivity for related cytokines. Recent findings continue to elucidate the expression patterns of interleukin receptors associated with a range of diseases, including... *read more*

PDB ID:	1GA3
MMDB ID:	184728
PDB Deposition Date:	2000/11/29
Updated in MMDB:	2020/02
Experimental Method:	solution nmr
Source Organism:	Homo sapiens
Similar Structures:	VAST+

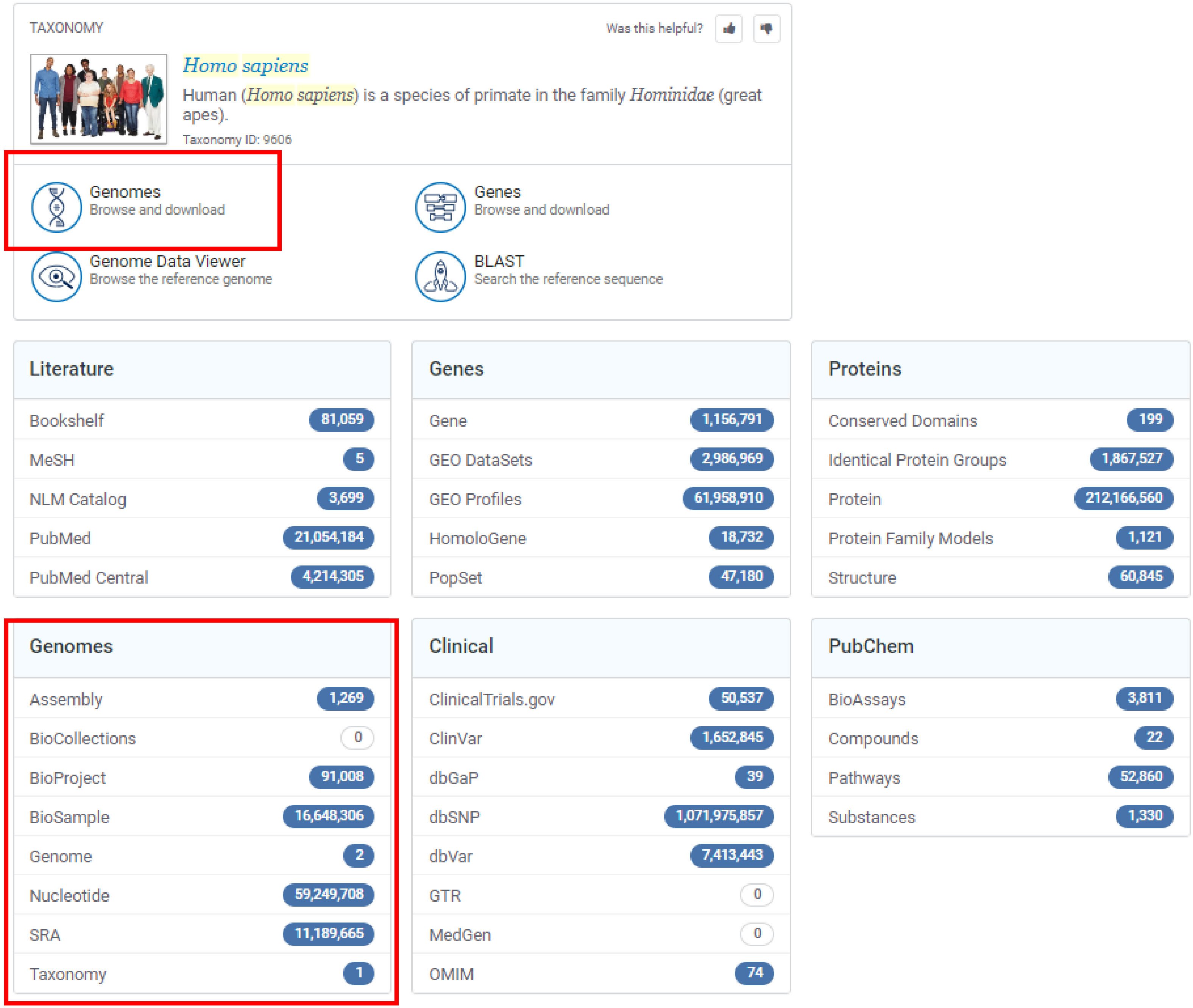
Biological Unit for 1GA3: monomeric; determined by author

Drag symbols to move
Double click symbols to explore molecules

Format: ASN.1 (Cn3D)
Data Set: Single 3D structure

NCBI DATABASES

1. Gene:
    - Gene sequences and annotations used as references for the study of orthologs structure, expression, and evolution
2. Protein:
    - Protein sequences, 3-D structures, and tools for the study of functional protein domains and active sites
3. **Genome:**
    - **Genome sequence assemblies, large-scale functional genomics data, and source biological samples**
4. Clinical:
    - Heritable DNA variations, associations with human pathologies, and clinical diagnostics and treatments



TAXONOMY

Was this helpful?

Homo sapiens

Human (*Homo sapiens*) is a species of primate in the family *Hominidae* (great apes).

Taxonomy ID: 9606

Genomes
Browse and download

Genes
Browse and download

Genome Data Viewer
Browse the reference genome

BLAST
Search the reference sequence

Literature	
Bookshelf	81,059
MeSH	5
NLM Catalog	3,699
PubMed	21,054,184
PubMed Central	4,214,305

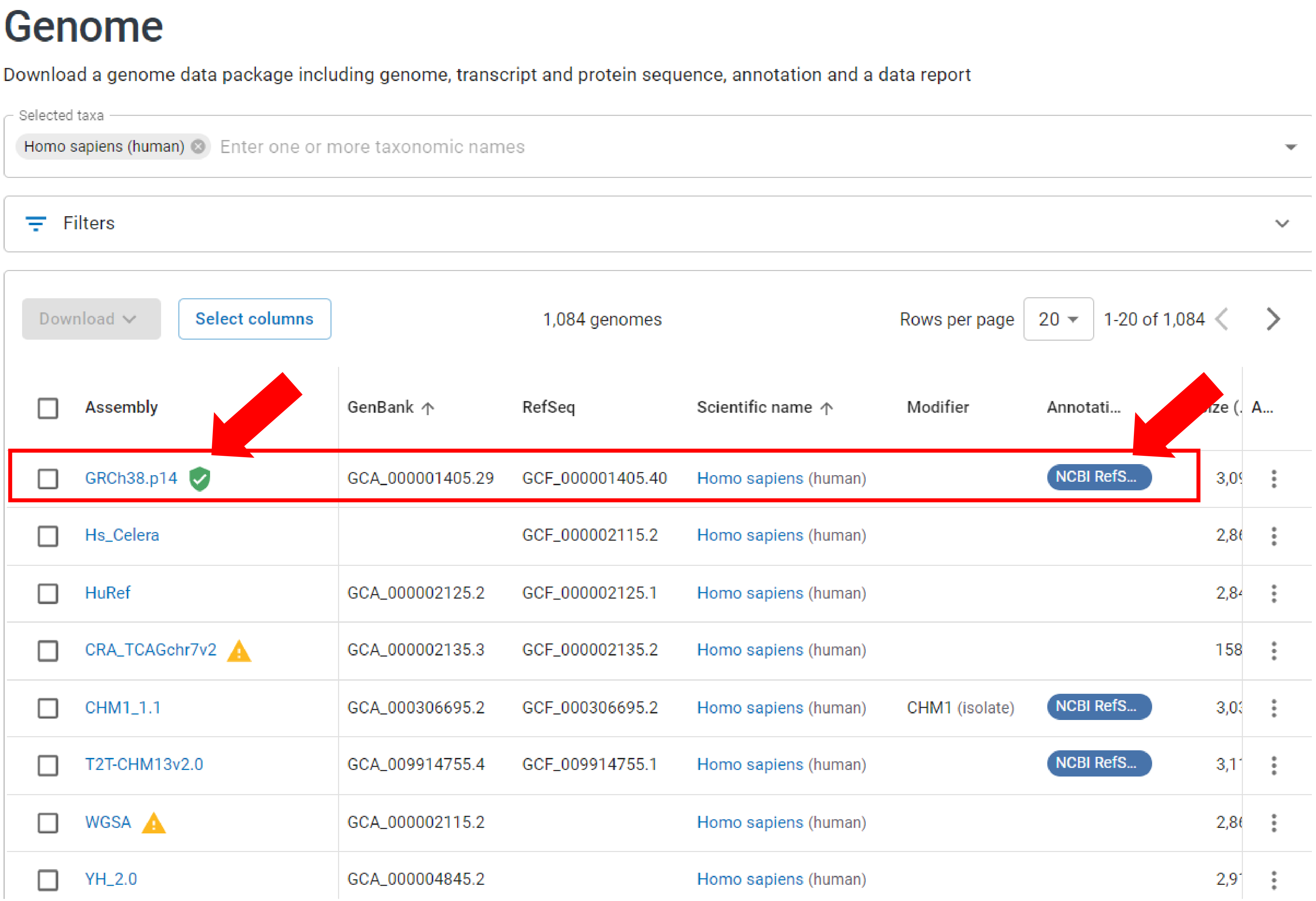
Genes	
Gene	1,156,791
GEO DataSets	2,986,969
GEO Profiles	61,958,910
HomoloGene	18,732
PopSet	47,180

Proteins	
Conserved Domains	199
Identical Protein Groups	1,867,527
Protein	212,166,560
Protein Family Models	1,121
Structure	60,845

Genomes	
Assembly	1,269
BioCollections	0
BioProject	91,008
BioSample	16,648,306
Genome	2
Nucleotide	59,249,708
SRA	11,189,665
Taxonomy	1

Clinical	
ClinicalTrials.gov	50,537
ClinVar	1,652,845
dbGaP	39
dbSNP	1,071,975,857
dbVar	7,413,443
GTR	0
MedGen	0
OMIM	74

