



BLAST Homepage and Selected Search Pages

Introducing the BLAST homepage and form elements/functions of selected search pages
<http://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 (<http://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 right hand column (B) provides links to a list of recently completed results, BLAST-related news and search tips.
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, "BLAST Assembled RefSeq Genomes" (C), "Basic BLAST" (D), and "Specialized BLAST" (E).
4. The search box (F) in the "BLAST Assembled Genomes" section takes the name of an organism as input and suggests a list of candidates. Selecting from the suggested list will locate the best genomic sequence dataset for BLAST alignment purposes.

The screenshot shows the BLAST homepage interface. Annotations are as follows:

- A**: Points to the top navigation bar containing 'Home', 'Recent Results', 'Saved Strategies', and 'Help'.
- B**: Points to the 'My NCBI' account management links ('Sign In', 'Register') in the top right corner.
- C**: Points to the 'BLAST Assembled Genomes' section.
- D**: Points to the 'Basic BLAST' section.
- E**: Points to the 'Specialized BLAST' section.
- F**: Points to the search input box in the 'BLAST Assembled Genomes' section.
- G**: Points to the BLAST logo in the footer.
- H**: Points to the 'Home' tab in the footer.
- I**: Points to the 'Recent Results' tab in the footer.
- J**: Points to the 'Saved Strategies' tab in the footer.
- K**: Points to the 'Help' tab in the footer.
- L**: Points to the 'My NCBI' account management links in the footer.

The common BLAST header

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

- The NCBI logo (G) links to the NCBI homepage (<http://www.ncbi.nlm.nih.gov>) to allow access to non-BLAST related functions and content.
- The "Home" tab (H) links to the BLAST homepage.
- BLAST search results are temporarily saved for 36 hours. The "Recent Results" tab (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.
- My NCBI [2] is a free account from NCBI, which allows users to customize their site preference and manage their works performed on the NCBI site. Login-related links for My NCBI (L) are at the right. BLAST searches performed while logged in to a My NCBI account enables access to BLAST search results for their full 36-hour life span through the "Recent Results" tab. Strategies saved will be saved permanently.

The “Recent Results” page

BLAST search results are available for 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. The My NCBI login from the header is shown as an insert at the upper right (A). Each result is given as a row in the table. The identifier in the Request ID (RID) column (B) provides an one-click access to the search result. The program, Title and Database column (C) combine to provide a summary for a specific search. Database restriction applied are indicated by “more ...” and the popup upon mouseover (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.

NCBI/ BLAST/ Recent Results

Links to your unexpired BLAST jobs appear below. [more...](#)

My NCBI Welcome ttao_phd. [Sign Out]

Lookup BLAST Job

Request ID: Go

Your Recent Results

(Click headers to sort columns)

Submitted at	Request ID	Status	Program	Title	Qlength	Database	Expires at	
08-11 12:23	0FEKPWZ1014	Running	d.megablast	segment of chromosome 1	50001	wgs more...	08-13 00:23	save download x
08-11 11:49	0FCKVM77015	Done	blastn	(3) - embjF12345 (327 letters)	327	wgs more...	08-12 23:49	save download x
08-11 00:23	0E4E3Y9C014	Done	megablast	(2) - embjF12345 (327 letters)	327	Genome (all accessible scaffolds) Homo sapiens included: Vicugna pacos (taxid:30538) 	08-12 12:23	save download x
08-11 00:22	0E4BKB2X014	Done	blastp	dbj BAG37749.1 (122 letters)	122	nr	08-12 12:22	save download x
08-10 23:12	0E089W23015	Done	blastn	embjF12345 (327 letters)	327	refseq_ma	08-12 11:12	save download x

NCBI/ BLAST/ Format Request

Request ID: [View report](#) Show results in a new window

Format

Show as Old View [Reset form to defaults](#)

Alignment View

Display Graphical Overview NCBI-gi CDS feature

Masking Character: Color:

Limit results Descriptions: Graphical overview: Line length:

Pairwise Organism Exclude +

Entrez query:

Expect Min: Expect Max:

Percent Identity Min: Percent Identity Max:

Format for PSI-BLAST with inclusion threshold:

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.

Clicking the “Go” button without an RID loads the format form (G) where the content (H) and format (I) of results can be adjusted.

The “Saved Strategies” page

The “Saved Strategies” tab (shown below) displays a list of search strategies. The first four columns (J) provide a good summary of the search settings for each saved entry. The “view” link (K) loads the settings in a search page, while the

“download” link (L) saves the settings in an ASN.1 formatted file for use with standalone BLAST or reloading on the web services using the “Choose File” and “View” button (M). Clicking the red X (N) removes the entry from the list.

