

# Metagenomes in IMG

- [https://gold.jgi.doe.gov/project\\_submission](https://gold.jgi.doe.gov/project_submission)

The screenshot shows a web browser window with the URL [gold.jgi.doe.gov/project\\_submission](https://gold.jgi.doe.gov/project_submission). The page is titled "JGI GOLD GENOMES ONLINE DATABASE". The top navigation bar includes links for Home, Search, Distribution Graphs, Biogeographical Metadata, SRA Explorer, Statistics, GOLD Usage Policy, Team, Help, News, Downloads, My Profile, and My Past Searches. A sidebar on the left displays statistics for various data types: Studies (45,769), Biosamples (101,560), Sequencing Projects (355,804), Analysis Projects (283,762), and Organisms (387,468). The main content area features a blue header bar with the text "Create Sequencing Projects, Analysis Projects, Studies and Biosamples". Below this, a red note states: "Please note: As of Feb 12, 2019 we are only accepting FASTA files (not Genbank formatted files) for processing when submitting data for analysis (see the [IMG FAQ page](#) for more information)." It also provides links to the "GOLD Project Entry Help Document" and the "GOLD Standardized Metagenome Naming Document". At the bottom of the content area, there are two blue buttons: "Create a new Sequencing Project in GOLD" and "Create a new Analysis Project for submission to IMG". A red arrow points to the second button. On the far right of the page, a welcome message reads "Welcome, Julie Huber".

# From GOLD to IMG/MER- 9/3/2020 mtg

- Rename analysis project name
- Study should come up with Axial
- Put in Assembly Method, etc
- Press SUBMIT, will get an analysis project number
- Then go to IMG/MER, Submit Data Set
- New Submission: “AP ID” is the number you got from GOLD
- Need to specify gene calling, mixed microbial community
- “Submit Dataset to IMG”
- Need to resolve lack of access to Axial study in GOLD

**The following worked on a test sample assembly  
(FS917) for Julie, 9/28/2020**

28 Google Calendar - Week of Septe ● JGI GOLD | Sequencing Project S ● +

gold.jgi.doe.gov/project\_submission

Apps MIT Library Load WHOI Int Web Calendar PO Orders Unanet My Website WHOI WiFi JP Morgan COVID Greenfield Others CDI Other bookmarks

Welcome, Julie Huber

**JGI** **GOLD**  
GENOMES ONLINE DATABASE

Home Search Distribution Graphs Biogeographical Metadata SRA Explorer Statistics GOLD Usage Policy Team Help News Downloads My Profile My Past Searches

Studies 45,791  
Biosamples 101,901  
Sequencing Projects 357,218  
Analysis Projects 284,645  
Organisms 387,971

Create Sequencing Projects, Analysis Projects, Studies and Biosamples

Please note: As of Feb 12, 2019 we are only accepting FASTA files (not Genbank formatted files) for processing when submitting data for analysis (see the [IMG FAQ page](#) for more information).

For information on the new structure of GOLD please [review the GOLD Project Entry Help Document](#)

For Metagenome naming help, please refer to: [GOLD Standardized Metagenome Naming Document](#)

My Data:

Studies 6  
Biosamples 147  
Sequencing Projects 153  
Analysis Projects 180

▶ Create a new Sequencing Project in GOLD

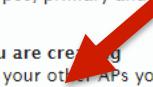
▼ Create a new Analysis Project for submission to IMG

For more information about Analysis Projects please review the [Gold Analysis Projects SOP](#)

There are two Analysis Project submission types, primary and reanalysis. Your first submission for any Sequencing Project will default to Primary. You can always update this later, [read more...](#)

1. Select the type of Analysis Project you are creating  
(depending upon the processing status of your other APs you may not be able to submit any new APs at this time):

[Next Step >>](#)



[Disclaimer](#)  
[Credits](#)  
[Contact](#)  
[Accessibility / Section 508 Statement](#)

©1997-2020 The Regents of the University of California



28 Google Calendar - Week of Septe x JGI GOLD | Select Project x +

gold.jgi.doe.gov/appprocessing

Apps MIT Library Load WHOI Int Web Calendar PO Orders Unanet My Website WHOI WiFi JP Morgan COVID Greenfield Others CDI Other bookmarks

Welcome, Julie Huber

JGI GOLD GENOMES ONLINE DATABASE

Home Search Distribution Graphs Biogeographical Metadata SRA Explorer Statistics GOLD Usage Policy Team Help News Downloads My Profile My Past Searches

Studies 45,791 Biosamples 101,901 Sequencing Projects 357,218 Analysis Projects 284,645 Organisms 387,971

My Data:

Studies 6 Biosamples 147 Sequencing Projects 153 Analysis Projects 180

Select Parent Project

You are creating a new Metagenome Analysis Project

(Note: A metagenome Analysis Project can only accept sequence data from shotgun sequencing using DNA as starting material. Do not pass on 16s surveys or iTags as metagenomes. Violators may risk losing GOLD/IMG privileges to submit further data sets.)

