BV-BRC Test Report

A2. Service - Genome Annotation - Bacteria

Item to test	Genome Annotation Service using bacterial contig files
URL	https://www.bv-brc.org/app/Annotation
Prerequisites	Bacterial 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 bacterial 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				
Mycobacterium tuberculosis H37Rv	Reference genome	Mycobacterium_tuberculosis_H37Rv.fast				
		a				
Escherichia coli MG1655	Reference genome	Escherichia_coli_MG1655.fasta				
Brucella suis 1330	Reference genome	Brucella_suis_1330.fasta				
Buchnera aphidicola Tuc7	Workshop example	Buchnera aphidicola strain				
		Tuc7_SRR4240359.fasta				
Escherichia coli MRSN388634	Workshop example	Escherichia_coli_MRSN388634.fasta				
Staphylococcus aureus VB4283	Workshop example	Staphylococcus_aureus_VB4283.fna				

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.

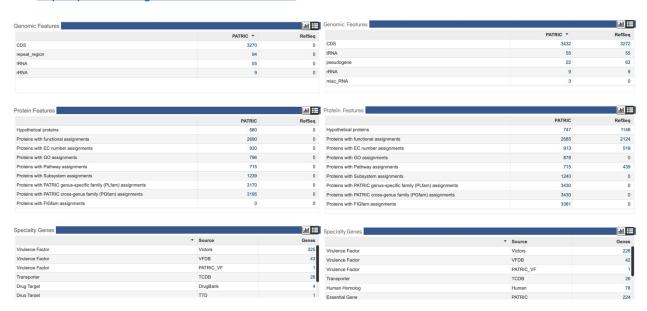
- All genomes were indexed successfully in the Solr database and were accessible from the BV-BRC website.
- For each of the genomes, the genome overview page, genome browser, and all other genome level tabs loaded successfully as expected.
- For each of the genomes, annotation of genomic features, protein functions, specialty genes was as expected when compared to those in corresponding public genomes 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/Bacteria
- 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.

Status	ID	Service	Output Name	Submit -	Start	Completed •
completed	7419704	Annotation	Buchnera aphidicola strain Tuc7	4/20/22, 2:48 PM	4/20/22, 2:48 PM	4/20/22, 2:53 PM
completed	7419707	Annotation	Escherichia coli MRSN388634	4/20/22, 2:48 PM	4/20/22, 2:49 PM	4/20/22, 2:58 PM
completed	7419711	Annotation	Staphylococcus aureus VB4283	4/20/22, 2:49 PM	4/20/22, 2:49 PM	4/20/22, 2:56 PM
completed	7419813	Annotation	Brucella suis 1330	4/20/22, 3:16 PM	4/20/22, 3:16 PM	4/20/22, 3:24 PM
completed	7419816	Annotation	Escherichia coli MG1655	4/20/22, 3:16 PM	4/20/22, 3:17 PM	4/20/22, 3:27 PM
completed	7419819	Annotation	Mycobacterium tuberculosis H37Rv	4/20/22, 3:17 PM	4/20/22, 3:17 PM	4/20/22, 3:27 PM

Result verification: Brucella suis 1330
 Job result directory: https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Bacteria/Brucella%20suis%201330

	Name	Size	Owner	Members	Created
t	Parent folder			-	
~	Brucella suis 1330.contigs.fasta	3.4 MB	me	Public	4/20/22, 3:21 PM
	Brucella suis 1330.embl	7.3 MB	me	Public	4/20/22, 3:22 PM
	Brucella suis 1330.feature_dna.fasta	3.4 MB	me	Public	4/20/22, 3:21 PM
	Brucella suis 1330.feature_protein.fasta	1.3 MB	me	Public	4/20/22, 3:21 PM
	Brucella suis 1330.features.txt	430.9 kB	me	Public	4/20/22, 3:21 PM
	Brucella suis 1330.gb	7.0 MB	me	Public	4/20/22, 3:21 PM
	Brucella suis 1330.genome	11.8 MB	me	Public	4/20/22, 3:21 PM
	Brucella suis 1330.gff	458.6 kB	me	Public	4/20/22, 3:22 PM
	Brucella suis 1330.merged.gb	7.0 MB	me	Public	4/20/22, 3:21 PM
	Brucella suis 1330.tar.gz	1.9 MB	me	Public	4/20/22, 3:22 PM
	Brucella suis 1330.txt	4.5 MB	me	Public	4/20/22, 3:22 PM
	Brucella suis 1330.xls	5.0 MB	me	Public	4/20/22, 3:21 PM
	GenomeReport.html	22.0 kB	me	Public	4/20/22, 3:21 PM
	genome_quality_details.txt	28.0 kB	me	Public	4/20/22, 3:21 PM
	load_files		me	Public	4/20/22, 3:22 PM
	quality.json	308.9 kB	me	Public	4/20/22, 3:24 PM