NCBI/ BLAST/ Saved Strategies

Browse and run your saved sets of BLAST search parameters. (Query sequences larger than 10K are not saved.) [more...](#)

Upload Search Strategy

Upload file: No file chosen

Your Saved Strategies

(Click headers to sort columns)

Program	Created	Title	Database	
blastp	2013-08-11 00:33:50	dbj BAG37749.1 (122 letters)	nr	view download x
megablast	2013-08-11 00:26:41	(2) - embjF12345 (327 letters)	GPIPE/9606/current/all_contig	view download x
blastn	2012-07-11 13:28:04	AP Biology Fossil Specimen Gene 3	nr	view download x
blastp	2010-04-09 16:01:46	(2) - NP_055023:dentin sialophosphoprotein preproprotein...	nr	view download x

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 combination	Alignment type	Programs & functions (default program in bold)
nucleotide blast	nucleotide vs nucleotide	nucleotide vs 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 vs protein	blastp : general sequence identification and similarity searches DELTA-BLAST [2] : protein similarity search with higher sensitivity than blastp PSI-BLAST : iterative search for position-specific score matrix (PSSM) construction or identification of distant relatives for a protein family PHI-BLAST : protein alignment with input pattern as anchor/constraint
blastx	nucleotide (translated) vs protein	protein vs protein	blastx : for identifying potential protein products encoded by a nucleotide query
tblastn	protein vs nucleotide (translated)	protein vs protein	tblastn : for identifying database sequences encoding proteins similar to the query
tblastx	nucleotide (translated) vs nucleotide (translated)	protein vs protein	tblastx : for identifying nucleotide sequences similar to the query based on their coding potential

Table 2. Contents of the common BLAST sequence databases

Database	Type	Content
nr (nt) default	Nucleotide	All GenBank + EMBL + DDBJ + PDB sequences, excluding sequences from PAT, EST, STS, GSS, WGS, TSA and phase 0, 1 or 2 HTGS sequences, mostly non-redundant.
refseq_rna	Nucleotide	Curated (NM_, NR_) plus predicted (XM_, XR_) sequences from NCBI Reference Sequence Project.
refseq_genomic	Nucleotide	Genomic sequences from NCBI Reference Sequence Project.
chromosome	Nucleotide	Complete genomes and complete chromosomes from the NCBI Reference Sequence project.
Human G+T	Nucleotide	The genomic sequences plus curated and predicted RNAs from the current build of the human genome.
Mouse G+T	Nucleotide	The genomic sequences plus curated and predicted RNAs from the current build of the mouse genome.
est	Nucleotide	Database of GenBank + EMBL + DDBJ sequences from EST division
HTGS	Nucleotide	Unfinished High Throughput Genomic Sequences; Sequences: phases 0, 1 and 2
wgs	Nucleotide	Assemblies of Whole Genome Shotgun sequences.
pat	Nucleotide	Nucleotides from the Patent division of GenBank.
pdb	Nucleotide	Nucleotide sequences from the 3-dimensional structure records from Protein Data Bank.
alu_repeats	Nucleotide	Selected Alu repeats from REPBASE, suitable for identifying Alu repeats from query sequences. See "Alu alert" by Claverie and Makalowski, Nature 371: 752 (1994).
TSA	Nucleotide	Transcriptome Shotgun Assemblies, assembled from RNA-seq SRA data
16S microbial	Nucleotide	16S Microbial rRNA sequences from Targeted Loci Project
nr default	Protein	Non-redundant GenBank CDS translations + RefSeq + PDB + SwissProt + PIR + PRF, excluding those in PAT, TSA, and env_nr.
refseq_protein	Protein	Protein sequences from NCBI Reference Sequence project.
swissprot	Protein	Last major release of the UniProtKB/SWISS-PROT protein sequence database (no incremental updates).
pat	Protein	Proteins from the Patent division of GenBank.
pdb	Protein	Protein sequences from the 3-dimensional structure records from the Protein Data Bank.
env_nr	Protein	Protein sequences translated from the CDS annotation of metagenomic nucleotide sequences.
tsa_nr	Protein	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 the 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 [1]. For a single query, “Query subrange” boxes (E) define a segment of the query to be used in the search. Query sequences saved in a plain text file can be uploaded using the “Choose File” button (F). The “Align two or more sequences” checkbox (G) changes the “Choose Search Set” sections below to “Enter Subject Sequence” sections below to “Enter Subject Sequence” to allow comparison between the query and the those in the subject input box (H).