Search for a Public Sequencing Project (you will be asked to assign a study on the next screen):

OR

Select one of your existing Sequencing Projects:

Select a sequencing project...

Gp0106823 - Marine hydrothermal diffuse flow vent microbial communities from the Mid-Cayman Rise, Caribbean Sea - Sample FS856  
Gp0107592 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS821\_Marshmallow\_DNA  
Gp0107594 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS828\_N3Area\_DNA  
Gp0107596 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS857\_Anemone\_DNA  
Gp0107598 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS861\_Marker13\_DNA  
Gp0107619 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS891\_Anemone\_DNA  
Gp0107621 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS896\_ElGuapo\_DNA  
Gp0107623 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS898\_N3Area\_DNA  
Gp0107625 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS900\_Dependable\_DNA  
Gp0107627 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS903\_Marker13\_DNA  
A Gp0107629 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS904\_Marker33\_DNA  
Gp0119879 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS906\_Marker13\_DNA  
Gp0119880 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS907\_Anemone\_DNA  
Gp0119881 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS908\_Marker33\_DNA  
Gp0121531 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample CTD1200\_DNA  
Gp0128098 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS915\_Marker13\_DNA  
Gp0128096 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS914\_Anemone\_DNA  
Gp0128110 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS917\_Marker33\_DNA  
Gp0128117 - Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS918\_NRZ\_DNA



Disclaimer Credits Contact Accessibility / Section 508 Statement

©1997-2020 The Regents of the University of California

U.S. DEPARTMENT OF ENERGY Office of Science

BERKELEY LAB

Type here to search 100% 8:33 AM 9/28/2020

**JGI GOLD GENOMES ONLINE DATABASE**

Home Search Distribution Graphs Biogeographical Metadata SRA Explorer Statistics GOLD Usage Policy Team Help News Downloads My Profile My Past Searches

Studies 45,791  
Biosamples 101,901  
Sequencing Projects 357,218  
Analysis Projects 284,645  
Organisms 387,971

My Data:  
Studies 9  
Biosamples 147  
Sequencing Projects 153  
Analysis Projects 180

**Enter and Submit Analysis Project Details**

You are creating a new: Metagenome Analysis Project

Please fill in the form below to create a new Analysis Project.

The Sequencing Project for this AP is: Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS917\_Marker33\_DNA

Analysis Project Type: Metagenome Analysis Reanalysis Project

**ANALYSIS PROJECT DESCRIPTION**

Analysis Project Name: Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS917\_Marker33\_DNA  
Suffix: SE\_2020\_test

Study: Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean

Analysis Project Description \*: SE\_2020\_IDBA\_UD\_test

Public Visibility: No

**ANALYSIS PROJECT METADATA**

Assembly Method \*: IDBA\_UD v 1.1.3

PI Name: Caroline Fortunato

PI Email: cfortunato@gmail.com

Scaffold Count:

Contig Count:

Estimated Size:

**EXTERNAL REFERENCES**

Genbank: e.g. CP00045, AWDB00000000

Assembly Accession: e.g. GCA\_000183345.1

This didn't work at first- I added a period after v and it did: "IDBA\_UD v. 1.1.3 and then it did work

[Disclaimer](#)  
[Credits](#)  
[Contact](#)  
[Accessibility / Section 508 Statement](#)

©1997-2020 The Regents of the University of California





JGI HOME LOG OUT

Welcome, Julie Huber

Home Search Distribution Graphs Biogeographical Metadata SRA Explorer Statistics GOLD Usage Policy Team Help News Downloads My Profile My Past Searches

Studies	<a href="#">45,791</a>
Biosamples	<a href="#">101,901</a>
Sequencing Projects	<a href="#">357,218</a>
Analysis Projects	<a href="#">284,646</a>
Organisms	<a href="#">387,971</a>

Processing Complete

Your Analysis Project was created in GOLD.

