Dec 03, 2025

Kraken2 - Taxonomic Classification Workflow on GalaxyTrakr

DOI
https://dx.doi.org/10.17504/protocols.io.yxmvmbbw6g3p/v1

Christopher Duda[1], Anna Brover[1], Padmini Ramachandran[1]

[1]FDA

GenomeTrakr
Tech. support email: genomeTrakr@fda.hhs.gov

Christopher Duda
FDA

DOI: https://dx.doi.org/10.17504/protocols.io.yxmvmbbw6g3p/v1

Protocol Citation: Christopher Duda, Anna Brover, Padmini Ramachandran 2025. Kraken2 - Taxonomic Classification Workflow on GalaxyTrakr. **protocols.io** https://dx.doi.org/10.17504/protocols.io.yxmvmbbw6g3p/v1

Protocol status: In development
We are still developing and optimizing this protocol

Created: October 24, 2025

Last Modified: December 03, 2025

Protocol Integer ID: 230684

Keywords: taxonomic classification workflow on galaxytrakr kraken2, taxonomic sequence classifier, taxonomic classification workflow, taxonomic label, taxonomic tree, single taxonomic label for the read, taxonomic tree of all genome, nodes in the taxonomic tree, single taxonomic label, galaxytrakr kraken2, kraken2, kraken, genomic library, genomic library to the lowest common ancestor, mer in kraken, genome, microbe, lowest common ancestor, set of lca taxa, lca taxa

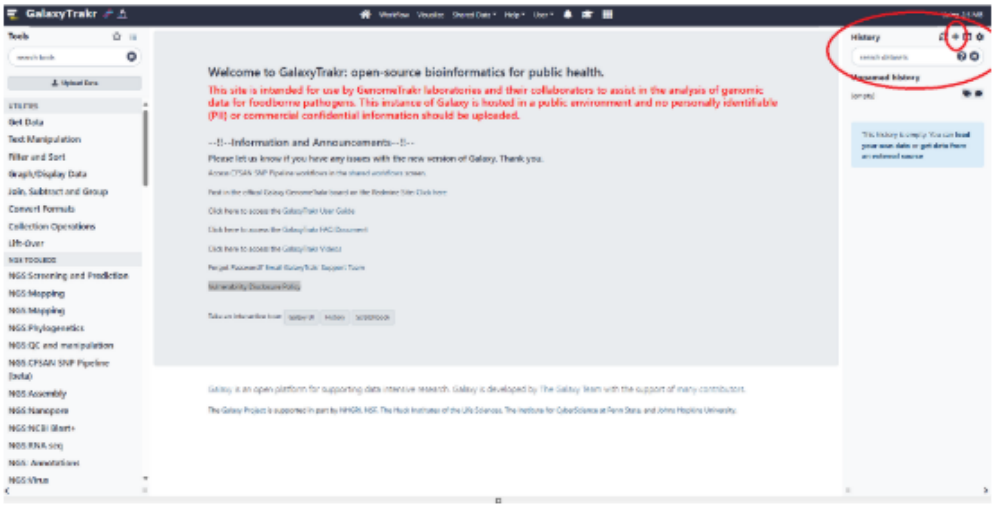
Abstract

Kraken2 is a taxonomic sequence classifier that assigns taxonomic labels to sequence data. Essentially profiling the microbes that is present in a sample. It does this by examining the k-mers within a read and querying a database with those k-mers. This database contains a mapping of every k-mer in Kraken's genomic library to the lowest common ancestor (LCA) in a taxonomic tree of all genomes that contain that k-mer. The set of LCA taxa that correspond to the k-mers in a read are then analyzed to create a single taxonomic label for the read; this label can be any of the nodes in the taxonomic tree. Kraken is designed to be rapid, sensitive, and highly precise.

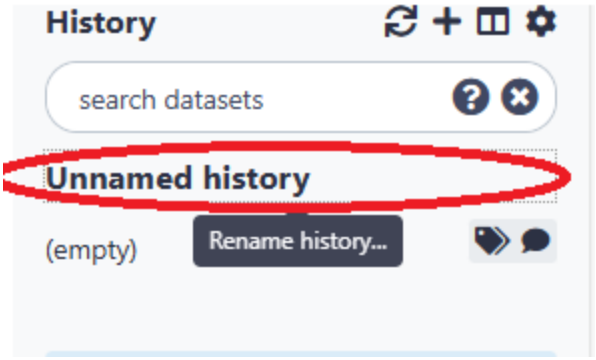
Pairing Fastq Datasets

1 Log into your GalaxyTrakr account at **Galaxy**

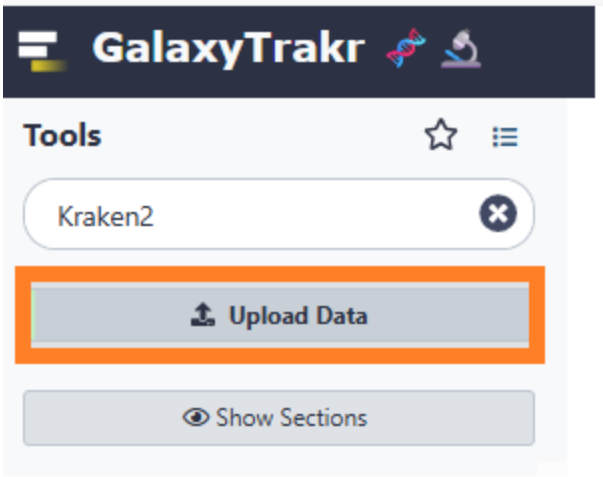
2 On the right-hand side of the screen, click the plus next to the History tab to create a new history.



