



BLAST Homepage and Search Pages

Introducing the BLAST homepage and design/function of search pages

<https://blast.ncbi.nlm.nih.gov>

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

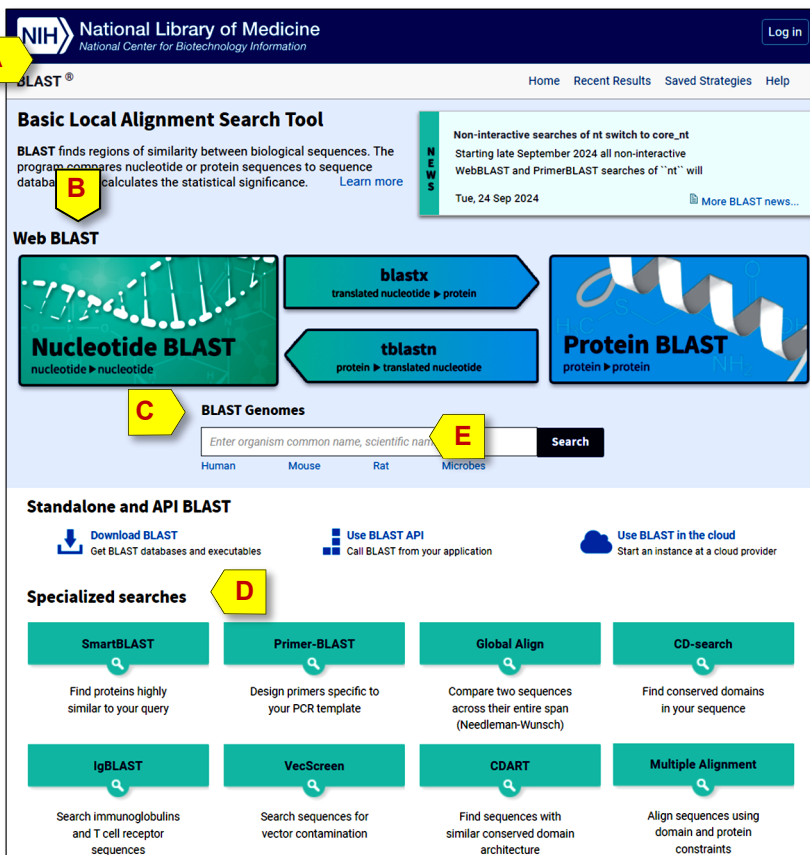
Introduction

BLAST is a suite of programs for nucleotide- and protein sequence similarity search. Developed by NCBI investigators [1], BLAST algorithms create local alignments between query sequences and sequences in a selected database. The NCBI BLAST homepage (<https://blast.ncbi.nlm.nih.gov/>) provides central access to the standard web search pages, specialized searches, standalone BLAST, and BLAST API. This document describes functionality of the home page and the five standard search pages.

The BLAST Homepage

This homepage consists of several sections, each provides a specific set of functions:

1. The common header (A) provides easy access to other content or functions not directly accessible from the homepage.
2. The page body organizes search pages into three categories, **Web BLAST** (B), **BLAST Genomes** (C), and **Specialized searches** (D). Each link or button points to a search page.
3. The **search box** (E) in the "BLAST Genomes" section takes an organism name and autosuggests a list of candidates. Selecting from the suggested list and clicking the "Search" button locates the best genomic sequence dataset for BLAST alignment purposes.



The BLAST Header

The common BLAST header provides a convenient way to access various BLAST pages with specific functions and to navigate among them.

- The logo (F) links to the NCBI homepage (<https://www.ncbi.nlm.nih.gov>) so that you can access non-BLAST related functions and content from other databases and services available from the NCBI site.
- The **Home** (G) links to the BLAST homepage, from other search or result pages.
- The BLAST service temporarily saves search results for up to 36 hours. The **Recent Results** (H) links to a page that keeps track of recently submitted search requests that have not expired. The Request ID uniquely assigned to a submitted search provides a one-click access to that result.
- The **Saved Strategies** (I) lists a set of search setups saved earlier. It allows the examination of search settings used, quick re-launch of these searches, and download of specific strategies for sharing or re-use in standalone BLAST.
- The **Help** (J) points to a page with links to help documents, tutorials, references and useful download directories on the BLAST ftp site (<https://ftp.ncbi.nlm.nih.gov/blast/>).
- The **Log in** button (K) at the top right points to the login page for the free My NCBI account [3]. That account allows you to customize your site preference and manage your works performed on the site. Upon successful login, this button will change to your login name. BLAST searches performed while logged in allows access to the search results for their full 36-hour life span through the **Recent Results** page, independent from the browser session. Search strategies saved will be saved permanently.



