



BLAST Homepage and Selected Search Pages

Introducing the BLAST homepage and form elements/functions of selected 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

Background

BLAST [1] is a suite of programs provided by NCBI for aligning query sequences against those present in a selected target database. The NCBI BLAST homepage (<https://blast.ncbi.nlm.nih.gov>) provides an access point for these tools to perform sequence alignment on the web.

The BLAST Homepage

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

1. The common header (A), present in most BLAST-related pages, provides easy access to other content or functions not directly accessible from the homepage.
2. The “Recent Results” (B) lists a set of links to recently completed search results.
3. Pages with web forms for submitting searches are listed as links in the body of the BLAST homepage. These links are organized into three categories, “Web BLAST” (C, used to be called “Basic BLAST”), “BLAST Genomes” (D), and “Specialized searches” (E).
4. The **search box** (F) in the “BLAST Genomes” section takes the name of an organism as input and suggests 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 Common BLAST Header

The common BLAST header provides a convenient way to navigate among different pages to access different contents or functions.

The screenshot shows the BLAST homepage with several callout boxes labeled A through L. Box A points to the top navigation bar. Box B points to the 'Recent Results' link. Box C points to the 'Basic Local Alignment Search Tool' section. Box D points to the 'Web BLAST' section. Box E points to the 'Specialized searches' section. Box F points to the search box in the 'BLAST Genomes' section. Box G points to the 'NCBI' link in the footer. Box H points to the 'Home' link in the footer. Box I points to the 'Recent Results' link in the footer. Box J points to the 'Saved Strategies' link in the footer. Box K points to the 'Help' link in the footer. Box L points to the 'Sign in to NCBI' link in the footer.

- The NCBI name (G) 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 site.
- The “Home” tab (H) links to the BLAST homepage, from other search or result pages.
- The BLAST service temporarily saves your search results for up to 36 hours. The “Recent Results” (I) 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” tab (J) 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” tab (K) points to page with a list of links to help documents, tutorials, references and useful download directories on the BLAST ftp site (<ftp://ftp.ncbi.nlm.nih.gov/blast/> or <https://ftp.ncbi.nlm.nih.gov/blast/>).
- My NCBI [2], a free account from NCBI, allows users to customize their site preference and manage their works performed on the site. Login-related links for My NCBI (L) are at the top right. 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. Strategies saved will be saved permanently.

The Recent Results Page

BLAST search results are available for up to 36 hours. The “Recent Results” tab displays a list of recently submitted search requests that have not expired. The list is session-specific and will be lost if session cookie is cleared upon browser exit. For this reason, it is recommended that BLAST searches be done with an active login to a My NCBI account such as the one shown below, with the My NCBI login indicated by an insert at the upper right (A). Each result is given as a row in the table, with the identifier in the Request ID (RID) column (B) providing an one-click access to the search result. The program, Title and Database columns (C) combine to provide a summary for a specific search. Database restriction applied are indicated by “more ...” and the popup upon hover (D). The “save”, “download” and the red “X” (E) allow saving the search strategy, downloading the search strategy, and removing the search from the list.

The input box (F) above the table is for retrieving other results using their assigned RIDs, such as those shared among colleagues, used as teaching or demonstration examples, or those with issues encountered and reported to NCBI’s blast-help group. The formatting options table is no longer available with the switch to updated BLAST results page.

The Saved Strategies Page

The “Saved Strategies” tab (right) displays a list of search strategies you saved earlier. The first four columns (J) provide a summary of the search settings for each saved entry. The “view” button (K) loads the settings in a search page, while the “download” link (L) saves the settings in an ASN.1 formatted text file for use with standalone BLAST or reloading on the web services using the “Choose File” and “View” button (M), a useful way to share settings.

Functions of BLAST Search Pages

There are five BLAST search pages, each performs a specific type of sequence alignment. These pages are the foundation for the NCBI BLAST service and will be described in more detail. Table 1 below summarizes key aspects of pages. These pages access a set of common databases, a summary of the contents for these databases are given in Table 2.