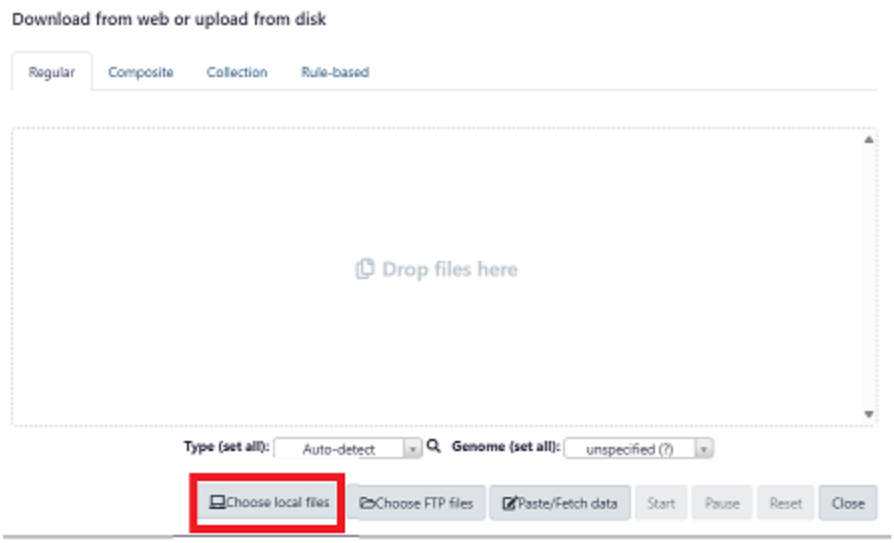
Click on "Unnamed History" and rename it according to your system.



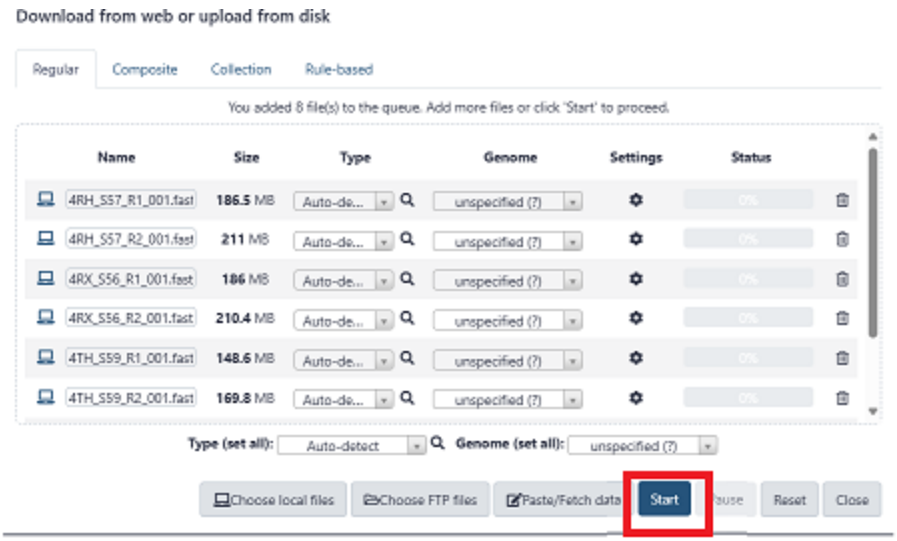
3 On the left side of the screen, click the "Upload data" button.



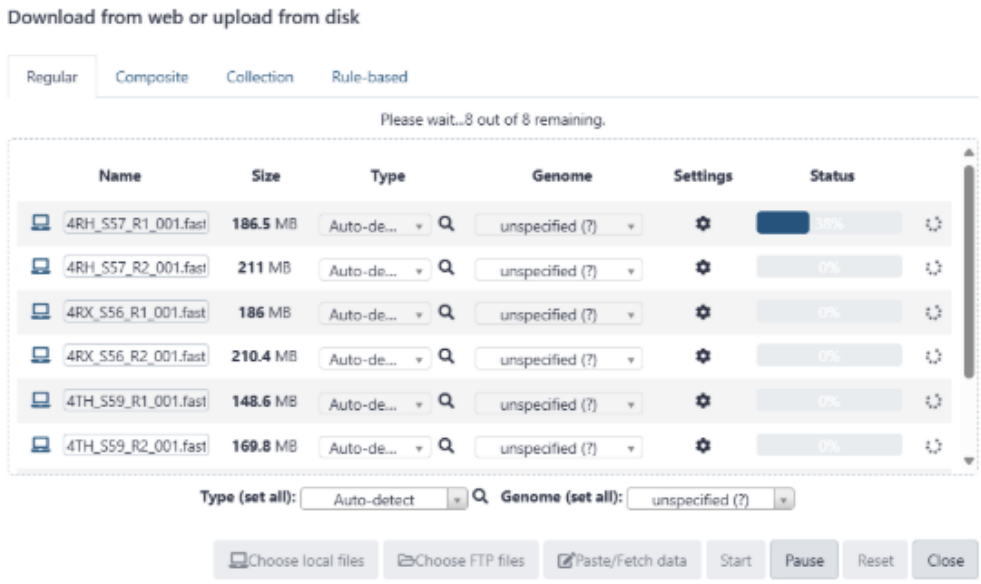
An interface where we can find and select files to upload will appear, for most users, the files will be on local files so choose that option unless otherwise noted.

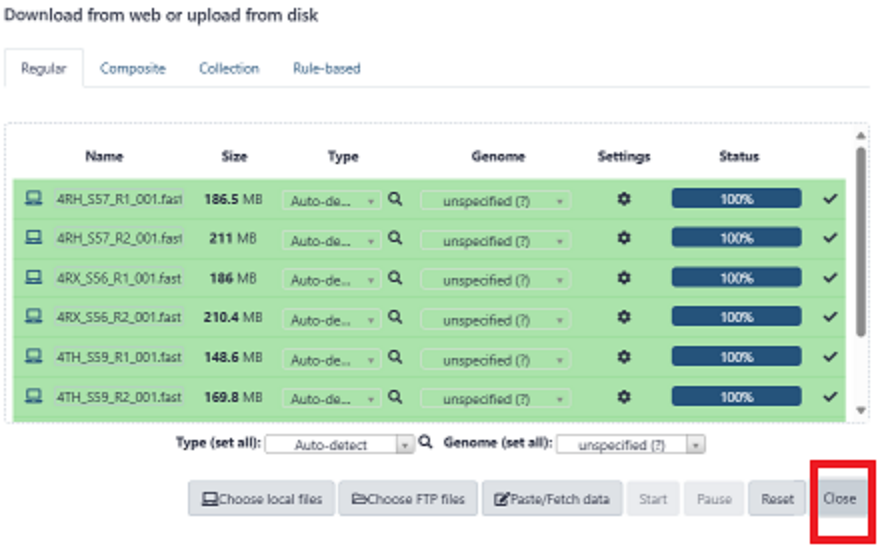


Select all files to be processed, It should resemble this:
Click "Start" to begin uploading.

