

Tutorial 2 – Overview of CHOgenome.org

The website for the Chinese hamster genome database is <http://www.CHOgenome.org>.

CHOgenome.org Home

General Info Genomes Resources Partners

News

Update: July 1, 2014
CHOgenome.org will be unavailable on Thursday, July 3rd from 6 am to 6 pm EST. Please plan accordingly and we apologize for any inconvenience.

Update: June 27, 2014
The BLAST server will periodically be unavailable for the next few days to permit updates to be made.

CHOStart
A crowdfunding request for PacBio hamster sequencing
May 29, 2014
The community plans to collect third generation sequencing data (PacBio) on the hamster to help create a new reference assembly. We expect the effort to cost \$60,000. We request support from the community. The suggested support level is \$5000 per organization, although amounts larger and smaller are welcome. Contributions from U.S. organizations are tax deductible. Our timeline is to collect all data by end of 2014 and to manually correct automated annotations through 2015. Any funds collected over our \$60,000 target will support the annotation effort.
Goal: \$60,000
Current Commitments: \$20,000
details
[see all](#)

Search RefSeq

Search the RefSeq CHO-K1 and Chinese hamster assemblies by gene name, symbol, or gene ID:

For detailed searches use [Advanced Search](#)

Welcome to the updated CHO Genome website! The updated website currently has all the information the previous CHO Genome website hosted, plus a few improved attributes. Tutorials for how to use this new website can be viewed here and we welcome any questions or suggestions you may have.

Gene Search
Search the RefSeq Chinese Hamster and CHO-K1 assemblies

BLAST
BLAST sequences against Chinese Hamster or CHO-K1 genomes

Genome Viewer
View the CHO-K1 or Chinese Hamster genomes

Proteome Browser
Search the 2-D gel and shotgun datasets

Downloads
Useful Chinese Hamster publications and files

Events

Upcoming Events

Cell Line Development and Engineering Conference
Berkeley, CA, USA
September 8 - September 10, 2014
Leading experts, vendors, scientists, research scholars, and other interested parties related to the field of cell line development and engineering, from all parts of the globe, will be coming together to share their experience and recent advancements regarding various aspects of cell line development and engineering.
[read more](#)

Previous Events

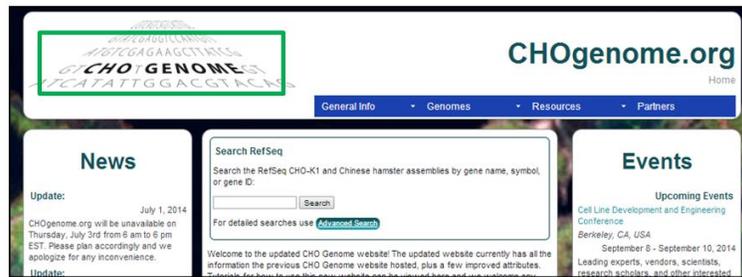
CHOgenome Workshop 2014
Vienna, Austria
March 13 - March 14, 2014
The workshop was a great success. The latest developments in CHO-related sequence information, crosstalk between the genomic and bioprocess research, and genome scale science were well presented.
[read more](#)

2013 ESACT Meeting
Lille, France
June 23 - June 26, 2013
The CHOgenome.org workshop was titled "The www.CHOgenome.org Resource for the International CHO Biotechnology Community". The

The homepage provides access to the underlying database, as well as to the Chinese hamster-specific genomic tools and community resources.

The Version 2.0 tutorials outline the CHOgenome.org website as of August 2014. The legacy website (Version 1.0) will remain available at www.CHOgenome.org/legacy.html, but will not be updated after August 2014.

Users can return to the homepage at any time by clicking on the Home heading on the far left of the navigational menu.



A **News** section on the left side of the homepage displays the most recent news items, while an **Events** section on the right side of the homepage displays the upcoming and most recent events. Additional events and news items can be viewed by selecting the [see all](#) link at the bottom of the respective section or by selecting the **News** or **Events** links from the **General Info** tab of the menu bar.