The Recent Results Page

BLAST search results are available for up to 36 hours. The **Recent Results** page (below) displays a list of recently submitted search requests that have not expired. The list is session-specific and will be lost when browser cookie is cleared. The best practice is to conduct BLAST searches while logged into your My NCBI account. The login status is indicated by the username at the upper right (insert, **A**). Each row in the table lists attributes for a single search, with the identifier in the Request ID (RID) column (**B**) providing an one-click access to its results. The Program, Title and Database columns (**C**) combine to provide a summary for a specific search. The “save”, “download” and the “X” (**D**) allow saving the search strategy, downloading the search strategy, and removing the search from the list, respectively. The input box (**E**) is for retrieving arbitrary search results using their assigned RIDs, such as those shared among colleagues, used as teaching or demonstration examples, or those problematic ones reported to NCBI’s blast-help team.

BLAST®

Home

Recent Results

Saved Strategies

Help

Lookup BLAST Job

E

Request ID:

Go

Available BLAST jobs

C

[more...](#)

A

ttao_phd

Your Available BLAST jobs

(Click headers to sort columns)

Submitted at	Request ID	Status	Program	Title	Qlength	Database	Expires at			
02-26 15:26	VY6UAMG013	Running	megablast	(2) - AE011191:Bacillus anthracis str. A2012 plasmid...	1671	wgs more...	02-28 03:26	save	download	✖
02-26 15:24	VY3B6SM016	Done	megablast	AE011191:Bacillus anthracis str. A2012 plasmid...	1671	core_nt	02-28 03:24	save	download	✖
02-26 09:46	VYCMJNN016	Done	blastp	NP_001814:creatine kinase B-type isoform 1...	381	refseq_select_prot	02-27 22:26	save	download	✖
02-26 09:46	VYAAWZBD016	Done	megablast	15 sequences (OTU_1 SIZE=1015688)	2002	core_nt	02-27 21:46	save	download	✖
02-26 08:44	VY6PNA5H016	Failed	blastp	wlinker	72348	refseq_select_prot	02-27 20:44	save	download	✖

The Saved Strategies Page

The **Saved Strategies** page (right) displays a list of search strategies saved earlier. The first four columns (**F**) provide a summary of the search settings for saved entries. The “view” link (**G**) loads a search page with saved custom settings, while the “download” link (**H**) saves the settings in an ASN.1 formatted text file for reloading on the web using the “Choose File” and “View” button (**I**). It is a useful way to share search settings with your students and colleagues.

BLAST®

Home

Recent Results

Saved Strategies

Help

I

Upload Search Strategy

Upload file :

Choose File

 No file chosen

View

F

Your Saved Search Strategies

(Click headers to sort columns)

Program	Created	Title	Database		
megablast	2025-02-26 15:28:23	AE011191:Bacillus anthracis str. A2012 plasmid...	core_nt	view	download
blastp	2025-02-26 15:21:00	NP_001814:creatine kinase B-type isoform 1...	refseq_select_prot	view	download
megablast	2025-01-22 15:13:50	unknown	core_nt	view	download
d.megablast	2021-09-20 13:44:18	F12345:HSC38H011 normalized infant brain cDNA...	genomic/8407/GCA_009802015.1	view	download

G

H

Functions of BLAST Search Pages

There are five standard BLAST search pages, each performing a specific type of sequence alignment. These pages are the foundation for the NCBI BLAST service. Table 1 below summarizes key aspects of these pages. These pages access a set of common databases as described in Table 2 on p.3.