You will need the following Analysis Project id when submitting your data set to IMG: Ga0454710

[Follow this link to access the IMG submission site.](#)

[Click here to return and start another Analysis Project](#)

Write this down and save in spreadsheet format or something

Then I clicked here



My Data:

Studies	<a href="#">6</a>
Biosamples	<a href="#">147</a>
Sequencing Projects	<a href="#">153</a>
Analysis Projects	<a href="#">181</a>

[Disclaimer](#)  
[Credits](#)  
[Contact](#)  
[Accessibility / Section 508 Statement](#)

©1997-2020 The Regents of the University of California



U.S. DEPARTMENT OF  
**ENERGY** | Office of  
Science

← → C [img.jgi.doe.gov/cgi-bin/submit/main.cgi?apld=Ga0454710](https://img.jgi.doe.gov/cgi-bin/submit/main.cgi?apld=Ga0454710)

Apps MIT Library Load WHOI Int Web Calendar PO Orders Unanet My Website WHOI WiFi JP Morgan COVID Greenfield Others CDI Other bookmarks

**JGI** EXPERT REVIEW DATA SUBMISSION

Set Cookie: CGISESSID=submit=51bde0115be067f33db8f3e80186a1de; path=/ Set Cookie: jgi\_return=https%3A%2F%2Fimg.jgi.doe.gov%2Fcgi-bin%2Fsubmit%2Fmain.cgi; domain=jgi.doe.gov; path=/ Date: Mon, 28 Sep 2020 12:43:22 GMT Content-Type: text/html; charset=ISO-8859-1

Submission Home Submitted Datasets New Submission Filter Statistics FAQ

New Submission

**hint:** As per the new four level classification system implemented by GOLD for genome and metagenome projects, all future submissions to IMG will be based on GOLD Analysis Projects. All IMG submissions now require an Analysis Project in GOLD. Please go to [GOLD](#) to define an Analysis Project. Refer to the GOLD's latest publication (<http://nar.oxfordjournals.org/content/early/2014/10/27/nar.gku950.full>) describing the four level classification system and/or the help document in how to define a new project and obtain analysis project id from GOLD.

AP ID:  This was already here

[Accessibility/Section 508](#)

Type here to search

100% 8:44 AM 9/28/2020

28 Google Calendar - Week of Septe X | JGI GOLD | Completion Message X | Expert Review Submission X | IDBA-UD X | IDBA-Bioinformatics Research Gro X | Release Release 1.1.3 · loneknight X | +

← → C img.jgi.doe.gov/cgi-bin/submit/main.cgi?section=Submission

Apps MIT Library Load WHOI Int Web 28 Calendar PO Orders Unanet My Website WHOI WiFi JP Morgan COVID Greenfield Others CDI Other bookmarks

**JGI** **IMG/ER & MER**  EXPERT REVIEW DATA SUBMISSION

Submission Home Submitted Datasets New Submission Filter Statistics FAQ

## Submit GOLD Analysis Project

GOLD AP ID	Ga0454710
Analysis Project Name	Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS917_Marker33_DNA
Analysis Project Type	Metagenome Analysis
Genome Type	metagenome
Domain	

Submit Analysis Project 

This was already here



[Submission Home](#)

## Submitted Datasets

New Submission

## Filter

## statistics

1

## IMG Expert Review Submission Agreement

Submission of genome datasets to IMG ER or metagenome datasets to IMG/M ER complies with the following conditions

## Metadata Requirements

Genome and metagenome datasets submitted for annotation and/or integration in IMG ER and IMG/M ER need to have a descriptive name- acronyms are not accepted. The following fields are **mandatory**: *project name; latitude; longitude; isolation (collection location); country; collection time; habitat; sequencing method; assembly; gene calling method; sequencing center*.

Submissions that are missing mandatory metadata fields will be **cancelled** without processing.

## Data Release & Replacement

Genome and metagenome datasets submitted for annotation and/or integration in IMG ER and IMG/M ER will be kept private (not available publicly) for 18 months for isolate genomes and 24 months for single cells and metagenomes from the date they become available for analysis, after which they will become automatically public.

For multiple submissions of the same genome or metagenome sample, the multiple versions will be kept into the system for no more than 6 months, after which only the latest version will be kept in IMG, with previous (older) versions automatically removed.

A genome or metagenome dataset submitted to IMG can be replaced by newer versions of the same genome/metagenome dataset, but cannot be removed in order to avoid making them public.

Note that **private** datasets in IMG are accessed by IMG team members mainly for data quality and validation purposes, and occasionally for data mining. Data validation and data mining results will not be used or published without agreement from the dataset owners as long as the datasets remain **private**.