PubChem	
BioAssays	3,811
Compounds	22
Pathways	52,860
Substances	1,330



Genome

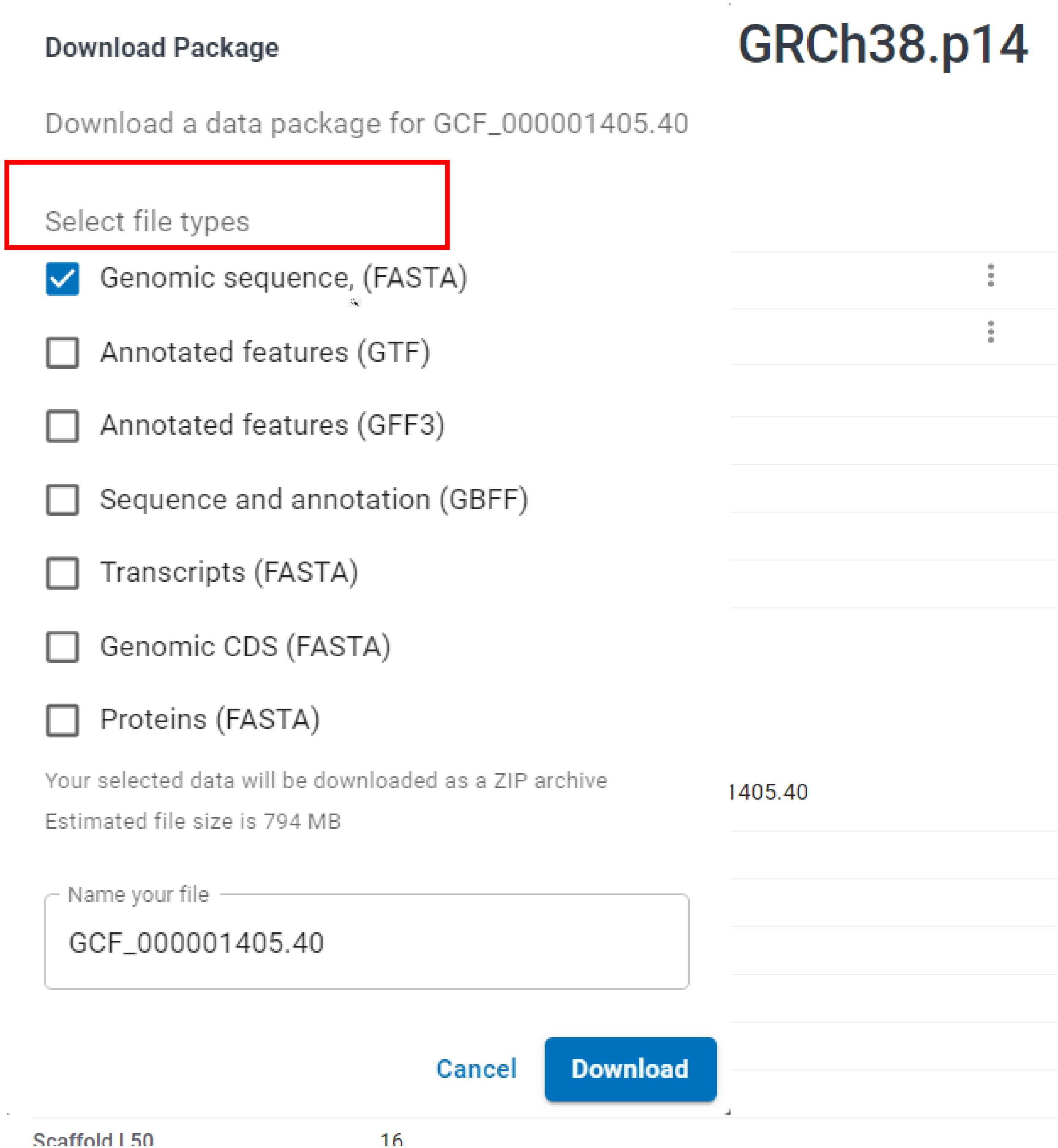
Download a genome data package including genome, transcript and protein sequence, annotation and a data report

Selected taxa: Homo sapiens (human) ⊗ Enter one or more taxonomic names

1,084 genomes

Rows per page 20 1-20 of 1,084

☐	Assembly	GenBank ↑	RefSeq	Scientific name ↑	Modifier	Annotati...	ize (.	A...
☐	GRCh38.p14 ✓	GCA_000001405.29	GCF_000001405.40	Homo sapiens (human)		NCBI RefS...	3,0	⋮
☐	Hs_Celera		GCF_000002115.2	Homo sapiens (human)			2,8	⋮
☐	HuRef	GCA_000002125.2	GCF_000002125.1	Homo sapiens (human)			2,8	⋮
☐	CRA_TCAGchr7v2 ⚠	GCA_000002135.3	GCF_000002135.2	Homo sapiens (human)			158	⋮
☐	CHM1_1.1	GCA_000306695.2	GCF_000306695.2	Homo sapiens (human)	CHM1 (isolate)	NCBI RefS...	3,0	⋮
☐	T2T-CHM13v2.0	GCA_009914755.4	GCF_009914755.1	Homo sapiens (human)		NCBI RefS...	3,1	⋮
☐	WGSA ⚠	GCA_000002115.2		Homo sapiens (human)			2,8	⋮
☐	YH_2.0	GCA_000004845.2		Homo sapiens (human)			2,9	⋮



Download Package

GRCh38.p14

Download a data package for GCF_000001405.40

Select file types

☑ Genomic sequence, (FASTA)

☐ Annotated features (GTF)

☐ Annotated features (GFF3)

☐ Sequence and annotation (GBFF)

☐ Transcripts (FASTA)

☐ Genomic CDS (FASTA)

☐ Proteins (FASTA)

Your selected data will be downloaded as a ZIP archive

Estimated file size is 794 MB

Name your file: GCF_000001405.40

SEQUENCE FEATURES AND ANNOTATION FILE

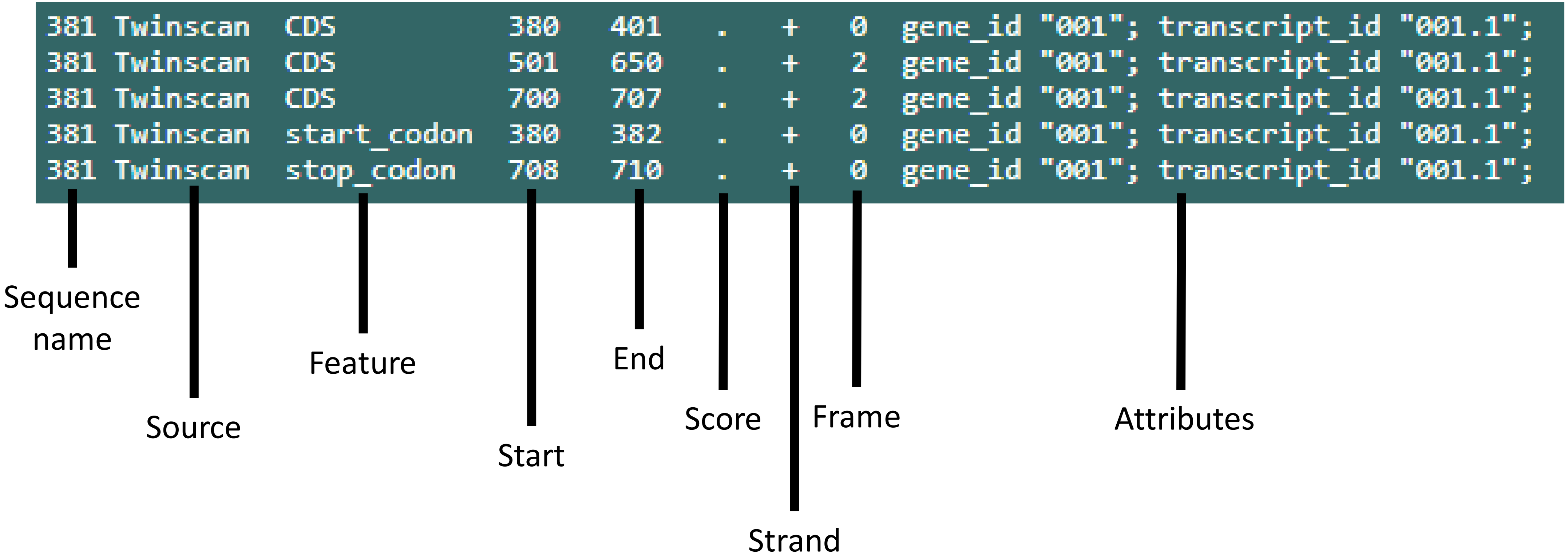
GFF3 – General Feature Format or Gene Finding Format

GTF – Gene Transfer Format

- Both are TAB-separated text files

Fields	Description
Sequence name	The name of the sequence (Commonly chromosome ID or contig ID)
Source	indicating where the annotations came from (commonly predicted program or public database)
Feature	"CDS", "start_codon", "stop_codon"
Start	Integer start coordinates of the feature
End	Integer end coordinates of the feature
Score	indicates a degree of confidence in the feature's existence and coordinates
Strand	Single character that indicates the strand of the feature; it can assume the values of "+" (positive, or 5'->3'), "-", (negative, or 3'->5'), "." (undetermined).
Frame	it can be either one of 0, 1, 2 (for CDS features) or "." (for everything else).
Attribute	All the other information pertaining to this feature.

GTF AND GFF3 GENERAL STRUCTURE



```
381 Twinscan  CDS          380   401   .   +   0  gene_id "001"; transcript_id "001.1";
381 Twinscan  CDS          501   650   .   +   2  gene_id "001"; transcript_id "001.1";
381 Twinscan  CDS          700   707   .   +   2  gene_id "001"; transcript_id "001.1";
381 Twinscan  start_codon  380   382   .   +   0  gene_id "001"; transcript_id "001.1";
381 Twinscan  stop_codon   708   710   .   +   0  gene_id "001"; transcript_id "001.1";
```

BIOPROJECTS AND BIOSAMPLES

Bioprojects https://www.ncbi.nlm.nih.gov/bioproject/

- a collection of biological data related to a single initiative, originating from a single organization or from a consortium
- a single place to find all data generated for that specific project
- “the raw data”

Biosamples https://www.ncbi.nlm.nih.gov/biosample/

- contains descriptions of biological source materials used in experimental assays
- “the metadata”

SEQUENCE READ ARCHIVE (SRA)