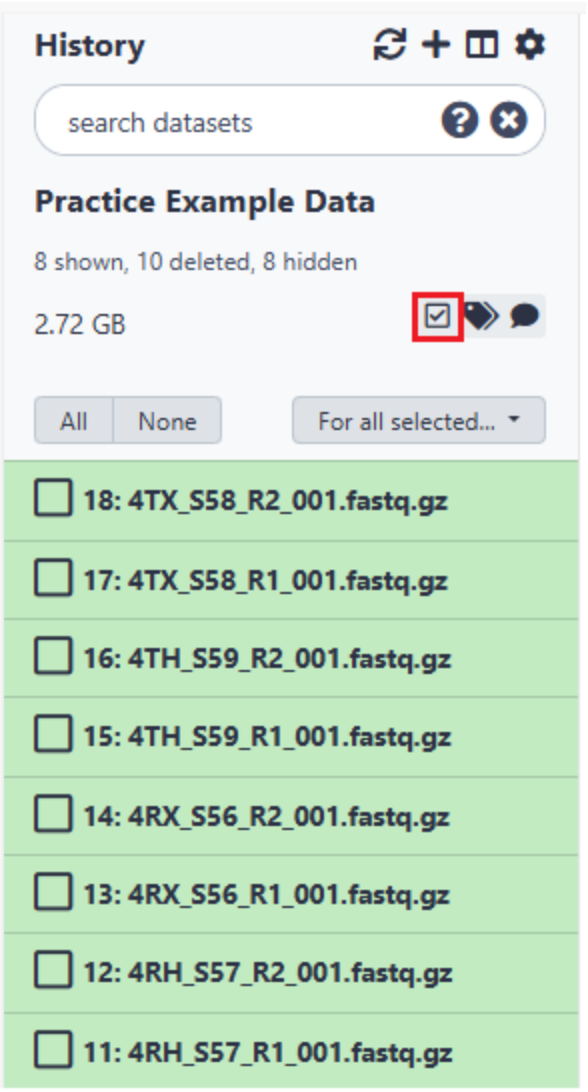


Note* Depending on the number of files, it can take a while to upload. This is what it looks like when uploading has started.

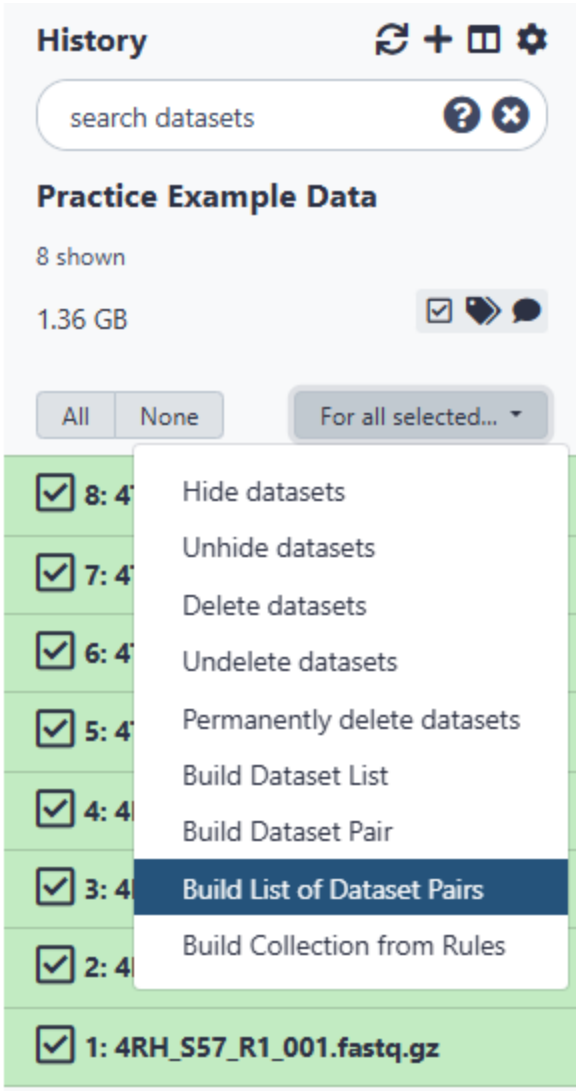




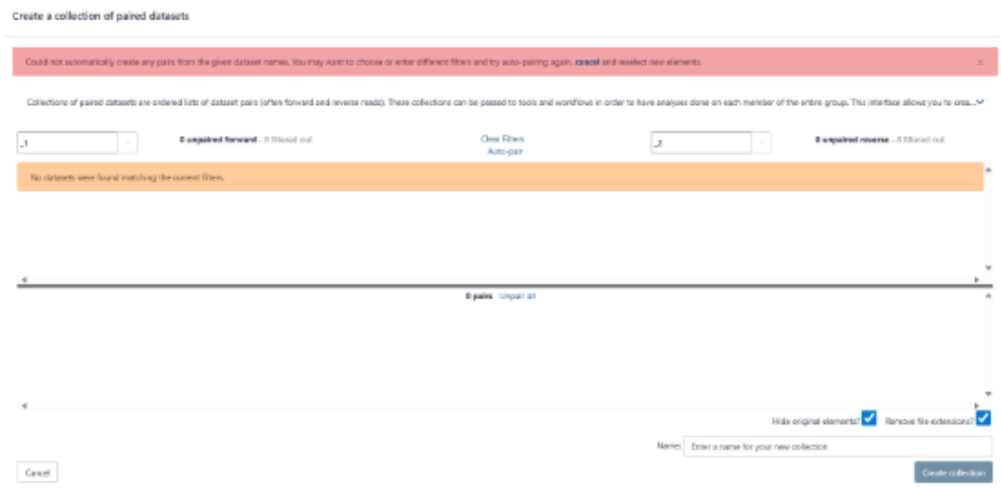
4 Before running "Kraken2", the dataset needs to be paired.
Click on the square icon, or the "Operations on multiple datasets" on the right side of the tab. select boxes will appear on the green file list. This will allow you to select all files in your dataset.



