# **BV-BRC Test Report**

## A8. Service - Similar Genome Finder - Bacteria

Item to test	Similar Genome Finder Service using bacterial genome, fasta and fastq files					
URL	https://www.bv-brc.org/app/GenomeDistance					
Prerequisites	Bacterial Fasta contig file and FastQ file in the workspace					
References	https://www.bv-brc.org/docs/quick_references/services/ similar_genome_finder_service.html https://www.bv-brc.org/docs/tutorial/similar_genome_finder/ similar_genome_finder.html					
Tester(s)	Rebecca Wattam, Maulik Shukla					
Test date	date 07-May-2022 (follow-up from original test)					
Test result	Passed					

#### Overview

- Test the Similar Genome Finder Service using exemplar bacterial genome.
- Test input options, i.e., genome id / genome name, fasta contig file, and fastq file as input.
- Test different databases, i.e., reference / representative genomes and all public genomes.
- For each job submitted, verify successful completion of the job and the quality of the search results.
- Test the selection and actions from the search result table.

#### **Test Data**

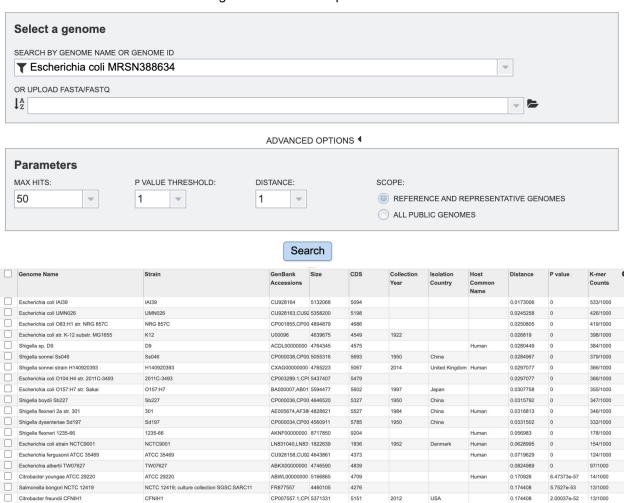
Dataset	Rational	Input Format	Input				
Escherichia coli MRSN388634	Workshop example	Genome Name	Escherichia coli MRSN388634				
Escherichia coli MRSN388634	Workshop example	Fasta file	Escherichia_coli_MRSN388634.f asta				
Escherichia coli MRSN388634	Workshop example	Fastq file	SRR3584989_1.fastq				

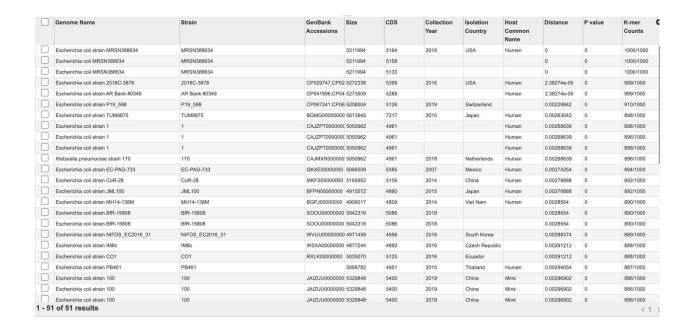
 All test datasets are available in the following public workspace:
 https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Similar%20Genome %20Finder

### **Test Results**

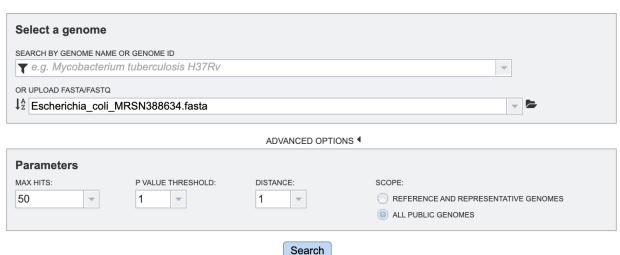
- All jobs completed successfully, without any errors.
- All jobs resulted in expected results, finding genomes that were most similar to the query genome.
- The results were presented as an interactive table and the table filters, selection and actions worked as expected.