Table 1. Functions of different BLAST pages: key features of the search pages in the “Web BLAST” category

Page name	Query vs database	Alignment	Programs & functions (default program in bold)
blastn (nucleotide blast)	nucl vs nucl	Nucleotide	megablast : sequence identification, intra-species comparison discontiguous megablast : cross-species comparison, searching with coding sequences blastn : searching with shorter queries, cross-species comparison
blastp (protein blast)	prot vs prot	protein	blastp : general sequence identification and similarity searches Quick BLASTP : with a kmer match to accelerate search speed for very similar proteins DELTA-BLAST [3] : protein similarity search with higher sensitivity than blastp PSI-BLAST : iterative search for position-specific score matrix (PSSM) to identify distant relatives for a protein family PHI-BLAST : protein alignment with input pattern as anchor/constraint
blastx	nucl(tr) vs prot	protein	blastx : identifying potential protein products encoded by a nucleotide query
tblastn	prot vs nucl (tr)	protein	tblastn : identifying database sequences encoding proteins similar to the query
tblastx	nucl (tr) vs nucl (tr)	protein	tblastx : identifying nucleotide sequences similar to the query based on their coding potential

Functions of BLAST Search Pages (cont.)

Table 2 describes the content of selected databases from the standard set to help you choose an appropriate database.

Table 2. Contents of BLAST Sequence Databases (default in bold)

Nucl. Database	Content
core_nt	All GenBank + EMBL + DDBJ + PDB sequences, excluding entries from PAT, EST, STS, GSS, WGS, TSA, HTGS phase 0-2, and eukaryotic genomic assemblies . Non-redundant, records with identical sequences collapsed into a single entry.
nt	Same as core_nt, but with eukaryotic genomic assemblies included.
refseq_rna	Curated (NM_, NR_) plus predicted (XM_, XR_) sequences from NCBI Reference Sequence Project.
refseq_select	NCBI RefSeq transcript sequences from human and mouse, restricted to the RefSeq Select set with one representative transcript per protein-coding gene
rRNA/ITS databases	A collection of four databases: a 16S Microbial rRNA sequences from NCBI's Targeted Loci Projects , an 18S and a 26S rRNA databases for fungi, plus an ITS database for fungi.
refseq_reference_genomes	Reference genomes selected from the NCBI RefSeq Genomes database. Entries in this database are among the best quality genomes available at NCBI. It is constructed with minimum redundancy in genome representation. For eukaryotes, one genome per organism is included. For other organisms, multiple genomes from diverse isolates of the same organism, such as <i>E. coli</i> , may be included.
refseq_genomes	This database contains NCBI RefSeq genomes across all taxonomy groups. It contains only the top-level sequences, <i>i.e.</i> the longest sequences representing any given part of the genomes, to reduce redundancy. Limit to specific subset required.
wgs	Assemblies of Whole Genome Shotgun sequences. Limit to specific subset required.
SRA	Nextgen sequences from NCBI's Sequence Read Archive (SRA). Limit to specific subset required.
G+T databases	The genomic sequences plus curated and predicted RNAs from the current build of the human or mouse genome.
TSA	Transcriptome Shotgun Assemblies, assembled from RNA-seq SRA data
pat	Nucleotides from the Patent division of GenBank.
pdb	Nucleotide sequences from the 3-dimensional structure records from Protein Data Bank.
est	Database of GenBank + EMBL + DDBJ sequences from EST division (stale without new entries).
HTGS	Unfinished High Throughput Genomic Sequences; Sequences: phases 0, 1 and 2 (stale without new entries).
Prot. Database	Content
nr (default)	Non-redundant GenBank CDS translations + RefSeq + PDB + SwissProt + PIR + PRF, excluding those in PAT, TSA, and env_nr.
refseq_protein	Protein sequences from NCBI Reference Sequence project.
refseq_select	RefSeq Select includes one representative protein per protein-coding gene for human and mouse, and RefSeq proteins annotated on reference and representative genomes for prokaryotes.
Landmark	The landmark database includes proteomes from representative genomes spanning a wide taxonomic range
swissprot	Last major release of the UniProtKB/SWISS-PROT protein sequence database (no incremental updates).
pataa	Proteins from the Patent division of GenBank.
pdb	Protein sequences from the 3-dimensional structure records from the Protein Data Bank.
env_nr	Protein sequences translated from the CDS annotation of metagenomic nucleotide sequences.
tsa_nr	Protein sequences translated from CDSs annotated on transcriptome shotgun assemblies.

