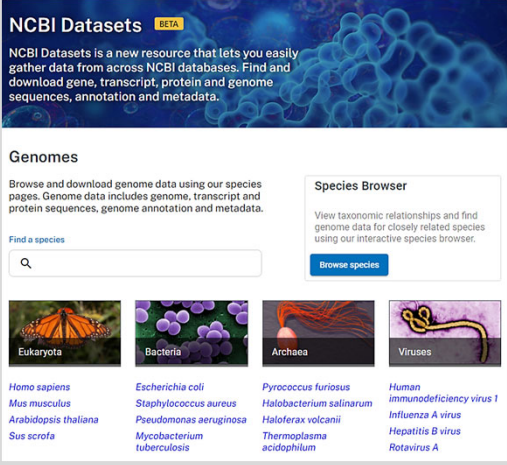
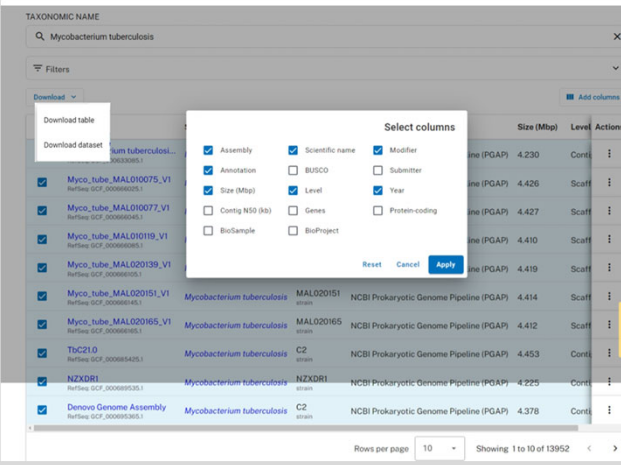


https://www.nlm.nih.gov/ncbi/workshops/2022-08_intro-to-pathogen-data/


New Datasets resource! Batch Genome & Gene downloads




The screenshot shows the NCBI Datasets homepage with a search bar, a 'Species Browser' section, and a 'Genomes' section. The 'Species Browser' lists various organisms like Homo sapiens, Mus musculus, and Escherichia coli. The 'Genomes' section provides a brief overview of the resource.



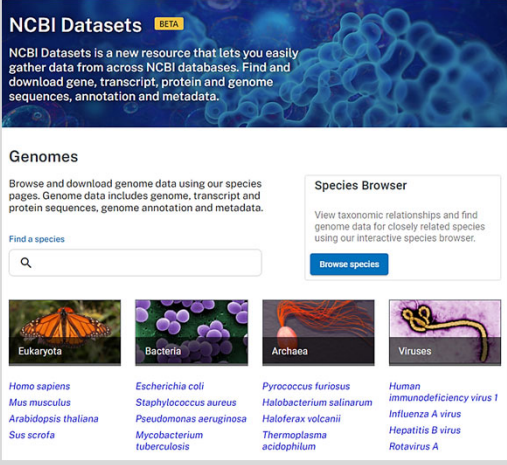
The screenshot shows search results for 'Mycobacterium tuberculosis'. A 'Download' dropdown menu is open, showing options for 'Download dataset' and 'Download table'. A 'Select columns' dialog box is also visible, with checkboxes for 'Assembly', 'Annotation', 'Size (Mbp)', 'Contig N50 (kb)', 'BioSample', 'Scientific name', 'BUSCO', 'Level', 'Genes', 'Protein-coding', 'Modifer', 'Year', and 'BioProject'.



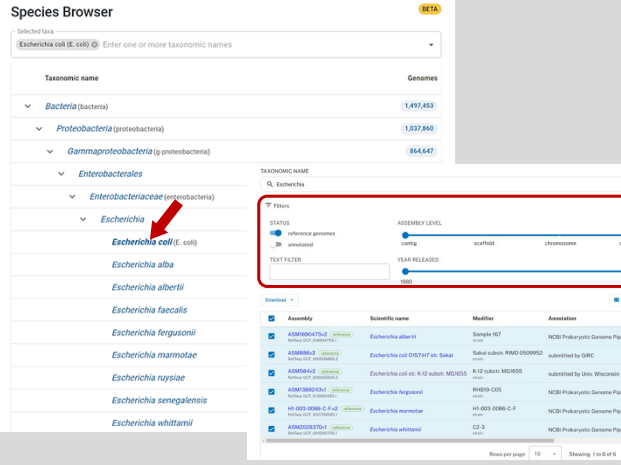
National Library of Medicine
National Center for Biotechnology Information

NCBI Workshop 


New Datasets resource! Batch Genome & Gene downloads Taxonomy & Genome pages




This is a duplicate of the homepage screenshot from the first slide, showing the NCBI Datasets interface with search and species browser options.



The screenshot shows the 'Species Browser' interface. The search bar contains 'Escherichia coli (E. coli)'. The taxonomic tree is expanded to show 'Escherichia coli (E. coli)'. A red arrow points to this entry. A 'Filters' dialog box is open, showing 'ASSEMBLY LEVEL' with options for 'contig', 'scaffold', and 'chromosome'. The 'Download' dropdown menu is also visible.



National Library of Medicine
National Center for Biotechnology Information

NCBI Workshop 

New Datasets resource!

Batch Genome & Gene downloads

Taxonomy & Genome pages

NCBI Datasets BETA

NCBI Datasets is a new resource that lets you easily gather data from across NCBI databases. Find and download gene, transcript, protein and genome sequences, annotation and metadata.

Genomes

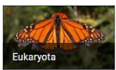
Browse and download genome data using our species pages. Genome data includes genome, transcript and protein sequences, genome annotation and metadata.

Species Browser

View taxonomic relationships and find genome data for closely related species using our interactive species browser.


Find a species

Browse species



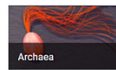
Eukaryota

[Homo sapiens](#)
[Mus musculus](#)
[Arabidopsis thaliana](#)
[Sus scrofa](#)




Bacteria

[Escherichia coli](#)
[Staphylococcus aureus](#)
[Pseudomonas aeruginosa](#)
[Mycobacterium tuberculosis](#)



Archaea

[Pyrococcus furiosus](#)
[Halobacterium salinarum](#)
[Haloflex volcanii](#)
[Thermoplasma acidophilum](#)



Viruses

[Human immunodeficiency virus 1](#)
[Influenza A virus](#)
[Hepatitis B virus](#)
[Rotavirus A](#)

Bacteria / Proteobacteria / Gammaproteobacteria / Enterobacterales / Enterobacteriaceae / Escherichia BETA

Escherichia coli ☆

E. coli (*Escherichia coli*) is a species of enterobacteria in the family *Enterobacteriaceae* (enterobacteria).

Browse taxonomy

Current scientific name
Escherichia coli

Common name
E. coli

Taxonomic rank
species

NCBI Taxonomy ID
562

For more details

Genome

Browse all 188,644 genomes

Reference genome ASM886v2	Reference genome ASM584v2
GIRC (2018). Strain: Sakai substr. RIMD 0509952. RefSeq OCF_000008865.2	Univ. Wisconsin (2013). Strain: K-12 substr. MG1655. RefSeq OCF_000005845.2
Download	Download
Genome size 5.6 Mb	Genome size 4.6 Mb
Contig N50 5.5 Mb	Contig N50 4.6 Mb
Genes 5,417	Genes 4,639
Annotation submitted by GIRC	Annotation submitted by Univ. Wisconsin