Questions or comments can be sent to the CHO genome website management team by clicking on the [Please report issues or provide feedback](#) button located at the bottom of all the CHO genome webpages. Please communicate any issues you encounter or pertinent feedback and a CHO genome representative will provide a timely response.

reference assembly. We expect the effort to cost \$60,000. We request support from the community. The suggested support level is \$5000 per organization, although amounts larger and smaller are welcome. Contributions from U.S. organizations are tax deductible. Our timeline is to collect all data by end of 2014 and to manually correct automated annotations through 2015. Any funds collected over our \$60,000 target will support the annotation effort.
Goal: \$60,000
Current Commitments: \$20,000
details

[see all](#)

Type	NCBI Link	Sequence	Graphics	Transcript #
gene			Graphic	
mRNA	XM_003497721.2	Download	Graphic	X4
- Protein	XP_003497769.1	Download	Graphic	X3 14 CDS results found
mRNA	XM_003497718.2	Download	Graphic	X5
- Protein	XP_003497769.1	Download	Graphic	X4 16 CDS results found
mRNA	XM_007616534.1	Download	Graphic	X1
- Protein	XP_007614724.1	Download	Graphic	X1 14 CDS results found
mRNA	XM_007616560.1	Download	Graphic	X3
- Protein	XP_007614750.1	Download	Graphic	X1 14 CDS results found
mRNA	XM_007616554.1	Download	Graphic	X8
- Protein	XP_007614744.1	Download	Graphic	X1 14 CDS results found
mRNA	XM_007616544.1	Download	Graphic	X2
- Protein	XP_007614734.1	Download	Graphic	X1 14 CDS results found

[Please report issues or provide feedback](#)

CHOgenome.org version 2.0b

The **Info** section in the center contains a comprehensive gene search box for all the NCBI RefSeq Chinese hamster genome databases, a welcome statement, and icons that serve as direct links to the genomic tools. The icons include: (1) a **Gene Search** icon that leads to the advanced RefSeq search page, (2) a **BLAST** icon that provides access to the 13 Chinese hamster databases hosted on the CHO BLAST server, (3) a **Genome Viewer** icon that leads to the GBrowse tool to allow users to see whole scaffolds with full annotations, (4) a **Proteome Browser** icon that provides access to the CHO proteome databases consisting of the 2D-PAGE and Shotgun results, and (5) a **Downloads** icon that leads to the Chinese hamster-specific data section.

Search Page

Search Term:

Genome:

Select at least one

- Symbol: CHO-K1 (RefSeq Assembly GCF_000223151.1 | 20ay2014 - Release 101)
- Gene Name: Chinese Hamster (RefSeq Assembly GCF_000419095.1 | 20ay2014 - Release 101)
- Gene ID: CHO-K1 (RefSeq Assembly GCF_000223151.1 | 10Mar2012 - Release 1)

CHOblast Search

BLAST Search - Required parameters

Enter query sequences here in FASTA format

Currently available nucleotide and protein databases

Nucleotide Databases:	Amino Acid Databases:
1: CHO-K1[TC2]_RefSeq_2014	1: CHO-K1[TC2]_RefSeq_2014
2: CHO_K1[TC2]	2: CHO-K1[TC2]
3: CHO-K1[TC2]_GenBank_2011	3: CHO_K1[TC2]
4: CHO-K1[TC2]_GenBank_2013	4: CHO-K1[TC2]_GenBank_2013
5: CHO-K1[TC2]_GenBank_2015	5: CHO-K1[TC2]_GenBank_2015
6: CHO-K1[TC2]_GenBank_2017	6: CHO-K1[TC2]_GenBank_2017
7: CHO-K1[TC2]_RefSeq_2010	7: CHO-K1[TC2]_RefSeq_2010
8: CHO-K1[TC2]_RefSeq_2012	8: CHO-K1[TC2]_RefSeq_2012
9: CHO-K1[TC2]_RefSeq_2014	9: CHO-K1[TC2]_RefSeq_2014

Overview

Region: 0k 10k 20k 30k 40k 50k 60k 70k 80k 90k 100k 110k 120k 130k 140k 150k 160k 170k 180k 190k 200k 210k

Details: 2 kbp

Annotations: DNAAG_Contest, Gene (LOC100769435), Exon, mRNA (NM_00512521.1), sorting nexin-3C-like

Genomic Data Resources

This is a temporary page containing links to relevant publications and public data sets. Links to next generation sequencing raw data in the Sequence Read Archive (SRA) are provided when available. Journal access may be limited to subscribed users.

