

# BV-BRC Test Report

## A18. Service – Metagenomic Read Mapping

<b>Item to test</b>	Metagenomic Read Mapping Service using bacterial and metagenomic reads
<b>URL</b>	<a href="https://www.bv-brc.org/app/MetagenomicReadMapping">https://www.bv-brc.org/app/MetagenomicReadMapping</a>
<b>Prerequisites</b>	Bacterial and metagenomic read files in Workspace
<b>References</b>	<a href="https://www.bv-brc.org/docs/quick_references/services/metagenomic_read_mapping_service.html">https://www.bv-brc.org/docs/quick_references/services/metagenomic_read_mapping_service.html</a> <a href="https://www.bv-brc.org/docs/tutorial/metagenomic_read_mapping/metagenomic_read_mapping.html">https://www.bv-brc.org/docs/tutorial/metagenomic_read_mapping/metagenomic_read_mapping.html</a>
<b>Tester(s)</b>	Rebecca Wattam, Maulik Shukla
<b>Test date</b>	08-May-2022 (follow-up from original test)
<b>Test result</b>	<b>Passed</b>

### Overview

- Test the Metagenomic Read Mapping Service using exemplar bacterial and metagenomic reads sets.
- Test input options, i.e., single-end / paired-end read files and SRA accessions.
- Search against antimicrobial resistance gene database (CARD) and virulence factor database (VFDB).
- For each job submitted, verify successful completion of the job and presence of output files.
- Review the list of AMR and virulence genes detected.

### Test Data

<b>Dataset</b>	<b>Rational</b>	<b>Input Format</b>	<b>Input</b>
Escherichia coli - SRR3584989	Workshop example	Fastq file, SRA accession	SRR3584989_1.fastq SRR3584989_2.fastq
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/Metagenomic%20Read%20Mapping>

### Test Results

- All jobs completed successfully, without any errors.
- All jobs resulted in expected output files in corresponding job output directory, providing list of AMR or virulence genes detected, their fasta sequences, read alignments, and an HTML report.

- The HTML report was informative and provided a concise summary of the AMR or virulence genes detected, corresponding score and p-value and alignment stats.
- 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/Metagenomic%20Read%20Mapping>
- 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 with summary of AMR and virulence genes detected.

**Input File**

PAIRED READ LIBRARY

READ FILE 1

READ FILE 2

SINGLE READ LIBRARY

READ FILE

SRA RUN ACCESSION

SRR3584989

**Selected libraries**

Place read files here using the arrow buttons.

SRR3584989

**Parameters**

PREDEFINED GENE SET NAME

CARD

OUTPUT FOLDER

Metagenomic Read Mapping

OUTPUT NAME

Ecoli - SRR3584989 - CARD1

Status	ID	Service	Output Name	Submit	Start	Completed
completed	7747921	MetagenomicReadMapping	Ecoli - SRR3584989 - CARD	5/8/22, 6:12 PM	5/8/22, 6:12 PM	5/8/22, 6:13 PM
completed	7747922	MetagenomicReadMapping	Ecoli - SRR3584989 - VFDB	5/8/22, 6:12 PM	5/8/22, 6:12 PM	5/8/22, 6:13 PM
completed	7747923	MetagenomicReadMapping	Ecoli - SRR3584989 - fastq	5/8/22, 6:15 PM	5/8/22, 6:15 PM	5/8/22, 6:16 PM
completed	7747924	MetagenomicReadMapping	Ecoli - SRR3584989 - fastq - CARD	5/8/22, 6:16 PM	5/8/22, 6:16 PM	5/8/22, 6:17 PM
completed	7747930	MetagenomicReadMapping	ICU Metagenome - CARD	5/8/22, 7:15 PM	5/8/22, 7:15 PM	5/8/22, 7:22 PM
completed	7747931	MetagenomicReadMapping	ICU Metagenome - VFDB	5/8/22, 7:15 PM	5/8/22, 7:15 PM	5/8/22, 7:25 PM

BVBRC / BVBRC Tests / Metagenomic Read Mapping (9 items)

Name	Size	Owner	Members	Created
Parent folder				
Ecoli - SRR3584989 - CARD	5.2 kB	me	Public	5/8/22, 6:13 PM
Ecoli - SRR3584989 - VFDB	5.2 kB	me	Public	5/8/22, 6:13 PM
SRR3584989_1.fastq	493.2 MB	me	Public	5/8/22, 6:14 PM
SRR3584989_2.fastq	495.8 MB	me	Public	5/8/22, 6:14 PM
Ecoli - SRR3584989 - fastq	5.4 kB	me	Public	5/8/22, 6:16 PM
Ecoli - SRR3584989 - fastq - CARD	5.4 kB	me	Public	5/8/22, 6:16 PM
UC.MICU.02.30.fastq	8.3 GB	me	Public	5/8/22, 7:14 PM
ICU Metagenome - CARD	5.3 kB	me	Public	5/8/22, 7:22 PM
ICU Metagenome - VFDB	5.3 kB	me	Public	5/8/22, 7:25 PM

