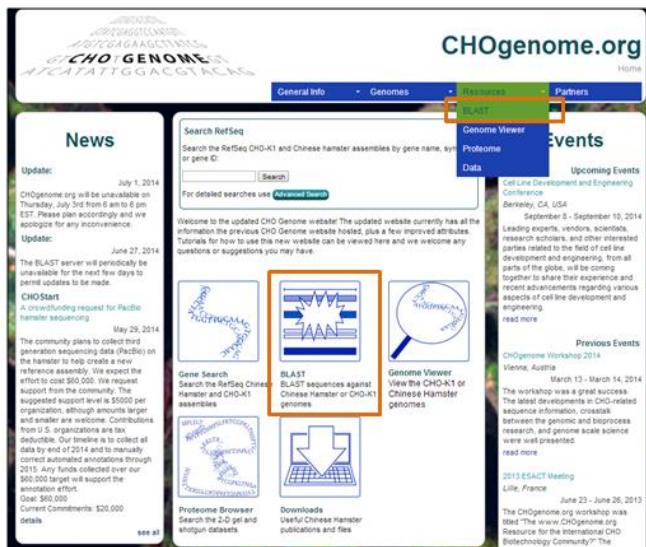


Tutorial 4 – BLAST Searching the CHO Genome

Accessing the CHO Genome BLAST Tool

The CHO BLAST server can be accessed by clicking on the BLAST button on the home page or by selecting “BLAST” from the menu bar under the Resources tab. This tab is available on all web pages within the CHO genome project. An additional link to the CHO genome BLAST web server is also provided on the CHO genome search pages, as well as a link to the NCBI BLAST web server.

Select the BLAST icon on the home page or from the resources tab



Click on the BLAST server link on the CHO genome search pages

This screenshot shows the 'Search Page' of the CHO genome database. It features a search form at the top and a list of search terms on the left. In the center, there is a section titled 'The CHO-K1 RefSeq database can be searched by:' with a numbered list of search criteria. One of the items in the list, 'BLAST the CHO-K1 RefSeq and Chinese Hamster RefSeq genome [here]' is highlighted with a red box. To the right of this list, there is a 'Tips for using the database:' section with a numbered list of tips, the fourth item of which also has a red box around it.

Search Term Genome

Select at least one Select at least one

Symbol CHO-K1 (RefSeq Assembly GCF_000223151.1 | 2May2014 - Release 101)

Gene Name Chinese Hamster (RefSeq Assembly GCF_000419365.1 | 2May2014 - Release 101)

Gene ID CHO-K1 (RefSeq Assembly GCF_000223151.1 | 15Mar2012 - Release 1)

The CHO-K1 RefSeq database can be searched by:

1. Gene name (i.e. Caspase 1)
2. Gene symbol (i.e. Casp1)
3. Gene ID (i.e. 100759171)

BLAST the CHO-K1 RefSeq and Chinese Hamster RefSeq genome [here] and at NCBI.

Tips for using the database:

1. Search by gene name, symbol, or ID to find individual gene pages.
2. Multiple genomes may be selected at once, but the time required for the query may increase.
3. Each gene, transcript, and protein has a unique, individual entry. To obtain the relevant protein information or download the protein sequence, select the gene or transcript entry of interest, scroll to the bottom of the "Gene Details" page, and select the protein entry associated with the relevant transcript in the "Gene Relations" table.
4. Many pseudogenes do not have a gene name or symbol, but all have a gene ID and may be searched.

Using the CHO Genome BLAST Tool

- 1) The CHO BLAST page allows for BLAST searches against the CHO and Chinese hamster (CH) genome databases.

