**Guidelines for SRA submission to NCBI**

* Login or sign up for NCBI account

Create a Bioproject:

A BioProject is a collection of biological data related to a single initiative, originating from a single organization or from a consortium of coordinating organizations.

Steps to create a bioproject ID:

First go to [**https://submit.ncbi.nlm.nih.gov/subs/bioproject/**](https://submit.ncbi.nlm.nih.gov/subs/bioproject/) and create a New Submission. This will give you submission Id.

* + **Submitter**: fill out your information about your name, email, organization etc. and click continues.
  + **Project Type**: specify your data type to be submitted. For example to submit SRA data, choose raw sequence reads.
  + **Target:** Give most descriptive organism name for your study.
  + **General Info**: Give a short description about the study goal of your study and relevance
  + **Publications:** this page collects publication information specific to the registered project. A publication identifier is required. A PubMed ID or DOI is required..
  + **Overview :** this page presents a summary of the provided information. Click the ‘Submit’ button at the bottom of the page to complete the submission.

Once you complete submission you receive a bioproject ID. The format of the BioProject Accession is five alpha-letters followed by one to six numbers. For example **PRJNA428490**

## BioSample

### Steps in Submission Process

Click on this link [**https://submit.ncbi.nlm.nih.gov/subs/biosample**](https://submit.ncbi.nlm.nih.gov/subs/biosample) and create a new submission

* **Submitter:** Please fill in all requested fields and click continues.
* **General info:** Set the release date and choose whether you want to select batch biosamples or single biosample. Click Continue once you are satisfied with your selections
* **BioSample Type:** Select the BioSample type that best describes your samples. For plasmodium falciparum select Microbe.

*You can look up the different BioSample types*[*here*](https://submit.ncbi.nlm.nih.gov/biosample/template/)

* **BioSample attributes:** Download the tab-separated BioSample worksheet and complete it for your samples. Upload the completed BioSample worksheet to this page.
* It is required that each sample in your worksheet has at least one unique attribute. Sample name, title, and description are not taken into account when searching for uniqueness.
* Additional attributes can be added by creating another column with a unique header
* Complete the template table. Save the worksheet as a Text (Tab-delimited) file. Upload the text file on the 'Attributes' tab of the BioSample Submission Portal at <https://submit.ncbi.nlm.nih.gov/subs/biosample/>
* **Overview**: Look over the submission and when you are satisfied “Click” submit button if everything is fine

Here is the example Biosample attribute file for BioProject ID: PRJNA428490

[**MaRS\_US\_1-9-18.txt**](https://submit.ncbi.nlm.nih.gov/api/2.0/files/bzlzwfuz/mars_us_1-9-18.txt/?format=attachment)

# Sequence Read Archive (SRA) submission

### Steps in Submission Process

* **Submitter:** Please fill in all requested fields.
* **General info**: BioProjects are selected by typing in the BioProject name (either the PRJNA#### accession or the title)
* **SRA metadata**: Download and complete the tab-separated SRA metadata table. Upload the completed SRA metadata table to this page.
* Paired-end data with separate forward and reverse read files should have 2 filename columns so that both files are listed on the same row of the spreadsheet
* Additional attributes can be added by creating another column with a header
* here is the example metadtata file for SRA submission for BioProject ID: PRJNA428490

Upload Metadata file.

* **Files**: is where you upload the files and this is done through the “Browse” button. The files can be transferred either through HTTP
* Each file must be listed in the [SRA metadata table that you uploaded.](https://submit.ncbi.nlm.nih.gov/api/2.0/files/?format=attachment) If you are uploading a tar archive, list each file name, not the archive name.
* Files can be compressed using gzip or bzip2, and may be submitted in a tar archive but archiving and/or compressing your files is not required. Do not uses zip!
* All file names must be unique.
* Use the preload option if you are uploading files over 10 GB or more than 300 files. All files for a submission must be uploaded into a single folder.

**FTP upload instructions**

* 1. Navigate to the source folder where the files for submission are;
  2. Establish an FTP connection using the credentials below:
* **Address: ftp –i ftp-private.ncbi.nlm.nih.gov (to turn off the interactive mode)**
* **Username: user ID (provided by ncbi)**
* **Password: provided by ncbi**
  1. Navigate to your account folder:
* **cd uploads/yyr4@cdc.gov\_ocC82XZX**
  1. Create a subfolder (required!) with a meaningful name:
* **mkdir new\_folder (name it according to your project)**
  1. Navigate to the target folder you just created:
* **cd new\_folder**
  1. Copy your files into the target folder:
* **mput -I file\_name (mput command for all(\*)files)**

If you upload your files in your root directory, you will not be able to see them or to select the folder during the submission.

Make a new subdirectory for each new submission. Your submission subfolder is a temporary holding area and it will be removed once the whole submission is complete.

Do not upload complex directory structures or files that do not contain sequence data.

Return back to this page and select preload folder, then continue submission.

Please note: it takes at least 10 minutes for uploaded files to become available for selection.

* **Overview**: Look over the submission and when you are satisfied “Click” submit button if everything is fine