Table 1. Key features of the BLAST search pages in the “Basic BLAST” category

Search page	Query & database	Alignment	Programs & functions (default program in bold)
nucleotide blast	nucleotide vs nucleotide	Nucleotide	megablast : for sequence identification, intra-species comparison discontiguous megablast : for cross-species comparison, searching with coding sequences blastn : for searching with shorter queries, cross-species comparison
protein blast	Protein vs protein	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 wgs score matrix (PSSM) to identify distant relatives for a protein family PHI-BLAST : protein alignment with input pattern as anchor/constraint
blastx	nucleotide (tr) vs protein	protein	blastx : for identifying potential protein products encoded by a nucleotide query
tblastn	protein vs nucleotide (tr)	protein	tblastn : for identifying database sequences encoding proteins similar to the query
tblastx	nucleotide (tr) vs nucleotide (tr)	protein	tblastx : for identifying nucleotide sequences similar to the query based on their coding potential

Functions of BLAST Search Pages (cont.)

Table 2 describes the contents of the standard set of databases, which will help you select the proper database to setup more efficient searches.

Table 2. Contents of the common BLAST sequence databases

Nucl. Database	Content
nt (default)	All GenBank + EMBL + DDBJ + PDB sequences, excluding sequences from PAT, EST, STS, GSS, WGS, TSA and phase 0, 1 or 2 HTGS sequences. Non-redundant, records with identical sequences collapsed into a single entry.
rRNA/ITS databases	A collection of four databases: a 16S Microbial rRNA sequences from NCBI's Targeted Loci Projects , an 18S and a 26S RNA rRNA databases for fungi, plus an ITS database for fungi.
refseq_rna	Curated (NM_, NR_) plus predicted (XM_, XR_) sequences from NCBI Reference Sequence Project.
refseq_representative_genomes	NCBI RefSeq Reference and Representative genomes across broad taxonomy groups including eukaryotes, bacteria, archaea, viruses and viroids. These genomes are among the best quality genomes available with minimum redundancy - one genome per species for eukaryotes and diverse isolates for the same species for others.
Refseq genome	This database contains NCBI RefSeq genomes across all taxonomy groups. It contains only the top-level sequences, i.e. the longest sequences representing any given part of the genomes, to reduce redundancy.
wgs	Assemblies of Whole Genome Shotgun sequences.
est	Database of GenBank + EMBL + DDBJ sequences from EST division
Human G+T	The genomic sequences plus curated and predicted RNAs from the current build of the human genome.
Mouse G+T	The genomic sequences plus curated and predicted RNAs from the current build of the mouse genome.
est	Database of GenBank + EMBL + DDBJ sequences from EST division
TSA	Transcriptome Shotgun Assemblies, assembled from RNA-seq SRA data
SRA	Nextgen sequences from NCBI's Sequence Read Archive (SRA), limit to specific subset required.
HTGS	Unfinished High Throughput Genomic Sequences; Sequences: phases 0, 1 and 2.
pat	Nucleotides from the Patent division of GenBank.
pdb	Nucleotide sequences from the 3-dimensional structure records from Protein Data Bank.
refseq_genomic	All genomic sequences from NCBI Reference Sequence Project, highly redundant.
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.
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).
pat	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.

Elements of the Standard Nucleotide BLAST Search Page

The "nucleotide-blast" link loads the "Standard Nucleotide BLAST" search page. The top of the page (below the common BLAST header) contains a breadcrumb indicating the page position in the site hierarchy (A), the page title, a set of tabs for quick navigation among the five core BLAST search pages (B), plus links to set the page back to default and to bookmark a search page with customized settings (C). The default display of the page contains three sections with the functions described below.

Enter Query Sequence

The main input box (D) takes nucleotide query sequences in various formats, such as accession (NM_000249) or FASTA [4]. For a single input, "Query subrange" boxes (E) define a segment of the query to use in the search. Use the "Choose File" button (F) to upload a plain text file with one or more sequences in FASTA format. The "Align two or more sequences" checkbox (G) changes the "Choose Search Set" sections below to "Enter Subject Sequence" to allow comparison of query against those in the subject input box.