Name	Size	Owner	Members	Created
Parent folder				
MetagenomicReadMappingReport.html	53.3 kB	me	Public	5/8/22, 6:13 PM
kma.aln	457.3 kB	me	Public	5/8/22, 6:13 PM
kma.frag.gz	4.0 MB	me	Public	5/8/22, 6:13 PM
kma.fsa	132.5 kB	me	Public	5/8/22, 6:13 PM
kma.res	15.8 kB	me	Public	5/8/22, 6:13 PM

## Results

The table below shows results from read mapping against the selected template database CARD using [KMA \(reference\)](#). The matching reference genes are shown as *Template* and the input reads are shown as *Query*. The table provides a summary of the most common statistics for each mapped template.

Template	Function	Genome	Score	Expected	Template length	Template Identity	Template Coverage	Query Identity	Query Coverage	Depth	q value	p value
<a href="#">CARDIAAC75136.1</a>	mdtB	<a href="#">Escherichia coli str. K-12 substr. MG1655</a>	188245	11188	3123	95.93	100.00	95.93	100.00	74.31	157189.18	1.0e-26
<a href="#">CARDIU00096.1</a>	Escherichia coli 16S rRNA mutation in the rrsB gene conferring resistance to paromomycin	<a href="#">Escherichia coli K-12</a>	549553	5306	1542	100.00	100.00	100.00	100.00	375.93	533835.56	1.0e-26
<a href="#">CARDIACI32877.1</a>	dfra14	<a href="#">Escherichia coli</a>	68111	1751	483	99.79	100.00	99.79	100.00	150.74	63033.46	1.0e-26
<a href="#">CARDIAE014075.1</a>	Escherichia coli 23S rRNA with mutation conferring resistance to clindamycin	<a href="#">Escherichia coli CFT073</a>	1426608	9018	2904	99.24	100.03	99.21	99.97	544.48	1399778.57	1.0e-26
<a href="#">CARDINP_414996.1</a>	Escherichia coli acrA	<a href="#">Escherichia coli str. K-12 substr. MG1655</a>	78071	4325	1194	99.08	100.00	99.08	100.00	71.96	66003.45	1.0e-26
<a href="#">CARDICD172759</a>	Escherichia coli PtsI with mutation conferring resistance to fosfomicin	<a href="#">Escherichia coli str. K-12 substr. MC4100</a>	120524	6232	1728	99.71	100.00	99.71	100.00	75.46	103051.75	1.0e-26
<a href="#">CARDIAAC76539.1</a>	mdtF	<a href="#">Escherichia coli str. K-12 substr. MG1655</a>	204002	11137	3114	98.04	100.00	98.04	100.00	74.94	172895.17	1.0e-26
<a href="#">CARDINP_415434.1</a>	msbA	<a href="#">Escherichia coli str. K-12 substr. MG1655</a>	131376	6301	1749	99.03	100.00	99.03	100.00	82.98	113625.99	1.0e-26
<a href="#">CARDIBAE78082.1</a>	mdtP	<a href="#">Escherichia coli str. K-12 substr. W3110</a>	86249	5310	1467	96.66	100.00	96.66	100.00	70.49	71550.50	1.0e-26
<a href="#">CARDINP_414995.1</a>	acrB	<a href="#">Escherichia coli str. K-12 substr. MG1655</a>	215686	11252	3150	99.08	100.00	99.08	100.00	75.69	184160.31	1.0e-26
<a href="#">CARDICD173208</a>	Escherichia coli UhpT with mutation conferring resistance to fosfomicin	<a href="#">Escherichia coli str. K-12 substr. MC4100</a>	101748	5030	1392	97.70	100.00	97.70	100.00	84.78	87605.27	1.0e-26

## Results

The table below shows results from read mapping against the selected template database VFDB using [KMA \(reference\)](#). The matching reference genes are shown as *Template* and the input reads are shown as *Query*. The table provides a summary of the most common statistics for each mapped template.