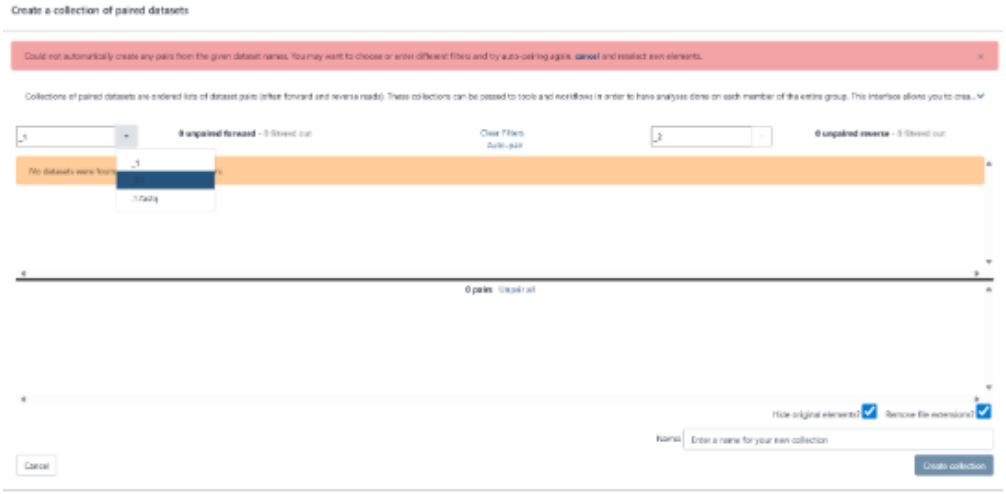
Select all files in the workspace, then click the dropdown that says for all selected. A menu will appear, select "Build List of Dataset Pairs".



5 A new interface will appear.



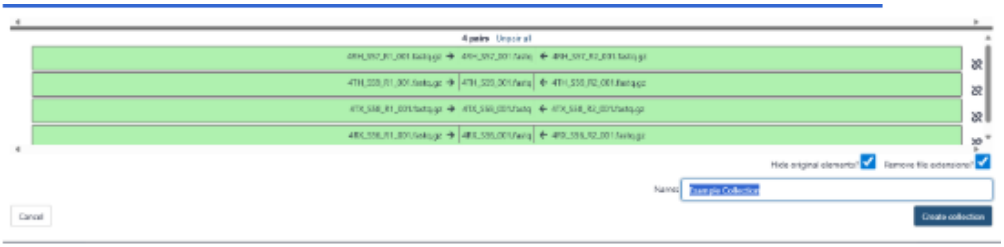
In the dropdown on the left, select _R1. This will auto select _R2 for the other dropdown.
`



You will see both columns populate with filenames. Double check that the files have the same name other than "R1" or "R2". Once you know the program has properly paired the

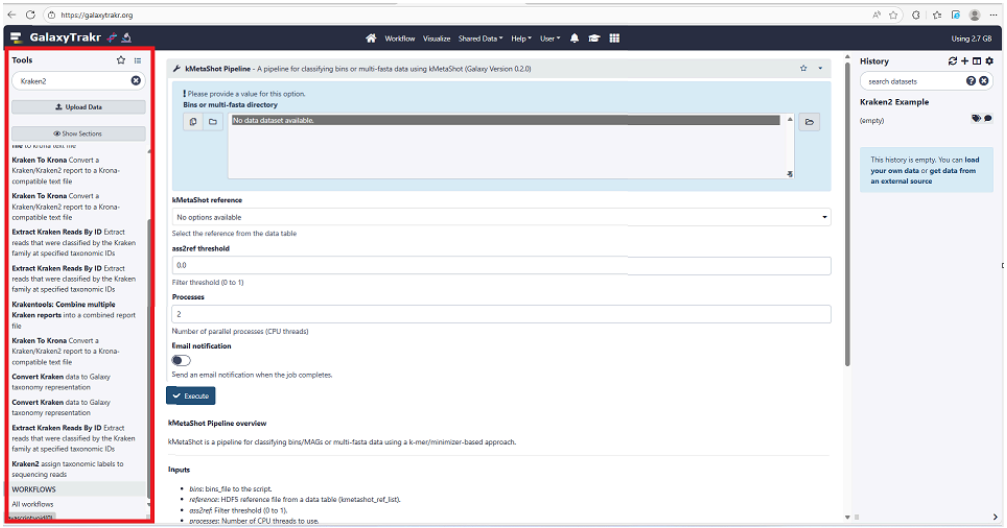
files, Click the "Pair these datasets" for all entries. There two columns will be empty if you do it correctly and they appear in green below the black bar.

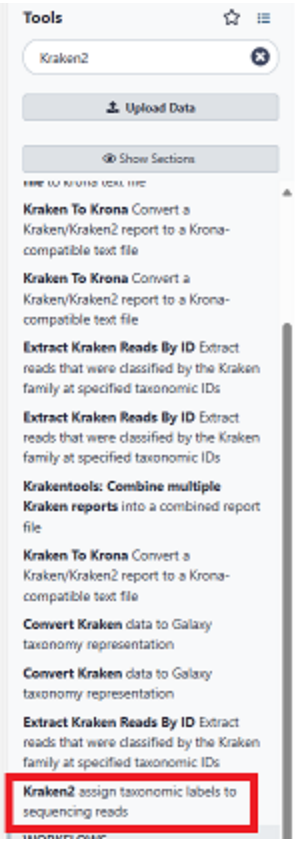
6 Create a name for your collection and click the"Create collection" button.