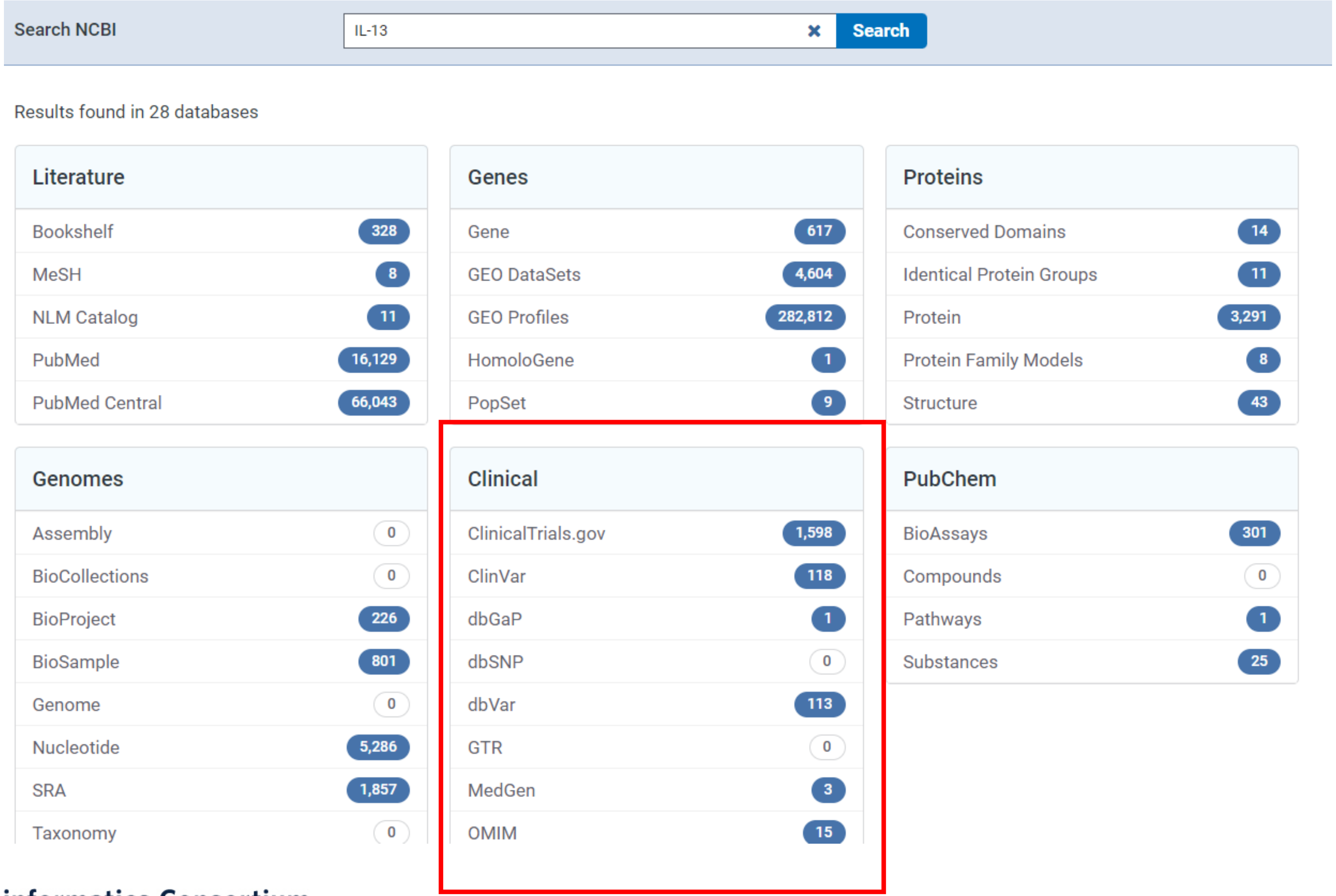
- the largest publicly available repository of high throughput sequencing data
- stores raw sequencing data and alignment information
- accepts data from all kinds of sequencing projects

https://www.ncbi.nlm.nih.gov/sra

NCBI DATABASES

1. Gene:
   - Gene sequences and annotations used as references for the study of orthologs structure, expression, and evolution
2. Protein:
   - Protein sequences, 3-D structures, and tools for the study of functional protein domains and active sites
3. Genome:
   - Genome sequence assemblies, large-scale functional genomics data, and source biological samples
4. **Clinical:**
   - **Heritable DNA variations, associations with human pathologies, and clinical diagnostics and treatments**

CLINICAL



Results found in 28 databases

Literature	
Bookshelf	328
MeSH	8
NLM Catalog	11
PubMed	16,129
PubMed Central	66,043

Genes	
Gene	617
GEO DataSets	4,604
GEO Profiles	282,812
HomoloGene	1
PopSet	9

Proteins	
Conserved Domains	14
Identical Protein Groups	11
Protein	3,291
Protein Family Models	8
Structure	43

Genomes	
Assembly	0
BioCollections	0
BioProject	226
BioSample	801
Genome	0
Nucleotide	5,286
SRA	1,857
Taxonomy	0

Clinical	
ClinicalTrials.gov	1,598
ClinVar	118
dbGaP	1
dbSNP	0
dbVar	113
GTR	0
MedGen	3
OMIM	15

PubChem	
BioAssays	301
Compounds	0
Pathways	1
Substances	25

DBSNP

https://www.ncbi.nlm.nih.gov/snp/

- human SNP (single nucleotide variations)
- microsatellites
- small-scale insertions and deletions along with publication
- population frequency
- genomic and RefSeq mapping information for both common variations and clinical mutations

NCBI TOOLS

BASIC LOCAL ALIGNMENT SEARCH TOOL (BLAST)

https://blast.ncbi.nlm.nih.gov/Blast.cgi

Overview:

- One of the most widely used sequence similarity search tool
- It provides non-change similarities between biological sequences (DNA, RNA, protein…)

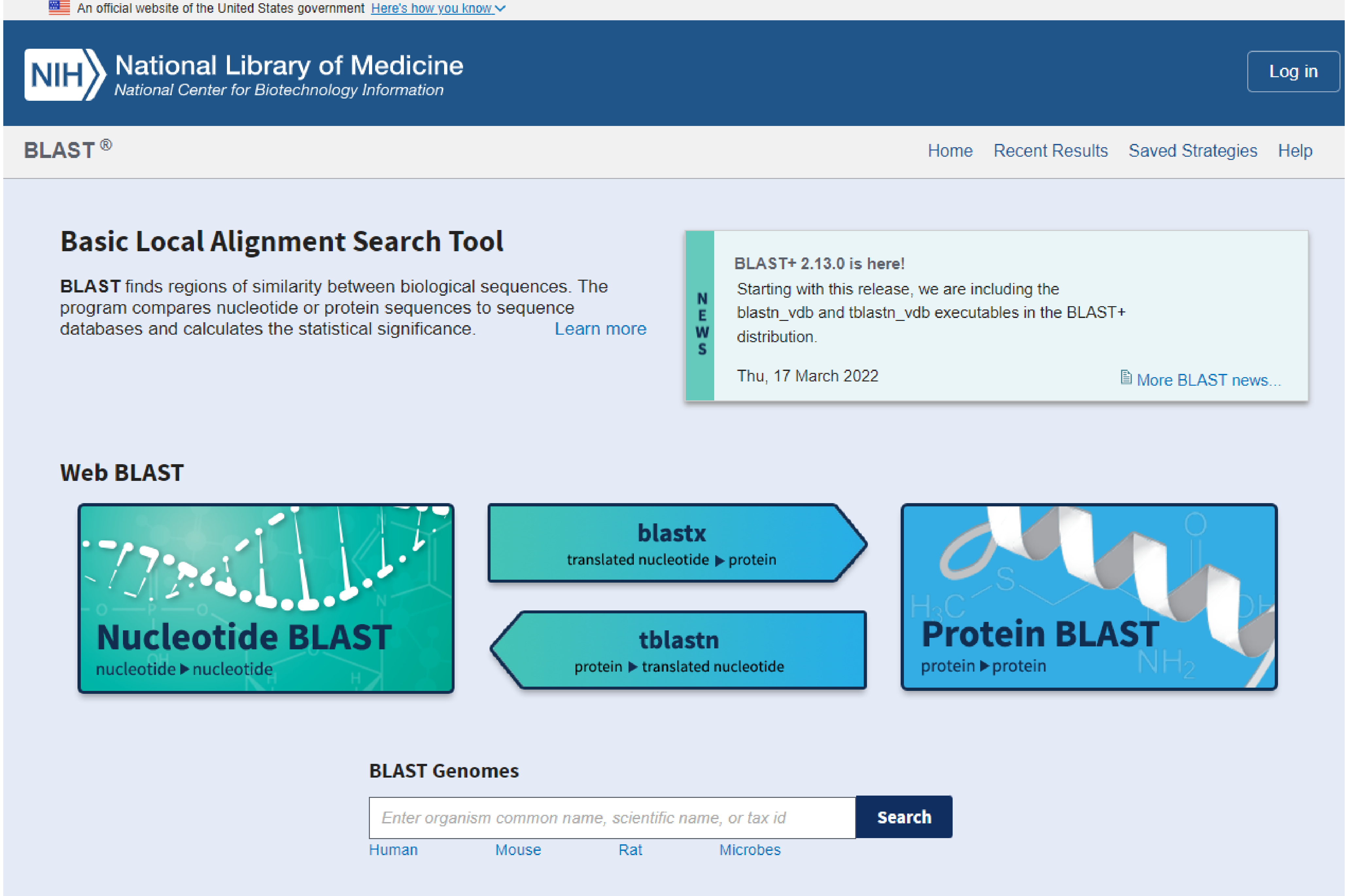
Purpose:

- The original purpose of BLAST is to study the homology sequences that are descended from a common ancestor
- Today people run BLAST to simply identify and annotate the sequences (*e.g.* possible contamination)

Different types of BLAST:

- BlastN – nucleotide to nucleotide
- BlastP– protein to protein
- BlastX – translated nucleotide to protein
- TBlastN – protein to translated nucleotide
- TBlastX– translated nucleotide to translated nucleotide

https://ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_BLASTGuide.pdf



Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. Learn more

NEWS

BLAST+ 2.13.0 is here!

Starting with this release, we are including the blastn_vdb and tblastn_vdb executables in the BLAST+ distribution.

Thu, 17 March 2022

More BLAST news...

