Mar 10, 2023

🌐 Submitting Genomes

DOI

dx.doi.org/10.17504/protocols.io.8epv5jdkdl1b/v1

Emma Pearce[1]

[1]UCSF

Emma Pearce
UCSF

Protocol status: In development
We are still developing and optimizing this protocol

Created: February 14, 2023

Last Modified: March 10, 2023

Protocol Integer ID: 77005

Keywords: bankit, submitting, ncbi, genome to ncbi, submitting genomes step, genomes step, ncbi, genome, bankit

Abstract

Steps to submit a genome to NCBI through Bankit

Troubleshooting

Submitting to Bankit

1   Log into Bankit with UCSF username and password:
    - Follow this link: https://www.ncbi.nlm.nih.gov/WebSub/?form=history&tool=genbank
    - Choose more login options and select University of California, San Francisco

2   Select "Start Bankit Submission"

3   Fill out the contact information

4   **Sequence Authors:**
    List all sequence authors (use the Add button for more than one)
    Reference Information #1:
    Select publication status (unpublished, in-press, or Published)
    - If in-press add reference title, journal title, year, volume, issue, and page numbers
    - If published add reference title, journal title, year, volume, issue, page numbers, and PubMed ID
    - Optional: if unpublished add reference title
    Select reference authors as "Same as Sequence Authors" or "Specify New Authors"
    - If specify new authors, list authors
    (Optional) Add Another Reference:
    Follow Reference Information #1 steps

5   **Sequencing Technology:**
    Select the sequencing technology (or other and specify)
    Select unassembled sequence reads or assembled sequences (consisting of two or more sequence reads)
    List assembly program and version or date

    Example: Illumina, Geneious v. Geneious Prime 2023.0.1

6   **Submission Release Date:**
    Select "Immediately after processing" or "Release date"
    - If Release Date is selected, specify date in DD-Mon-YYYY format
    Sequence(s) and Definition Line(s):
    Choose molecule type, linear or circular, if it is a complete sequence, and fasta or alignment format
    Either download sequence file or copy and paste it in the box
    - Note: make sure the first line follows this format; make everything in bold specific to your submission
        >Seq**#**[organism=**genus species**] **species straininfo**

Example: >Seq1[organism=Enterovirus rhinovirus] Rhinovirus C17 strain RvC17/USA/2019

7   **Submission Category:**
Select "Original" or "Third Party Annotation"

8   **Source Modifiers:**
Choose organelle/location if applicable
Choose source modifiers and list information in value column, add as many as desired

Example: Strain = RvC17/USA/2019, host = homo sapiens, tissue_type = Lung, country = USA, collection_date = 2019

9   **Features (Overview):**
Select "Add features by completing input forms"
Select "Coding Region (CDS) / Gene / mRNA"
Select "providing intervals"
Click Add

10   **Features (Detail):**
Choose if the CDS is on the + or - strand
If the CDS is not compete, select 5' or 3' for which end it is incomplete at
Select "Entire Sequence" or "Specific Spans" for Nucleotide Interval Spans
    - If Specific Spans, specify start and stop and which strand it is on
List protein name (optional: also list protein description and EC number)
Click accept

Example: specific spans: start = 19 stop = 6498 strand = -minus; protein name = polyprotein

11   **Review Submission:**
Make sure everything looks good
If resubmitting check the box under the resubmission section