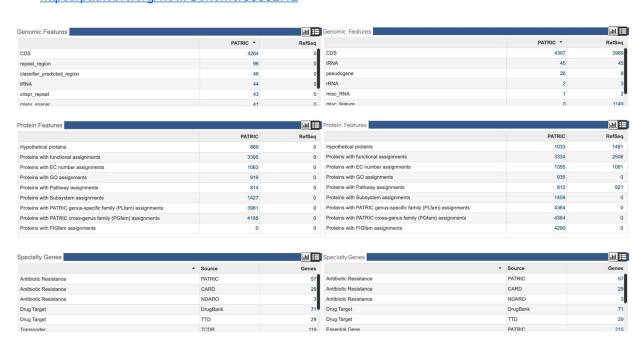
 Annotation comparison with the public reference genome in PATRIC: https://patricbrc.org/view/Genome/204722.5



Result verification: Mycobacterium tuberculosis H37Rv
 Job result directory: https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Bacteria/Mycobacterium%20tuberculosis%20H37Rv

				1	1	
	Name	•	Size	Owner	Members	Created
t	Parent folder				2	
	GenomeReport.html		12.7 kB	me	Public	4/20/22, 3:25 PM
-	Mycobacterium tuberculosis H37Rv.contigs.fasta		4.5 MB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.embl		9.8 MB	me	Public	4/20/22, 3:25 PM
	Mycobacterium tuberculosis H37Rv.feature_dna.fasta		4.7 MB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.feature_protein.fasta		1.9 MB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.features.txt		566.2 kB	me	Public	4/20/22, 3:25 PM
	Mycobacterium tuberculosis H37Rv.gb		9.4 MB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.genome		15.8 MB	me	Public	4/20/22, 3:25 PM
	Mycobacterium tuberculosis H37Rv.gff		599.7 kB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.merged.gb		9.4 MB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.tar.gz		2.5 MB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.txt		6.1 MB	me	Public	4/20/22, 3:26 PM
	Mycobacterium tuberculosis H37Rv.xls		6.8 MB	me	Public	4/20/22, 3:26 PM
	genome_quality_details.txt		54.8 kB	me	Public	4/20/22, 3:25 PM
	load_files			me	Public	4/20/22, 3:26 PM
	quality.json		554.7 kB	me	Public	4/20/22, 3:27 PM

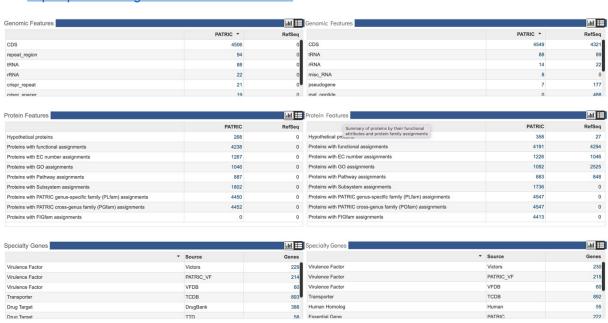
Annotation comparison with the public reference genome in PATRIC: https://patricbrc.org/view/Genome/83332.12



Result verification: Escherichia coli MG1655
 Job result directory: https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Bacteria/Escherichia%20coli%20MG1655

	News		Size	0	Members	Created C
	Name	•	Size	Owner	Members	Created
1	Parent folder				-	
7	Escherichia coli MG1655.contigs.fasta		4.7 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.embl		10.2 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.feature_dna.fasta		4.8 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.feature_protein.fasta		1.9 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.features.txt		597.6 kB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.gb		9.8 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.genome		16.6 MB	me	Public	4/20/22, 3:23 PM
	Escherichia coli MG1655.gff		636.8 kB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.merged.gb		9.8 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.tar.gz		2.7 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.txt		6.4 MB	me	Public	4/20/22, 3:24 PM
	Escherichia coli MG1655.xls		7.1 MB	me	Public	4/20/22, 3:24 PM
	GenomeReport.html		6.0 kB	me	Public	4/20/22, 3:23 PM
	genome_quality_details.txt		44.9 kB	me	Public	4/20/22, 3:23 PM
	load_files			me	Public	4/20/22, 3:24 PM
	quality.json		448.0 kB	me	Public	4/20/22, 3:27 PM