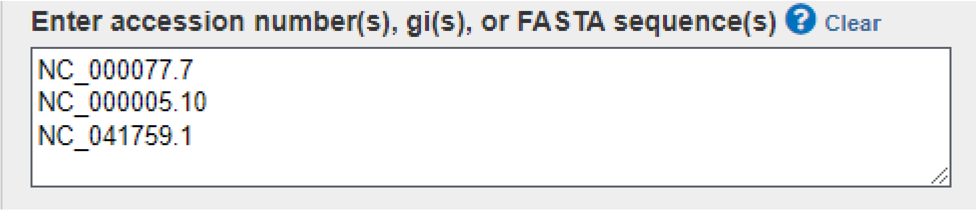
Web BLAST

BLAST Genomes

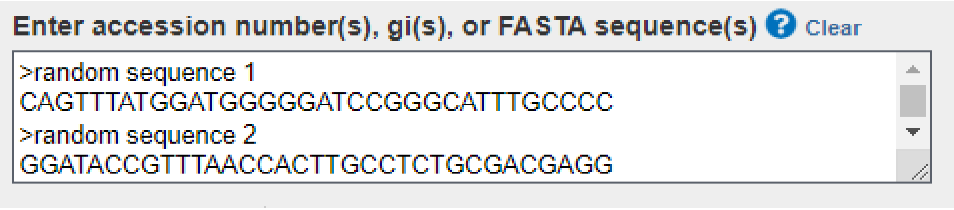
BASIC LOCAL ALIGNMENT SEARCH TOOL (BLAST)

Input:

- Sequence of interest (FASTA format, or accession #)

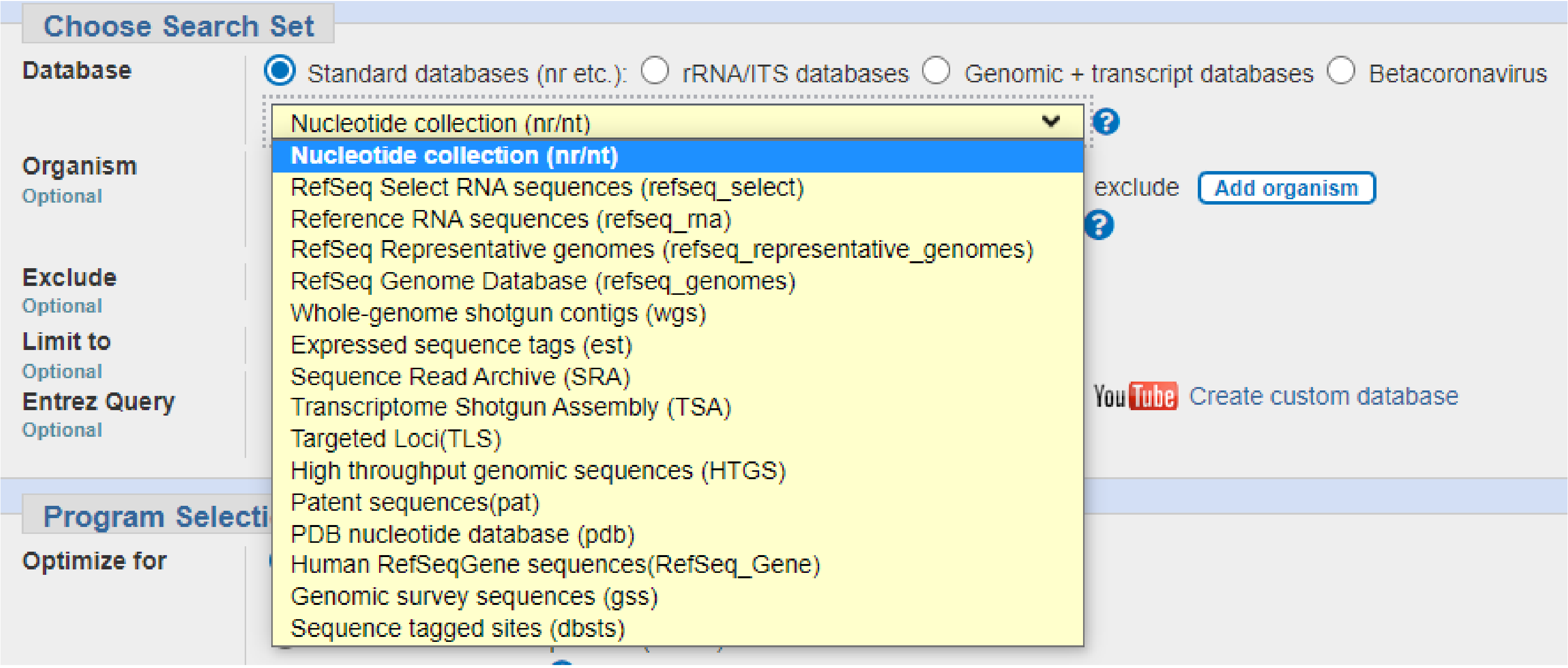


```
NC_000077.7
NC_000005.10
NC_041759.1
```



```
>random sequence 1
CAGTTTATGGATGGGGATCCGGGCATTTGCCCC
>random sequence 2
GGATACCGTTTAACCACTTGCCTCTGCGACGAGG
```

- Reference database



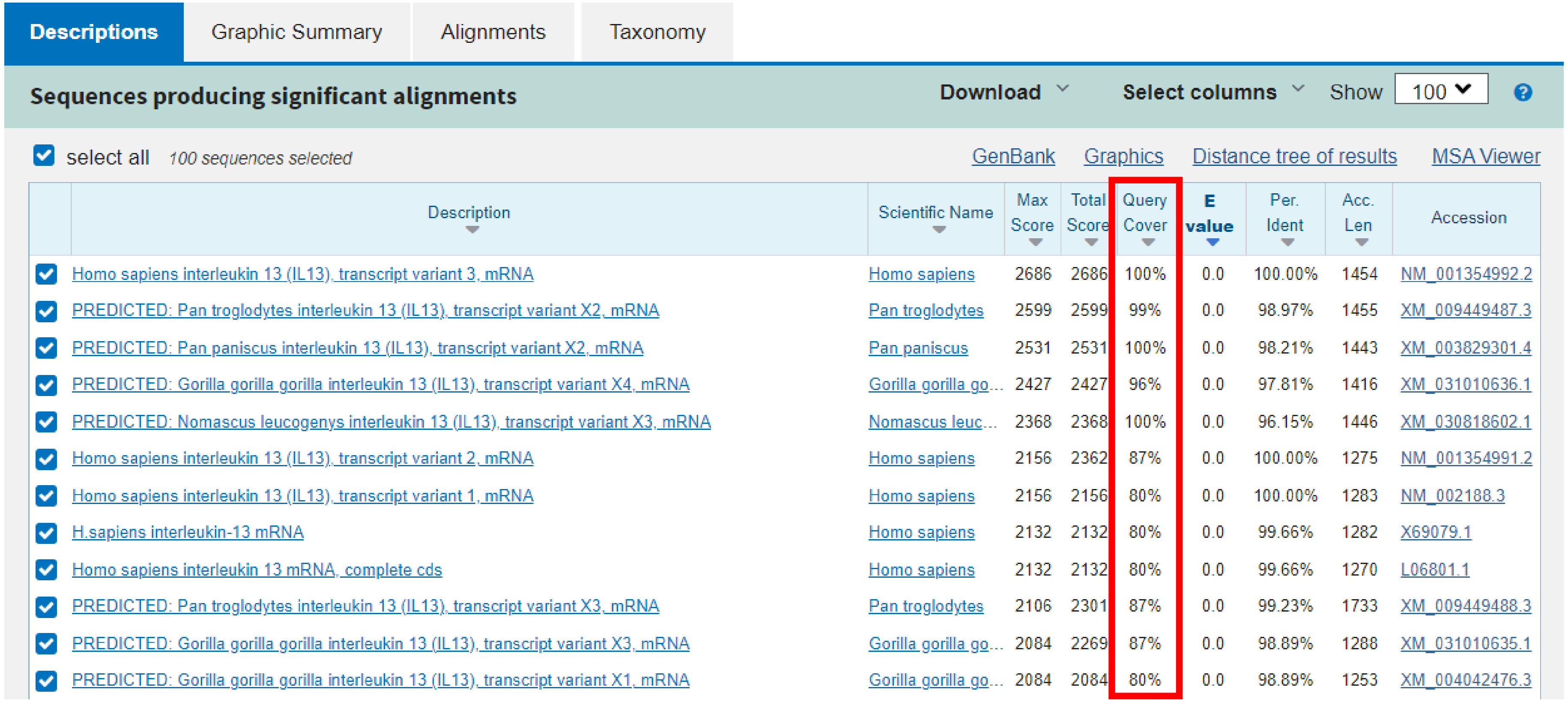
Database: ● Standard databases (nr etc.):

Nucleotide collection (nr/nt)

BASIC LOCAL ALIGNMENT SEARCH TOOL (BLAST)

Query Coverage:

- the % of the sequence length (query) that aligns with the NCBI hit (subject)
- less query coverage % = less alignment