CHOblast Search

BLAST Search - Required parameters

Enter query sequences here in Fasta format

Or upload fasta file: No file chosen

Algorithm:

Database(s)

Genome (Scaffolds)

- 1) CHO-K1[ATCC]_RefSeq_2014
- 2) CH_RefSeq_2014
- 3) CHO-K1[ATCC]_GenBank_2011
- 4) CH_GenBank_2013
- 5) CH-17A/GY_Chr_GenBank_2013

Transcripts (RNA)

- 6) CHO-K1[ATCC]_RefSeq_2014
- 7) CHO-K1[ATCC]_RefSeq_2012
- 8) CH_RefSeq_2014

Currently available nucleotide and protein databases

Nucleotide Databases: <ul style="list-style-type: none"> Genome (Scaffolds) 1) CHO-K1[ATCC]_RefSeq_2014 2) CH_RefSeq_2014 3) CHO-K1[ATCC]_GenBank_2011 4) CH_GenBank_2013 5) CH-17A/GY_Chr_GenBank_2013 <ul style="list-style-type: none"> Transcripts (RNA) 6) CHO-K1[ATCC]_RefSeq_2014 7) CHO-K1[ATCC]_RefSeq_2012 8) CH_RefSeq_2014 	Amino Acid Databases: <ul style="list-style-type: none"> Proteins 1) CHO-K1[ATCC]_RefSeq_2014 2) CHO-K1[ATCC]_RefSeq_2012 3) CH_RefSeq_2014 4) CHO-K1[ATCC]_GenBank_2011 5) CH-17A/GY_Chr_GenBank_2013
---	---

Assembly Color Key:

RefSeq Assembly CH Chinese hamster ovary cell line
 GenBank Assembly CH Chinese hamster cell
 CH(O)-xxxxx Strain definition
 [xxxx] Source of cells

Assembly ID Key:

CHO-K1 RefSeq (GCF_000223135.1)
 CH RefSeq (GCF_000419365.1)
 CHO-K1 GenBank (GCA_000223135.1)
 CH GenBank (GCA_000419365.1)
 CH-17A/GY GenBank (GCA_000448345.1)

Database Naming Convention:

CHO Chinese hamster ovary cell line
 CH Chinese hamster cell
 CH(O)-xxxxx Strain definition
 [xxxx] Source of cells
 genbank GenBank assembly
 refseq RefSeq assembly
 chr Chromosomal identification

BLAST Search - Other parameters

Expect threshold: <input type="text" value="10"/> Word size: <input type="button" value="11"/> Max target sequences: <input type="button" value="50"/> Match/Mismatch scores: <input type="button" value="2.-3"/> Gap costs: <input type="button" value="Existence: 5, Extension: 2"/>	Filter: <input checked="" type="checkbox"/> Low complexity regions Mask: <input checked="" type="checkbox"/> Mask for lookup table only <input type="checkbox"/> Mask for lower case letters Alignment: <input type="checkbox"/> Perform ungapped alignment Alignment output format: <input type="button" value="pairwise"/> Other parameters: <input type="text"/>
--	--

Hosted by Delaware Biotechnology Institute / CBCB at the University of Delaware
 BLAST tool adapted from ViroBLAST v2.2 © 2005-2010 University of Washington. All rights reserved. ([Terms of Service](#))

The nucleotide and amino acid databases hosted on the Chinese hamster genome database are listed to the right of the Basic Search panel. The nucleotide databases are divided into Genome (scaffold) and Transcript (RNA) databases, while the amino acid databases consist only of protein databases. The organism or cell line of origin is listed first, followed by the type of assembly (RefSeq or GenBank), and finally the year of release. The keys for the abbreviations and naming conventions are listed below these database lists.

2) Additional details regarding the multiple BLAST programs and databases are available. Clicking on the **Algorithm** link provides a brief description of the BLAST programs.

Programs available for CHOBlast	
blastn	compares a nucleotide query sequence against a nucleotide sequence database
blastp	compares an amino acid query sequence against a protein sequence database
blastx	compares a nucleotide query sequence translated in all reading frames against a protein sequence database
tblastn	compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames
tblastx	compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database

Clicking on the **Database(s)** link or the **details** button in the “Currently available nucleotide and protein databases” section title bar brings up a webpage with a brief description of the databases currently available for BLAST searching, including the name, version, date, and a link to the original publication article.