The Nucleotide BLAST Search Page

The “nucleotide-blast” link loads the standard “Nucleotide BLAST” search page, with the breadcrumb at the top indicating the page position within the site hierarchy (A). A set of tabs for quick navigation among the five flavors of BLAST search pages (B) are under the descriptive page title. Two buttons to the right (C) reset the search settings to default and bookmark the custom settings, respectively. The page contains three sections, whose functions are described below.

Enter Query Sequence

This input box (D) takes nucleotide query sequences in accession (NM_000249) or FASTA [4] format. For a single input, “Query subrange” boxes (E) define a segment of that input to use in the search. The “Choose File” button (F) upload a plain text file with one or more sequences in FASTA format. The “Align two or more sequences” checkbox (G) changes the database list in “Choose Search Set” to a “Enter Subject Sequence” input box to allow the entering of another set of sequences for custom comparison [4].

The screenshot shows the NCBI BLAST Standard Nucleotide BLAST search page. Annotations are as follows:

- A**: Breadcrumb at the top right: "BLAST® » blastn suite".
- B**: Navigation tabs: "blastn", "blastp", "blastx", "tblastn", "tblastx".
- C**: "Reset page" and "Bookmark" buttons at the top right.
- D**: "Enter Query Sequence" input box for "Enter accession number(s), gi(s), or FASTA sequence(s)".
- E**: "Query subrange" input boxes for "From" and "To".
- F**: "Choose File" button for uploading a FASTA file.
- G**: "Align two or more sequences" checkbox.

The Nucleotide BLAST Search Page (cont.)

Choose Search Set

You can change the default database using the pull-down menu (A). To focus a search, you can restrict the lookup to a subset of the database by typing the name of the organism in the "Organism" textbox and selecting from the suggested list, using the checkbox to exclude the selection, and clicking the "Add organism" button to add extra input boxes to broaden the selection (B). Use checkboxes in Exclude line (C) to exclude sequences with low information content from the hits, or enter query terms in the "Entrez Query" textbox (D) to restrict a search to entries satisfying the specified criteria for non-wgs and non-sra databases under the "Standard databases" list. For example, "500:1000[slen]" restricts a search to database entries with length between 500 and 1000.

Program Selection

Three programs (E and Table 1 in p.2) with different search speed and sensitivity are available for nucleotide searches. The default megablast is better for tasks such as identifying input queries or searching with a large genomic entry; discontiguous megablast works better in finding related sequences from other species; while blastn works better for cross-species searches than megablast and for identifying short matches.

Algorithm parameters

Open Algorithm Parameters (F) to access additional parameter settings.

Under General Parameters, the "Max target sequences" (G) sets the upper limit on database sequences with matches BLAST will save and report back. The checked "Short queries" (H) allows BLAST to optimize settings for queries 30 bases/residues or shorter. The "Expect threshold" (I) filters out less significant matches, with Expect value above the setting. The "Word size" (J) controls the size of the initial seed match, with smaller settings more sensitive. The "Max matches in a query range" (K) limits the matches saved to a given region of the query (such as a repeat region) so matches to other regions of the query can be reported. The default setting of "0" means no limit.

Under Scoring Parameters, the "Match/Mismatch Scores" (L) specifies the reward assigned to an exact match and penalty assigned to a mismatch. The "Gap Costs" (M) specifies how penalties for gaps introduced in the alignment. For megablast, the default is linear, i.e., no penalty for opening a gap and a linear penalty proportional to the length of the gap upon gap extension.

Parameters under Filter and Masking specify what query masking should be used. Default is to mask low complexity regions in the query and only at the seed match stage (N). For genomic sequence, make sure to activate "species-specific repeats for:" option (O) and select the correct species from the pulldown list.