Sequences producing significant alignments

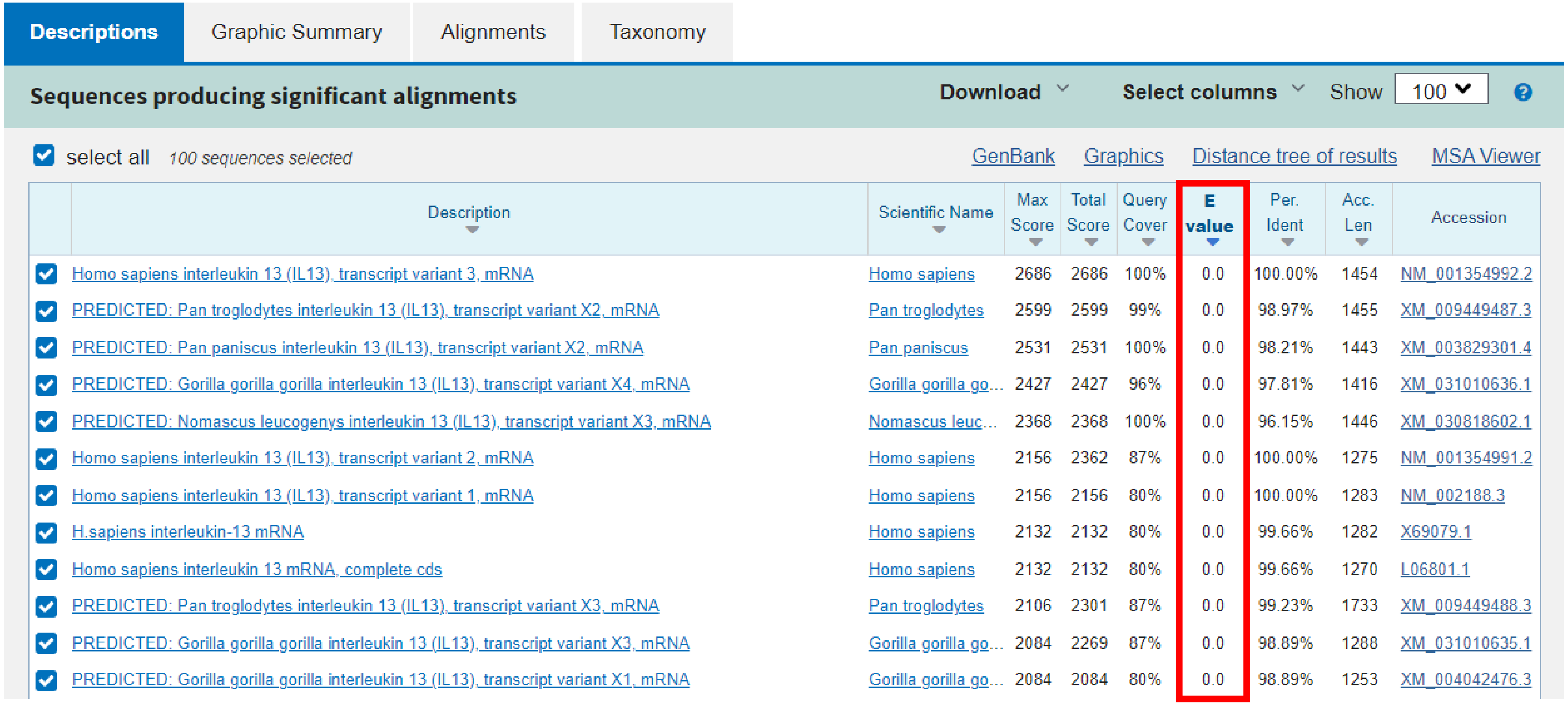
☑ select all *100 sequences selected*

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Homo sapiens interleukin 13 (IL13), transcript variant 3, mRNA	Homo sapiens	2686	2686	100%	0.0	100.00%	1454	NM_001354992.2
☑	PREDICTED: Pan troglodytes interleukin 13 (IL13), transcript variant X2, mRNA	Pan troglodytes	2599	2599	99%	0.0	98.97%	1455	XM_009449487.3
☑	PREDICTED: Pan paniscus interleukin 13 (IL13), transcript variant X2, mRNA	Pan paniscus	2531	2531	100%	0.0	98.21%	1443	XM_003829301.4
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X4, mRNA	Gorilla gorilla go...	2427	2427	96%	0.0	97.81%	1416	XM_031010636.1
☑	PREDICTED: Nomascus leucogenys interleukin 13 (IL13), transcript variant X3, mRNA	Nomascus leuc...	2368	2368	100%	0.0	96.15%	1446	XM_030818602.1
☑	Homo sapiens interleukin 13 (IL13), transcript variant 2, mRNA	Homo sapiens	2156	2362	87%	0.0	100.00%	1275	NM_001354991.2
☑	Homo sapiens interleukin 13 (IL13), transcript variant 1, mRNA	Homo sapiens	2156	2156	80%	0.0	100.00%	1283	NM_002188.3
☑	H.sapiens interleukin-13 mRNA	Homo sapiens	2132	2132	80%	0.0	99.66%	1282	X69079.1
☑	Homo sapiens interleukin 13 mRNA, complete cds	Homo sapiens	2132	2132	80%	0.0	99.66%	1270	L06801.1
☑	PREDICTED: Pan troglodytes interleukin 13 (IL13), transcript variant X3, mRNA	Pan troglodytes	2106	2301	87%	0.0	99.23%	1733	XM_009449488.3
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X3, mRNA	Gorilla gorilla go...	2084	2269	87%	0.0	98.89%	1288	XM_031010635.1
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X1, mRNA	Gorilla gorilla go...	2084	2084	80%	0.0	98.89%	1253	XM_004042476.3

BASIC LOCAL ALIGNMENT SEARCH TOOL (BLAST)

E value (expect value):

- The number of alignment hits one would “expect” to see by chance against a particular database
- Ranges from zero to the size of the reference database
- Lower e-value = more signficant



Sequences producing significant alignments

Download | Select columns | Show 100

☑ select all *100 sequences selected*

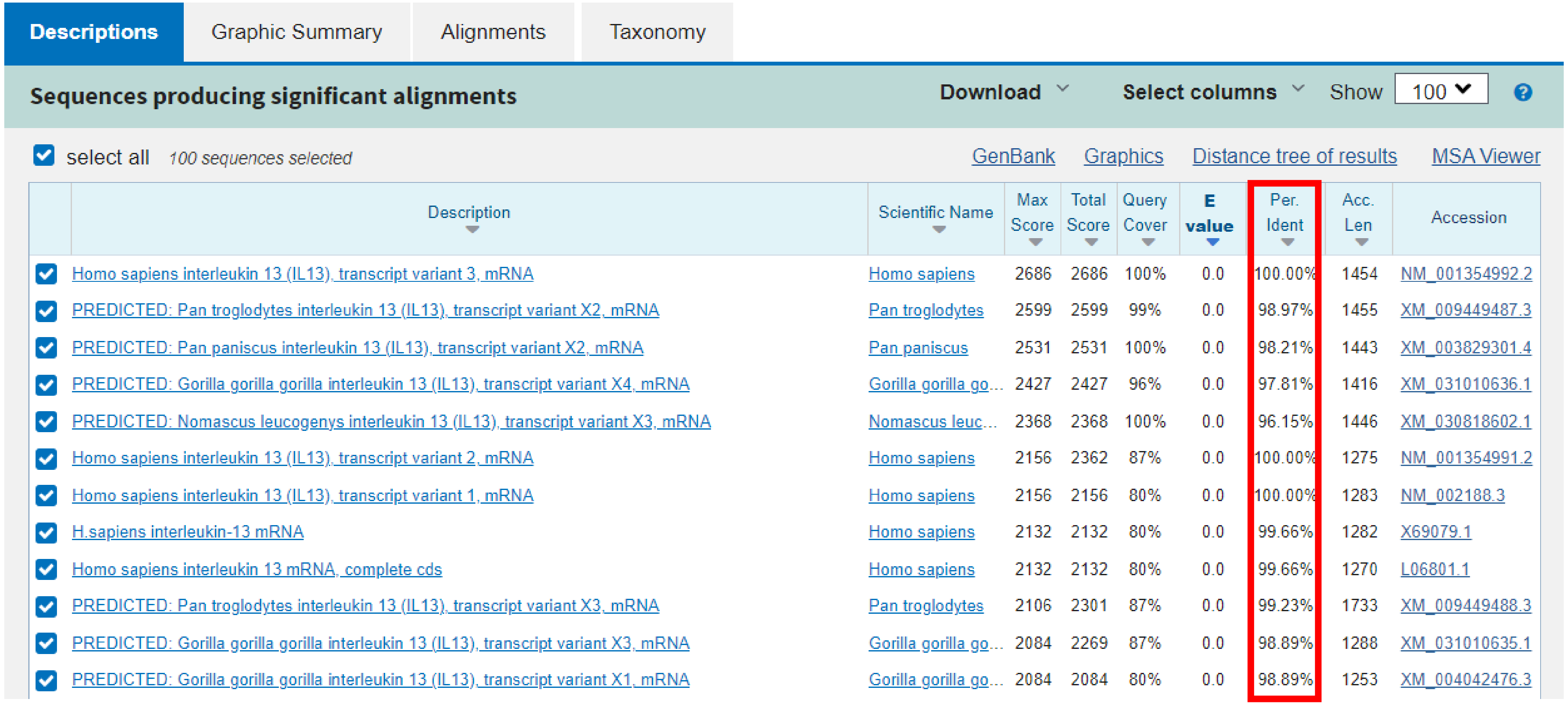
GenBank | Graphics | Distance tree of results | MSA Viewer

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Homo sapiens interleukin 13 (IL13), transcript variant 3, mRNA	Homo sapiens	2686	2686	100%	0.0	100.00%	1454	NM_001354992.2
☑	PREDICTED: Pan troglodytes interleukin 13 (IL13), transcript variant X2, mRNA	Pan troglodytes	2599	2599	99%	0.0	98.97%	1455	XM_009449487.3
☑	PREDICTED: Pan paniscus interleukin 13 (IL13), transcript variant X2, mRNA	Pan paniscus	2531	2531	100%	0.0	98.21%	1443	XM_003829301.4
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X4, mRNA	Gorilla gorilla go...	2427	2427	96%	0.0	97.81%	1416	XM_031010636.1
☑	PREDICTED: Nomascus leucogenys interleukin 13 (IL13), transcript variant X3, mRNA	Nomascus leuc...	2368	2368	100%	0.0	96.15%	1446	XM_030818602.1
☑	Homo sapiens interleukin 13 (IL13), transcript variant 2, mRNA	Homo sapiens	2156	2362	87%	0.0	100.00%	1275	NM_001354991.2
☑	Homo sapiens interleukin 13 (IL13), transcript variant 1, mRNA	Homo sapiens	2156	2156	80%	0.0	100.00%	1283	NM_002188.3
☑	H.sapiens interleukin-13 mRNA	Homo sapiens	2132	2132	80%	0.0	99.66%	1282	X69079.1
☑	Homo sapiens interleukin 13 mRNA, complete cds	Homo sapiens	2132	2132	80%	0.0	99.66%	1270	L06801.1
☑	PREDICTED: Pan troglodytes interleukin 13 (IL13), transcript variant X3, mRNA	Pan troglodytes	2106	2301	87%	0.0	99.23%	1733	XM_009449488.3
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X3, mRNA	Gorilla gorilla go...	2084	2269	87%	0.0	98.89%	1288	XM_031010635.1
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X1, mRNA	Gorilla gorilla go...	2084	2084	80%	0.0	98.89%	1253	XM_004042476.3

BASIC LOCAL ALIGNMENT SEARCH TOOL (BLAST)

Percent Identity:

- the % of bases in the query that are identical to the subject
- However, low % identity still could be a real hit
- Need to take the E value and query coverage into account to draw conclusion



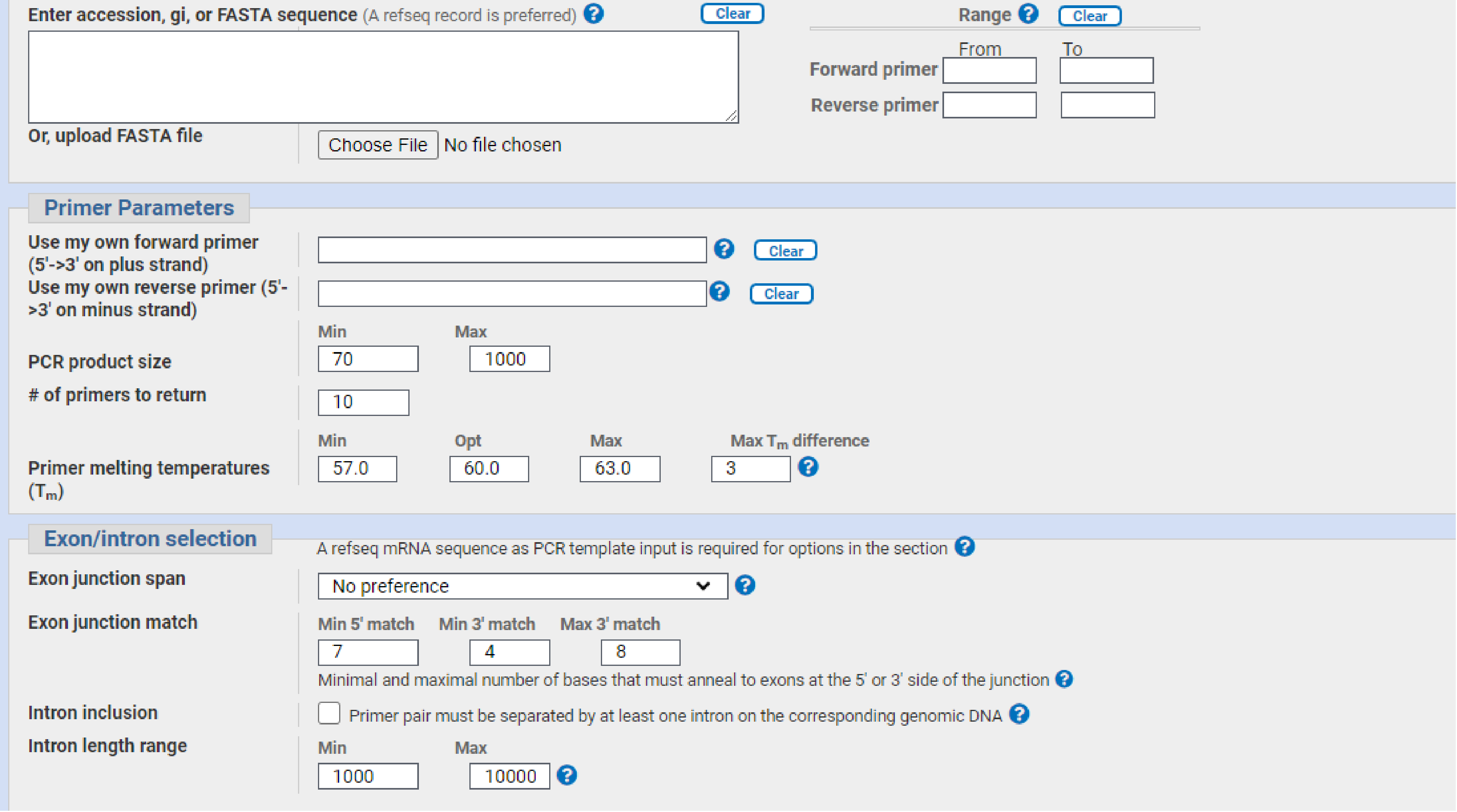
Sequences producing significant alignments