Template	Function	Genome	Score	Expected	Template length	Template Identity	Template Coverage	Query Identity	Query Coverage	Depth	q value	p value
<a href="#">VFDBIVFG000871</a>	Type 1 fimbriae Regulatory protein fimB	<a href="#">Escherichia coli CFT073</a>	39056	707	603	98.51	100.00	98.51	100.00	72.13	36984.37	1.0e-26
<a href="#">VFDBIVFG000872</a>	Type 1 fimbriae Regulatory protein fimE	<a href="#">Escherichia coli CFT073</a>	39774	700	597	98.99	100.00	98.99	100.00	73.25	37721.93	1.0e-26
<a href="#">VFDBIVFG000873</a>	Type-1 fimbrial protein, A chain precursor	<a href="#">Escherichia coli CFT073</a>	30922	712	606	92.08	100.00	92.08	100.00	67.43	28849.90	1.0e-26
<a href="#">VFDBIVFG000874</a>	Fimbrin-like protein fimI precursor	<a href="#">Escherichia coli CFT073</a>	40540	633	540	99.26	100.00	99.26	100.00	82.89	38679.33	1.0e-26
<a href="#">VFDBIVFG000875</a>	Chaperone protein fimC precursor	<a href="#">Escherichia coli CFT073</a>	45624	850	726	99.04	100.00	99.04	100.00	68.63	43134.95	1.0e-26
<a href="#">VFDBIVFG000876</a>	Outer membrane usher protein fimD precursor	<a href="#">Escherichia coli CFT073</a>	188689	2996	2637	99.05	100.00	99.05	100.00	79.05	179888.11	1.0e-26
<a href="#">VFDBIVFG000877</a>	FimF protein precursor	<a href="#">Escherichia coli CFT073</a>	29487	627	534	98.88	100.00	98.88	100.00	61.35	27656.38	1.0e-26
<a href="#">VFDBIVFG000878</a>	FimG protein precursor	<a href="#">Escherichia coli CFT073</a>	26438	592	504	98.81	100.00	98.81	100.00	58.31	24711.70	1.0e-26
<a href="#">VFDBIVFG000879</a>	FimH protein precursor	<a href="#">Escherichia coli CFT073</a>	50252	1067	912	97.48	100.00	97.48	100.00	63.53	47138.87	1.0e-26
<a href="#">VFDBIVFG000916</a>	heme oxygenase ChuS	<a href="#">Escherichia coli CFT073</a>	46288	1205	1029	97.96	100.00	97.96	100.00	50.67	42794.50	1.0e-26
<a href="#">VFDBIVFG000917</a>	Outer membrane heme/hemoglobin receptor ChuA	<a href="#">Escherichia coli CFT073</a>	68461	2270	1947	98.46	100.00	98.46	100.00	39.36	61941.64	1.0e-26
<a href="#">VFDBIVFG000918</a>	periplasmic heme-binding protein ChuT	<a href="#">Escherichia coli CFT073</a>	57517	1160	993	98.49	100.00	98.49	100.00	63.89	54127.83	1.0e-26
<a href="#">VFDBIVFG000919</a>	Putative oxygen independent heme/hemoglobin III oxidase	<a href="#">Escherichia coli CFT073</a>	84046	1554	1338	98.65	100.00	98.65	100.00	70.39	79494.82	1.0e-26

### Input File ?

PAIRED READ LIBRARY +

↓ READ FILE 1 📁

↓ READ FILE 2 📁

SINGLE READ LIBRARY +

↓ UC.MICU.02.30.fastq 📁

SRA RUN ACCESSION +

SRR

### Selected libraries ?

Place read files here using the arrow buttons.

S(UC.MICU.02.30.fastq) i x

### Parameters ?

PREDEFINED GENE SET NAME

CARD ▼

OUTPUT FOLDER

↓ Metagenomic Read Mapping 📁

OUTPUT NAME

ICU Metagenome - CARD1

Reset
Submit

### Input data

- Single-end libraries:
- [/BVBRCC@patricbrc.org/BVBRCC/Tests/Metagenomic Read Mapping/UC.MICU.02.30.fastq](#)

### Results

The table below shows results from read mapping against the selected template database CARD using [KMA \(reference\)](#). The matching reference genes are shown as *Template* and the input reads are shown as *Query*. The table provides a summary of the most common statistics for each mapped template.

