



JAN 16, 2023

OPEN ACCESS

Protocol Citation: Paul Lorenzo A Gaite, Dr Ritchie Mae T Gamot, Dr Lyre Anni E Murao 2023. Submission of sequence and contextual data to GISAID, INSDC repositories, or other databases.

protocols.io
<https://protocols.io/view/submission-of-sequence-and-contextual-data-to-gisa-cgh2tt8e>

License: This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

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

Created: Sep 13, 2022

Last Modified: Jan 16, 2023

PROTOCOL integer ID:
 69914

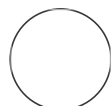
Submission of sequence and contextual data to GISAID, INSDC repositories, or other databases

Paul Lorenzo A Gaite¹, Dr Ritchie Mae T Gamot^{1,2},
 Dr Lyre Anni E Murao^{1,2}

¹Philippine Genome Center Mindanao;

²University of the Philippines Mindanao

PHA4GE Subgrant - Philippines



phagesubgrantph

ABSTRACT

Timely submission of viral sequence and corresponding contextual data by public health laboratories is an essential step to SARS-CoV-2 biosurveillance. This enables real-time updating of repositories of sequence data, hence real-time tracking of information on the virus as gathered from the sequence data and metadata as well. There are various sequence repositories or databases (with a section dedicated to SARS-CoV-2 viral sequences) that are publicly-available online, such as GISAID and the INSDC repositories.

The previous project involved submission of the sequences generated by the PGC Mindanao workflow to the GISAID database. In collaboration with BugSeq on the PHA4GE subgrant, PGC Mindanao was also able to upload and submit the previously-generated SARS-CoV-2 sequences to NCBI, a database that is part of the INSDC, which in turn subsequently released these sequences to its public database successfully.

1 Abstract/Introduction

Timely submission of viral sequence and corresponding contextual data by public health laboratories is an essential step to SARS-CoV-2 biosurveillance. This enables real-time updating of repositories of sequence data, hence real-time tracking of information on the virus as gathered from the sequence data and metadata as well. There are various sequence repositories or

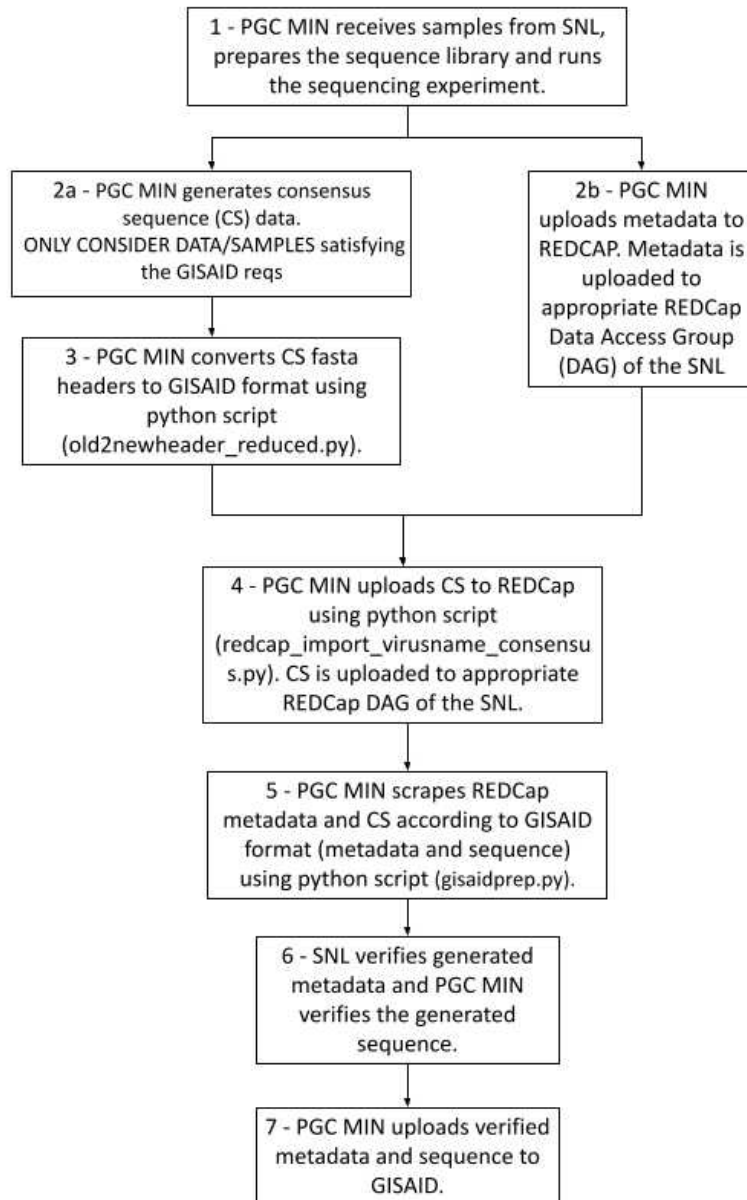
databases (with a section dedicated to SARS-CoV-2 viral sequences) that are publicly-available online, such as GISAID and the INSDC repositories.

