BV-BRC Test Procedure

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
Author(s)	Rebecca Wattam, Ron Kenyon
Last update	01-Apr-2022

Overview

This Test Procedure is intended to confirm proper operation and returned results for the Genome Annotation Service using the bacterial (RASTtk) recipe with an exemplar fasta file.

Test Data

- Brucella_suis_test__contigs.fasta', available in the BV-BRC Test Data / Genome Annotation / Bacteria public workspace

Expected Results

- "Brucella suis 7121680725-test" validated job results and Genome Report in the public workspace

- Annotated Brucella suis 7121680725 genome in BV-BRC

Procedure

1. Sign in to BV-BRC.

2. Select 'Annotation' from the 'Tools & Services' top menu (https://www.bv-brc.org/app/Annotation).

3. In the 'Parameters' section of the Genome Annotation Service input form, use the following:

    a. CONTIGS: Use 'Brucella_suis_test__contigs.fasta', available in the BV-BRC Test Data / Genome Annotation / Bacteria public workspace. Public workspaces are available from the 'Workspaces' top menu.

    b. ANNOTATION RECIPE: Choose 'BACTERIA/ARCHAEA.'

    c. TAXONOMY NAME: Type "Brucella suis" and select "[species] Brucella suis" from the dropdown list. TAXONOMY ID should automatically display "29461."

    d. MY LABEL: Enter a name for the genome of your own choosing, e.g., "annotation test." This name will be concatenated with the TAXONOMY NAME and displayed in the OUTPUT NAME box, e.g., "Brucella suis annotation test."

    e. OUTPUT FOLDER: Choose or create a folder in the workspace to where you want the results of the annotation job to be stored.

4. The 'Annotate' button at the bottom of the input form should now be activated. Click it to start the annotation job.

5. View the completed annotation job from the Job List.

Test Verification

Verify correct output by comparing your job results (which will be a private genome) with the "Brucella suis 7121680725-test" validated job results in the public workspace:

- Click the 'VIEW' icon at the top right of your job results page and compare your private genome to the corresponding annotated Brucella suis 7121680725 genome in BV-BRC. Compare the counts for the following in the Overview tab. Results should be nearly identical (some variation may occur over time due to updates in the service or underlying data).
    - Annotation Statistics
    - Genome Quality
    - Genome Features
    - Protein Features
    - Specialty Genes
- Click the 'CDS' icon from the top right of your job results page and compare counts of predicted CDS with the "Brucella suis 7121680725-test" validated job results in the public workspace. Spot check results via keyword and filters.
    - CDS count
    - Filtered counts
- Double-click the GenomeReport.html in the list of files at the bottom of your job results page and compare results with the validated Genome Report.
    - Summary results
    - Problematic roles

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

Notes

- The Genome Annotation Service Tutorial provides detailed step-by-step instructions and screenshots for using the Genome Annotation Service.
- Exemplar results have been validated via comparison with output from the RASTtk computational service that is used to annotate all BV-BRC bacterial and archaeal genomes:
    - Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA 3rd, Stevens R, Vonstein V, Wattam AR, Xia F. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep. 2015 Feb 10;5:8365. doi: 10.1038/srep08365. PMID: 25666585; PMCID: PMC4322359.