Template	Function	Genome	Score	Expected	Template length	Template Identity	Template Coverage	Query Identity	Query Coverage	Depth	q value	p value
<a href="#">CARDINP_461214.1</a>	Salmonella enterica gyrA conferring resistance to fluoroquinolones	<a href="#">Salmonella enterica subsp. enterica serovar Typhimurium str. LT2</a>	97010	36766	2637	77.82	88.21	88.22	113.37	52.30	27129.31	1.0e-26
<a href="#">CARDIBAJ09383.1</a>	qacA	<a href="#">Staphylococcus aureus</a>	2805	21588	1545	77.86	77.99	99.83	128.22	1.82	14463.45	1.0e-26
<a href="#">CARDIAE014075.1_237159_240063</a>	Escherichia coli 23S rRNA with mutation conferring resistance to clindamycin	<a href="#">Escherichia coli CFT073</a>	3759780	36397	2904	96.63	100.03	96.59	99.97	1454.55	3651982.75	1.0e-26
<a href="#">CARDINZ_CP009828.1_497113_500039</a>	Staphylococcus aureus 23S rRNA with mutation conferring resistance to linezolid	<a href="#">Staphylococcus aureus</a>	548507	40291	2926	96.96	100.03	96.93	99.97	257.29	438660.22	1.0e-26
<a href="#">CARDICDJ72759</a>	Escherichia coli PtsI with mutation conferring resistance to fosfomycin	<a href="#">Escherichia coli str. K-12 substr. MC4100</a>	143472	24053	1728	86.92	97.11	89.51	102.98	113.91	85126.09	1.0e-26
<a href="#">CARDIAAO04716.1</a>	dfrC	<a href="#">Staphylococcus epidermidis ATCC 12228</a>	381	6788	486	75.51	79.63	94.83	125.58	0.89	5726.54	1.0e-26
<a href="#">CARDINP_414995.1</a>	acrB	<a href="#">Escherichia coli str. K-12 substr. MG1655</a>	106228	43916	3150	81.56	95.43	85.46	104.79	55.39	25859.77	1.0e-26
<a href="#">CARDINZ_CP018138.1_1692408_1695312</a>	Streptococcus pneumoniae 23S rRNA mutation conferring resistance to macrolides	<a href="#">Streptococcus pneumoniae</a>	20444341	17741	2904	96.32	99.93	96.38	100.07	7943.34	20391179.40	1.0e-26

# Metagenomic Read Mapping Report for job 7747931

## Input data

Single-end libraries:

- [/BVBRC@patricbrc.org/BVBRC Tests/Metagenomic Read Mapping/UC.MICU.02.30.fastq](#)

## Results

The table below shows results from read mapping against the selected template database VFDB using [KMA \(reference\)](#). The matching reference genes are shown as *Template* and the input reads are shown as *Query*. The table provides a summary of the most common statistics for each mapped template.

Template	Function	Genome	Score	Expected	Template length	Template Identity	Template Coverage	Query Identity	Query Coverage	Depth	q value	p value
<a href="#">VFDB VFG001443</a>	outer membrane protein A	<a href="#">Escherichia coli O18:K1:H7 str. RS218</a>	34666	2862	1041	86.07	101.15	85.09	98.86	54.79	26952.30	1.0e-26
<a href="#">VFDB VFG044326</a>	yersiniabactin ABC transporter ATP-binding/permease protein YbtP	<a href="#">Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044</a>	191261	4887	1803	99.56	99.72	99.83	100.28	107.55	177084.71	1.0e-26
<a href="#">VFDB VFG044334</a>	salmochelin receptor IroN	<a href="#">Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044</a>	72525	5961	2175	79.40	92.41	85.92	108.21	54.84	56451.88	1.0e-26
<a href="#">VFDB VFG044335</a>	glucosyltransferase IroB	<a href="#">Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044</a>	44219	3065	1116	75.36	89.61	84.10	111.60	67.54	35816.39	1.0e-26
<a href="#">VFDB VFG048409</a>	2,3-dihydroxybenzoate-2,3-dehydrogenase	<a href="#">Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044</a>	59903	2156	786	90.97	100.38	90.62	99.62	101.47	53734.40	1.0e-26
<a href="#">VFDB VFG048419</a>	2,3-dihydro-2,3-dihydroxybenzoate synthetase, isochroismatase	<a href="#">Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044</a>	90079	2330	852	94.37	100.00	94.37	100.00	127.54	83321.95	1.0e-26
<a href="#">VFDB VFG048429</a>	enterobactin synthase subunit E	<a href="#">Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044</a>	152838	4374	1608	92.41	100.00	92.41	100.00	122.72	140201.57	1.0e-26
<a href="#">VFDB VFG048439</a>	isochorismate synthase	<a href="#">Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044</a>	68045	3256	1188	87.37	100.00	87.37	100.00	87.73	58869.88	1.0e-26
		<a href="#">Klebsiella pneumoniae</a>										

## References

- [Metagenomic Read Mapping Quick Reference Guide](#)
- [Metagenomic Read Mapping Tutorial](#)