Databases available for CHOBlast		
Nucleotide sequence databases (blastn, tblastn, tblastx):		
Genome (Scaffolds)		
Name	Release Date/ Assembly Version	Reference
1) CHO-K1(ATCC)_RefSeq_2014	08 May 2014 Assembly v1.0	Xu X, Nagarajan H, Lewis NE et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. <i>Nature Biotechnology</i> , 29(6), 735-741 (2011). [LINK]
2) CH_RefSeq_2014	08 May 2014 Assembly v1.0	Lewis NE, Liu X, Li Y et al. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the <i>Cricetulus griseus</i> draft genome. <i>Nature Biotechnology</i> , 31(8), 759-765 (2013). [LINK]
3) CHO-K1(ATCC)_GenBank_2011	26 August 2011 Assembly v1.0	Xu X, Nagarajan H, Lewis NE et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. <i>Nature Biotechnology</i> , 29(6), 735-741 (2011). [LINK]
4) CH_GenBank_2013	12 July 2013 Assembly v1.0	Lewis NE, Liu X, Li Y et al. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the <i>Cricetulus griseus</i> draft genome. <i>Nature Biotechnology</i> , 31(8), 759-765 (2013). [LINK]
5) CH-17A/GY_Chr_GenBank_2013	29 August 2013 Assembly v1.0	Brinkhoff K, Rupp O, Laux H et al. Chinese hamster genome sequenced from sorted chromosomes. <i>Nature Biotechnology</i> , 31(8), 694-695 (2013). [LINK]
Transcripts (RNA)		
Name	Release Date/ Assembly Version/ Annotation Version	Reference
6) CHO-K1(ATCC)_RefSeq_2014	08 May 2014 Assembly v1.0 Annotation v101	Xu X, Nagarajan H, Lewis NE et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. <i>Nature Biotechnology</i> , 29(6), 735-741 (2011). [LINK]
7) CHO-K1(ATCC)_RefSeq_2012	15 March 2012 Assembly v1.0 Annotation v1	Xu X, Nagarajan H, Lewis NE et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. <i>Nature Biotechnology</i> , 29(6), 735-741 (2011). [LINK]
8) CH_RefSeq_2014	08 May 2014 Assembly v1.0 Annotation v101	Lewis NE, Liu X, Li Y et al. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the <i>Cricetulus griseus</i> draft genome. <i>Nature Biotechnology</i> , 31(8), 759-765 (2013). [LINK]
Amino Acid Sequence Databases (blastp, blastx):		
Proteins		
Name	Release Date/ Assembly Version/ Annotation Version	Reference
1) CHO-K1(ATCC)_RefSeq_2014	08 May 2014 Assembly v1.0 Annotation v101	Xu X, Nagarajan H, Lewis NE et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. <i>Nature Biotechnology</i> , 29(6), 735-741 (2011). [LINK]
2) CHO-K1(ATCC)_RefSeq_2012	15 March 2012 Assembly v1.0 Annotation v1	Xu X, Nagarajan H, Lewis NE et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. <i>Nature Biotechnology</i> , 29(6), 735-741 (2011). [LINK]
3) CH_RefSeq_2014	08 May 2014 Assembly v1.0 Annotation v101	Lewis NE, Liu X, Li Y et al. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the <i>Cricetulus griseus</i> draft genome. <i>Nature Biotechnology</i> , 31(8), 759-765 (2013). [LINK]
4) CHO-K1(ATCC)_GenBank_2011	26 August 2011 Assembly v1.0	Xu X, Nagarajan H, Lewis NE et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. <i>Nature Biotechnology</i> , 29(6), 735-741 (2011). [LINK]
5) CH-17A/GY_Chr_GenBank_2013	29 August 2013 Assembly v1.0	Brinkhoff K, Rupp O, Laux H et al. Chinese hamster genome sequenced from sorted chromosomes. <i>Nature Biotechnology</i> , 31(8), 694-695 (2013). [LINK]
Database Naming Convention:		
CHO Chinese hamster ovary cell line	Assembly Color Key:	
CH Chinese hamster cell	RefSeq Assembly	
CHO-<xxxx> Strain definition	GenBank Assembly	
<xxxx> Source of cells	Assembly ID Key:	
GenBank GenBank assembly	CHO-K1 RefSeq (GCF_000223195.1)	
RefSeq RefSeq assembly	CH RefSeq (GCF_000419365.1)	
Chr Chromosomal identification	CHO-K1 GenBank (GCA_000223195.1)	
	CH GenBank (GCA_000419365.1)	
	CH-17A/GY GenBank (GCA_000448345.1)	

3) Query sequences in FASTA format can be pasted into the search box at the top of the page or uploaded as a FASTA file. Multiple query sequences may be entered for each search.