## Data Distribution

Genome and metagenome datasets submitted for annotation and/or integration in IMG ER and IMG/M ER are distributed solely through individual genome and metagenome data portals and are limited to assembled and annotated datasets; no other type of data distribution (data downloads) is provided.

IMG is not depositing raw/assembled/annotated datasets into Genbank

I Agree





Submission Home Submitted Datasets New Submission Filter Statistics FAQ

## New Metagenome Dataset Submission

Submit metagenome for inclusion into IMG/M ER system.

All fields marked with (\*) are required fields.

[Submission Information](#) [Submit sequence file](#)



I did not change anything here, then clicked on "Submit Sequence File"

**hint:** Please provide general information about your submitted project.

Submission project	Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample FS917_Marker33_DNA
ER submission sample	Sample Name
GOLD ID	Ga0454710
Target ER System (*)	IMG/M ER
Is public in IMG database? (*)	No
Sequencing Status (*)	Draft
Replacing Taxon OID	<input type="text"/>

Go to 'Submit sequence file' section to submit you [fasta](#) format file(s).

[Submit](#) [Cancel](#) [Reset](#)

[Logout](#)

[Accessibility/Section 508](#)



Submission Home Submitted Datasets New Submission Filter Statistics FAQ

## New Metagenome Dataset Submission

Submit metagenome for inclusion into IMG/M ER system.

All fields marked with (\*) are required fields.

[Submission Information](#) [Submit sequence file](#)

**hint:** Please submit a [fasta](#) format file. Compressed files (.gz or .zip) are also accepted.

Assembled sequences 	<input type="button" value="Choose File"/> contig.fa
Sequence coverage file for assembled sequences 	<input type="button" value="Choose File"/> No file chosen
Topology (*) 	<input type="button" value="linear"/>

I uploaded the contig file, didn't do anything else, and pressed submit.  
Took about 5ish minutes and then next screen appeared

[Accessibility/Section 508](#)

← → C img.jgi.doe.gov/cgi-bin/submit/main.cgi?section=Submission

Apps MIT Library Load WHOI Int Web Calendar PO Orders Unanet My Website WHOI WiFi JP Morgan COVID Greenfield Others CDI Other bookmarks

**JGI**  **IMG/ER & MER**   
EXPERT REVIEW DATA SUBMISSION

Submission Home Submitted Datasets New Submission Filter Statistics FAQ

## Check Submission

New submission is added to the database (submission ID: 240800)

Use gene calling method: Metagenome Gene Calling  
assembled\_seq\_file: contig.fa

OK  I guess this means it was accepted

---

Logout

[Accessibility/Section 508](#)

Type here to search

8:58 AM 9/28/2020



Submission Home Submitted Datasets New Submission Filter Statistics FAQ

You log in as User jhuber. You can switch 'Data Type' to see different types of submissions.

Data Type: Metagenome

## Metagenome Submissions

Selected Submission Count: 57 (Total: 57)

Note: Check Status only applies to the first selected submission.)

Filter column: Submission ID Filter text:  Apply ?

Export Page 1 of 1 << first < prev 1 next > last >> All

Column Selector Select Page Deselect Page

Select	Submission ID	Analysis Project ID	Analysis Project Type	Submission Type	GOLD Review Status	Project/Study Name	Seq Status	Sequencing Project ID	IMG Taxon ID	Analysis Product Name	IMG Contact Name(s)	Submitter	Seq Center	Submission Date	St
<input type="checkbox"/>	240800	Ga0454710	Metagenome Analysis	Reanalysis	In Review	Diffuse hydrothermal flow volcanic vent microbial communities from Axial Seamount, northeast Pacific ocean - Sample ES917, Marker33, DNA	Draft	Gp0128110		jhuber	-	jhuber	-	2020-09-28	Sub Pen ann
<input type="checkbox"/>	91321	Ga0132240	Metatranscriptome Analysis	Primary	Approved	Aquatic microbial communities from different depth of meromictic Siders Pond, Falmouth, Massachusetts; Cast 7, 12m depth; RNA IDBA-UD	Draft	Gp0151168	3300009941	jhuber	-	ereddingtonSSOAPI	Marine Biological Laboratory	2016-06-30	Finis with nam assi
<input type="checkbox"/>	91320	Ga0132239	Metatranscriptome Analysis	Primary	Approved	Aquatic microbial communities from different depth of meromictic Siders Pond, Falmouth, Massachusetts; Cast 7, 8m depth; RNA IDBA-UD	Draft	Gp0151167	3300009917	jhuber	-	ereddingtonSSOAPI	Marine Biological Laboratory	2016-06-30	Finis with nam assi

AND THERE IT IS!