Choose Search Set

BLAST database can be selected from the standard list using the pull-down menu (I). A search can be restricted to a subset of entries in the selected database using the “Organism” field by typing the name of the species, strains, or taxonomic group in the text-box and selecting from the suggested list (J). Checking the exclusion box to the right excludes sequences from the selected organism from the search. Multiple organisms can be selected by adding extra input box using the “+” button. Specific types of relatively low value sequences can be excluded using the checkboxes below (K). For certain databases, entering custom queries in the “Entrez Query” textbox (L) will restrict 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 (M and Table 1) with different speed and sensitivity are available for nucleotide vs nucleotide sequence alignment. The default megablast is better for certain tasks, such as identifying the input query

and searching with large genomic query; discontinuous megablast works better in finding related sequences from other organisms; 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 (N) submits the search to BLAST server for processing. Results will be automatically displayed when completed.

“Algorithm parameters” link (O) opens a normally collapsed section that allows access to other parameter settings, including adjustment of number of alignments saved, search stringencies, scoring systems (score matrix) and gap penalties, as well as query filtering. Detailed descriptions are given next.

Elements of the Standard Nucleotide BLAST search page (cont.)

General Parameters

Parameters in this section specify the search sensitivity. The “Max target sequences” (A) sets the maximum database matches BLAST saves for a query. The checked “Short queries” checkbox (B) allows BLAST to automatically optimize settings for short input queries 50 nucleotides or less. The “Expect threshold” (C) filters out matches that are less significant with Expect value above the setting. The “Word size” (D) set the size of the initial seed match phase, smaller settings are more sensitive. The “Max matches in a query range” (E) limits the matches saved to a given region of the query (such as from repeats) so matches to other region of the query are not crowded out. The default setting of “0” means no limit.

Scoring Parameters

Parameters in this section specify 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.

Filters and Masking

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

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. In addition, checking the “Align two or more sequences” will change the “Program Selection” section to leave blastp as the only choice.

Choose Search Set

Most of the components are similar to the “Standard Nucleotide BLAST” page (pg.3). The main difference is that the database pull-down menu contains a smaller list of protein databases (J).

Program Selection

Four different programs (K and Table 1) are available to satisfy various search needs. 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 and for 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 (L). Parameters specific to protein blast are described next.

Elements of the Standard Protein BLAST search page (cont.)

The “Algorithm parameters” portion of the “Standard Protein BLAST” search page 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 available (A). The default BLOSUM62 matrix 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 pull-down menu (B). Scores for protein alignment can be adjusted

to account for the bias in sequence composition using various approaches as indicated by the “Compositional adjustments” setting. Other options including “no adjustment” can be selected using the pull-down menu (C).

Parameters in the section specify whether low complexity and whether to filter only at the seed lookup stage (D). Lower case letters in the query (provided as a mixed upper and lower case letters in FASTA, representing custom features) can also be masked (E). These settings are not needed when compositional adjustments are used.

Filters and Masking

Parameters in the section specify whether low complexity and whether to filter only at the seed lookup stage (D). Lower case letters in the query (provided as a mixed upper and lower case letters in FASTA, representing custom features) can also be masked (E). These settings are 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 BLAST: blastx

In the “Enter Query Sequence” section, a “Genetic code” field (F) 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 BLAST: tblastn

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 BLAST: tblastx

The layout differences are the presence of the “Genetic code” field (F), which is also present in the blastx page, and the databases listed under the Database pull-down menu. In addition, the main input box in the “Enter Query Sequence” takes a nucleotide query.

Other search pages

BLAST search pages under the “BLAST Assembled RefSeq Genomes” category differ from these under the “Basic BLAST” category only in the databases they access. The link names clearly indicate the source organism(s) for the database sequences against which the query will be searched.

The “Specialized BLAST” category contains different types of search pages. Those using the core BLAST programs and the same general layout described about are summarized in Table 3.

Table 3. Function of Specialized BLAST pages following the standard layout

Page name	Searching against specialized databases
Search trace archives	unannotated nucleotide capillary sequencing trace reads from various organisms
Search sequences that have gene expression profiles (GEO)	nucleotide sequences with expression information
Align two (or more) sequences using BLAST (bl2seq)	direct comparison of two groups of sequences. This is the “Basic BLAST” page with “Align two or more sequences” checked
Search protein or nucleotide targets in PubChem BioAssay,	nucleotide or protein sequences with associated chemical activity assay data from PubChem
Search SRA by experiment	raw nucleotide reads from the next generation sequencing technology
Search RefSeqGene	nucleotide sequences representing selected loci from the RefSeqGene project

Other search pages (cont.)

