

# Using NCBI Blast to ID sequences

Navigate to the NCBI Nucleotide blast homepage

The screenshot shows the NCBI BLAST homepage. At the top, there is a navigation bar with the NIH logo and the text 'National Library of Medicine National Center for Biotechnology Information'. Below this, the page title is 'BLAST® » blastn suite'. The main content area is titled 'Standard Nucleotide BLAST'. Underneath, there are tabs for 'blastn', 'blastp', 'blastx', 'tblastn', and 'tblastx'. The 'blastn' tab is selected. The main form area is titled 'Enter Query Sequence'. It contains a text input field for 'Enter accession number(s), gi(s), or FASTA sequence(s)', a 'Query subrange' section with 'From' and 'To' input fields, and an 'Or, upload file' section with a 'Choose File' button and 'No file chosen' text. Below this is a 'Job Title' input field and a checkbox for 'Align two or more sequences'. The 'Choose Search Set' section includes a 'Database' dropdown menu set to 'Standard databases (nr etc.)', a 'Nucleotide collection (nr/nt)' dropdown menu, and an 'Organism' input field. An orange arrow points from the 'Choose File' button to the 'Job Title' input field.

Select “Choose File” to upload your consensus sequence

Scroll to the bottom of the page and make sure “Highly similar sequences” is selected

The screenshot shows the bottom portion of the NCBI BLAST homepage. The 'Program Selection' section is highlighted. It contains a 'BLAST' button and a 'Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences)' option. Below this, there is a checkbox for 'Show results in a new window'. The 'Program Selection' section includes a 'Choose BLAST algorithm' section with three radio buttons: 'Highly similar sequences (megablast)', 'More dissimilar sequences (discontiguous megablast)', and 'Somewhat similar sequences (blastn)'. The 'Highly similar sequences (megablast)' option is selected. An orange arrow points from the 'Highly similar sequences (megablast)' radio button to the 'BLAST' button. Below the 'BLAST' button, there is a note: 'Note: Parameter values that differ from the default are highlighted in yellow and marked with a sign'. At the bottom, there is a blue bar with a plus sign and the text '+ Algorithm parameters'.

Click on BLAST  
Your results will be found here

Job Title **Nucleotide Sequence**

RID [7NACEUVH016](#) Search expires on 05-12 02:06 am [Download All](#) ▾

Program BLASTN [Citation](#) ▾

Database nt [See details](#) ▾

Query ID lc|Query\_24739

Description None

Molecule type dna

Query Length 451

Other reports [Distance tree of results](#) [MSA viewer](#) ?

**Filter Results**

Organism only top 20 will appear  exclude

Type common name, binomial, taxid or group name

[+ Add organism](#)

Percent Identity  to  E value  to  Query Coverage  to

[Filter](#) [Reset](#)

**Descriptions** Graphic Summary Alignments Taxonomy

**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
<input checked="" type="checkbox"/>	<a href="#">Quillaja saponaria voucher Y114277-KUN plastid, complete genome</a>	<a href="#">Quillaja saponaria</a>	645	645	100%	1e-180	91.80%	160392	<a href="#">NC_047356.1</a>
<input checked="" type="checkbox"/>	<a href="#">Quillaja saponaria plastid, complete genome</a>	<a href="#">Quillaja saponaria</a>	645	645	100%	1e-180	91.80%	160392	<a href="#">MN709839.1</a>
<input type="checkbox"/>	<a href="#">Quillaja saponaria ribulose-bisphosphate carboxylase large subunit (rbcL), partial cds: chloroplast gene for c...</a>	<a href="#">Quillaja saponaria</a>	645	645	100%	1e-180	91.80%	1398	<a href="#">U06822.1</a>

## Downloading a sequence from GenBank

Navigate to the NCBI GenBank website and enter species along with gene of interest in search bar

An official website of the United States government [Here's how you know](#) ▾

**NIH** National Library of Medicine  
National Center for Biotechnology Information

GenBank Nucleotide  [Search](#)

GenBank ▾ Submit ▾ Genomes ▾ WGS ▾ Metagenomes ▾ TPA ▾ TSA ▾ INSDC ▾ Other ▾

**GenBank Overview**

**What is GenBank?**

GenBank® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research, 2013 Jan 41\(D1\):D36-42](#)). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An [annotated sample GenBank record](#) for a *Saccharomyces cerevisiae* gene demonstrates many of the features of the GenBank flat file format.

