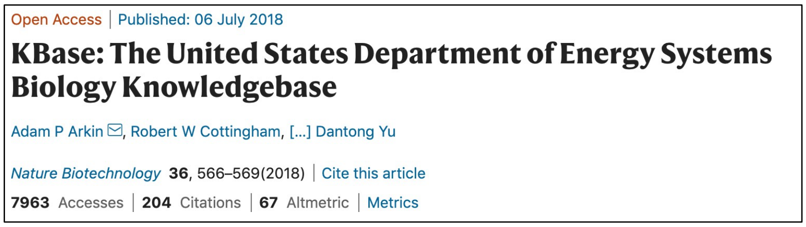


Open Access | Published: 06 July 2018

KBase: The United States Department of Energy Systems Biology Knowledgebase

Adam P Arkin, Robert W Cottingham, […] Dantong Yu

Nature Biotechnology **36**, 566–569(2018) | Cite this article

7963 Accesses | **204** Citations | **67** Altmetric | Metrics

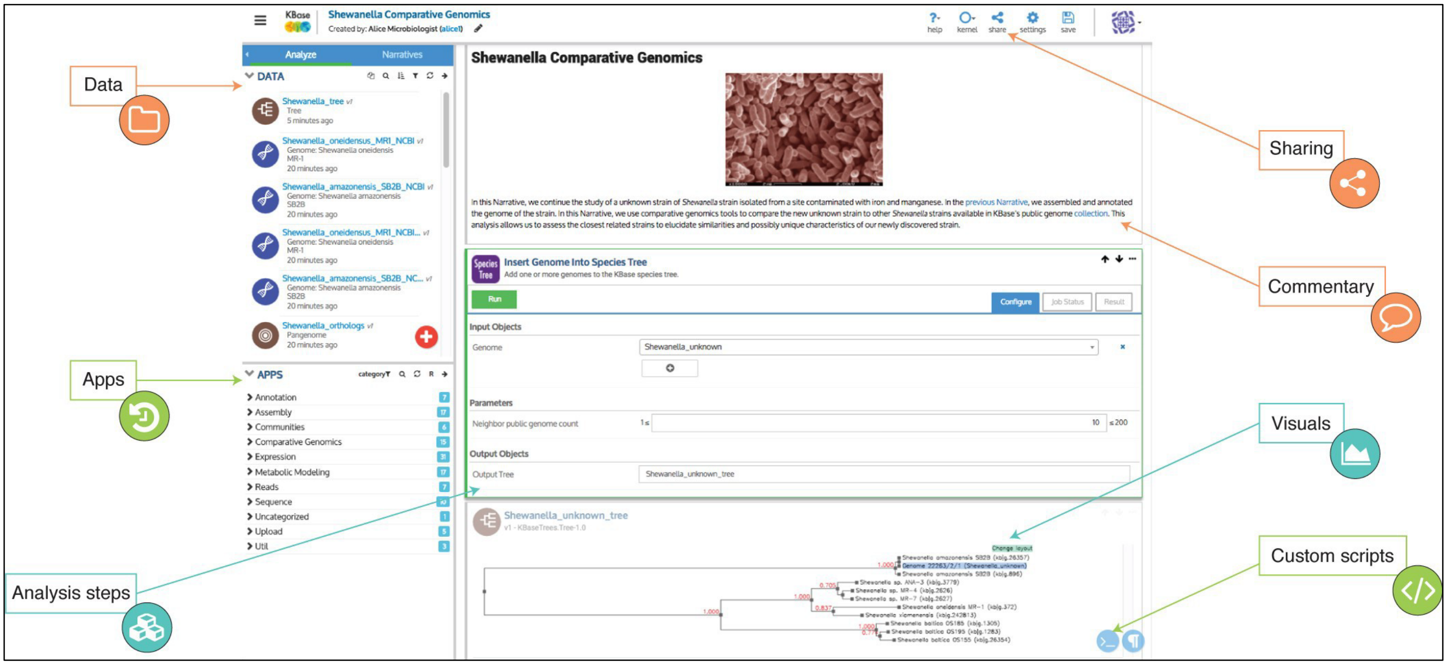
The United States Department of Energy Systems Biology Knowledgebase

What is KBase?

Open-source software and data platform with an graphical interface that enables data sharing and analysis of microbes and microbial communities…

To use KBase you can login/register through a google account.

Overall look of KBase:



Shewanella Comparative Genomics

In this Narrative, we continue the study of a unknown strain of *Shewanella* strain isolated from a site contaminated with iron and manganese. In the previous Narrative, we assembled and annotated the genome of the strain. In this Narrative, we use comparative genomics tools to compare the new unknown strain to other *Shewanella* strains available in KBase's public genome collection. This analysis allows us to assess the closest related strains to elucidate similarities and possibly unique characteristics of our newly discovered strain.

Insert Genome Into Species Tree

Add one or more genomes to the KBase species tree.