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

- A:** Breadcrumb navigation: NCBI/BLAST/blastn suite
- B:** Navigation tabs: blastn, blastp, blastx, tblastn, tblastx
- C:** Action links: Reset page, Bookmark
- D:** Main input box: Enter Query Sequence
- E:** Subrange boxes: From, To
- F:** File upload button: Choose File
- G:** Checkboxes: Align two or more sequences

Elements of the Standard Nucleotide BLAST Search Page (cont.)

Choose Search Set

BLAST database can be selected from the standard list using the pull-down menu (A). A search can be restricted to a subset of entries in the selected database by typing the name of the species, strains, or taxonomic group in the "Organism" textbox and selecting from the suggested list (B). The exclusion box to the right removes sequences from the selected organism during the search. Clicking the "+" button adds extra input boxes for broader organism selection. Sequences with low information content can be excluded using the checkboxes below (C). For certain databases, entering custom queries in the "Entrez Query" textbox (D) restricts a search to entries satisfying the specified criteria. For example, entering "[biomol_mrna\[prop\] AND 500:1000\[slen\]](#)" will restrict a search to mRNA entries 500 to 1000 bases long.

Program Selection

Three programs (E and Table 1 in p.2) with different speed and sensitivity are available for nucleotide vs nucleotide searches. The default [megablast](#) is better for certain tasks, such as identifying the input query and searching with large genomic query; [discontiguous megablast](#) works better in finding related sequences from other species; while [blastn](#) works better for short input queries and identifying short matches, it also works better for cross-species searches than megablast.

Clicking the "BLAST" button (F) submits the search to BLAST server for processing. Results will be automatically displayed when completed. "Algorithm parameters" link (G) opens a normally collapsed section allowing access to additional parameter settings.

General Parameters

Parameters in this section specify the search sensitivity. The "Max target sequences" (H) sets the maximum database sequences with matches BLAST saves for a given query. The checked "Short queries" checkbox (I) allows BLAST to automatically optimize settings for queries 30 bases/residues or shorter. The "Expect threshold" (J) filters out matches that are less significant, with Expect value above the setting. The "Word size" (K) set the size of the initial seed match, smaller settings are more sensitive. The "Max matches in a query range" (L) limits the matches saved to a given region of the query (such as from repeats) so matches to other region of the query can be reported. The default setting of "0" means no limit.

Scoring Parameters

Parameters here also affect the search sensitivity. The "Match/Mismatch Scores" (F) specifies the reward assigned to exact match and penalty assigned to a mismatch. The "Gap Costs" (G) field specifies how gaps introduced in the alignment should be penalized. For megablast, the default is linear, no penalty for opening a gap, while extending a gap assumes a linear penalty proportional to the length of the gap. For both parameters, non-default settings can be selected using the pull-down menu.

Elements of the Standard Nucleotide BLAST Search Page (cont.)

Filters and Masking

Parameters here specify whether low complexity sequences and organism-specific repeats should be filtered (**H**), and whether to filter only at the initial seed match stage (**B**, Mask for lookup table only) or during alignment extension as well. Lower case letters in the FASTA query (provided as a mixed upper and lower case letters, representing custom features) can also be masked.

Make sure species-specific repeats for the source organism is selected when searching with genomic sequences to help avoid spurious hits from repeats overwhelming the BLAST search leading to search errors.

Elements of the Standard Protein BLAST Search Page

The “protein-blast” link in the “Basic BLAST” links to the “Standard Protein BLAST” search page. The top of this page has the same tab and links found in the “Standard Nucleotide BLAST” search page (pg. 4) that provide the same functions. The default page display contains three sections with the functions described below.

Enter Query Sequence

Refer to the description for “Standard Nucleotide BLAST” (pg.4) for details. Checking “Align two or more sequences” will remove the “Program Selection” section to add subject input box for direct comparison of sequences in the query and subject input box. Only blastp program is available for that.

