

BV-BRC Test Report

A19. Service – Taxonomic Classification

Item to test	Taxonomic Classification Service using metagenomic read sets
URL	https://www.bv-brc.org/app/TaxonomicClassification
Prerequisites	Metagenomic reads as SRA accession or fastq files
References	https://www.bv-brc.org/docs/quick_references/services/taxonomic_classification_service.html https://www.bv-brc.org/docs/tutorial/taxonomic_classification/taxonomic_classification.html
Tester(s)	Rebecca Wattam, Maulik Shukla
Test date	08-May-2022 (follow-up from original test)
Test result	Passed

Overview

- Test the Taxonomic Classification Service using exemplar metagenomic reads sets.
- Test input options, i.e., single-end / paired-end read files and SRA accessions.
- For each job submitted, verify successful completion of the job, presence of output files and their format.
- Review and verify the taxonomic classification results using tabular report.
- Review and verify the results using interactive taxonomic classification viewer, Krona.

Test Data

Dataset	Rational	Input Format	Input
Louse Fly Metagenome	Workshop example	Fastq file, SRA accession	ERR969522 ERR969522_1.fastq.gz ERR969522_2.fastq.gz
ICU metagenome sample	Workshop example	Fastq file	UC.MICU.02.30.fastq

- All test datasets and corresponding job results are available in the following public workspace:
<https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Taxonomic%20Classification>

Test Results

- All jobs completed successfully, without any errors.
- All jobs resulted in expected output files in corresponding job output directory, including taxonomic classification report in text and HTML formats and classified reads as a fastq file.
- The HTML report was informative and provided a concise summary of the count and percentage of reads mapped to every matching taxonomy level.

- The interactive pie chart viewer, Krona, provided nice visual summary of the taxa in the metagenomic sample. Progressive zoom in function allowed further exploration of the bacterial, viral and eukaryotic taxa in the mixed samples.
- All test datasets and corresponding job results are available in the following public workspace: <https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Taxonomic%20Classification>
- Below are a series of screenshots showing successful completion of the jobs, availability of the result files in the workspace, excerpts of the HTML report and viewer.

Start With: READ FILE ASSEMBLED CONTIGS

Input File

PAIRED READ LIBRARY

ERR969522_1.fastq.gz

ERR969522_2.fastq.gz

SINGLE READ LIBRARY

READ FILE

SRA RUN ACCESSION

SRR

Selected libraries

Place read files here using the arrow buttons.

P(ERR96...fq.gz, ERR96...fq.gz) i x

Parameters

ALGORITHM

Kraken2

DATABASE

All genomes

SAVE CLASSIFIED SEQUENCES

NO YES

SAVE UNCLASSIFIED SEQUENCES

NO YES

OUTPUT FOLDER

Taxonomic Classification

OUTPUT NAME

Output Name

Your job has been submitted successfully. Please visit your [Jobs List](#) to check the status of your job and access the results.

[Reset](#) [Submit](#)

Status	ID	Service	Output Name	Submit	Start	Completed
completed	7747932	TaxonomicClassification	Louse fly metagenome - ERR969522 - fastq	5/8/22, 8:12 PM	5/8/22, 8:12 PM	5/8/22, 8:59 PM
completed	7747933	TaxonomicClassification	Louse fly metagenome - ERR969522 - SRA	5/8/22, 8:13 PM	5/8/22, 8:13 PM	5/8/22, 9:28 PM
completed	7747934	TaxonomicClassification	ICU metagenome	5/8/22, 8:14 PM	5/8/22, 8:14 PM	5/8/22, 9:40 PM

BVBRC / BVBRC Tests / Taxonomic Classification (6 items)