Input Objects

Genome: Shewanella_unknown

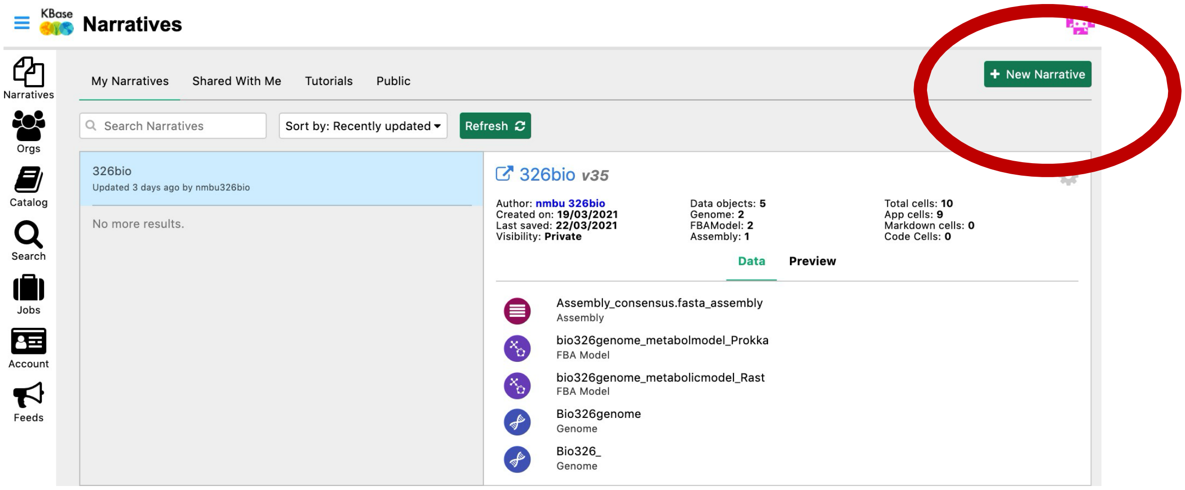
Parameters

Neighbor public genome count: 1 ≤ 10 ≤ 200

Output Objects

Output Tree: Shewanella_unknown_tree

When you first start, you will need to make a “Narrative”.
Narrative: is a “notebook” that holds your computational experiment.



326bio
Updated 3 days ago by nmbu326bio

No more results.

326bio *v35*

Author: **nmbu 326bio**
Created on: **19/03/2021**
Last saved: **22/03/2021**
Visibility: **Private**

Data objects: **5**
Genome: **2**
FBAModel: **2**
Assembly: **1**

Total cells: **10**
App cells: **9**
Markdown cells: **0**
Code Cells: **0**

Assembly_consensus.fasta_assembly
Assembly

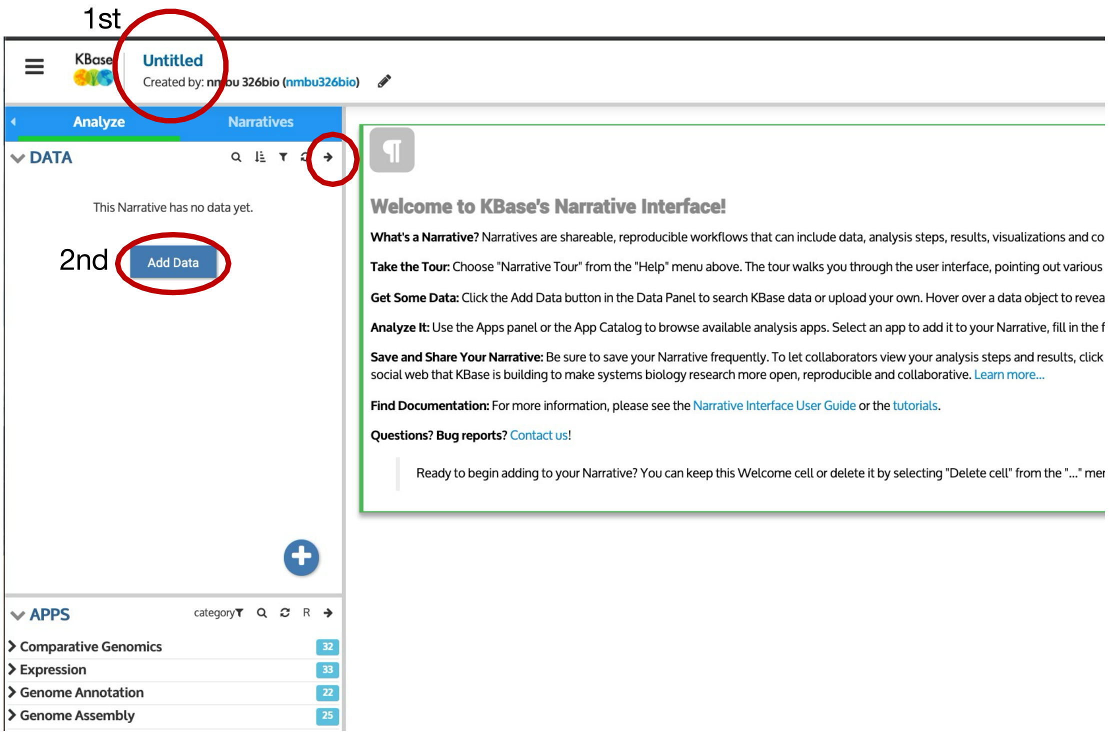
bio326genome_metabolmodel_Prokka
FBA Model

bio326genome_metabolicmodel_Rast
FBA Model

Bio326genome
Genome

Bio326_
Genome

It will load a new Narrative that is automatically called Untitled. First thing you should do is rename it, you can do this by clicking “Untitled”



Created by: nmbu 326bio (nmbu326bio)

This Narrative has no data yet.