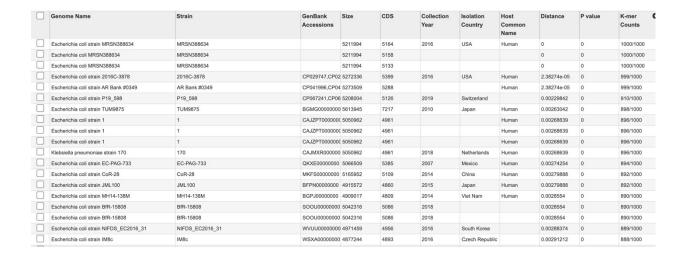
- For the same query genome, input using genome name, contigs, and fastq file produced similar results.
- Below are a series of screenshots showing submission forms, parameters values and results from various searches.
- Escherichia coli MRSN388634 genome name as input



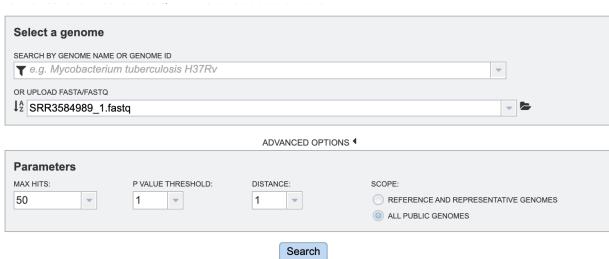


Escherichia coli MRSN388634 – fasta file as input





Escherichia coli MRSN388634 – fastaq file as input



	Genome Name	Strain	GenBank Accessions	Size	CDS	Collection Year	Isolation Country	Host Common Name	Distance	P value	K-mer Counts
	Escherichia coli strain MRSN388634	MRSN388634		5211994	5164	2016	USA	Human	0.0489202	0	218/1000
	Escherichia coli strain 2016C-3878	2016C-3878	CP029747,CP02	5272336	5399	2016	USA	Human	0.0489202	0	218/1000
	Escherichia coli MRSN388634	MRSN388634		5211994	5158				0.0489202	0	218/1000
	Escherichia coli strain AR Bank #0349	AR Bank #0349	CP041996,CP04	5273509	5288			Human	0.0489202	0	218/1000
	Escherichia coli MRSN388634	MRSN388634		5211994	5133				0.0489202	0	218/1000
	Escherichia coli strain 100	100	JAIZUU0000000	5329848	5400	2019	China	Mink	0.0515304	0	204/1000
	Escherichia coli strain 100	100	JAIZUU0000000	5329848	5400	2019	China	Mink	0.0515304	0	204/1000
	Escherichia coli strain 100	100	JAIZUU0000000	5329848	5400	2019	China	Mink	0.0515304	0	204/1000
	Escherichia coli strain CoR-28	CoR-28	MKFS00000000	5165952	5109	2014	China	Human	0.052117	0	201/1000
	Escherichia coli strain B3866	B3866	RWYJ00000000	5206483	5324				0.052117	0	201/1000
	Escherichia coli strain P19_598	P19_598	CP067241,CP06	5208004	5126	2019	Switzerland		0.052117	0	201/1000
	Escherichia coli strain SYAB3	SYAB3	LNUE00000000	5241410	5221	2012	China	Human	0.0523149	0	200/1000
	Escherichia coli strain SYAB1	SYAB1	CM003776,LNU	5241449	5213	2012	China	Human	0.0523149	0	200/1000
	Escherichia coli strain PB185	PB185		5145397	5088	2015	Thailand	Human	0.0523149	0	200/1000
	Escherichia coli strain ZTA15/00685EB1	ZTA15/00685EB1		5202454	5211	2015	Spain	Pig	0.0523149	0	200/1000
	Escherichia coli strain SYAU2	SYAU2	CM003779,LNU	5224350	5160	2012	China	Human	0.0525139	0	199/1000
	Escherichia coli strain SYAU3	SYAU3	CM003777,LNU	5174089	5119	2012	China	Human	0.0525139	0	199/1000
	Escherichia coli strain SYAU1	SYAU1	LNUG00000000	5187611	5140	2012	China	Human	0.0525139	0	199/1000
	Escherichia coli strain SYAB2	SYAB2	CM003775,LNU	5237295	5209	2012	China	Human	0.0525139	0	199/1000
	Escherichia coli strain SYAB4	SYAB4	CM003774,LNU	5229355	5208	2012	China	Human	0.0525139	0	199/1000
	Escherichia coli strain MRY15-117	MRY15-117	AP017617,AP01	5295071	5224	2012	Japan	Cow	0.0525139	0	199/1000
	Escherichia coli strain B222	B222	NRTO00000000	5321608	5341	2016	China	Human	0.0525139	0	199/1000
	Escherichia coli strain PB356	PB356		5155675	5125	2015	Thailand	Human	0.0525139	0	199/1000
- 5	Escherichia coli strain NIHE15-1569	NIHE15-1569	BPUH00000000	5190348	5205	2015	Viet Nam	Human	0.0525139	0	199/1000

## References

- Similar Genome Finder Service Quick Reference Guide
- Similar Genome Finder Service Tutorial