UPLoad ADD FOLDER SHOW HIDDEN

Name	Size	Owner	Members	Created
Parent folder				
UC.MICU.02.30.fastq	8.3 GB	me	Public	5/8/22, 8:07 PM
ERR969522_1.fastq.gz	3.9 GB	me	Public	5/8/22, 8:08 PM
ERR969522_2.fastq.gz	3.6 GB	me	Public	5/8/22, 8:08 PM
Louse fly metagenome - ERR969522 - fastq	7.9 kB	me	Public	5/8/22, 8:59 PM
Louse fly metagenome - ERR969522 - SRA	7.7 kB	me	Public	5/8/22, 9:28 PM
ICU metagenome	7.7 kB	me	Public	5/8/22, 9:40 PM

Results (Click here to view Interactive Chart)

The top hits seen in the input data are as follows. This chart contains any taxonomy entry which has one percent or more of the total hits. The complete data may be found in the file [report.txt](#). The output file [full_report.txt](#) contains the same data but also includes taxonomy entries for which there were zero hits.

In addition, the output file [output.txt.gz](#) contains information about each input sequence. Documentation on this format is available [here](#).

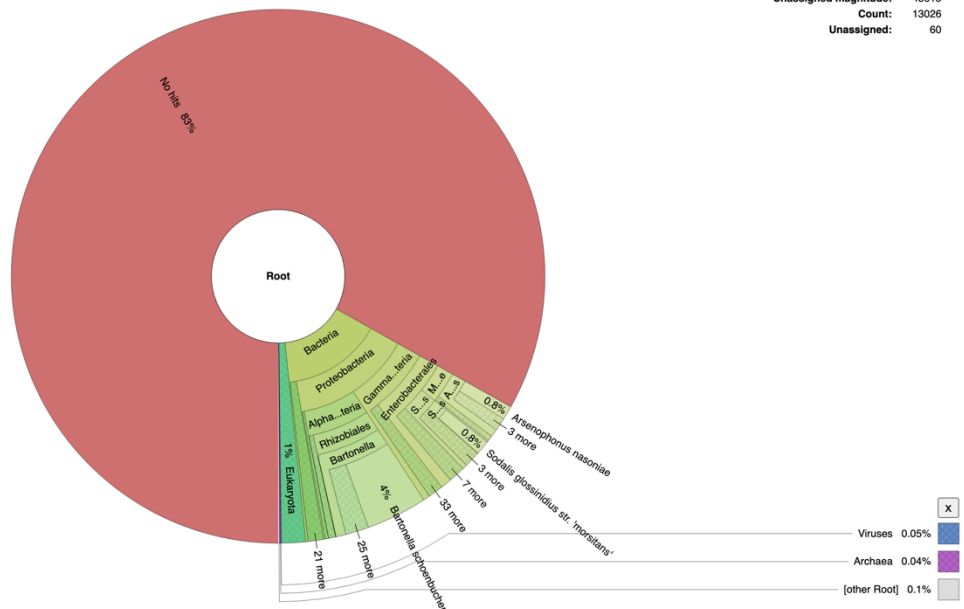
Pct Coverage	Fragments in Clade	Fragments in Taxon	Rank	NCBI Taxon ID	Scientific Name
83.13	36100208	36100208	U	0	
16.87	7323908	15241	R	1	
16.78	7286289	30939	R1	131567	cellular organisms
15.22	6607053	61965	D	2	Bacteria
14.14	6140233	40752	P	1224	Proteobacteria
7.67	3332194	175786	C	1236	Gammaproteobacteria
6.35	2758430	294000	O	91347	Enterobacterales
2.54	1102648	46250	F	1903414	Morganeliaceae
2.08	905276	206576	G	637	Arsenophonus
2.15	934409	899	F	1903410	Pectobacteriaceae
2.13	924785	122658	G	84565	Sodalis
6.17	2677102	2140	C	28211	Alphaproteobacteria
5.73	2487317	1162	O	356	Rhizobiales
5.70	2475898	0	F	772	Bartonellaceae
5.70	2475898	250526	G	773	Bartonella
3.70	1605124	0	S	165694	Bartonella schoenbuchensis
3.70	1605124	1605124	S1	687861	Bartonella schoenbuchensis R1
1.45	631579	9109	D	2759	Eukaryota

html file: [chart.html](#)