Welcome to KBase's Narrative Interface!

What's a Narrative? Narratives are shareable, reproducible workflows that can include data, analysis steps, results, visualizations and co

Take the Tour: Choose "Narrative Tour" from the "Help" menu above. The tour walks you through the user interface, pointing out various

Get Some Data: Click the Add Data button in the Data Panel to search KBase data or upload your own. Hover over a data object to revea

Analyze It: Use the Apps panel or the App Catalog to browse available analysis apps. Select an app to add it to your Narrative, fill in the f

Save and Share Your Narrative: Be sure to save your Narrative frequently. To let collaborators view your analysis steps and results, click social web that KBase is building to make systems biology research more open, reproducible and collaborative. Learn more...

Find Documentation: For more information, please see the Narrative Interface User Guide or the tutorials.

Questions? Bug reports? Contact us!

> Ready to begin adding to your Narrative? You can keep this Welcome cell or delete it by selecting "Delete cell" from the "..." mer

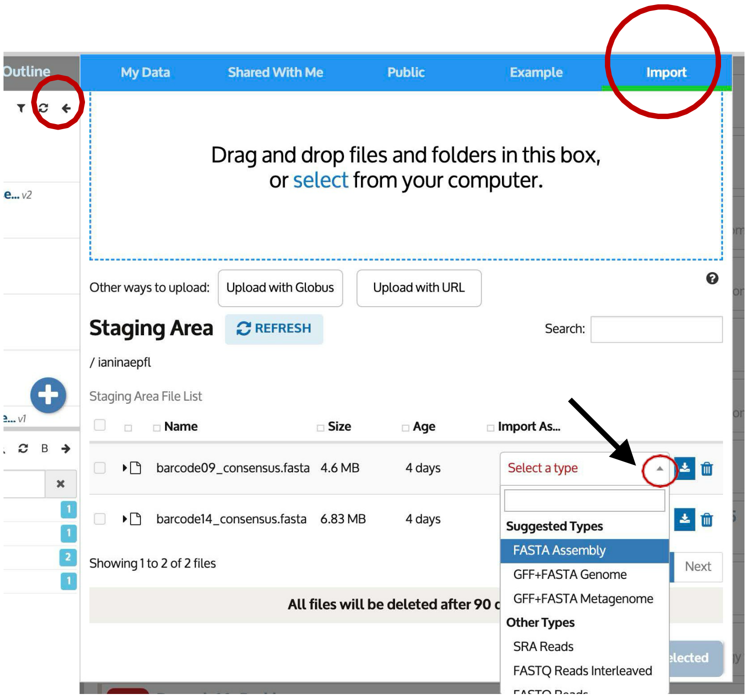
Comparative Genomics	32
Expression	33
Genome Annotation	22
Genome Assembly	25

Getting you data on KBase

You are directed to the import tab. In general, you need to first upload the files to your account, and then
import into your specific Narrative.

I have the *fasta* file “Assembly_consensus.fasta”. You will need to find a genome on NCBI. Don't worry the .fa and .fasta files are interchangeable

Once uploaded, the file will go into the “Staging Area”.



Drag and drop files and folders in this box,
or select from your computer.

Name	Size	Age	Import As...
barcode09_consensus.fasta	4.6 MB	4 days	Select a type
barcode14_consensus.fasta	6.83 MB	4 days	

Showing 1 to 2 of 2 files

All files will be deleted after 90 d

To import the assembly into your Narrative, you will need to scroll down to get the “FASTA Assembly” option and then hit the ‘ blue up arrow button’.

Once data is uploaded you can use the Apps

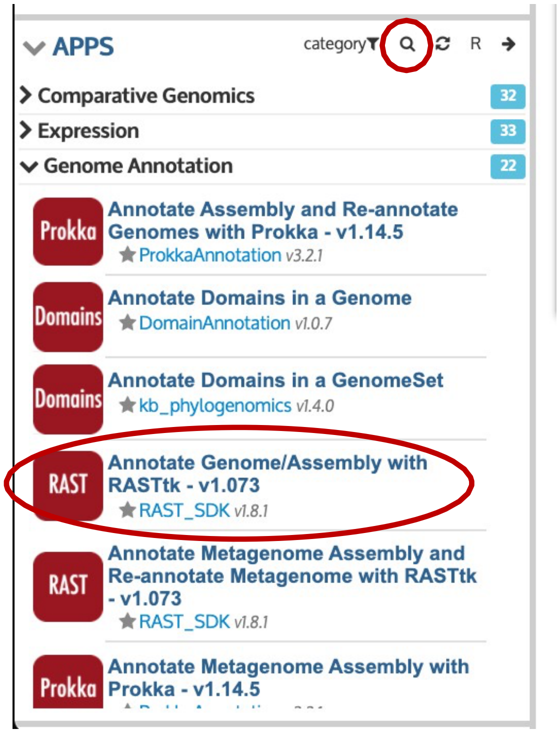
Lots of Apps including ones for:

Genome annotation: **RAST, Prokka**, **DRAM** you can check how the results differ when annotating with these different methods.

