

BV-BRC Test Report

A3. Service - Genome Annotation - Phage

Item to test	Genome Annotation Service using phage contig files
URL	https://www.bv-brc.org/app/Annotation
Prerequisites	Phage Fasta contig files in Workspace
References	https://www.bv-brc.org/docs/quick_references/services/genome_annotation_service.html https://www.bv-brc.org/docs/tutorial/genome_annotation/genome_annotation.html
Tester(s)	Rebecca Wattam, Maulik Shukla
Test date	21-Apr-2022 (follow-up from original test)
Test result	Passed

Overview

- Test the Genome Annotation Service using exemplar genome sequences for phage genomes.
- Test input options, i.e., contig file from the user's machine or from the user's workspace.
- For each job submitted, verify successful completion of the job, availability of the output files in the workspace, and quality of the annotations by comparing them with the same, or closely related, public genome in PATRIC.
- Verify successful integration of the genome in BV-BRC by reviewing the genome overview pages and other genome-level tabs.

Test Data

Genome Name	Rational	Input File
Bacillus phage phi105	Reference genome	Bacillus_phage_phi105.fasta
Escherichia phage ST31	Reference genome	Escherichia_phage_ST31.fasta
Pseudomonas phage PHB09	Reference genome	Pseudomonas_phage_PHB09.fasta

- 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/Genome%20Annotation/Phage>

Test Results

- All annotation jobs completed successfully, without any errors.
- All jobs resulted in expected output files in corresponding job output directory, providing annotations in various standard file formats.
- All genomes were indexed successfully in the Solr database and were accessible from the BV-BRC website.
- For each of the genomes, the genome overview page, genome browser, and all other genome level tabs loaded successfully as expected.
- For each of the genomes, annotation of genomic features, protein functions, specialty genes was as expected when compared to those in corresponding public genome in PATRIC.

- 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/Genome%20Annotation/Phage>
- Below are a series of screenshots showing successful completion of the jobs, availability of the result files in the workspace, and comparison of the annotation between test genome and corresponding public genome.
- Job status: successful completion of all test jobs

Status	ID	Service	Output Name	Submit	Start	Completed
completed	7420392	Annotation	Viruses Escherichia phage ST31	4/21/22, 5:38 AM	4/21/22, 5:38 AM	4/21/22, 5:41 AM
completed	7419859	Annotation	Influenza A virus A Puerto Rico 8 1934	4/20/22, 4:03 PM	4/20/22, 4:03 PM	4/20/22, 4:06 PM
completed	7419852	Annotation	Viruses Pseudomonas phage PHB09	4/20/22, 3:50 PM	4/20/22, 3:51 PM	4/20/22, 3:56 PM

- Result verification: Bacillus phage phi105
Job result directory: <https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Phage/Viruses%20Bacillus%20phage%20phi105>

Name	Size	Owner	Members	Created
Parent folder			-	
Viruses Bacillus phage phi105.contigs.fasta	40.0 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.embl	102.6 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.feature_dna.fasta	44.8 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.feature_protein.fasta	19.2 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.features.txt	7.2 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.gb	99.3 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.genome	150.9 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.gff	6.9 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.merged.gb	99.3 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.tar.gz	27.8 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.txt	57.7 kB	me	Public	4/20/22, 3:51 PM
Viruses Bacillus phage phi105.xls	70.7 kB	me	Public	4/20/22, 3:51 PM
load_files		me	Public	4/20/22, 3:51 PM

- Annotation comparison with the public reference genome in PATRIC:

Genomic Features			Genomic Features		
	PATRIC	RefSeq		PATRIC	RefSeq
CDS	69	0	CDS	69	51
			regulatory	0	53
			misc_feature	0	2

Protein Features			Protein Features		
	PATRIC	RefSeq		PATRIC	RefSeq
Hypothetical proteins	29	0	Hypothetical proteins	28	0
Proteins with functional assignments	40	0	Proteins with functional assignments	41	51
Proteins with EC number assignments	1	0	Proteins with EC number assignments	1	0
Proteins with GO assignments	1	0	Proteins with GO assignments	1	0
Proteins with Pathway assignments	0	0	Proteins with Pathway assignments	0	0
Proteins with Subsystem assignments	0	0	Proteins with Subsystem assignments	0	0
Proteins with PATRIC genus-specific family (PLfam) assignments	0	0	Proteins with PATRIC genus-specific family (PLfam) assignments	69	0
Proteins with PATRIC cross-genus family (PGfam) assignments	50	0	Proteins with PATRIC cross-genus family (PGfam) assignments	53	0
Proteins with FIGfam assignments	0	0	Proteins with FIGfam assignments	40	0