The BLAST program and database are then selected from the currently available options.

For example, to BLAST the most recent CH genome protein database, select the **blastp** program and the **CH_RefSeq_2014** database from the “Proteins” list.

BLAST Search - Required parameters

Enter query sequences here in Fasta format

```
>gi|186198305
MAWDILRAKRKRFQFINSVSIQTINGLLDELLERKVLNQGRENCKIRKLANITAMDFARDLCDHVSKRGPFQASQ
IIFTIVCINEDCYLAGILELGQSAPSAETPATEDSKSGHMPSSSETEEQNKEDGTGFLTGTLKPCFLEKA
QKLWKEHPSEIYPIMITITRLLALIICHTEPQLSPRVGAQVDLRENKLLLEDGYTVVKVKENLTALEK
VEEVKEFAACPERKTSOSTPLVFMSHQIQEICGTYSVNEVSQDILKVDTIFQMOBTLEKPSLKDKK&KVII
IQACRGERKQGVVLLKEDSVRDSEEDPLTDIAPEDGQIKKANIEKDIFIAFCGSSTPONVSMRSPVRGSLFIES
LIKHOKEYAMSCDLEDIFREVRFSFEPFRLQMPTADRVTLTERFYLFPPGH
```

Or upload fasta file: No file chosen

Algorithm:

Database(s):

To perform a basic BLAST search, click the **BLAST** button after all the above information is entered and selected. If you wish to perform a more advanced search, do not hit basic search yet and proceed to instruction #4.

4) In the BLAST Search – Other parameters section, the default BLAST parameters can be varied to perform an altered, more advanced BLAST search.

BLAST Search - Other parameters

Expect threshold:

Word size:

Max target sequences:

Match/Mismatch scores:

Gap costs:

Filter: Low complexity regions

Mask: Mask for lookup table only
 Mask for lower case letters

Alignment: Perform ungapped alignment

Alignment output format:

Other parameters:

Clicking on the highlighted blue terms (such as **Expect threshold**, etc.) will provide a brief description of each advanced search parameter that can be varied.

To perform an advanced BLAST search, click the **BLAST** button once all the required information is entered and the advanced parameters are altered.

5) The results of the BLAST alignment are summarized in a table with the query sequence name, the subject sequence name, the bit score, the identity length, the identity percentage, and the *E*-value.

The results can be filtered by score (showing only the top 1, 5, or 10 alignments), by Similarity Cutoff Percentage, or by BLAST Bit Score.

After entering the filter parameter, click either the “Filter” or the “Parse again” buttons to refresh the results table.

To view the RefSeq/GenBank entry for each subject sequence, click on the sequence name in the Subject column (*i.e.* [gi|625278770|ref|XP_007631029.1|](#)).

To view the pair-wise alignment for a specific alignment, click on the value in the Score column for any alignment (*i.e.* [598](#)).