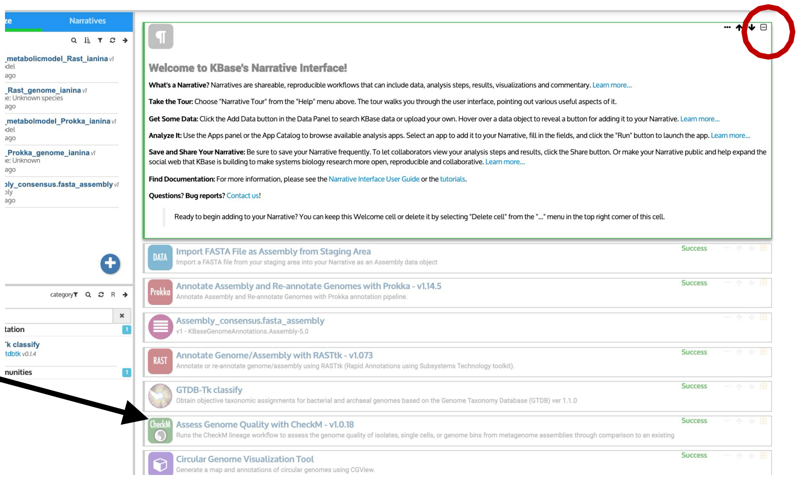
Genome assessment: **GTDB-tk, CheckM**

Metabolic Modeling: you can **Build Metabolic Model** with with the assembly/genome you annotated with either RAST or Prokka. Why not check how the different annotations effect the metabolic modelling results?

Also Apps for: (meta)genome assembly, Contig binning, Community modelling, etc……..



Every action/App performed or initiated by a user appears as a 'cell' in the Narrative. You can expand the cells by clicking on the + and - signs on the very right of the cell.



Welcome to KBase's Narrative Interface!

What's a Narrative? Narratives are shareable, reproducible workflows that can include data, analysis steps, results, visualizations and commentary. Learn more...

Take the Tour: Choose "Narrative Tour" from the "Help" menu above. The tour walks you through the user interface, pointing out various useful aspects of it.

Get Some Data: Click the Add Data button in the Data Panel to search KBase data or upload your own. Hover over a data object to reveal a button for adding it to your Narrative. Learn more...

Analyze It: Use the Apps panel or the App Catalog to browse available analysis apps. Select an app to add it to your Narrative, fill in the fields, and click the "Run" button to launch the app. Learn more...

Save and Share Your Narrative: Be sure to save your Narrative frequently. To let collaborators view your analysis steps and results, click the Share button. Or make your Narrative public and help expand the social web that KBase is building to make systems biology research more open, reproducible and collaborative. Learn more...

Find Documentation: For more information, please see the Narrative Interface User Guide or the tutorials.

Questions? Bug reports? Contact us!

> Ready to begin adding to your Narrative? You can keep this Welcome cell or delete it by selecting "Delete cell" from the "..." menu in the top right corner of this cell.

Import FASTA File as Assembly from Staging Area — Success
Import a FASTA file from your staging area into your Narrative as an Assembly data object

Annotate Assembly and Re-annotate Genomes with Prokka - v1.14.5 — Success
Annotate Assembly and Re-annotate Genomes with Prokka annotation pipeline.

Assembly_consensus.fasta_assembly
v1 - KBaseGenomeAnnotations.Assembly-5.0

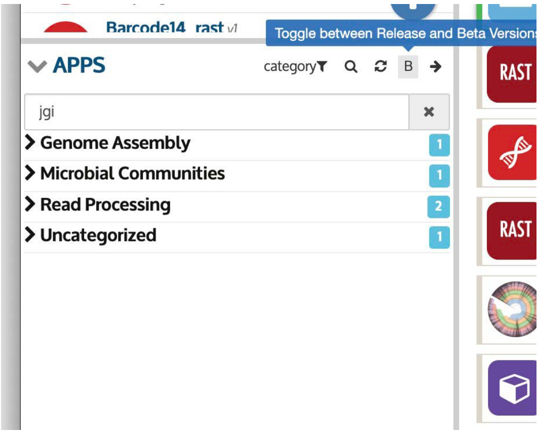
Annotate Genome/Assembly with RASTtk - v1.073 — Success
Annotate or re-annotate genome/assembly using RASTtk (Rapid Annotations using Subsystems Technology toolkit).