- Result verification: Escherichia phage ST31

Job result directory: <https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Phage/Viruses%20Escherichia%20phage%20ST31>

Name	Size	Owner	Members	Created
Parent folder			-	
Viruses Escherichia phage ST31.contigs.fasta	40.4 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.embl	97.8 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.feature_dna.fasta	43.6 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.feature_protein.fasta	18.2 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.features.txt	6.2 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.gb	94.5 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.genome	135.5 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.gff	6.1 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.merged.gb	94.5 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.tar.gz	25.6 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.txt	56.0 kB	me	Public	4/21/22, 5:39 AM
Viruses Escherichia phage ST31.xls	67.6 kB	me	Public	4/21/22, 5:39 AM
load_files		me	Public	4/21/22, 5:39 AM

Annotation comparison with the public reference genome in PATRIC:
<https://patricbrc.org/view/Genome/1983789.5>

Genomic Features			Genomic Features		
	PATRIC	RefSeq		PATRIC	RefSeq
CDS	56	0	CDS	56	45
tRNA	1	0			

Protein Features			Protein Features		
	PATRIC	RefSeq		PATRIC	RefSeq
Hypothetical proteins	12	0	Hypothetical proteins	12	29
Proteins with functional assignments	44	0	Proteins with functional assignments	44	16
Proteins with EC number assignments	5	0	Proteins with EC number assignments	5	0
Proteins with GO assignments	4	0	Proteins with GO assignments	4	0
Proteins with Pathway assignments	2	0	Proteins with Pathway assignments	2	0
Proteins with Subsystem assignments	0	0	Proteins with Subsystem assignments	0	0
Proteins with PATRIC genus-specific family (PLfam) assignments	0	0	Proteins with PATRIC genus-specific family (PLfam) assignments	56	0
Proteins with PATRIC cross-genus family (PGfam) assignments	30	0	Proteins with PATRIC cross-genus family (PGfam) assignments	30	0
Proteins with FIGfam assignments	0	0	Proteins with FIGfam assignments	19	0

- Result verification: Pseudomonas phage PHB09

Job result directory: <https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Phage/Viruses%20Pseudomonas%20phage%20PHB09>

Name	Size	Owner	Members	Created
Parent folder			-	
Viruses Pseudomonas phage PHB09.contigs.fasta	96.4 kB	me	Public	4/20/22, 3:53 PM
Viruses Pseudomonas phage PHB09.embl	264.8 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.feature_dna.fasta	111.6 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.feature_protein.fasta	49.6 kB	me	Public	4/20/22, 3:53 PM
Viruses Pseudomonas phage PHB09.features.txt	23.0 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.gb	257.0 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.genome	378.6 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.gff	22.4 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.merged.gb	257.0 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.tar.gz	65.2 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.txt	141.1 kB	me	Public	4/20/22, 3:52 PM
Viruses Pseudomonas phage PHB09.xls	170.0 kB	me	Public	4/20/22, 3:52 PM
load_files		me	Public	4/20/22, 3:53 PM

Annotation comparison with the public reference genome in PATRIC:

<https://patricbrc.org/view/Genome/511145.12>

Genomic Features			Genomic Features		
	PATRIC	RefSeq		PATRIC	RefSeq
CDS	211	0	CDS	211	185
tRNA	15	0	tRNA	18	0
			gene	0	185

Protein Features			Protein Features		
	PATRIC	RefSeq		PATRIC	RefSeq
Hypothetical proteins	202	0	Hypothetical proteins	202	164
Proteins with functional assignments	9	0	Proteins with functional assignments	9	21
Proteins with EC number assignments	2	0	Proteins with EC number assignments	0	0
Proteins with GO assignments	2	0	Proteins with GO assignments	2	0
Proteins with Pathway assignments	2	0	Proteins with Pathway assignments	0	0
Proteins with Subsystem assignments	0	0	Proteins with Subsystem assignments	0	0
Proteins with PATRIC genus-specific family (PLfam) assignments	0	0	Proteins with PATRIC genus-specific family (PLfam) assignments	10	0
Proteins with PATRIC cross-genus family (PGfam) assignments	15	0	Proteins with PATRIC cross-genus family (PGfam) assignments	15	0
Proteins with FIGfam assignments	0	0	Proteins with FIGfam assignments	0	0

References

- [Genome Annotation Protocol](#)
- [RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes](#)
- [Genome Annotation Service](#)
- [Genome Annotation Service Quick Reference Guide](#)
- [Genome Annotation Service Tutorial](#)