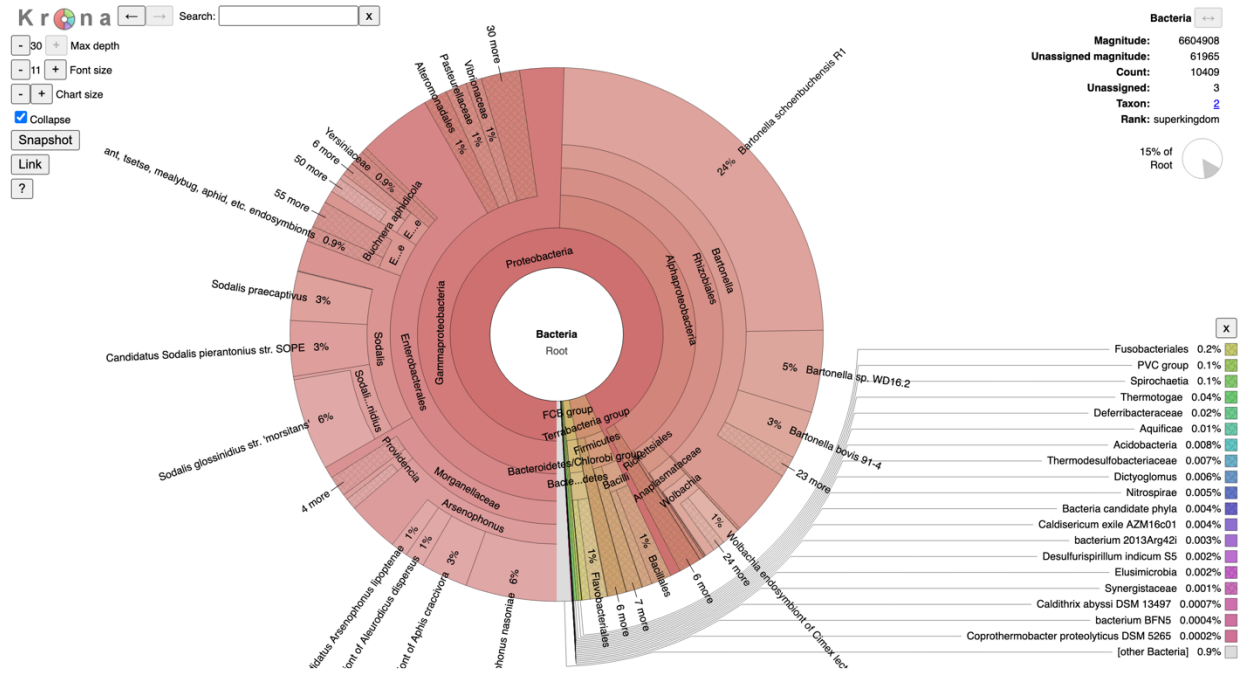
Krona

30 Max depth
 11 Font size
 Chart size
 Collapse

Root
 Magnitude: 43424116
 Unassigned magnitude: 48615
 Count: 13026
 Unassigned: 60