BLAST Results					
Query	Subject	Bit_Score	Identity (Query_Len)	Similarity	E-Value
gi 86198305	gi 625278770 ref XP_007631029.1 PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]	598	284/363 (402)	78	0.0
gi 86198305	gi 625278768 ref XP_007631028.1 PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]	664	318/402 (402)	79	0.0
gi 86198305	gi 625278766 ref XP_007631027.1 PREDICTED: caspase-12-like isoform X2 [Cricetulus griseus]	280	172/437 (402)	39	3e-89
gi 86198305	gi 625260408 ref XP_007621635.1 PREDICTED: caspase-14 [Cricetulus griseus]	135	65/89 (402)	73	7e-37
gi 86198305	gi 625260409 ref XP_007621635.1 PREDICTED: caspase-14 [Cricetulus griseus]	72.8	54/195 (402)	28	2e-14
gi 86198305	gi 625241519 ref XP_007611988.1 PREDICTED: caspase-6 isoform X1 [Cricetulus griseus]	68.9	67/230 (402)	29	6e-13
gi 86198305	gi 625241513 ref XP_007611987.1 PREDICTED: caspase-6 isoform X1 [Cricetulus griseus]	68.9	65/229 (402)	28	9e-13
gi 86198305	gi 625249026 ref XP_007615799.1 PREDICTED: caspase-7 isoform X3 [Cricetulus griseus]	67.4	65/246 (402)	26	4e-12
gi 86198305	gi 625249255 ref XP_007615919.1 PREDICTED: caspase-9-like isoform X2, partial [Cricetulus griseus]	62.4	46/153 (402)	30	3e-11
gi 86198305	gi 625256876 ref XP_007619604.1 PREDICTED: putative caspase-16 [Cricetulus griseus]	48.5	47/164 (402)	29	6e-06
gi 86198305	gi 625249026 ref XP_007615798.1 PREDICTED: caspase-14 [Cricetulus griseus]	43.0	50/238 (402)	26	1e-04

BLAST Results

Query	Subject	Bit_Score	Identity (Query_Len)	Similarity	E-Value
gi 86198305	gi 625278770 ref XP_007631029.1 PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]	598	284/363 (402)	78	0.0
gi 86198305	gi 625278768 ref XP_007631028.1 PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]	664	318/402 (402)	79	0.0
gi 86198305	gi 625278766 ref XP_007631027.1 PREDICTED: caspase-12-like isoform X2 [Cricetulus griseus]	280	172/437 (402)	39	3e-89
gi 86198305	gi 625278791 ref XP_007631040.1 PREDICTED: caspase-1-like [Cricetulus griseus]	135	65/89 (402)	73	7e-37
gi 86198305	gi 625260408 ref XP_007621635.1 PREDICTED: caspase-14 [Cricetulus griseus]	72.8	54/195 (402)	28	2e-14

PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]

NCBI Reference Sequence: XP_007631029.1
[FASTA](#) [Graphics](#)

[Go to: ▾](#)

Locus: XP_007631029 363 aa linear ROD 30-AFR-2014
 Definition: PREDICTED: caspase-1 isoform X2 [Cricetulus griseus].
 Accession: XP_007631029
 Version: XP_007631029.1 GI:625278770
 DBLINK: BioProject: PRJNA239316
 DBSOURCE: REFSEQ: accession XM_007632839.1

> gi|86198305 on [gi|625278770|ref|XP_007631029.1|](#) PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]
 Length=363

Score = 598 bits (1542), Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 284/363 (78%), Positives = 317/363 (87%), Gaps = 0/363 (0%)

Query 40 MDKIKLANTAMKARDLCDHVSKKGPFQASIFITYICNEDCYLAGILELSQAPSATFV 99
 M++IK N T MKKARDLCD V+KKGP ASQI ITYIC EDCYLAG+LEI+S P AE +
 Sbjct 1 MERIKCINATVMDKARDLCDSVTRKGFLASQICITYICEDCYLAGVLELESQPPAENSM 60

Query 100 ATEDSKGGHPSSSETKEEQNKEKGDTFFGGLTGTILKFCPLEKAQKLWNKENPSEIYFIMNTTT 159
 T+D +G+PSSSETKEEQ KE GT PG +G+LK C LE AQR+ KENPSEIYFIM-T+T
 Sbjct 61 RTDDFGGGYFSSSETKEEQKKEGGTCPGPSGSSSLKLCSELAQKIRKENPSEIYFIMDTST 120

6) To download the resulting BLAST sequences, click the “Check here to download all sequences” box or select individual subject boxes within the table and then hit the “Download” button. A text file of the selected BLAST sequence(s) in the FASTA format will open in a new window.

