Jun 13, 2025

🌐 NCBI_rRNA_Submission

DOI

dx.doi.org/10.17504/protocols.io.dm6gprqppvzp/v1

Avery S Hiley[1]

[1]UCSD - Scripps Institution of Oceanography

Rouse Lab

Avery S Hiley
UCSD - Scripps Institution of Oceanography

DOI: dx.doi.org/10.17504/protocols.io.dm6gprqppvzp/v1

Protocol Citation: Avery S Hiley 2025. NCBI_rRNA_Submission. **protocols.io**
https://dx.doi.org/10.17504/protocols.io.dm6gprqppvzp/v1

Protocol status: Working
We use this protocol and it's working

Created: May 09, 2020

Last Modified: June 13, 2025

Protocol Integer ID: 36813

Abstract

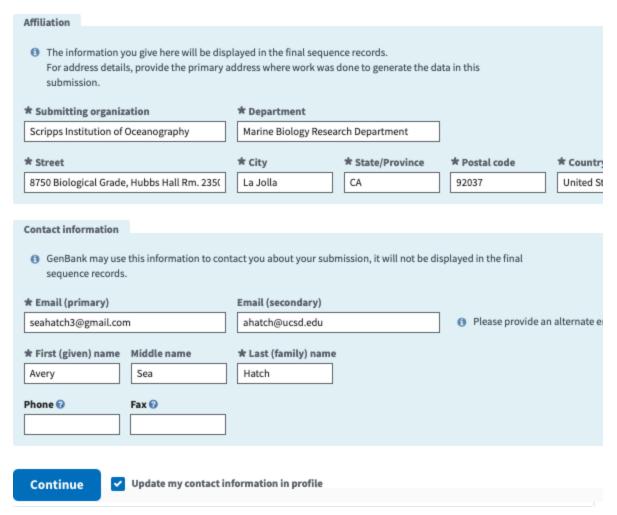A short tutorial on uploading rRNA sequences to genBank.

1   First and foremost, gather your rRNA (e.g. mitochondrial 16S, 12S and nuclear 18S, 28S) sequences that need to be uploaded and align them in either Geneious, Mesquite, or using the online MAFFT version 7 server: https://mafft.cbrc.jp/alignment/server/.

    *Note: Each rRNA gene must be aligned and submitted separately.

2   Blast one of the sequences and make sure that the first record it matches with on NCBI says 'Plus/Plus'. This indicates that the sequence is in the correct forward 5′ direction. If it says 'Plus/Minus', then reverse complement your entire alignment.

3   Important note: It is normal for a rRNA gene alignment to have gaps in it, so don't worry if this is the case.

    However, please make sure there are no ambiguities in your sequences. If a sequence has lots of bases that aren't ACGT (e.g. N, K, M, other IUPAC codes, etc.):

3.1 Please look at your original De Novo assembly in Geneious and check if you can make a nucleotide base call. Compare the amplitudes of the Forward/Reverse peaks at the corresponding nucleotide positions in the chromatogram (.ab1 files), and choose the nucleotide with the greater amplitude.

3.2 Check if the ends of the sequence assembly need to be trimmed due to low/poor quality.

4   Make any nucleotide edits and/or deletions as necessary before proceeding.

5   Once the rRNA sequences do not have ambiguities, export the final alignment in Mesquite as a fasta file. Do not 'include gaps'!

6   Open this fasta file in TextWrangler or BBEdit. Edit all of your sequence headers to match this format:

    >A9919 [organism=Peinaleopolynoe mineoi] Peinaleopolynoe mineoi voucher SIO:BIC:A9919 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial

6.1 In this example, "A9919" is the BIC number; use either the BIC number or a different institution's catalog number here to represent the sequence ID. If catalog numbers are not assigned yet, you may use a lab code (e.g. S25170) as the sequence ID here.

6.2 "[organism=Peinaleopolynoe mineoi]" is obviously where you insert the species name for the corresponding sequence ID.

6.3 "Peinaleopolynoe mineoi": Repeat the species name here.

6.4 "voucher SIO:BIC:A9919": Follow this proper format to identify the corresponding BIC number with our institution abbreviations. If the specimen is deposited at a different institution, this will instead mimic the abbreviations of that place before listing the catalog number (e.g. MNHN for Muséum national d'Histoire naturelle). If your catalog numbers aren't assigned yet, exclude this section altogether.

6.5 "cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial" should be listed at the end of the aforementioned information.