The Protein BLAST Search Page

In the “Standard Protein BLAST” search page, it has the same tab and links at the top as seen in the “Standard Nucleotide BLAST” search page (p. 4) and its default page display contains three similar sections.

Enter Query Sequence

Refer to the description on “standard nucleotide BLAST” on p.4.

Choose Search Set

An experimental clustered nr database (A) is available, which offers several features and advantages [5], with the checkbox below activating side-by-side searches against nr and clustered nr for result comparison purposes. The database list for protein BLAST is a much shorter list with the Organism and Exclude options functioning the same as that for the “standard nucleotide BLAST” page. Entrez query box is not available for protein database due to version 5 switch [6].

Program Selection

Five different programs (B, Table 1 on p.2) are available to satisfy diverse search need. The Quick BLASTP is a kmer-based faster scan for very similar matches. The default blastp is a general-purpose alignment program for identifying a protein sequence or finding others similar to it. PSI-BLAST is for finding more distant relatives through iterative custom PSSM construction. PHI-BLAST does protein alignment with a pattern in the query as a constraint, general for use as the first iteration of PSI-BLAST. DELTA-BLAST uses conserved domain matches the query to build a more sensitive PSSM for the match evaluation.

More complex searches may require adjustment of other search settings, toggle open the block under “Algorithm parameters” (C) to access them.

Algorithm parameters

Settings under General Parameters are essentially the same as that in the “Standard Nucleotide BLAST” (p. 4), except for the word size settings.

Settings under Scoring Parameters provides access to eight score matrices from two families (D), with default BLOSUM62 being the best general purpose one to use. For short queries adjust, PAM30 is selected. Each matrix has its own set of supported gap penalties under the Gap Costs menu (E). Protein alignment scores are adjusted to account for biases in composition through the “Compositional adjustments” (F) setting, choose “No adjustment” to disable.

Settings under Filters and Masking (G) specify whether low complexity should be filtered, only at the seed lookup stage, and if lower case letters a query (with mixed upper and lower case letters format, representing custom features) should be masked, respectively. These settings are generally not needed when compositional adjustments are used.

Translated Search Pages

The page layout for translated BLAST search pages is the same as “Standard Protein BLAST.” However, they do contain a few program-specific parameters.

- **Translated blastx search:** In the “Enter Query Sequence” section, a “Genetic code” field (A) is present under the “Choose File” button, specifying the codon table for the translation of input nucleotide queries. Choose a code appropriate for the source of the query sequence. The remaining sections are the same as the “Standard Protein BLAST” page.
- **Translated tblastn search:** The page layout is the same as the “Standard Protein BLAST” search page. The key difference is that the Database field lists available nucleotide databases instead.
- **Translated tblastx search:** The layout differences are the presence of the “Genetic code” field similar to the blastx page (A). The databases listed under the Database pull-down menu are the nucleotide databases. Note, the “Enter Query Sequence” takes nucleotide queries.

Other Search Pages

BLAST Genomes Section

BLAST search pages under the “BLAST Genomes” category (B) differ from these under the “Web BLAST” category only in the databases they access. The link names are for frequently targeted organisms. Most will be custom pages generated upon user input (C) with the best available genome dataset for the input target organism.

Specialized BLAST Section

This section lists different types of search pages designed for specialized tasks. Table 3 below summarizes the functions these pages provide. Most of the search pages have simple structure and will not be described separately.

Table 3. Function of Specialized BLAST pages not following the standard layout

Page link name	Functions
Primer-BLAST [7, 8]	Designing primers using the primer3 algorithm and checking their template specificity using BLAST against selected sequence collection
SmartBLAST [9]	SmartBLAST processes your protein query to present a concise summary of the five best protein matches from well-studied reference species in the landmark database.
IgBLAST [10]	Searching immunoglobulin or T cell receptor sequences against germline databases for annotation of the input immunoglobulin sequences
VecScreen	Screening input nucleotide sequences against a library of known vector and other artificial sequences to identify contaminations
CD-search	Searching an protein sequence against a database of curated domains for functional analysis. This search is performed for all protein-blast requests.
CDART	Identifying conserved domains present in the input protein sequence followed by finding other sequences containing these identified domains
Multiple Alignment [11]	Using Constraint Based Protein Multiple Alignment Tool (COBALT) to align multiple protein sequences with search link available in all protein BLAST result pages
Global Align	NCBI's implementation of the Needleman-Wunch global pair-wise alignment tool for nucleotide or protein queries
Align two or more sequences [12]	A common type of need that is not a listed entry. Use a relevant Web BLAST page and activate the mode by checking the “align two or more sequences” checkbox.