☑ select all *100 sequences selected*

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Homo sapiens interleukin 13 (IL13), transcript variant 3, mRNA	Homo sapiens	2686	2686	100%	0.0	100.00%	1454	NM_001354992.2
☑	PREDICTED: Pan troglodytes interleukin 13 (IL13), transcript variant X2, mRNA	Pan troglodytes	2599	2599	99%	0.0	98.97%	1455	XM_009449487.3
☑	PREDICTED: Pan paniscus interleukin 13 (IL13), transcript variant X2, mRNA	Pan paniscus	2531	2531	100%	0.0	98.21%	1443	XM_003829301.4
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X4, mRNA	Gorilla gorilla go...	2427	2427	96%	0.0	97.81%	1416	XM_031010636.1
☑	PREDICTED: Nomascus leucogenys interleukin 13 (IL13), transcript variant X3, mRNA	Nomascus leuc...	2368	2368	100%	0.0	96.15%	1446	XM_030818602.1
☑	Homo sapiens interleukin 13 (IL13), transcript variant 2, mRNA	Homo sapiens	2156	2362	87%	0.0	100.00%	1275	NM_001354991.2
☑	Homo sapiens interleukin 13 (IL13), transcript variant 1, mRNA	Homo sapiens	2156	2156	80%	0.0	100.00%	1283	NM_002188.3
☑	H.sapiens interleukin-13 mRNA	Homo sapiens	2132	2132	80%	0.0	99.66%	1282	X69079.1
☑	Homo sapiens interleukin 13 mRNA, complete cds	Homo sapiens	2132	2132	80%	0.0	99.66%	1270	L06801.1
☑	PREDICTED: Pan troglodytes interleukin 13 (IL13), transcript variant X3, mRNA	Pan troglodytes	2106	2301	87%	0.0	99.23%	1733	XM_009449488.3
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X3, mRNA	Gorilla gorilla go...	2084	2269	87%	0.0	98.89%	1288	XM_031010635.1
☑	PREDICTED: Gorilla gorilla gorilla interleukin 13 (IL13), transcript variant X1, mRNA	Gorilla gorilla go...	2084	2084	80%	0.0	98.89%	1253	XM_004042476.3

PRIMER-BLAST

- Allows users to design new target-specific PCR primers or check primer specificity



Enter accession, gi, or FASTA sequence (A refseq record is preferred)

Or, upload FASTA file: No file chosen

Range

	From	To
Forward primer		
Reverse primer		

Primer Parameters

Use my own forward primer (5'->3' on plus strand):

Use my own reverse primer (5'->3' on minus strand):

	Min	Opt	Max	Max T_m difference
PCR product size	70		1000	
# of primers to return	10			
Primer melting temperatures (T_m)	57.0	60.0	63.0	3

Exon/intron selection

A refseq mRNA sequence as PCR template input is required for options in the section

Exon junction span: No preference

Exon junction match: Min 5' match 7, Min 3' match 4, Max 3' match 8

Minimal and maximal number of bases that must anneal to exons at the 5' or 3' side of the junction

Intron inclusion: ☐ Primer pair must be separated by at least one intron on the corresponding genomic DNA

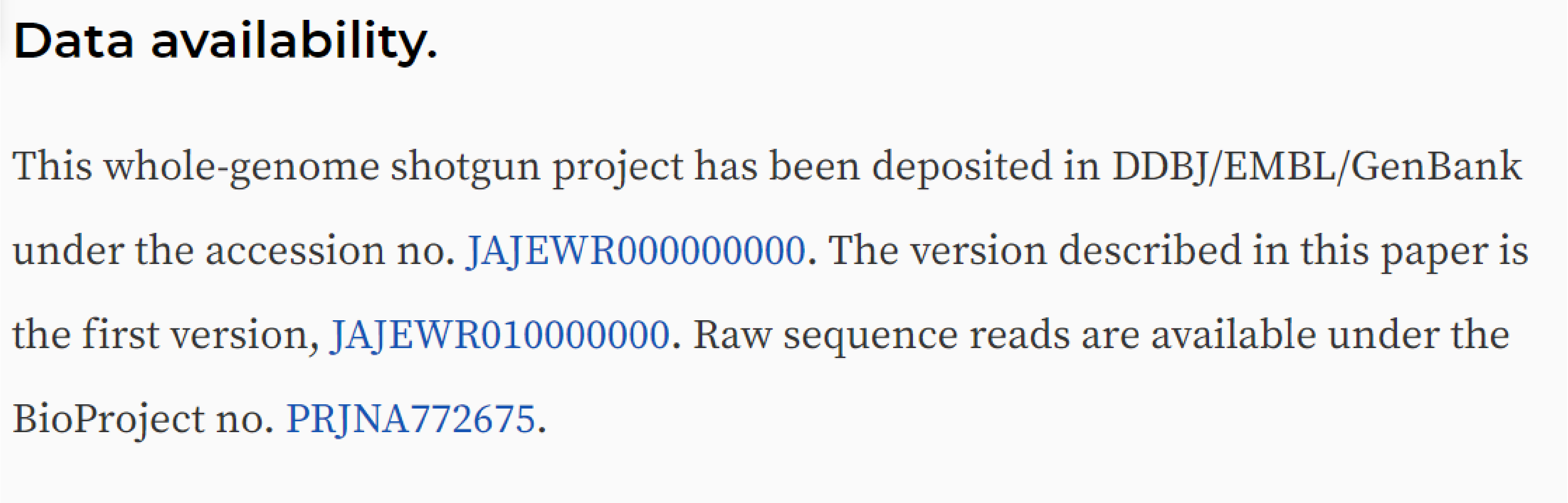
Intron length range: Min 1000, Max 10000

https://www.ncbi.nlm.nih.gov/tools/primer-blast/

DATA SUBMISSION TO NCBI

WHERE TO ACCESS SUBMITTED DATA FROM PUBLICATIONS

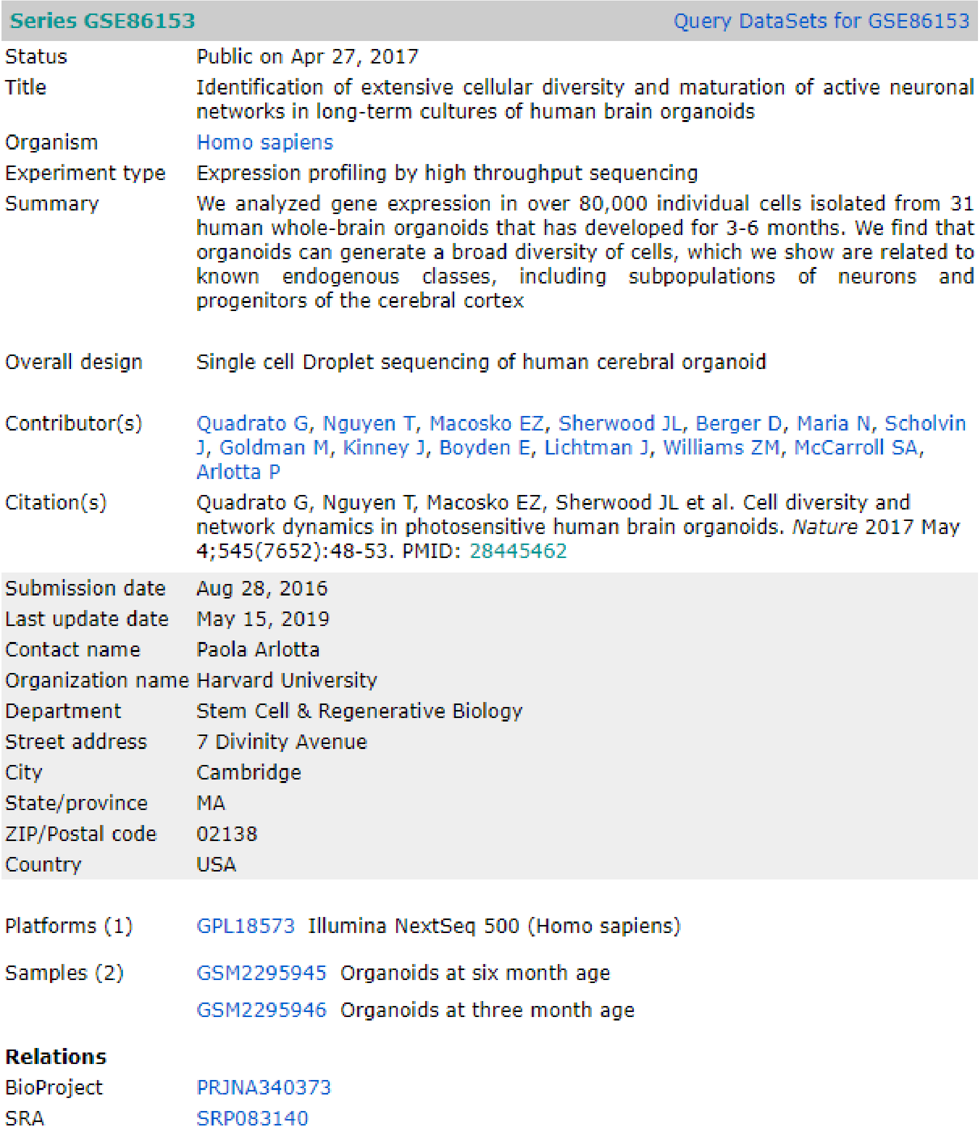
- Publications should have a data availability section (or similar) to reference where the sequence information was deposited
- Search NCBI based on the unique identifiers given