BLAST Results

Query	Subject	Bit_Score	Identity (Query_Len)	Similarity	E-Value
gi 86198305	<input checked="" type="checkbox"/> gi 625278770 ref XP_007631029.1 PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]	598	284/363 (402)	78	0.0
gi 86198305	<input checked="" type="checkbox"/> gi 625278768 ref XP_007631028.1 PREDICTED: caspase-1 isoform X1 [Cricetulus griseus]	664	318/402 (402)	79	0.0
gi 86198305	<input type="checkbox"/> gi 625278766 ref XP_007631027.1 PREDICTED: caspase-12-like isoform X2 [Cricetulus griseus]	280	172/437 (402)	39	3e-89
gi 86198305	<input type="checkbox"/> gi 625278791 ref XP_007631040.1 PREDICTED: caspase-1-like [Cricetulus griseus]	135	65/89 (402)	73	7e-37
	<input type="checkbox"/> gi 25260409 ref XP_007621625.1 PREDICTED:				

```
>gi|625278768|ref|XP_007631028.1| PREDICTED: caspase-1 isoform X1 [Cricetulus griseus]
MADNLIRAKRKQFIRSVGAATGTINGLDELLEKVLNQEEMERIKCINAVTMDKARDLCDSVTKKGFLASQ
ICITYCKEDCYLAGVLELESGPAAENSMRTEDDFGGGYSSSETKEECKEGGTCPFGSGSLKLCSLETA
QRIRKENPSEIYIMDTSTRTRLALIICTFEYLPRDGADVDLREMSLLQDGLYTVEKENLTALEM
ANAVYKEFADCPHEHTSDTFLVPMHSQHIGEIGCKGSYSD
IACRGENRGRVVLVKDSVEDTGGKFLVDADLEDDGIKAHIEKFIAFCSTSFPDVNSWRHPLKGSLFIVE
LHKMQMEYANSCLDEDIFRKVRFRSFEPSPSYSLQMPTIERVLTKRFYLFPGH

>gi|625278770|ref|XP_007631029.1| PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]
MERIKCINAVTMDKARDLCDSVTKKGFLASQICITYCKEDCYLAGVLELESGPAAENSMRTEDDFGGGYSP
SSSETKEECKEGGTCPFGSGSLKLCSLETAQKIRKENPSEIYIMDTSTRTRLALIICTFEYLPRD
GADVDLREMSLLQDGLYTVEKENLTALEMANAVKEFADCPHEHTSDTFLVPMHSQHIGEIGCKGSYSD
KVADFLVK/DTIFRMNLNTLKCPSLKDCKVII IQACRGENRGRVVLVKDSVEDTGGKFLVDADLEDDGIKKA
HIEKFIAFCSTSFPDVNSWRHPLKGSLFIVELKQMEYANSCLDEDIFRKVRFRSFEPSPSYSLQMPTIER
VLTKRFYLFPGH
```

7) Multiple query sequences can be BLAST searched at one time. The results from all query sequences will be displayed in a single results table. To inspect all the BLAST pair-wise alignments, click on [Raw BLAST Output Report](#), the link located under the “Download BLAST Results” heading. Links provided at the top of the page, in the red outlined box, can be used to quickly navigate between the alignment results for each, individual query sequence.

BLAST Search - Required parameters

Enter query sequences here in Fasta format

```
gi|86198305
>gi|625278770|ref|XP_007631028.1| PREDICTED: caspase-1 isoform X1 [Cricetulus griseus]
>gi|625278768|ref|XP_007631029.1| PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]
>gi|625278766|ref|XP_007631027.1| PREDICTED: caspase-12-like isoform X2 [Cricetulus griseus]
>gi|625278791|ref|XP_007631040.1| PREDICTED: caspase-1-like [Cricetulus griseus]
```

Or upload fasta file: [Choose File](#) No file chosen

Algorithm: blastp - Protein Database

Database(s):

Proteins: 1) CHO-K1(TCC)_RefSeq_2014
2) CHO-K1(TCC)_RefSeq_2012
3) Ch_RefSeq_2014
4) CHO-K1(TCC)_GenBank_2011
5) Ch-17A9_C_RefSeq_2010