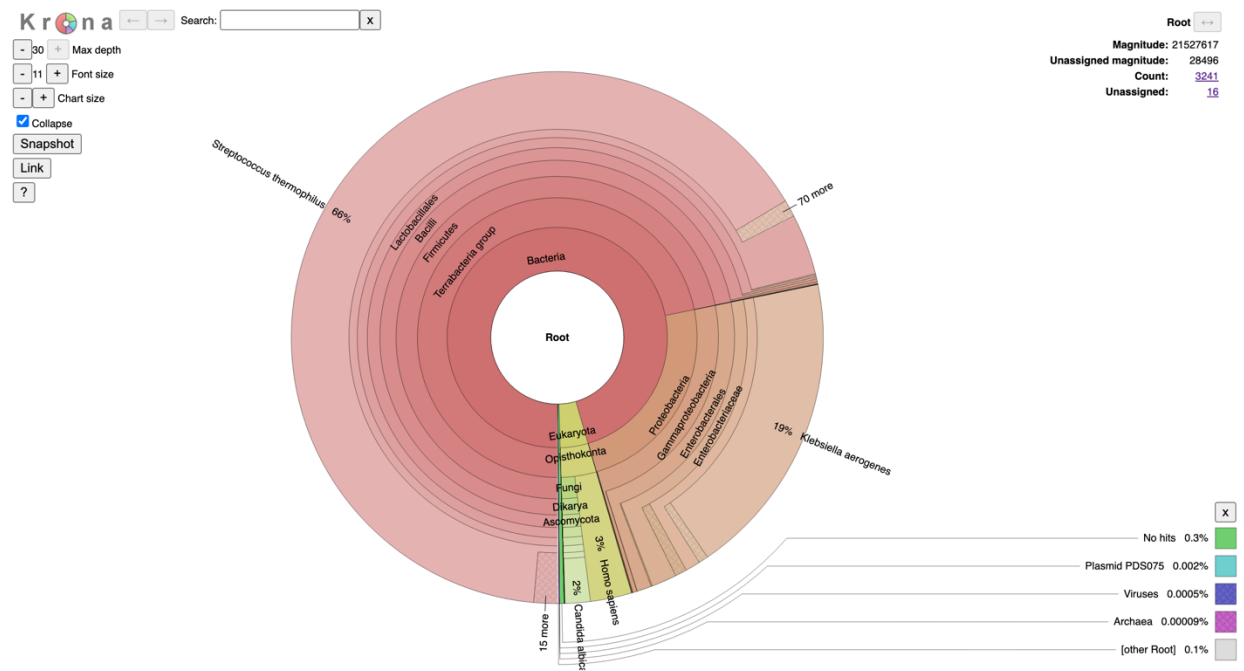
html file: chart.html



ICU Metagenome Analysis

BVBC / BVBC Tests / Taxonomic Classification / ICU metagenome / chart.html

html file: chart.html

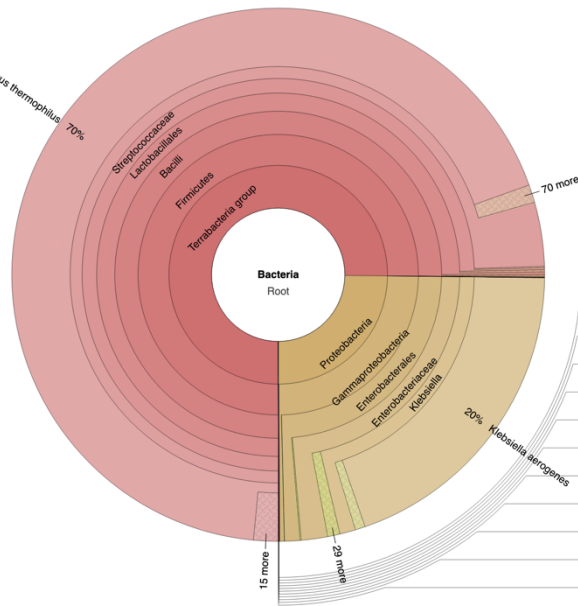


html file: chart.html

Krona Search: X

-30 + Max depth
-11 + Font size
- + Chart size

Collapse
Snapshot
Link
?