GTDB-Tk classify — Success
Obtain objective taxonomic assignments for bacterial and archaeal genomes based on the Genome Taxonomy Database (GTDB) ver 1.1.0

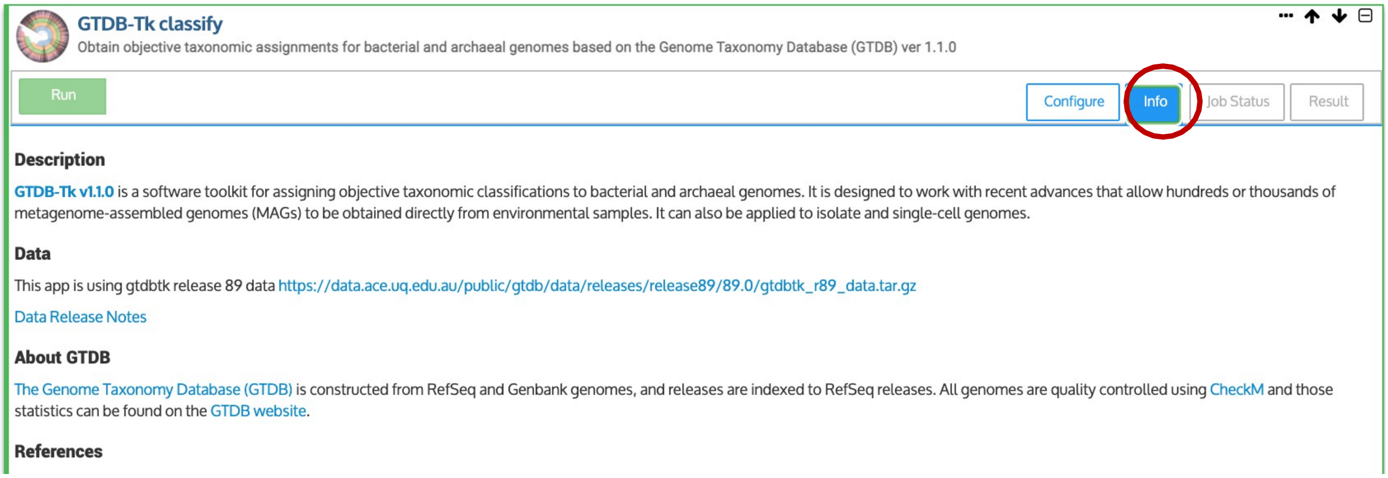
Assess Genome Quality with CheckM - v1.0.18 — Success
Runs the CheckM lineage workflow to assess the genome quality of isolates, single cells, or genome bins from metagenome assemblies through comparison to an existing

Circular Genome Visualization Tool — Success
Generate a map and annotations of circular genomes using CGView.

You can look for Apps in the bottom left corner of KBase. Once you click on an App it will appear as a cell in the Narrative. Once in the cell, each App will have several tabs on it. You can click the "Info" tab to see what the App does and have access to additional references. The "Configure" tab is where you can select what data/file you want to use the App with. In your case it would likely be your imported assembly file or the annotated genome. Sometimes you can also select other options in the "Configure", I recommend reading about the App and just trying a few things. Not all Apps will work with all data types. Click "Run" button to start the analysis.



Switch to be able to see the apps that are still in Beta mode.



GTDB-Tk classify
Obtain objective taxonomic assignments for bacterial and archaeal genomes based on the Genome Taxonomy Database (GTDB) ver 1.1.0

Description

GTDB-Tk v1.1.0 is a software toolkit for assigning objective taxonomic classifications to bacterial and archaeal genomes. It is designed to work with recent advances that allow hundreds or thousands of metagenome-assembled genomes (MAGs) to be obtained directly from environmental samples. It can also be applied to isolate and single-cell genomes.

Data

This app is using gtdbtk release 89 data https://data.ace.uq.edu.au/public/gtdb/data/releases/release89/89.0/gtdbtk_r89_data.tar.gz

Data Release Notes

About GTDB

The Genome Taxonomy Database (GTDB) is constructed from RefSeq and Genbank genomes, and releases are indexed to RefSeq releases. All genomes are quality controlled using CheckM and those statistics can be found on the GTDB website.

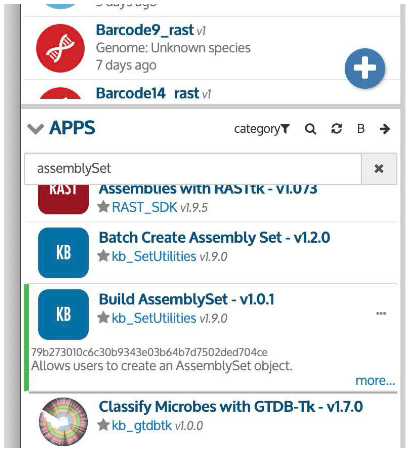
References

Results:

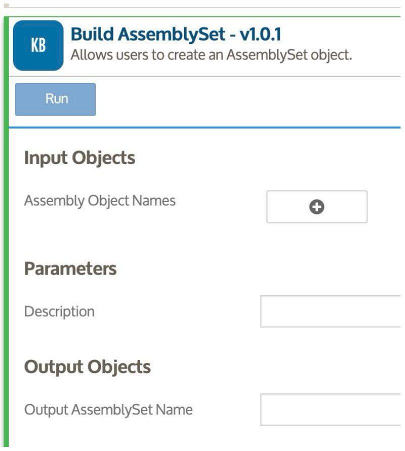
Once the App is finished running, the results are either in the cell under the “Results” tab of the App you ran, or they are created as objects in the ‘Data’ panels. If you object is in the Data panel you can click on it to explore the data in the cell, or on the ‘binocular’ looking button to <u>explore</u> the results in a separate window. Try checking what the other buttons do.