Description	Title	Reference
A chromosome sorting approach was used to facilitate genome assembly from short-read sequences of the Chinese hamster genome	Chinese hamster genome sequenced from sorted chromosomes	Brinkhoff K et al. Nature Biotechnol (2013) 31, 694-695
Draft genomic sequence of the Chinese hamster and resequencing of CHO-K1, CHO4, and CHO-K cell lines	Genomic landscapes of Chinese hamster ovary cell lines as revealed by the <i>Cricetulus crinitus</i> draft genome	Lewis N et al. Nature Biotechnol (2013) 31, 759-765

Proteome

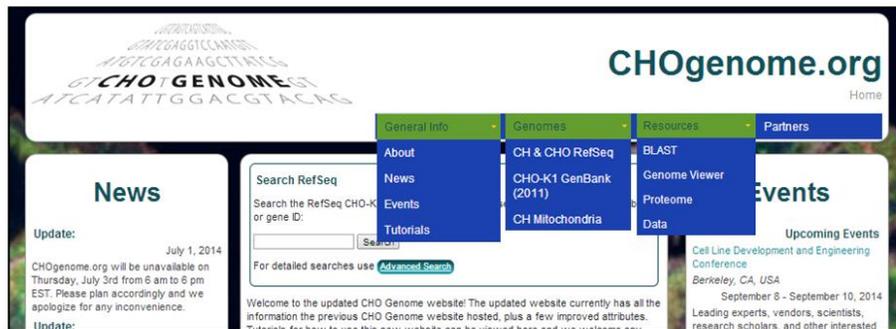
The CHO Proteome Database hosts proteomic data generated from CHO cell lines. This database hosts protein data from 2D polyacrylamide gel electrophoresis (2D PAGE) gels and shotgun proteomics analysis. To have your data included, please contact us.

2D PAGE Results

Shotgun Results

Navigation Menu

The blue navigational menu located at the top of each page on the CHO genome website enables quick access to the Chinese hamster-specific tools and resources. Several of the main sections (**General Info**, **Genomes**, and **Resources**) are drop down menus with multiple options that can be accessed by moving the cursor over the main section title.



General Information Tab

The **General Info** section contains direct links to the **About**, **News**, **Tutorials**, and **Events** pages.

- The **About** page contains a brief history of the CHO genome community, a brief introduction to the goals of this resource, the current and future plans, instructions for citing this resource, and the disclaimer.
- The **News** page contains news updates since 2012.
- The **Events** page contains the complete list of upcoming conferences of interest to the CHO community and highlights of past workshops and conferences.
- The **Tutorials** page contains six tutorials that can be downloaded and used to assist users with navigating and using CHOgenome.org. The six tutorials include: (1) an overview of the CHO genome resource, (2) an overview of CHOgenome.org, (3) assistance with searching the Chinese hamster genome databases, (4) BLAST searching the Chinese hamster genome, (5) viewing the Chinese hamster genomes, and (6) searching the Chinese hamster proteome databases.

Genomes Tab

Search pages for the RefSeq Chinese hamster (CH) and CHO cell line, GenBank CHO-K1, and CH mitochondrial genomes are each accessible from this tab.

For additional details, please see Tutorial #3, *Searching the Chinese Hamster Genome Database*.

- The comprehensive **CH & CHO RefSeq** search page enables users to select the RefSeq Chinese hamster genome(s) to be searched and the search terms to search against.

Search Page

Search Term <small>Select at least one</small>	Genome <small>Select at least one</small>	
<input checked="" type="checkbox"/> Symbol	<input checked="" type="checkbox"/> CHO-K1	(RefSeq Assembly GCF_000223151.1 2May2014 - Release 101)
<input checked="" type="checkbox"/> Gene Name	<input checked="" type="checkbox"/> Chinese Hamster	(RefSeq Assembly GCF_000419365.1 2May2014 - Release 101)
<input checked="" type="checkbox"/> Gene ID	<input checked="" type="checkbox"/> CHO-K1	(RefSeq Assembly GCF_000223151.1 15Mar2012 - Release 1)

- The **CHO-K1 GenBank (2011)** genome can be searched by accession number, gene name or symbol, or GO term.

Search CHO-K1 GenBank

Search CHO-K1 GenBank (2011) assembly by accession number, gene name, symbol, or GO Term:

This genome assembly corresponds to the CriGri_1.0 genome assembly released in August 2011 (GenBank Assembly ID GCA_000223135.1).

The CHO-K1 GenBank database can be searched by:

1. Accession number (i.e. EGV99227)
2. Gene name or symbol (i.e. Transcription factor E2F3 or E2F3)
3. GO term (i.e. GO:0003700 or Transcription factor activity)

There are currently 24,240 entries in the database. To display all database records, use % in the search field.

[BLAST the CHO-K1 genome here](#) and at NCBI.

- The **CH Mitochondria** can only be searched by gene symbol.

Search Chinese Hamster Mitochondria

Search Chinese hamster mitochondria assembly by gene symbol:

Gene Symbol:

This genome assembly corresponds to the complete nucleotide sequence of the Chinese hamster mitochondrial DNA published in 2007 (RefSeq Assembly ID GCF_000055695.1).