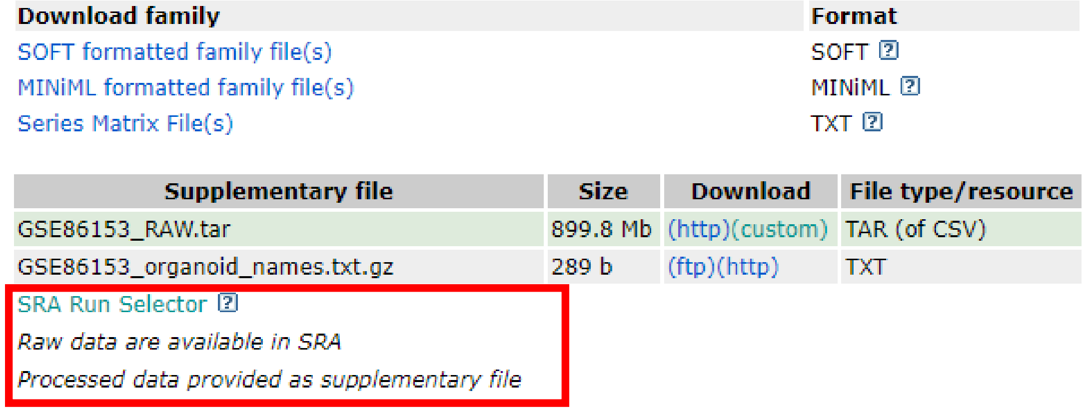
Data availability.

This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no. JAJEWR000000000. The version described in this paper is the first version, JAJEWR010000000. Raw sequence reads are available under the BioProject no. PRJNA772675.

WHERE TO ACCESS SUBMITTED DATA FROM PUBLICATIONS



Series GSE86153	Query DataSets for GSE86153
Status	Public on Apr 27, 2017
Title	Identification of extensive cellular diversity and maturation of active neuronal networks in long-term cultures of human brain organoids
Organism	Homo sapiens
Experiment type	Expression profiling by high throughput sequencing
Summary	We analyzed gene expression in over 80,000 individual cells isolated from 31 human whole-brain organoids that has developed for 3-6 months. We find that organoids can generate a broad diversity of cells, which we show are related to known endogenous classes, including subpopulations of neurons and progenitors of the cerebral cortex
Overall design	Single cell Droplet sequencing of human cerebral organoid
Contributor(s)	Quadrato G, Nguyen T, Macosko EZ, Sherwood JL, Berger D, Maria N, Scholvin J, Goldman M, Kinney J, Boyden E, Lichtman J, Williams ZM, McCarroll SA, Arlotta P
Citation(s)	Quadrato G, Nguyen T, Macosko EZ, Sherwood JL et al. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* 2017 May 4;545(7652):48-53. PMID: 28445462
Submission date	Aug 28, 2016
Last update date	May 15, 2019
Contact name	Paola Arlotta
Organization name	Harvard University
Department	Stem Cell & Regenerative Biology
Street address	7 Divinity Avenue
City	Cambridge
State/province	MA
ZIP/Postal code	02138
Country	USA
Platforms (1)	GPL18573 Illumina NextSeq 500 (Homo sapiens)
Samples (2)	GSM2295945 Organoids at six month age GSM2295946 Organoids at three month age
Relations	
BioProject	PRJNA340373
SRA	SRP083140



Download family	Format
SOFT formatted family file(s)	SOFT ?
MINiML formatted family file(s)	MINiML ?
Series Matrix File(s)	TXT ?

Supplementary file	Size	Download	File type/resource
GSE86153_RAW.tar	899.8 Mb	(http)(custom)	TAR (of CSV)
GSE86153_organoid_names.txt.gz	289 b	(ftp)(http)	TXT

SRA Run Selector ?

Raw data are available in SRA

Processed data provided as supplementary file

WHERE TO SUBMIT DATA FOR YOUR PUBLICATION

Most journals require DNA and amino acid sequences to be submitted to a public sequence repository

When you submit a sequence to GenBank, you will be provided an accession number to be included in your manuscript

Several options for preparing and submitting data to GenBank:

1. Web-based submission tools:
    - Submission Portal
    - BankIT
2. Submission preparation tools:
    - Table2asn
    - Genome Workbench

WEB-BASED SUBMISSION TOOLS

Submission Portal (https://submit.ncbi.nlm.nih.gov/)

- Has specialized wizards for specific data types to streamline the submission procession
  - SARS CoV-2, Influenza, Norovirus, Dengue
  - Prokaryotic rRNA
  - Organellar rRNA
  - Eukaryotic rRNA-ITS
  - Metazoan COX1
- Has other wizards for submitting:
  - Large genomes (Genomes-WGS)
  - Transcriptomes (TSA)
  - RNA-Seq, Chip-Seq, other gene expression and epigenomic datasets (GEO)
  - Raw sequence reads (SRA)

WEB-BASED SUBMISSION TOOLS

BankIT (https://www.ncbi.nlm.nih.gov/WebSub/)

- Can be used to submit most types of data:
    - mRNA or genomic DNA
    - Nuclear or organellar sequences
    - Single genes
    - Complete plasmids/organelles

WHERE TO SUBMIT DATA FOR YOUR PUBLICATION

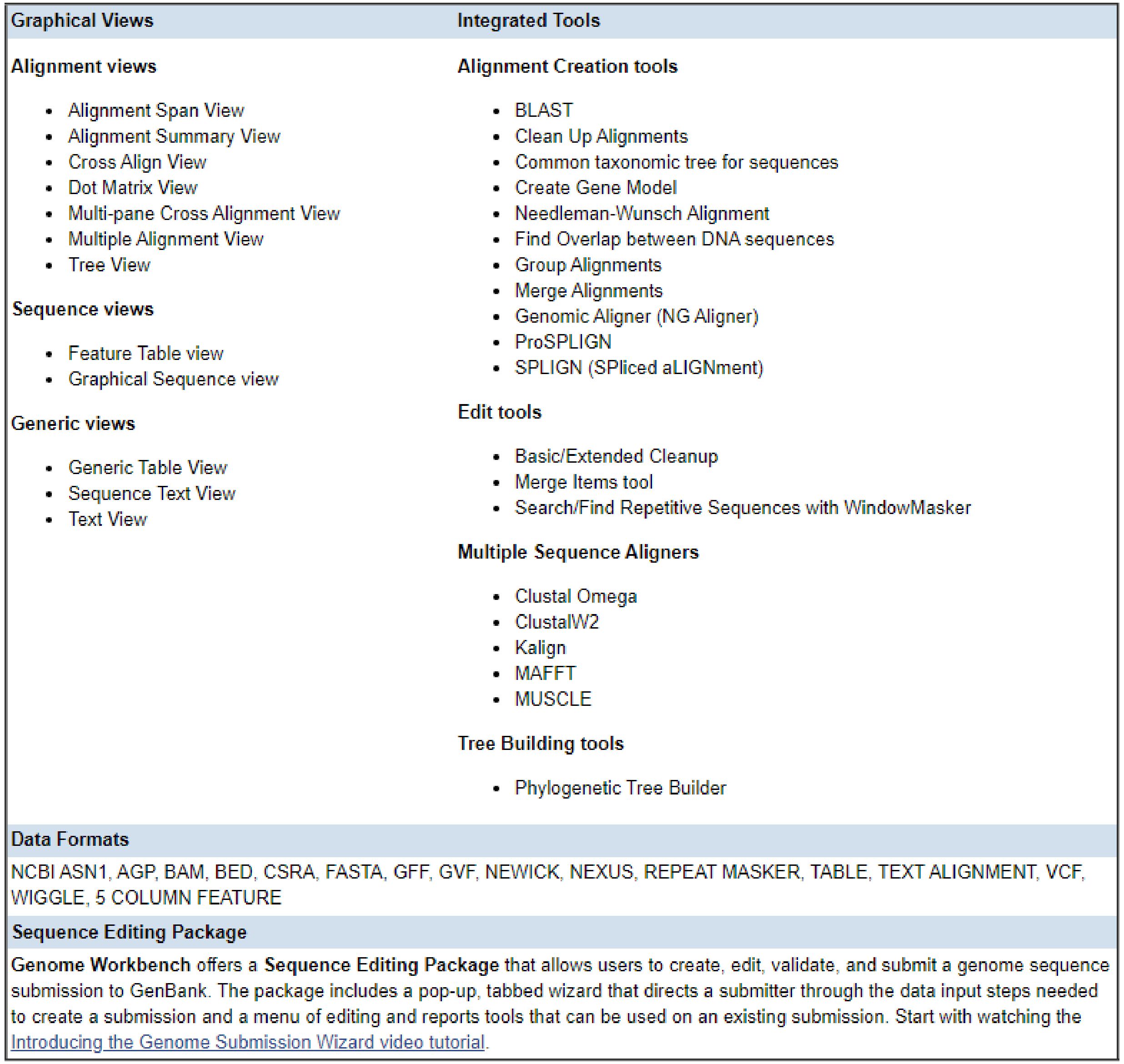
- Most journals require DNA and amino acid sequences to be submitted to a public sequence repository

- When you submit a sequence to GenBank, you will be provided an accession number to be included in your manuscript

- Several options for preparing and submitting data to GenBank:
  1. Web-based submission tools:
     - BankIt
     - Submission Portal
  2. **Submission preparation tools:**
     - Table2asn – command-line program to create sequence records
     - **Genome Workbench**

GENOME WORKBENCH

- Genome Workbench offers a set of integrated tools for studying and analyzing genetic data



Graphical Views

Alignment views

- Alignment Span View
- Alignment Summary View
- Cross Align View
- Dot Matrix View
- Multi-pane Cross Alignment View
- Multiple Alignment View
- Tree View

Sequence views

- Feature Table view
- Graphical Sequence view

Generic views

- Generic Table View
- Sequence Text View
- Text View

Integrated Tools

Alignment Creation tools

- BLAST
- Clean Up Alignments
- Common taxonomic tree for sequences
- Create Gene Model
- Needleman-Wunsch Alignment
- Find Overlap between DNA sequences
- Group Alignments
- Merge Alignments
- Genomic Aligner (NG Aligner)
- ProSPLIGN
- SPLIGN (SPliced aLIGNment)