Other search pages in the “Specialized BLAST” category have non-standard page layouts. These services use BLAST or other alignment programs in combination with other tools to accomplish specialized tasks. Table 4 provides a functional summary of these pages.

Table 4. Function of Specialized BLAST pages not following the standard layout

Page link name	Functions
Make specific primers with Primer-BLAST [4, 5]	Designing primers using the primer3 algorithm and checking their template specificity using BLAST
Search immunoglobulins and T cell receptor sequences (IgBLAST) [6]	Searching immunoglobulin or T cell receptor sequences against germline databases for annotation of the input immunoglobulin sequences
Screen sequence for vector contamination (vecscreen)	Screening input nucleotide sequences against a library of known vector and other artificial sequences to identify contaminations
Find conserved domains in your sequence (cdfs) [7]	Searching an protein sequence against a database of curated domains for functional analysis. This search is performed for all protein-blast requests.
Find sequences with similar conserved domain architecture (cdart)	Identifying conserved domains present in the input protein sequence followed by finding other sequences containing these identified domains
Constraint Based Protein Multiple Alignment Tool [8]	A protein multiple sequence alignment tool from NCBI. COBALT search can also be launched from any protein BLAST search results.
Needleman-Wunsch Global Sequence Alignment Tool	NCBI implementation of the Needleman-Wunch global pair-wise alignment tool for nucleotide or protein queries
Cluster multiple sequences together with their database neighbors (MOLE-BLAST)	This tool takes multiple related input sequences, finds their nearest neighbors from the database using BLAST, and then cluster these sequences according to their sequence similarities.

Other ways to access NCBI web BLAST services

In addition to access through a web browser, BLAST web services described above, with the exception for those listed in Table 4, can also be accessed using alternative venues. These venues include the “-remote” option in different standalone BLAST+ programs, the RESTful BLAST service (known as QBLAST or BLAST URLAPI), and BLAST SOAP services. The features of these venues are summarized in Table 5 below.

Table 5. 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 Versatility: available option enables searching against custom sequences Job Limitation: Not meant for high throughput searches with 1 hour CPU time limit Data partition: Access to different database requires different search pages
Standalone BLAST+ (-remote option) [9]	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) [10]	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 [11]	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 if not prefetched 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

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.

Information needed when reporting problems encountered during web BLAST searches are:

- A description on the goal of the search
- The RID of the search
- The detailed error message
- The search page and settings used along with a summary of the input query, particularly if the RID was not issued

CPU related errors are caused when searches exceed the processing time limit. Repeat the search using “Edit and re-submit” link to get back to the search page with the following adjustments will help resolve the issue:

- Reduce the number 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 focused 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

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 and SOAP 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. <http://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. doi: 10.1186/1745-6150-7-12.
4. [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.
5. Factsheet: Primer-BLAST. ftp://ftp.ncbi.nih.gov/pub/factsheets/HowTo_PrimerBLAST.pdf
6. [IgbLAST: an immunoglobulin variable domain sequence analysis tool](#). Ye J, et. al. Nucleic Acids Res. 2013 Jul 1;41 (Web Server issue):W34-40.
7. [CDD: specific functional annotation with the Conserved Domain Database](#). Marchler-Bauer A, et. al. Nucleic Acids Res. 2009 Jan;37(Database issue):D205-10.
8. [COBALT: constraint-based alignment tool for multiple protein sequences](#). Papadopoulos JS, Agarwala R. Bioinformatics. 2007 May 1;23(9):1073-9.
9. BLAST® Help manual. <http://www.ncbi.nlm.nih.gov/books/NBK1762/>.
10. QBLast's URL API User Guide. <http://www.ncbi.nlm.nih.gov/blast/Doc/urlapi.html>
11. Using SRA Toolkit to BLAST SRA Data Locally: ftp://ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_Local_SRA_BLAST.pdf
12. Entrez Programming Utilities Help. <http://www.ncbi.nlm.nih.gov/books/NBK25501>.
13. Blastdbinfo: API access to a database of BLAST databases. NCBIINSIGHT blog entry <http://ncbiinsights.ncbi.nlm.nih.gov/2013/03/19/blastdbinfo-api-access-to-a-database-of-blast-databases/>