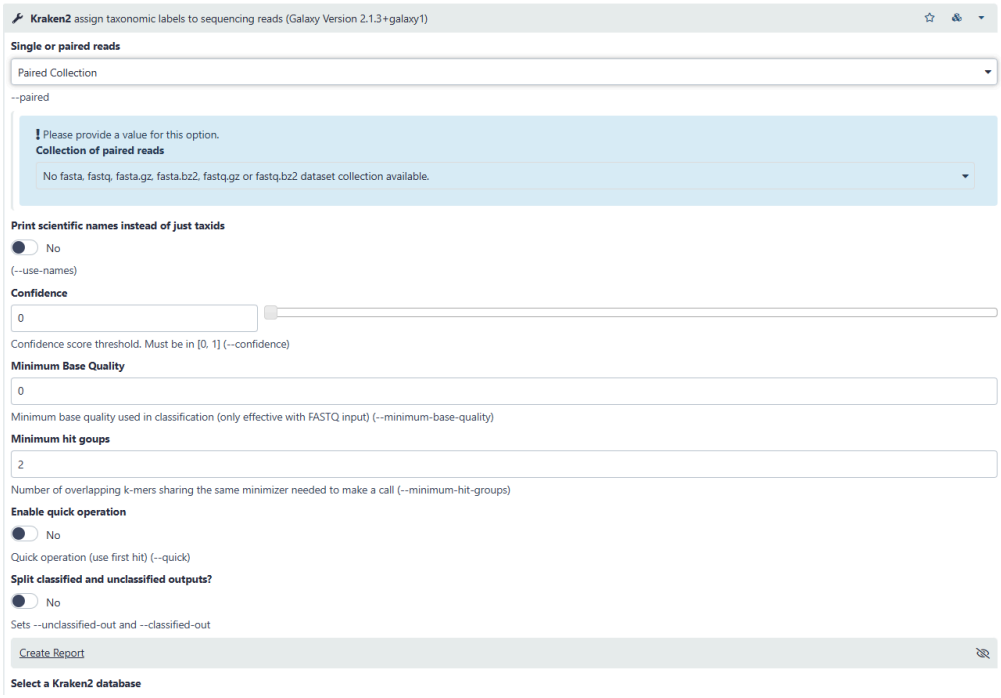
Creating Graphical Reports

6.1 On the left-hand side of the screen, click the search bar in the "tools" section and search for
"Kraken2".

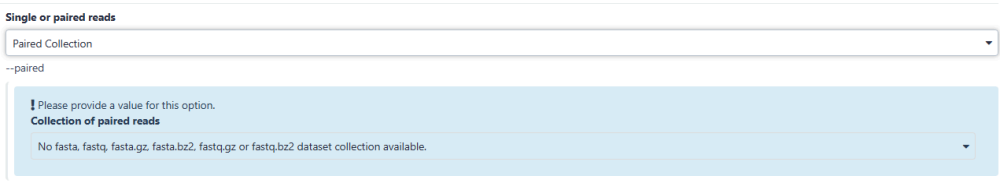




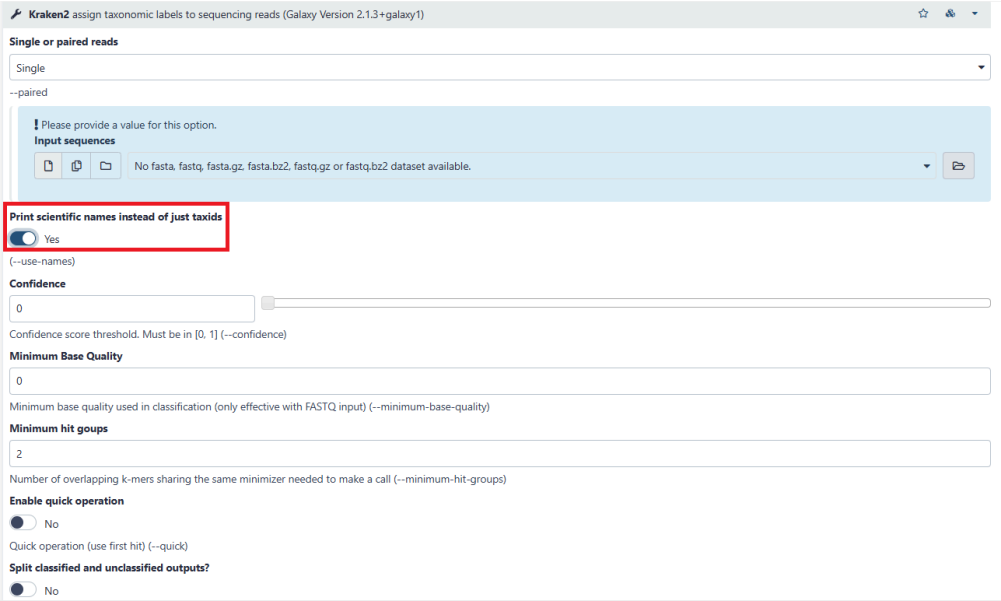
6.2 An interface will appear looking like this.



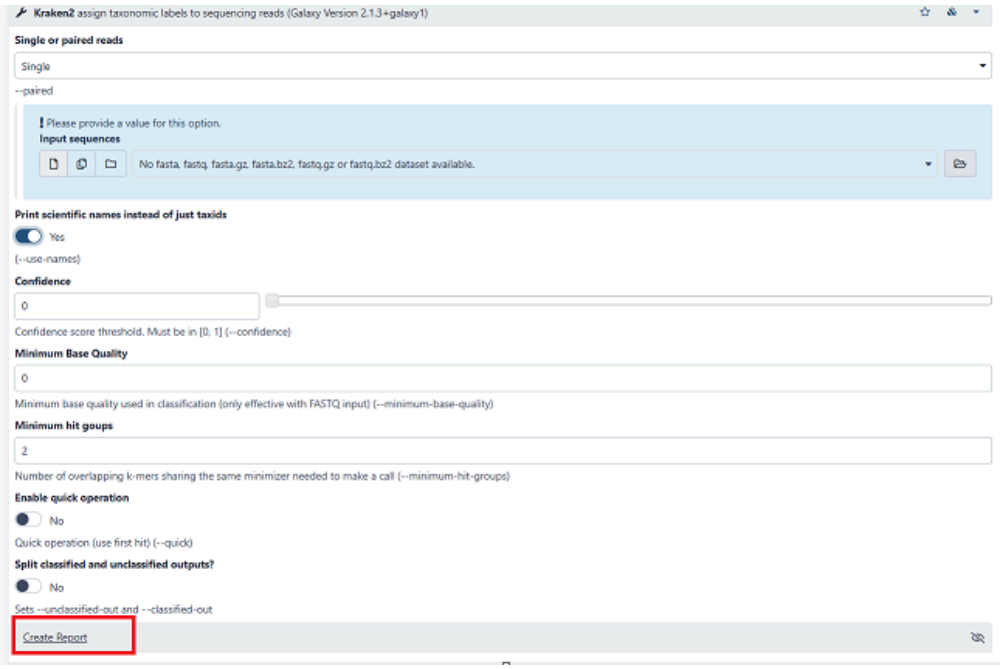
Change the "Select to read collection type" to "paired collection".



6.3 Click on the slider under "Print scientific name instead of just taxids" so it says yes and is a solid dark blue color.