The Chinese hamster mitochondria database can be searched by:

1. Gene symbol (i.e. Cox3)

There are currently 13 entries in the database. To display all database records, use % in the search field.

Resources Tab

The tools and resources relevant to the CHO Community are accessible from this menu.

- The **BLAST** section contains 13 nucleotide and amino acid CH databases for BLAST searches.
 - o For additional details, please see Tutorial #4, *BLAST Searching the CHO Genome*.

CHOblast Search

BLAST Search - Required parameters [help](#)

Enter query sequences here in Fasta format

Or upload fasta file: No file chosen

Algorithm: blastn - Nucleotide Database

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_Chrom_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 [details](#)

Nucleotide Databases:

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_Chrom_2013

Transcripts (RNA)

6) CHO-K1[ATCC]_RefSeq_2014

7) CHO-K1[ATCC]_RefSeq_2012

8) CH_RefSeq_2014

Amino Acid Databases:

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_Chrom_2013

Assembly Color Key:

RefSeq Assembly

GenBank Assembly

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)-xxxx Strain definition

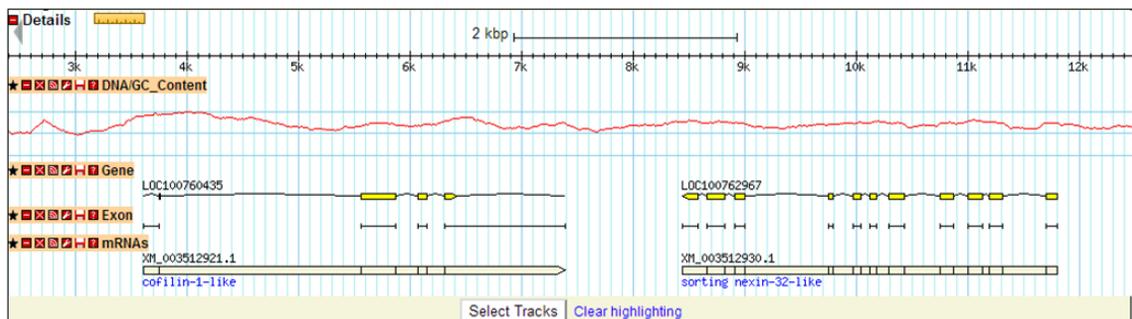
[xxxx] Source of cells

genbank GenBank assembly

refseq RefSeq assembly

chr Chromosomal identification

- The **GBrowse** section contains a genome viewer tool for the CHO-K1 RefSeq (2012) assembly, which enables users to visualize searched genes and scaffolds.
 - o For additional details, please see Tutorial #5, *Viewing the CHO & CH Genomes*.

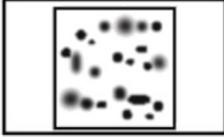


- The **Proteome Browser** section contains 2D PAGE and Shotgun databases of proteins previously identified by the Chinese hamster community.
 - o For additional details, please see Tutorial #6, *Searching the Proteome Database*.

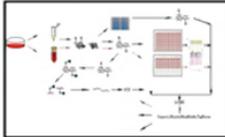
Proteome

The CHO Proteome Database hosts proteomic data generated from CHO cell lines. This database hosts protein data from 2D polyacrylamide gel electrophoresis (2D PAGE) gels and shotgun proteomics analysis. To have your data included, please [contact us](#).

2D PAGE Results



Shotgun Results



The **Data** section provides links to relevant publications and public data sets.

Genomic Data Resources

This is a temporary page containing links to relevant publications and public data sets. Links to next generation sequencing raw data in the Sequence Read Archive (SRA) are provided when available. Journal access may be limited to subscribed users.

Genome

Description	Title	Reference
A chromosome sorting approach was used to facilitate genome assembly from short-read sequences of the Chinese hamster genome	Chinese hamster genome sequenced from sorted chromosomes	Brinkrolf K et al. <i>Nature Biotechnol</i> (2013) 31, 694-695
Draft genomic sequence of the Chinese hamster and resequencing of CHO-K1, DG44, and CHO-K cell lines	Genomic landscapes of Chinese hamster ovary cell lines as revealed by the <i>Cricetulus griseus</i> draft genome	Lewis N et al. <i>Nature Biotechnol</i> (2013) 31, 759-765

Partners Tab

The **Partners** section provides a visual list of the agencies, corporations, and universities actively supporting this CHO genome initiative. Each logo serves as a direct link to the website associated with the respective partner.

- To become a partner, select the Contact Us button and a CHO genome representative will respond.