Edit tools

- Basic/Extended Cleanup
- Merge Items tool
- Search/Find Repetitive Sequences with WindowMasker

Multiple Sequence Aligners

- Clustal Omega
- ClustalW2
- Kalign
- MAFFT
- MUSCLE

Tree Building tools

- Phylogenetic Tree Builder

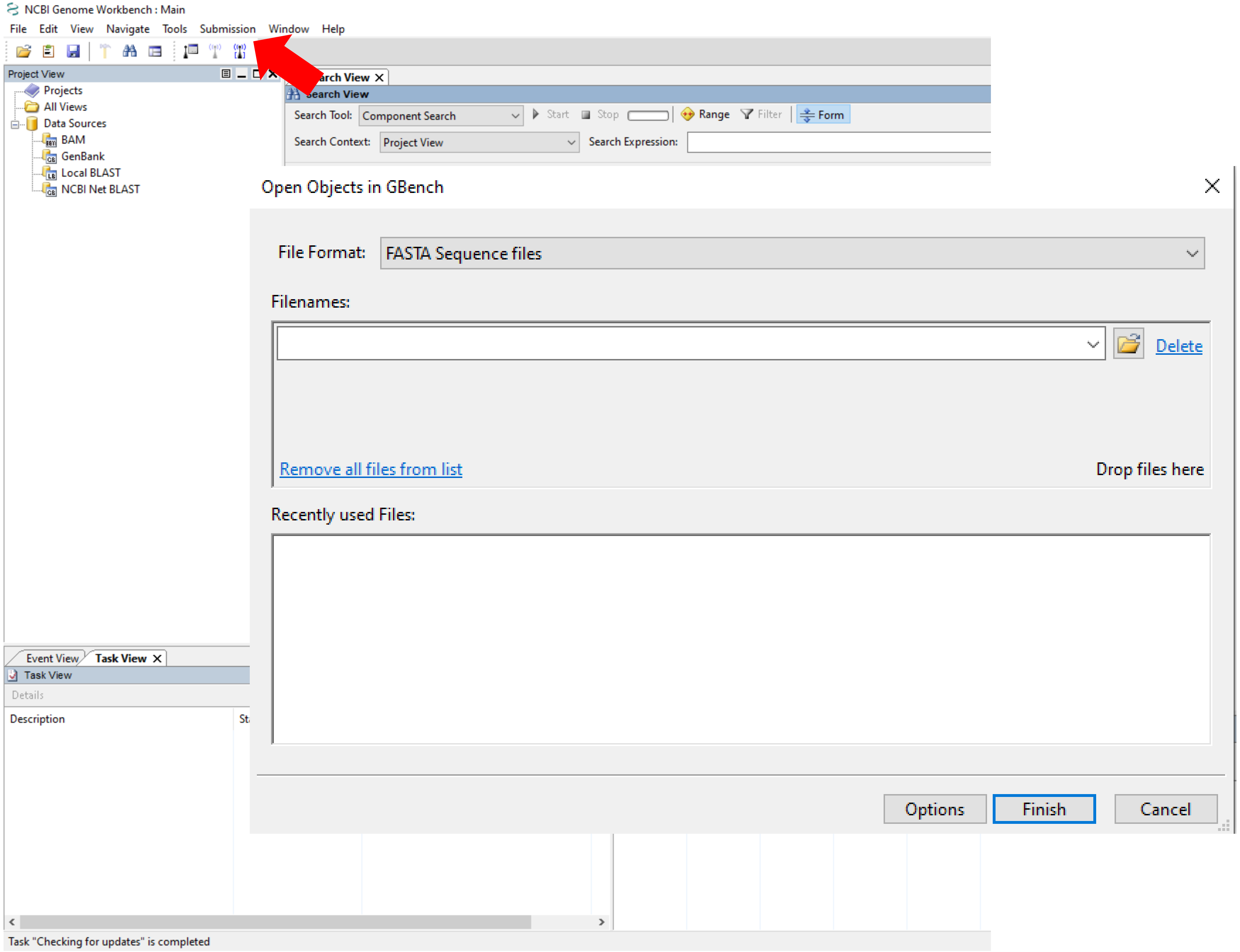
Data Formats

NCBI ASN1, AGP, BAM, BED, CSRA, FASTA, GFF, GVF, NEWICK, NEXUS, REPEAT MASKER, TABLE, TEXT ALIGNMENT, VCF, WIGGLE, 5 COLUMN FEATURE

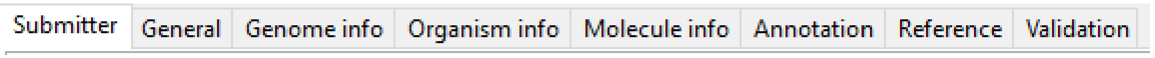
Sequence Editing Package

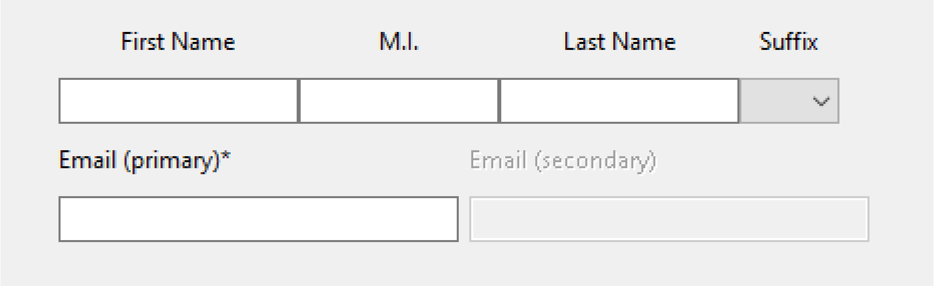
Genome Workbench offers a **Sequence Editing Package** that allows users to create, edit, validate, and submit a genome sequence submission to GenBank. The package includes a pop-up, tabbed wizard that directs a submitter through the data input steps needed to create a submission and a menu of editing and reports tools that can be used on an existing submission. Start with watching the Introducing the Genome Submission Wizard video tutorial.

https://www.ncbi.nlm.nih.gov/tools/gbench/



https://www.youtube.com/watch?v=fnpf9PVAS1w&ab_channel=NationalLibraryofMedicine

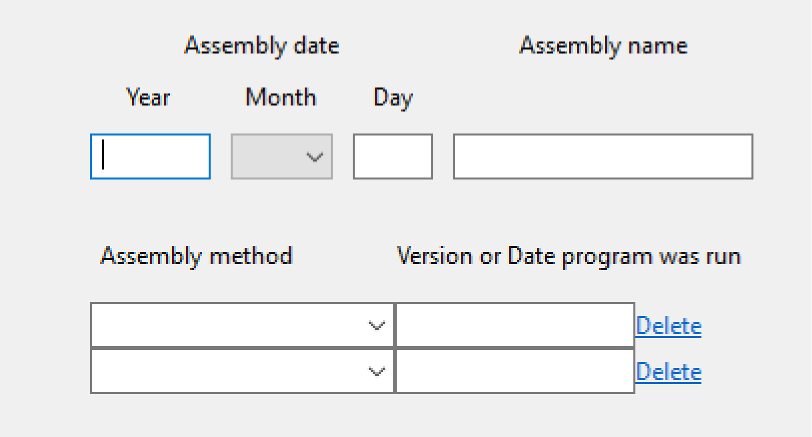




First Name | M.I. | Last Name | Suffix

Email (primary)*

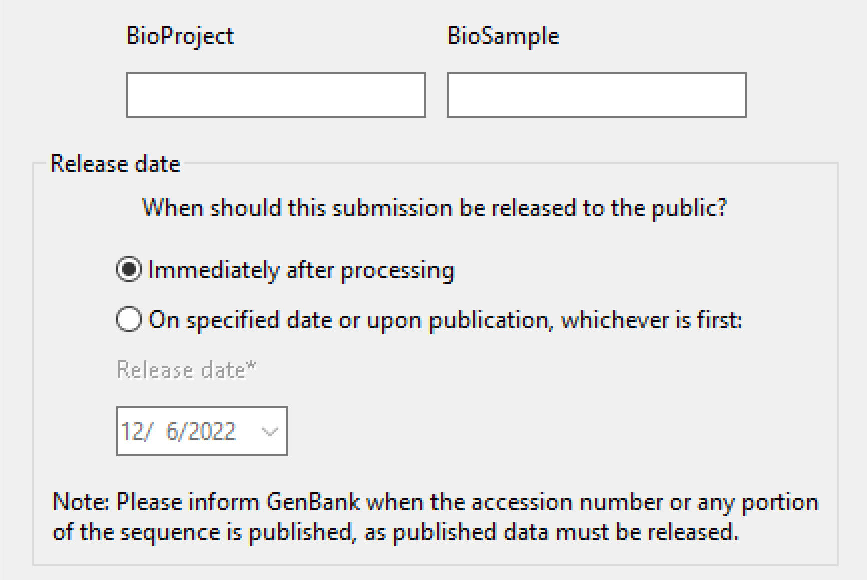
Email (secondary)



Assembly date: Year | Month | Day

Assembly name

Assembly method | Version or Date program was run



BioProject

BioSample

Release date

When should this submission be released to the public?

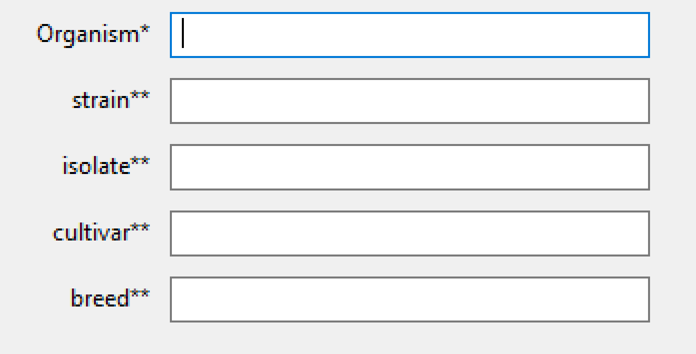
● Immediately after processing

○ On specified date or upon publication, whichever is first:

Release date*

12/ 6/2022

Note: Please inform GenBank when the accession number or any portion of the sequence is published, as published data must be released.



Organism*

strain**

isolate**

cultivar**

breed**

https://www.youtube.com/watch?v=fnpf9PVAS1w&ab_channel=NationalLibraryofMedicine

FINAL REMARKS

- NCBI is a powerful resource for biologists
- There are ~40 databases and many tools available to analyze the data
  - This webinar barely scratched the surface!
- Ensure you know what data you are using for your research, keeping track of details for reproducibility (ie. accession numbers, versions)
- When submitting to NCBI, ensure your data is accurate!

QUESTIONS?

THE UNIVERSITY OF BRITISH COLUMBIA