7 After all of your sequence headers follow this format, save the fasta file again.

8 Create an online account for NCBI's **Submission Portal** platform: **https://submit.ncbi.nlm.nih.gov/**.

9 Start a new submission.

10 **Submission Type**: 'Metazoan (multicellular animal) Mitochondrial COX1'.

11 **Submitter**: Fill out the corresponding information (see screenshot below for our lab details). Make sure to check 'Update my contact information in profile' in order for future submissions to use this specific info by default.

Submitter

Affiliation

ⓘ The information you give here will be displayed in the final sequence records.
For address details, provide the primary address where work was done to generate the data in this submission.

★ Submitting organization
Scripps Institution of Oceanography

★ Department
Marine Biology Research Department

★ Street
8750 Biological Grade, Hubbs Hall Rm. 235(

★ City
La Jolla

★ State/Province
CA

★ Postal code
92037

★ Country
United St

Contact information

ⓘ GenBank may use this information to contact you about your submission, it will not be displayed in the final sequence records.

★ Email (primary)
seahatch3@gmail.com

Email (secondary)
ahatch@ucsd.edu

ⓘ Please provide an alternate e

★ First (given) name
Avery

Middle name
Sea

★ Last (family) name
Hatch

Phone ❓

Fax ❓

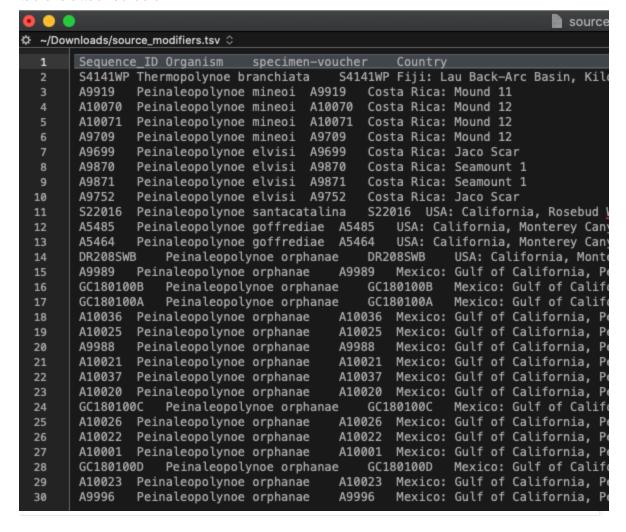[Continue] ☑ Update my contact information in profile

12 **Sequencing Technology**: Select 'Sanger dideoxy sequencing' & 'Assembled sequences (each sequence was assembled from two or more overlapping sequence reads)'.

13 **Sequences**: Select 'Release on specified date or upon publication, whichever is first' or release immediately if you are late in uploading your sequences (this should almost never be the case). Typically you will want to choose a year in advance to be safe. Next, upload your COI fasta file that was completed in step 16.

14 **Source Info**: Under 'Do your sequence IDs represent one of these?', select 'Specimen-Voucher' if you followed the headers format in step 15. If catalog numbers aren't assigned yet and your sequence IDs were listed as lab codes (e.g. S25147), select 'Isolate' here.

15 **Source Modifiers**: This is the section where you add key details that you would like to be attached to the sequences (e.g. locality and depth). However, only the organism name and specimen-voucher are required for submission. You should have already included

these in the previous steps, so you may continue if you do not wish to apply additional source modifiers.

∗Since all the specifics will be in your publication, you may typically stick with locality (column name = 'Country') and depth (column name = 'Altitude').

∗If you would like to add a couple key source modifiers, then choose one of the following two options under 'How do you want to apply source modifiers?':

15.1 Option 1. 'Upload a tab-delimited table': Proceed to download the source modifier template table. You can edit this file in several programs (TextEdit, TextWrangler, BBEdit, or even Microsoft Excel). In a text edit program, separate the information in each column by inserting 1 tab. The locality information in the 'Country' column must go from broad to specific. For example, "Costa Rica: Mound 12" follows this format. The depths entered in the 'Altitude' column must be negative (e.g. "-1800 m"). An example of a tab-delimited table is attached below:

15.2   Option 2. 'Use an editable table': This is self explanatory. You can manually add columns with corresponding information for each sequence ID via the online interface. However, please copy and paste, and **save** your work somewhere else for your records.

16   **References**: Under 'Sequence authors', add yourself (unless someone else did the lab work). Select the corresponding 'Publication status' that applies to you. Add the title of your paper and 'Specify new authors' to add the names of all authors on your paper.

17   **Review & Submit**: Check the details of your submission here. If everything looks good, submit your sequences.