Different Ways to Access NCBI Web BLAST Services

In addition to web browser, you can also access BLAST services described above (with the exception for those listed in Table 3) using the “-remote” option in different standalone BLAST+ programs or the RESTful BLAST service. Table 4 below summarizes characteristics of these venues. You can obtain available BLAST databases using the Entrez Programming Utilities [13] by querying the blastdbinfo database as described in an NCBIInsight blogpost [14].

Table 4. Features of available methods to access NCBI web BLAST services

Venue	Features
Web browser	Intuitive: graphical user interfaces and result presentation Convenience: ease of searching with single or small batch of query sequences Speed: fast turnaround from the distributed computing system [15] Versatility: “BLAST two or more sequences” option enables searching against custom sequences Job Limitation: Not meant for high throughput searches with limit on CPU time to ensure fair public access Data partition: Access to different database requires different search pages
Standalone BLAST+ (-remote option) [16]	Comprehensive: more options available than on the Web for customizing and fine tuning the search Batch processing: search with large query sequences by submitting them in smaller batches automatically Less manual intervention: option for saving output in various formats Workflow incorporation: input and output can be integrated in custom workflow Extra requirements: installing standalone BLAST+ package and configuring it properly, plus database maintenance
RESTful BLAST (Qblast, BLAST URLAPI) [17]	Comprehensive: more available options to customize and fine tune the search than on the Web Batch processing: search with large query sequences possible through batching (subject to the same web limit) Workflow incorporation: input and output can be integrated in custom workflow Extra requirements: efficient usage requires scripting/programming for requesting URL construction and result checking
blastn_vdb & tblastn_vdb [18]	Similarity: commands similar to BLAST+ counterparts, for SRR, WGS, and TSA datasets stored in vcb format Comprehensive: more options available than on the Web for customizing and fine tuning the search Built-in client function: automatically streams the data file to blast Less manual intervention: option for saving output in various formats Workflow incorporation: input and output can be integrated in custom workflow Extra requirements: installing BLAST+ and configuring it properly, installing sratoolkit to complement

Technical Assistance

NCBI provides technical assistance to BLAST users through its blast-help team. Please report problems or bugs, submit suggestions or feature requests, and other BLAST-related questions to blast-help@ncbi.nlm.nih.gov.

Submitting detailed information along with the problem report will help expedite the investigation and response. When reporting problems encountered during web BLAST searches (and URLAPI searches), make sure you include:

- A description on the goal of the search
- The valid RID(s) of relevant searches (they expired 36 hours after issuance)
- The detailed error message
- The search page and settings used, along with a summary of the input query, particularly if RIDs were not issued

CPU related errors are caused when searches exceed the allocated time limit. You can follow the steps below to redo the search:

- click on “Edit search” to get back to the search page
- Reduce the number of the input query sequences or use subsequence for a large single query
- Add database limit using Organism or Entrez query box to target a focused subset within the target database
- Increase the search stringency by using a combination of a lower Expect value, a larger Word size, filters and repeat masking, and a lower number for Maximum target sequences (results saved)

For errors occurred from using “-remote” option of the standalone BLAST+ package, as well as standalone BLAST+ package for local searches, consider providing the following pieces of information:

- A description on the goal of the search
- The platform and version of the installed BLAST+ package
- The complete error message
- The complete command issued of the search
- A summary and a small sample set of the input query file
- BLAST server returned RIDs (for -remote search) if available