This is example of output from the GTDB app that can taxonomically classify your organism.

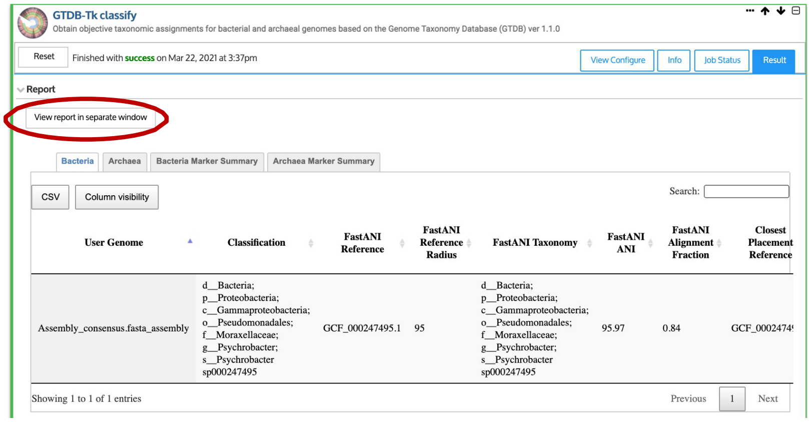
Note: To run GTDB you need to first convert your assembly/genome into an AssemblySet.



Allows users to create an AssemblySet object.



Allows users to create an AssemblySet object.



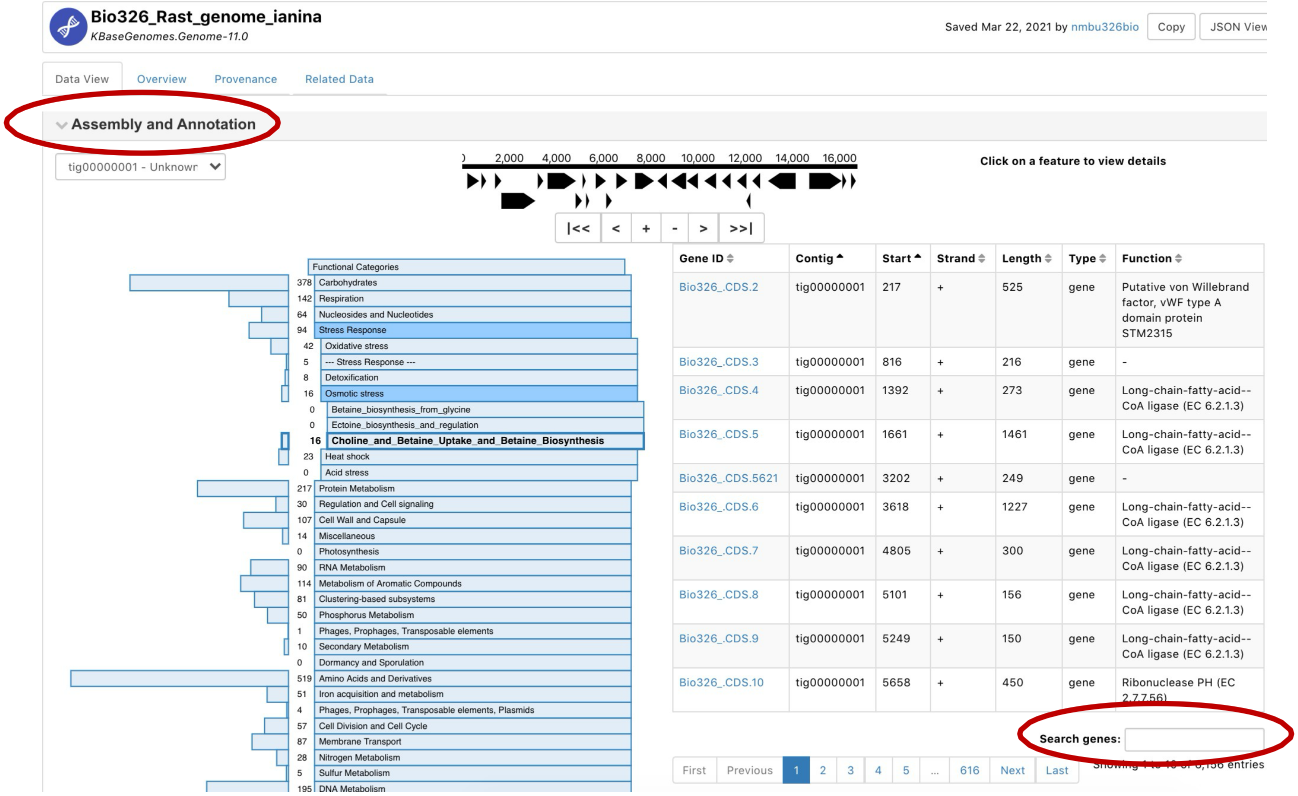
Obtain objective taxonomic assignments for bacterial and archaeal genomes based on the Genome Taxonomy Database (GTDB) ver 1.1.0

Finished with success on Mar 22, 2021 at 3:37pm

User Genome	Classification	FastANI Reference	FastANI Reference Radius	FastANI Taxonomy	FastANI ANI	FastANI Alignment Fraction	Closest Placement Reference
Assembly_consensus.fasta_assembly	d__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Moraxellaceae; g__Psychrobacter; s__Psychrobacter sp000247495	GCF_000247495.1	95	d__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Moraxellaceae; g__Psychrobacter; s__Psychrobacter sp000247495	95.97	0.84	GCF_0002474

Showing 1 to 1 of 1 entries

Here is example output from annotation of the assembly/genome with RAST. You need to scroll down to get to the Assembly and Annotation part.