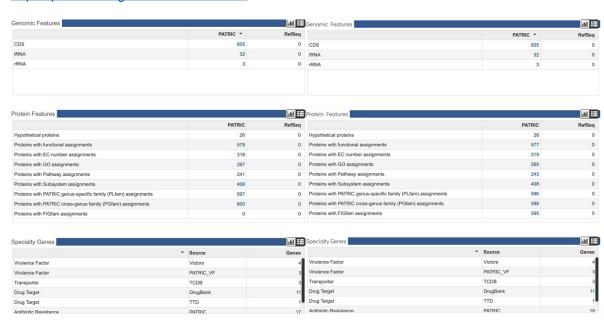
Annotation comparison with the public reference genome in PATRIC: https://patricbrc.org/view/Genome/511145.12



Result verification: Buchnera aphidicola strain Tuc7
 Job result directory: https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Bacteria/Buchnera%20aphidicola%20strain%20Tuc7

				_			•
	Name	•	Size	Owner	Members	Created	0
1	Parent folder				-		
-	Buchnera aphidicola strain Tuc7.contigs.fasta		671.1 kB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.embl		1.4 MB	me	Public	4/20/22, 2:51 PM	1
	Buchnera aphidicola strain Tuc7.feature_dna.fasta		660.8 kB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.feature_protein.fasta		260.6 kB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.features.txt		91.7 kB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.gb		1.4 MB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.genome		2.8 MB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.gff		101.3 kB	me	Public	4/20/22, 2:51 PM	1
	Buchnera aphidicola strain Tuc7.merged.gb		1.4 MB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.tar.gz		371.6 kB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.txt		897.2 kB	me	Public	4/20/22, 2:50 PM	1
	Buchnera aphidicola strain Tuc7.xls		993.8 kB	me	Public	4/20/22, 2:50 PM	1
	GenomeReport.html		8.1 kB	me	Public	4/20/22, 2:50 PM	1
	genome_quality_details.txt		27.8 kB	me	Public	4/20/22, 2:50 PM	1
	load_files			me	Public	4/20/22, 2:51 PM	1
	quality.json		283.4 kB	me	Public	4/20/22, 2:53 PM	1

Annotation comparison with corresponding public genome in PATRIC: https://patricbrc.org/view/Genome/9.164

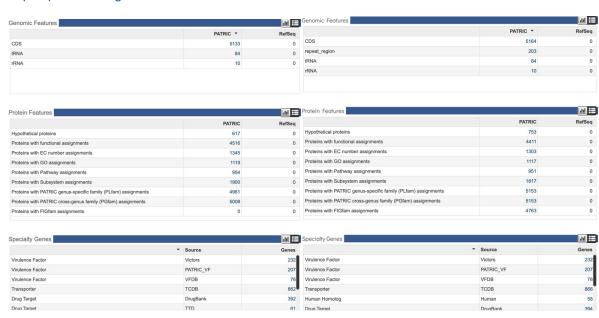


Result verification: Escherichia coli MRSN388634

Job result directory: https://www.bv-brc.org/workspace/BVBRC@patricbrc.org/BVBRC%20Tests/Genome%20Annotation/Bacteria/Escherichia%20coli%20MRSN388634

	Name	^	Size	Owner	Members	Created
t	Parent folder				-	
~	Escherichia coli MRSN388634.contigs.fasta		5.3 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.embl		11.5 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.feature_dna.fasta		5.3 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.feature_protein.fasta		2.1 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.features.txt		723.2 kB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.gb		11.1 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.genome		18.5 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.gff		750.8 kB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.merged.gb		11.1 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.tar.gz		2.9 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.txt		7.1 MB	me	Public	4/20/22, 2:56 PM
	Escherichia coli MRSN388634.xls		7.9 MB	me	Public	4/20/22, 2:56 PM
	GenomeReport.html		27.1 kB	me	Public	4/20/22, 2:56 PM
	alignments			me	Public	4/20/22, 2:56 PM
	genome_quality_details.txt		44.9 kB	me	Public	4/20/22, 2:56 PM
	load_files			me	Public	4/20/22, 2:56 PM

Annotation comparison with corresponding public genome in PATRIC: https://patricbrc.org/view/Genome/562.10576



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