Choose Search Set

Most of the components are similar to the “Standard Nucleotide BLAST” page (pg.4). The main difference is that the database pull-down menu has a smaller list of protein databases (**D**). Entrez query box is no longer available due to database switching to version 5 [5].

Program Selection

Four different programs (**E** and Table 1 on p. 2) are available to satisfy various search. The Quick BLASTP is for faster scan for very similar matches. The default blastp is a general purpose protein alignment program for identifying a sequence or finding others similar to it. PSI-BLAST is for finding more distant relatives through custom PSSM construction. PHI-BLAST does protein alignment with a pattern in the query as a constraint. DELTA-BLAST is a more sensitive search using conserved domain matches the query to build a PSSM for the match evaluation. More complex searches may require adjustment of other search settings listed under the “Algorithm parameters” link (**F**), which is organized in a similar manner to that for the “Standard Nucleotide BLAST.”

- **General Parameters:** This section is the same as that in the “Standard Nucleotide BLAST” (pg. 4).

- **Scoring Parameters:** Eight score matrices from two families are supported (**G**). The default BLOSUM62 is the best general purpose matrix. For short queries, PAM30 is often selected. Each matrix has its own set of supported gap penalties under the Gap Costs menu (**H**). Protein alignment scores can be adjusted to account for biases in composition using options under the “Compositional adjustments” (**I**).

- **Filters and Masking:** Parameters here specify whether low complexity should be filtered and only at the seed lookup stage (**J**). Lower case letters in the FASTA query (in a mixed upper and lower case letters format, representing custom features) should also be masked (**K**). These settings are generally not needed when compositional adjustments are used.

Items Unique to 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 specify the codon table used in the translation of the input nucleotide query. 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 (A) similar to the blastx page. 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 search pages under the “BLAST Genomes” category (B) differ from these under the “Basic BLAST” category only in the databases they access. The link names clearly indicate the source organism of the database the search will target. Most will be custom pages generated upon user input (C) with the best available genome dataset for the input target organism.

The “Specialized BLAST” category contains different types of search pages, designed for specific tasks. Table 3 below summarizes the functions these pages provide.

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

Page link name	Functions
Primer-BLAST [6, 7]	Designing primers using the primer3 algorithm and checking their template specificity using BLAST against selected sequence collection
SmartBLAST [8]	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 [9]	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 [10]	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 (COBAL) 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
MOLE-BLAST	Identifying neighbors for input nucleotide sequences from selected target database (using BLAST), and then cluster the collection according to their sequence similarities using multiple alignment (MUSCLE).
Targeted Loci	Search against curated nucleotide sequences from bacterial and archaeal 16S rRNA, or fungal 18S, 28S, and ITS for species identification need
Align two or more sequences	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.

Ways to Access NCBI Web BLAST Services

In addition to web browser, BLAST services described above, with the exception for those listed in Table 3, can also be accessed using the “-remote” option in different standalone BLAST+ programs, or the RESTful BLAST service (**Qblast** or BLAST URLAPI). Features of these venues are summarized in Table 5 below. Available BLAST database can be obtained using the Entrez Programming Utilities [12] by querying the blastdbinfo database as described in a blog post [13].

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 (splitd) Versatility: available 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) [14,15]	Comprehensive: more options available than on the Web for customizing and fine tuning the search Batch processing: search with large query sequences by submitting then 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
RESTful BLAST (Qblast, BLAST URLAPI) [16]	Comprehensive: more available options to customize and fine tune the search than the Web Batch processing: search with large query sequences possible through batching 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 Program similar to their counterparts from standalone blast+ that are included in the SRAToolkit [17]	Similar to their BLAST+ counterparts, but access SRR, WGS, and TSA files stored in vcb format. Comprehensive: more options available than on the Web for customizing and fine tuning the search Built-in client function: automatically downloads 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 SRAToolkit and configuring it properly

Setting NCBI BLAST Searches Locally

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 (through **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**).