Bio326_Rast_genome_ianina

KBaseGenomes.Genome-11.0

Saved Mar 22, 2021 by nmbu326bio

Assembly and Annotation

Click on a feature to view details

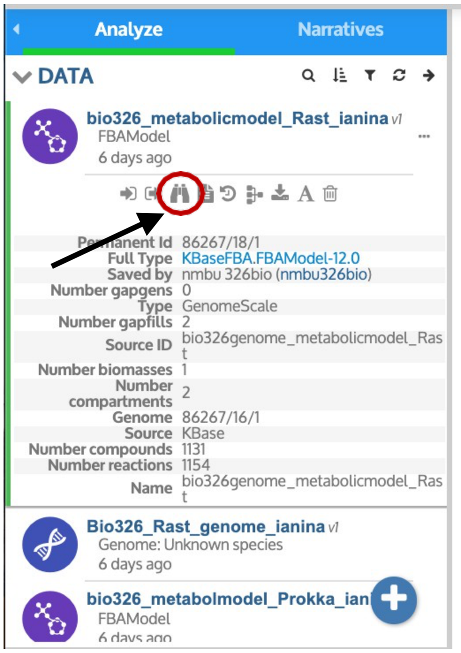
	Functional Categories
378	Carbohydrates
142	Respiration
64	Nucleosides and Nucleotides
94	Stress Response
42	Oxidative stress
5	--- Stress Response ---
8	Detoxification
16	Osmotic stress
0	Betaine_biosynthesis_from_glycine
0	Ectoine_biosynthesis_and_regulation
16	**Choline_and_Betaine_Uptake_and_Betaine_Biosynthesis**
23	Heat shock
0	Acid stress
217	Protein Metabolism
30	Regulation and Cell signaling
107	Cell Wall and Capsule
14	Miscellaneous
0	Photosynthesis
90	RNA Metabolism
114	Metabolism of Aromatic Compounds
81	Clustering-based subsystems
50	Phosphorus Metabolism
1	Phages, Prophages, Transposable elements
10	Secondary Metabolism
0	Dormancy and Sporulation
519	Amino Acids and Derivatives
51	Iron acquisition and metabolism
4	Phages, Prophages, Transposable elements, Plasmids
57	Cell Division and Cell Cycle
87	Membrane Transport
28	Nitrogen Metabolism
5	Sulfur Metabolism
195	DNA Metabolism

Gene ID	Contig	Start	Strand	Length	Type	Function
Bio326_.CDS.2	tig00000001	217	+	525	gene	Putative von Willebrand factor, vWF type A domain protein STM2315
Bio326_.CDS.3	tig00000001	816	+	216	gene	-
Bio326_.CDS.4	tig00000001	1392	+	273	gene	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
Bio326_.CDS.5	tig00000001	1661	+	1461	gene	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
Bio326_.CDS.5621	tig00000001	3202	+	249	gene	-
Bio326_.CDS.6	tig00000001	3618	+	1227	gene	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
Bio326_.CDS.7	tig00000001	4805	+	300	gene	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
Bio326_.CDS.8	tig00000001	5101	+	156	gene	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
Bio326_.CDS.9	tig00000001	5249	+	150	gene	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
Bio326_.CDS.10	tig00000001	5658	+	450	gene	Ribonuclease PH (EC 2.7.7.56)

Search genes:

Showing 1 to 10 of 6,156 entries

Here is example output from the Build Metabolic Model app using the assembly/ genome that was annotated with RAST as an input for this app.