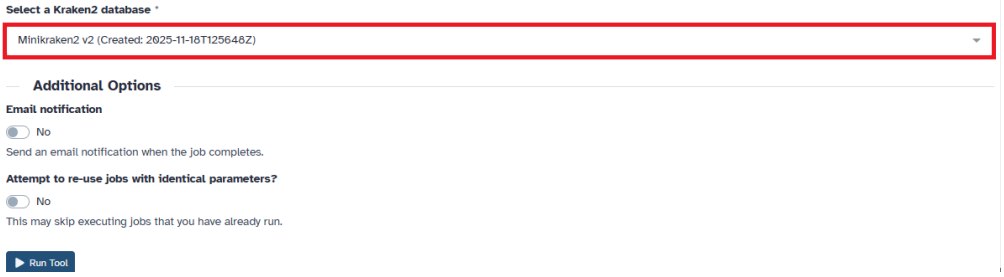


6.4 Click on "Create Report"

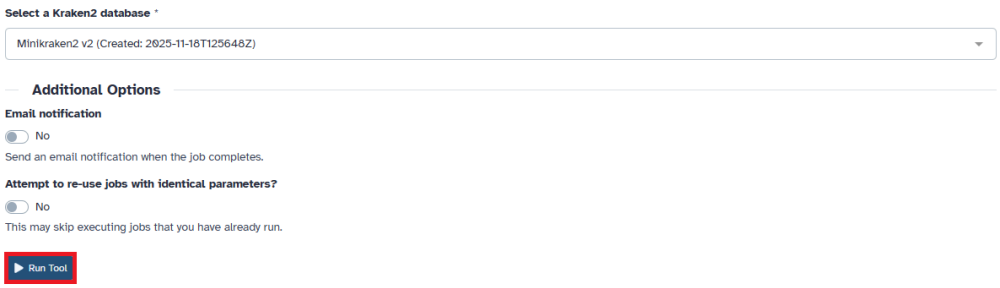


Click on the slider under "Print a report with aggregate counts/clades to a file" so it says yes and is a solid dark blue color.

6.5 Next you need to pick a kraken database. click on the dropdown under **Select a Kraken2 database** and select "Minikraken2 v2".



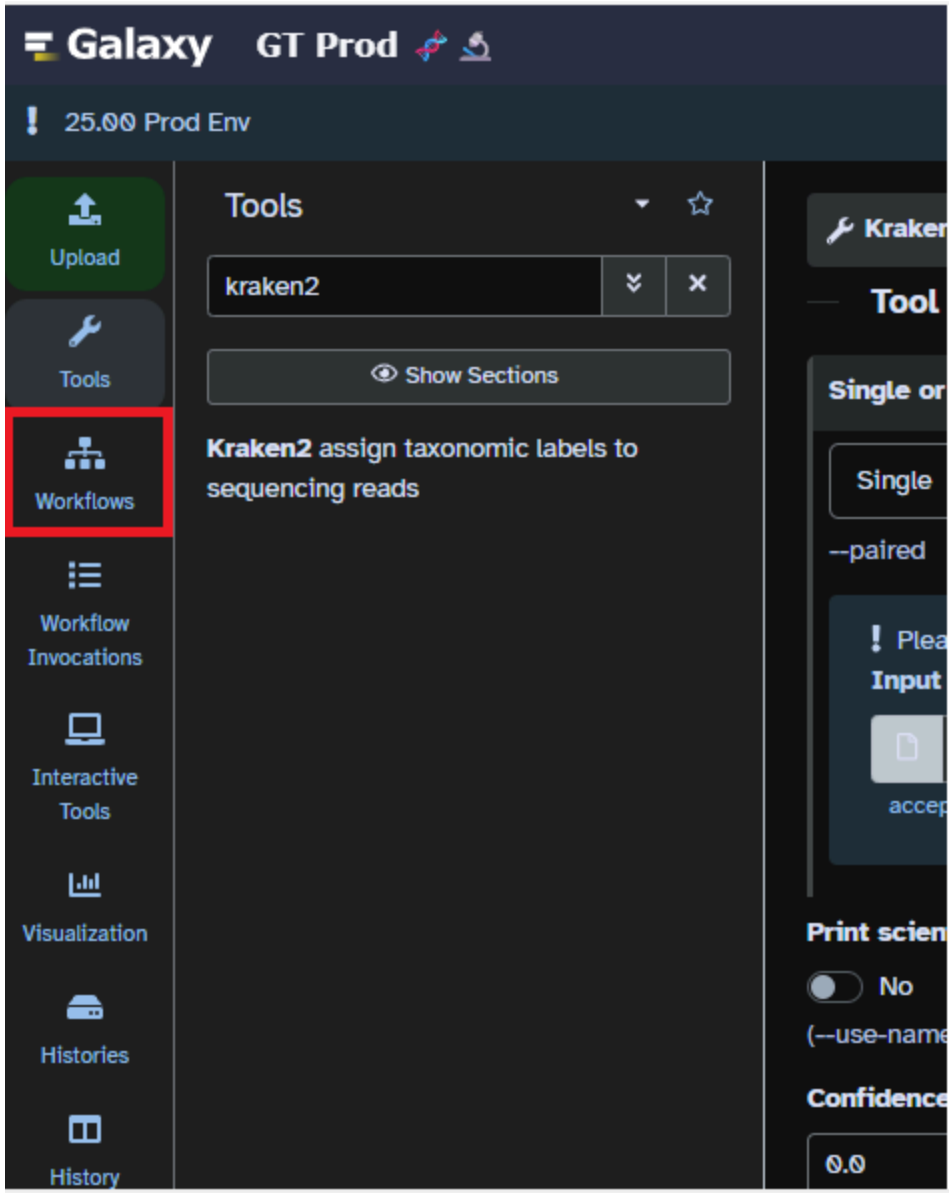
6.6 Click on the "Run Tool" button.