**Access to GenBank**

There are several ways to search and retrieve data from GenBank:

- Search GenBank for sequence identifiers and annotations with [Entrez Nucleotide](#).
- Search and align GenBank sequences to a query sequence using [BLAST](#) (Basic Local Alignment Search Tool). See [BLAST info](#) for more information about the numerous BLAST databases.
- Search, link, and download sequences programmatically using [NCBI e-utils](#).
- The ASN.1 and flatfile formats are available at NCBI's anonymous FTP server: <ftp://ftp.ncbi.nlm.nih.gov/ncbi-asn1> and <ftp://ftp.ncbi.nlm.nih.gov/genbank>

**GenBank Resources**

- [GenBank Home](#)
- [Submission Types](#)
- [Submission Tools](#)
- [Search GenBank](#)
- [Update GenBank Records](#)

Click on search

Scroll down to results and pick the entry you want to download

Herseq Sequences

Items: 1 to 20 of 31

1.  [Ailanthus altissima voucher bal30 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit \(rbcl\) gene, partial cds; chloroplast](#)  
552 bp linear DNA  
Accession: MF694658.1 GI: 1474380175  
[Taxonomy](#)  
[GenBank](#) [FASTA](#) [Graphics](#)

2.  [Ailanthus altissima clone Kb28 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit \(rbcl\) gene, partial cds; chloroplast](#)  
713 bp linear DNA  
Accession: MH536595.1 GI: 1561762343  
[Taxonomy](#)  
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

3.  [Ailanthus altissima chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit \(rbcl\) gene, complete cds](#)  
1,428 bp linear DNA  
Accession: U02726.1 GI: 433094  
[Taxonomy](#)  
[GenBank](#) [FASTA](#) [Graphics](#)

4.  [Ailanthus altissima voucher H. George s.n., CM atpB-rbcl intergenic spacer, partial sequence; and ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit \(rbcl\) gene, partial cds; chloroplast](#)  
744 bp linear DNA  
Accession: JX307353.1 GI: 476002269

Click on "Send to" and make sure "Complete Record", "File", and "FASTA" are selected

Nucleotide

Advanced

GenBank

**Ailanthus altissima voucher bal30 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcl) gene, partial cds; chloroplast**

GenBank: MF694658.1

[FASTA](#) [Graphics](#)

Go to: (v)

LOCUS MF694658 552 bp DNA linear PLN 15-SEP-2018

DEFINITION Ailanthus altissimus voucher bal30 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcl) gene, partial cds; chloroplast.

ACCESSION MF694658

VERSION MF694658.1

KEYWORDS .

SOURCE chloroplast Ailanthus altissimus

ORGANISM [Ailanthus altissimus](#)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae; Pentapetalae; rosids; malvids; Sapindales; Simaroubaceae; Ailanthus.

REFERENCE 1 (bases 1 to 552)  
AUTHORS Tahir,A. and Jamil,A.  
TITLE DNA barcoding and phylogenetic analysis of indigenous medicinal plants of Pakistan  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 552)  
AUTHORS Tahir,A. and Jamil,A.  
TITLE Direct Submission

Send to: (v)

- Complete Record
- Coding Sequences
- Gene Features

Choose Destination

- File
- Clipboard
- Collections
- Analysis Tool

Download 1 item

- Summary
- GenBank
- GenBank (full)
- FASTA**
- ASN.1
- XML
- INSDSeq XML
- TinySeq XML
- Feature Table
- Accession List
- GI List
- GFF3

Finally, select "Create File" to download the sequence as a .fasta file