BLAST Results

Query	Subject	Bit_Score	Identity (Query_Len)	Similarity	E-Value
gi 86198305	<input type="checkbox"/> gi 625278770 ref XP_007631029.1 PREDICTED: caspase-1 isoform X2 [Cricetulus griseus]	598	284/363 (402)	78	0.0
gi 86198305	<input type="checkbox"/> gi 625278768 ref XP_007631029.1 PREDICTED: caspase-1 isoform X1 [Cricetulus griseus]	664	318/402 (402)	79	0.0
gi 86198305	<input type="checkbox"/> gi 625278766 ref XP_007631027.1 PREDICTED: caspase-12-like isoform X2 [Cricetulus griseus]	280	172/437 (402)	39	3e-89
gi 86198305	<input type="checkbox"/> gi 625278791 ref XP_007631040.1 PREDICTED: caspase-1-like [Cricetulus griseus]	135	65/89 (402)	73	7e-37
gi 86198305	<input type="checkbox"/> gi 25260409 ref XP_007621625.1 PREDICTED:				
	<input type="checkbox"/> gi 25260409 ref XP_007621635.1 PREDICTED: caspase-14 [Cricetulus griseus]	72.8	54/195 (402)	28	2e-14
	<input type="checkbox"/> gi 25260409 ref XP_007621636.1 PREDICTED: caspase-8 isoform X2 [Cricetulus griseus]	68.9	67/230 (402)	29	6e-13
	<input type="checkbox"/> gi 25241513 ref XP_007611987.1 PREDICTED: caspase-8 isoform X1 [Cricetulus griseus]	68.9	65/229 (402)	28	9e-13
	<input type="checkbox"/> gi 25249028 ref XP_007611579.1 PREDICTED: caspase-7 isoform X3 [Cricetulus griseus]	67.4	65/246 (402)	26	4e-12
	<input type="checkbox"/> gi 25249025 ref XP_007611591.1 PREDICTED: caspase-14 isoform X1 [Cricetulus griseus]	62.4	46/153 (402)	30	3e-11
	<input type="checkbox"/> gi 25250576 ref XP_007611980.1 PREDICTED: putative caspase-16 [Cricetulus griseus]	48.5	47/164 (402)	29	6e-06
	<input type="checkbox"/> gi 25249020 ref XP_007611579.1 PREDICTED:				

BLAST Results

Filter BLAST results by bit score rank: Ascending Descending Sort by E-value Sort by Query length

Filter BLAST results by cutoff criterion: Similarity Cutoff: % Bit Score Cutoff:

Download BLAST Results

Examine and Download Full BLAST Report or Raw BLAST Output Report

Retrieve and download sequences in FASTA format

To download All sequences or Select particular sequences of interest below

Check here to download All sequences or Select particular sequences of interest below

BLAST searching the CHO-K1 Genome at NCBI

A link to the NCBI BLAST web server is also provided on the CHO-K1 genome search pages. To BLAST the CHO genome using the NCBI BLAST web server, enter the required BLAST information and select the “*Cricetulus griseus* WGS” database under the “Choose Search Set” menu.

The screenshot shows the NCBI BLAST search interface. On the left, there's a "Search Page" sidebar with options for "Search Term" (Genome, Symbol, Gene Name, Gene ID) and "Organism" (CHO-K1, Chinese Hamster, CHO-K1). Below this is a section for "The CHO-K1 RefSeq database can be searched by:" with numbered tips. The main search area has tabs for "blastn", "blastp", "blastx", "tblastn", and "tblastx". It includes fields for "Enter Query Sequence", "Or, upload file", "Job Title", and "Align two or more sequences". The "Choose Search Set" section is expanded, showing "Database" (Human genomic + transcript, Mouse genomic + transcript, Others), "Organism" (dropdown set to "Cricetulus griseus (taxid: 10029)", with an "Exclude" checkbox), "Exclude" (checkboxes for Models, Uncultured/environmental sample sequences, Sequences from type material), "Limit to" (checkbox for Entrez Query), and "Program Selection" (radio buttons for Highly similar sequences (megablast), More dissimilar sequences (discontiguous megablast), Somewhat similar sequences (blastn), and a "Choose a BLAST algorithm" dropdown). At the bottom, there's a "BLAST" button and a note about the search database.