References

1. [Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.](#) Altschul SF, et. al. Nucleic Acids Res. 1997 Sep 1;25(17):3389-402.
2. [NCBI BLAST: a better web interface.](#) Johnson M et. al. Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W5-9.
3. My NCBI Help Manual. <https://www.ncbi.nlm.nih.gov/books/NBK3843/>
4. Align Two Sequences Using NCBI BLAST. <https://ncbiinsights.ncbi.nlm.nih.gov/2013/03/19/blastdbinfo-api-access-to-a-database-of-blast-databases/>
5. New ClusteredNR database: faster searches and more informative BLAST results. https://ncbiinsights.ncbi.nlm.nih.gov/2022/05/02/clusterednr_1/
6. BLAST+ now has a better BLAST database. <https://ftp.ncbi.nlm.nih.gov/blast/db/v5/blastdbv5.pdf>
7. [Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction.](#) Ye J, et. al. (2012). BMC Bioinformatics. 13:134.
8. NCBI Primer-BLAST. https://ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_PrimerBLAST.pdf
9. SmartBLAST: A Tool For Rapid Protein Identification. https://ftp.ncbi.nlm.nih.gov/pub/factsheets/Factsheet_SmartBLAST.pdf
10. [IgBLAST: an immunoglobulin variable domain sequence analysis tool.](#) Ye J, et. al. Nucleic Acids Res. 2013 Jul;41 (Web Server issue):W34-40. doi: 10.1093/nar/gkt382.
11. [COBALT: constraint-based alignment tool for multiple protein sequences.](#) Papadopoulos JS and Agarwala R (2007). Bioinformatics 23:1073-79.
12. Align Two Sequences Using NCBI BLAST. <https://www.ncbi.nlm.nih.gov/books/NBK279690/>
13. Entrez Programming Utilities Help. <https://www.ncbi.nlm.nih.gov/books/NBK279690/>
14. Blastdbinfo: API access to a database of BLAST databases. <https://ncbiinsights.ncbi.nlm.nih.gov/2013/03/19/blastdbinfo-api-access-to-a-database-of-blast-databases/>
15. [A Fault-Tolerant Parallel Scheduler for BLAST.](#)
16. Standalone BLAST+ Cookbook. <https://www.ncbi.nlm.nih.gov/books/NBK279696>
17. Common URL API. https://ftp.ncbi.nlm.nih.gov/pub/factsheets/Factsheet_SmartBLAST.pdf
18. BLAST Search WGS, TSA and SRA Data Locally. https://ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_LocalSRABlast_StepByStep.pdf

Appendix: Local Installation of NCBI BLAST Packages

NCBI also makes the BLAST programs available as standalone packages for local installation. The blast+ packages for common platforms are available from the following FTP directory:

<https://ftp.ncbi.nlm.nih.gov/blast/executables/LATEST/>

The package also contains utilities for downloading the standard set of BLAST databases (**update_blastdb.pl**), in preformatted form from the BLAST db ftp directory (<https://ftp.ncbi.nlm.nih.gov/blast/db/>), and for extracting sequences from those databases (**blastdbcmd**). Note that executable programs within the package have no graphical user interface, so you need to interact with them through terminal console by issuing text commands.

To handle datasets stored in vdb format initially adopted by the NCBI SRA database, blastn_vdb and tblastn_vdb are made available and integrated into the BLAST+ package. These two programs can access SRA, WGS, and TSA sequence data stored in vdb format natively as databases, making the FASTA dumping and database formatting unnecessary. To enhance their functionality, consider install the sratoolkit to complement these tools:

<https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software>

Two of the specialized BLAST tools, i.e., IgBLAST and COBALT multiple protein sequence alignment tool, are available for local setup. They are downloadable from the following FTP directories:

<https://ftp.ncbi.nlm.nih.gov/blast/executables/igblast/release/>.

<https://ftp.ncbi.nlm.nih.gov/pub/cobalt/>

Magic-blast, a read aligner from NCBI, can also access reads from SRA directly as input query. You can download this tool from NCBI at: <https://ftp.ncbi.nlm.nih.gov/blast/executables/magicblast/>

Processing of large datasets generated by nextgen sequencing technology often requires large computational setup in the cloud. NCBI makes a BLAST+ based package available as a Docker image. Refer to this page for more details:

https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=CloudBlast