Bacteria ++

Magnitude: 20558204
Unassigned magnitude: 5431
Count: 2733
Unassigned: 3
Taxon: 2
Rank: superkingdom

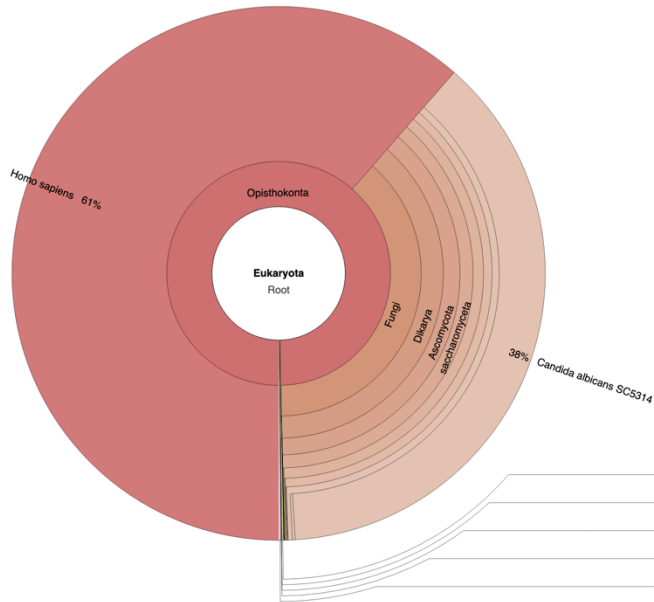
95% of Root

html file: chart.html

Krona Search: X

-30 + Max depth
-11 + Font size
- + Chart size

Collapse
Snapshot
Link
?




Eukaryota ++

Magnitude: 879781
Unassigned magnitude: 1467
Count: 351
Unassigned: 1
Taxon: 2759
Rank: superkingdom

4% of Root


html file: chart.html 

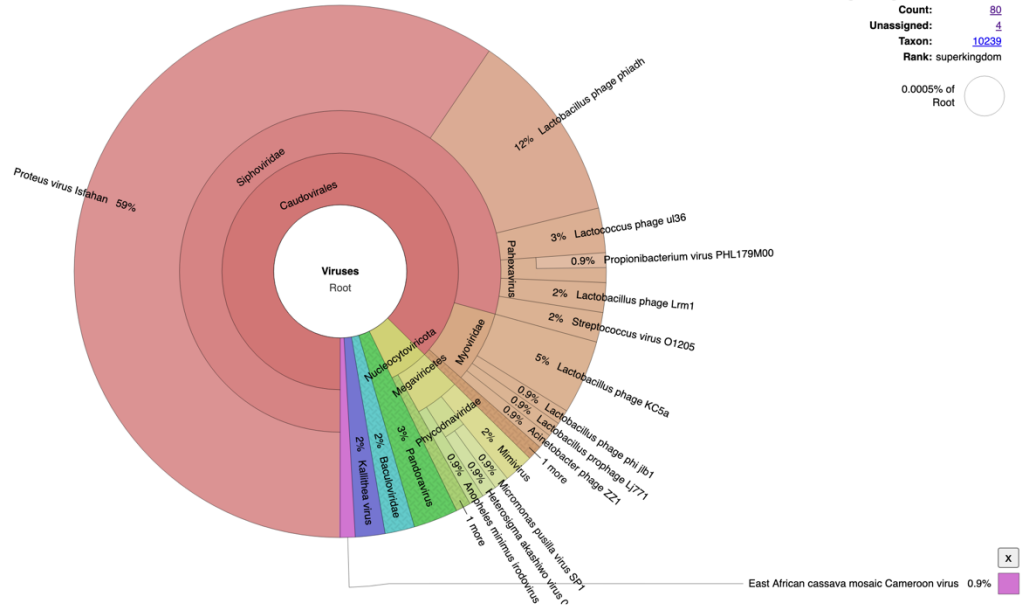
Krona 

Max depth
 Font size
 Chart size
 Collapse

Viruses

Magnitude: 111
Unassigned magnitude: 0
Count: 80
Unassigned: 4
Taxon: 10239
Rank: superkingdom

0.0005% of Root 



References

- [Taxonomic Classification Quick Reference Guide](#)
- [Taxonomic Classification Tutorial](#)