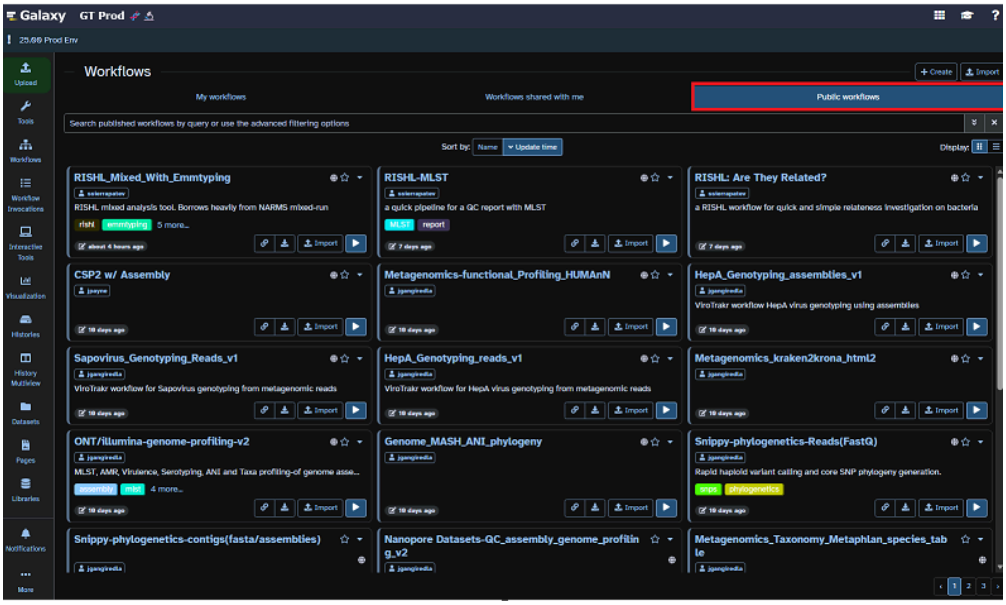
The report is generating once a green box appears. It should look similar to this.

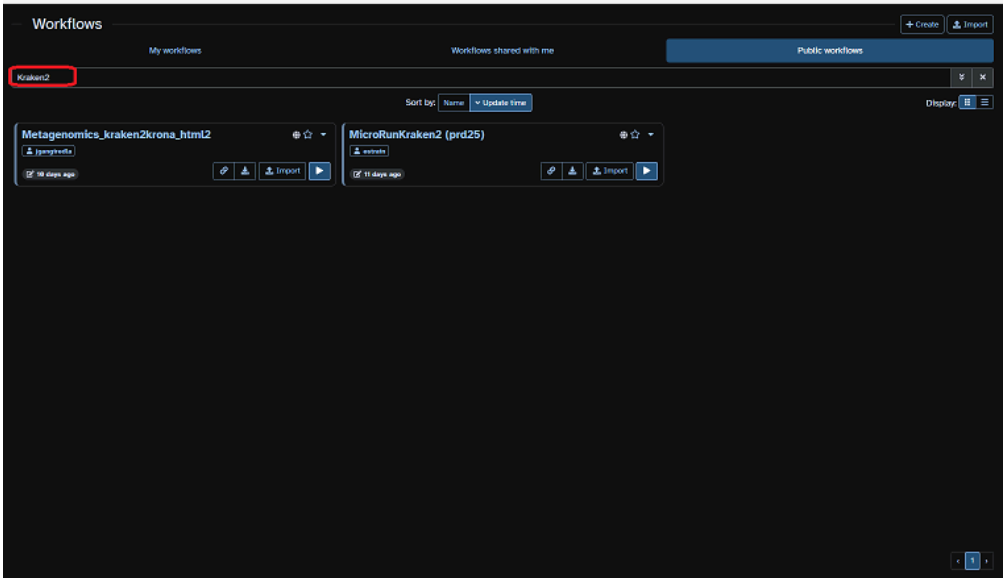
This process usually takes over an hour to generate the report.