To handle datasets stored in vdb format initially adopted by the NCBI SRA database, blastn_vdb and tblastn_vdb are made available. These two programs can access SRA, WGS, and TSA sequence data stored in vdb format natively as databases, making the FASTA dumping and blast database formatting unnecessary. Install the sratoolkit to gain access to these tools:

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

Magic-blast, a read aligner from NCBI BLAST group, can also access reads from SRA directly as input query. See:

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

Specialized BLAST are generally not available for local setup, with the exception of IgBLAST:

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

Processing of large datasets generated by nextgen sequencing technology often requires large computational setup in the cloud. NCBI makes BLAST+ 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

Technical Assistance

NCBI provides technical assistance to the BLAST user community through its blast-help group. Problem and bug reports, suggestions and feature requests, as well as other questions related to BLAST usage should be addressed to the group (blast-help@ncbi.nlm.nih.gov).

Submitting detailed information along with the problem report will help expedite the investigation and response. Information needed when reporting problems encountered during web BLAST searches are:

- A description on the goal of the search
- The valid RID(s) of relevant searches (note that they expired 36 hours after the search)
- 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 processing time. Repeat the search using “Edit and resubmit” link to get back to the search page with the following adjustments will help resolve the issue:

- Reduce the number and/or size of the input query sequence(s), use subsequence for large single query if possible
- Add database limit using Organism or Entrez query box to search a more focused smaller subset
- Increase the search stringency by using
 - ◊ A lower Expect value
 - ◊ A larger Word size
 - ◊ Filters and repeat masking
 - ◊ 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, the following pieces of information should be provided:

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

For RESTful BLAST, the following pieces of information should be provided:

- A description on the goal of the search
- The platform and relevant code used to call the service
- The complete error message
- A summary and small sample of the input query file

References

1. [NCBI BLAST: a better web interface](#). NCBI BLAST: a better web interface. Johnson M et. al. Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W5-9.
2. My NCBI Help. <https://www.ncbi.nlm.nih.gov/books/NBK3843/>
3. [Domain enhanced lookup time accelerated BLAST](#). Boratyn GM, et. al. Biol Direct. 2012 Apr 17;7:12.
4. FASTA format. https://en.wikipedia.org/wiki/FASTA_format
5. BLAST DBv5. <https://ftp.ncbi.nlm.nih.gov/blast/db/v5/blastdbv5.pdf>
6. [Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction](#). Ye J, et. al. BMC Bioinformatics. 2012 Jun 18;13:134. doi: 10.1186/1471-2105-13-134.
7. Factsheet: Primer-BLAST. https://ftp.ncbi.nih.gov/pub/factsheets/HowTo_PrimerBLAST.pdf
8. Factsheet: SmartBLAST. https://ftp.ncbi.nlm.nih.gov/pub/factsheets/Factsheet_SmartBLAST.pdf
9. [IgBLAST: an immunoglobulin variable domain sequence analysis tool](#). Ye J, et. al. Nucleic Acids Res. 2013 Jul 1;41(Web Server issue):W34-40.
10. [CDD: specific functional annotation with the Conserved Domain Database](#). Marchler-Bauer A, et. al. Nucleic Acids Res. 2009 Jan;37(Database issue):D205-10.
11. [COBALT: constraint-based alignment tool for multiple protein sequences](#). Papadopoulos JS, Agarwala R. Bioinformatics. 2007 May 1;23(9):1073-9.
12. Entrez Programming Utilities Help. <https://www.ncbi.nlm.nih.gov/books/NBK25501>.
13. Blastdbinfo: API access to a database of BLAST databases. NCBIINSIGHT blog entry <https://go.usa.gov/xp7gR>
14. Installation of standalone BLAST+. https://ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_blast_setup.pdf
15. Standalone BLAST+ Cookbook. <https://www.ncbi.nlm.nih.gov/books/NBK279690/>
16. QBLAST's URL API User Guide. <https://ncbi.github.io/blast-cloud/dev/api.html>
17. Local SRA BLAST: https://ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_Local_SRA_BLAST.pdf