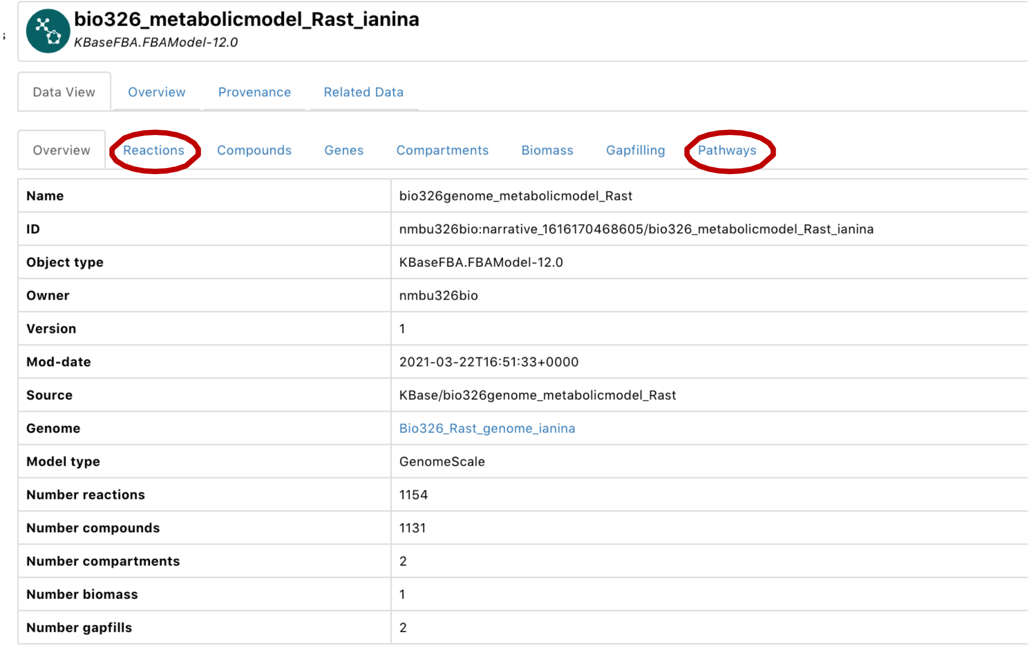


bio326_metabolicmodel_Rast_ianina v1
FBAModel
6 days ago

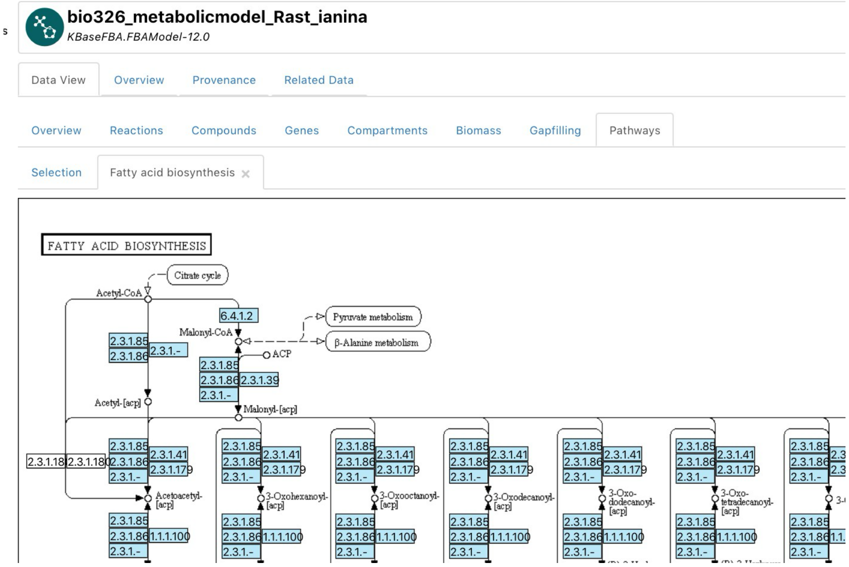
Permanent Id 86267/18/1
Full Type KBaseFBA.FBAModel-12.0
Saved by nmbu 326bio (nmbu326bio)
Number gapgens 0
Type GenomeScale
Number gapfills 2
Source ID bio326genome_metabolicmodel_Rast
Number biomasses 1
Number compartments 2
Genome 86267/16/1
Source KBase
Number compounds 1131
Number reactions 1154
Name bio326genome_metabolicmodel_Rast

Bio326_Rast_genome_ianina v1
Genome: Unknown species
6 days ago

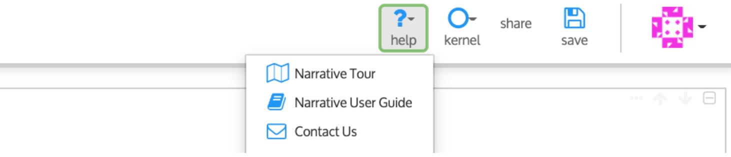
bio326_metabolmodel_Prokka_ian
FBAModel
6 days ago

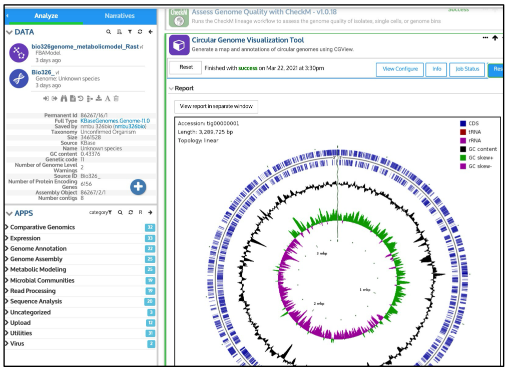


Name	bio326genome_metabolicmodel_Rast
ID	nmbu326bio:narrative_1616170468605/bio326_metabolicmodel_Rast_ianina
Object type	KBaseFBA.FBAModel-12.0
Owner	nmbu326bio
Version	1
Mod-date	2021-03-22T16:51:33+0000
Source	KBase/bio326genome_metabolicmodel_Rast
Genome	Bio326_Rast_genome_ianina
Model type	GenomeScale
Number reactions	1154
Number compounds	1131
Number compartments	2
Number biomass	1
Number gapfills	2



If you get stuck/lost KBase has a useful help section.





Have fun **exploring the Apps** with your data for the report!