6.7 Generating graphical reports.
On the left-hand side of the screen, click the tab titled "workflows".

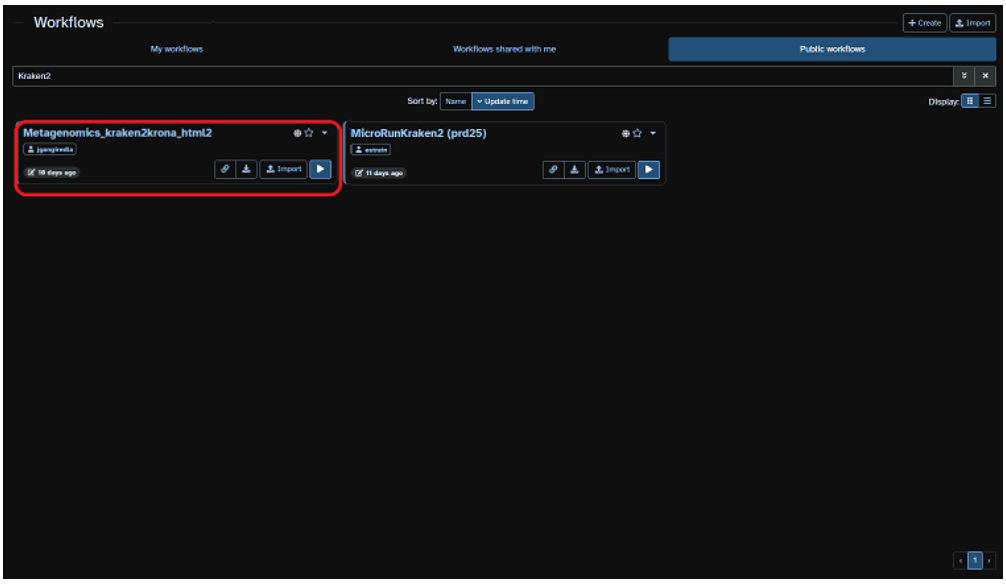


6.8 Click on the "Public workflows" tab on the top right and input kraken2 in the search bar.

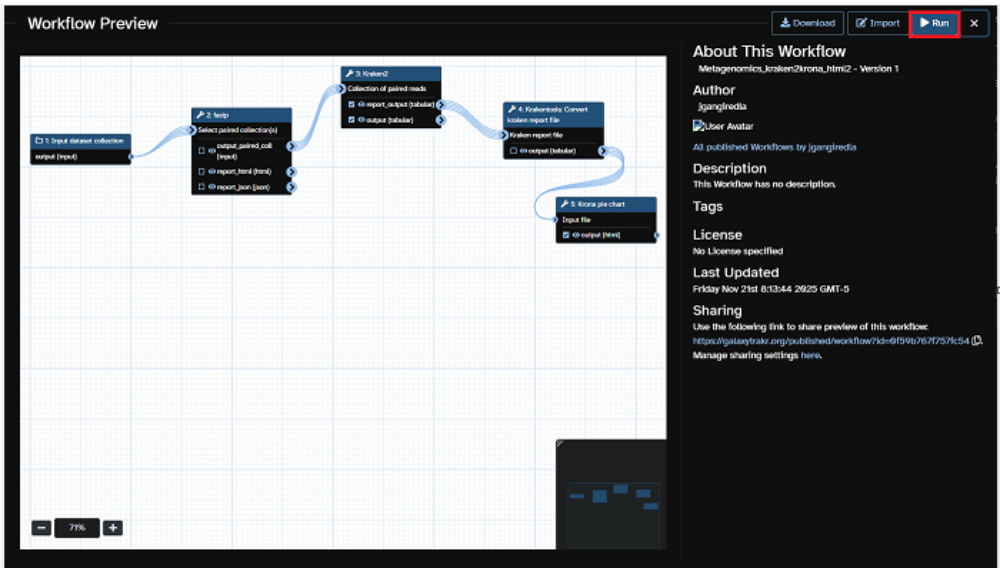




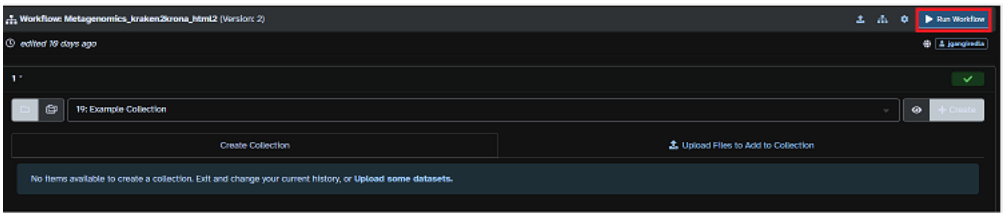
6.9 Select the workflow titled "**Metagenomics_kraken2krona_html2".**



6.10 Click the "Run" button on the top right of the window that appears.

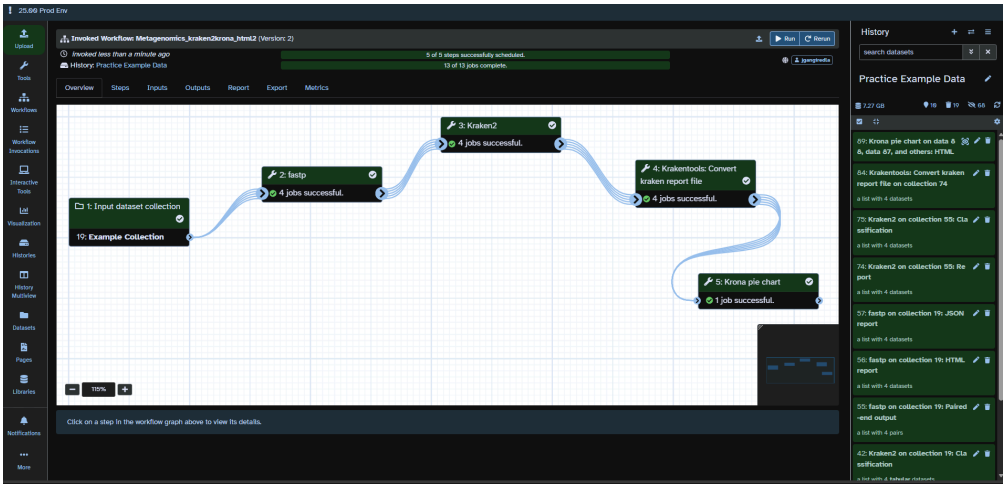


6.11 Make sure your paired fastq datasets are selected and click the "Run Workflow" button.

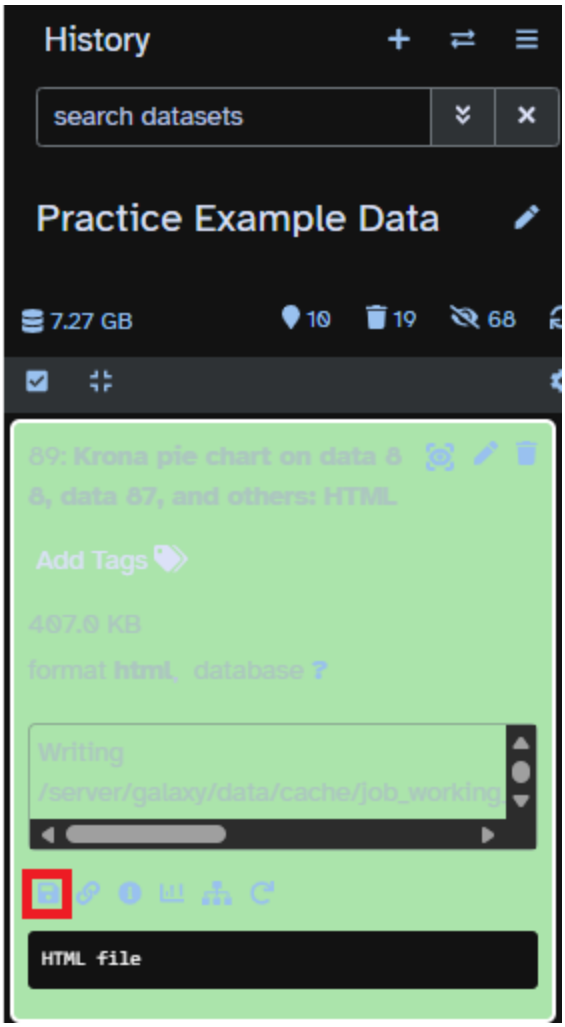


This will start the process for generating graphical reports. This can take over an hour to generate depending on size and traffic.

It will look like this when it is finished:



7 To access the figures, click on the desired file under the history and click download.



This will download a .zip file containing the figure for your data.

Protocol references

@article{Wood_2014,
journal = {Genome Biology},
number = {3},
publisher = {Springer Science and Business Media LLC},
title = {Kraken: ultrafast metagenomic sequence classification using exact alignments},
url = {http://dx.doi.org/10.1186/gb-2014-15-3-r46},
volume = {15},
author = {Wood, Derrick E and Salzberg, Steven L},
date = {2014-03},
year = {2014},
month = {3},
}