The previous project involved submission of the sequences generated by the PGC Mindanao workflow (refer to protocol on "[Introduction and Lineage Assignment of Assembled Sequences](#)") to the GISAID database. In collaboration with BugSeq on the PHA4GE subgrant, PGC Mindanao was also able to upload and submit the previously-generated SARS-CoV-2 sequences to NCBI, a database that is part of the INSDC, which in turn subsequently released these sequences to its public database successfully.

The protocol below outlines the PGC Mindanao workflow for the submission of sequence and contextual data to GISAID (Section 2). A short section shows how the PHA4GE contextual data package was used in the submission process (Section 3). The process of submission of sequence and contextual data by PGC Mindanao to the NCBI database is also outlined (Section 4). A comparison of results from GISAID and NCBI submissions is shown (Section 5).

2 PGC Mindanao workflow

This section outlines the PGC Mindanao workflow for the submission of sequence and contextual data to a public database. Figure 1 shows an overview of the entire workflow. The workflow ultimately deposits the sequence and contextual data to the online public database GISAID.



Note: Original python scripts provided by Paco Polotan of RITM and modified by Bioinfo Division PGCMin

Figure 1. Overview of the PGC Mindanao workflow for submission of sequence and contextual data to the GISAID database

After assembling and generating the consensus sequences from the viral samples, the sequences are assessed and filtered for the number of ambiguous bases and presence of frameshift mutations and unexpected stop codons in the sequences as confirmed by the sequencing data and not from sequencing artifacts. Sequencing having less than 50% ambiguous bases and have frameshift mutations and unexpected stop codons that are confirmed by the sequencing data will be included in the GISAID submission and those otherwise were not included

Sequences were uploaded to REDCap using the script provided by the previous project (succeeding steps were also performed with scripts provided by the previous project). Similar to the metadata, the sequence data were uploaded to REDCap according to the appropriate data access group of each SNL. Both sequence data and metadata were scraped from the REDCap database using the provided Python script (Section 2.2). PGC Mindanao verified the sequences and the corresponding metadata. When verification was done, the sequences and the metadata were uploaded to GISAID through their submission protocol (Section 2.3).

Various scripts were created and provided by the previous project to format heading of FASTA file to standard form and other process. The provided custom script "old2newheader_reduced.py" converted the header/s of the sequence FASTA file to conform to GISAID convention (Figure 2).

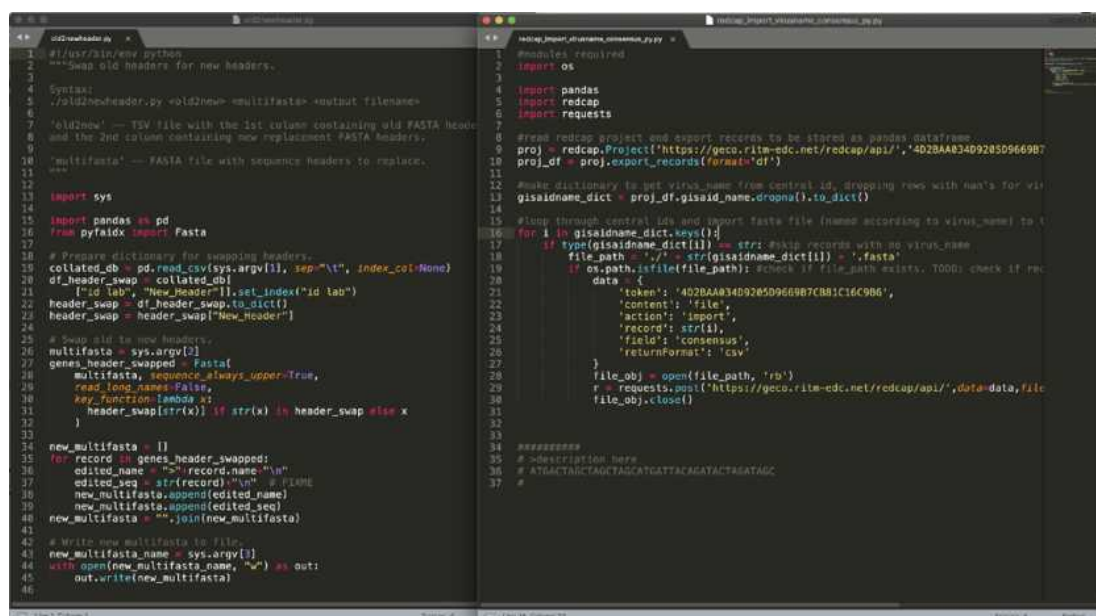


Figure 2. Screenshots of two of the scripts used in the workflow, "old2newheader_reduced.py" and "redcap_import_virusname_consensus.py"

2.2 Data entry into, and data "scraping" from REDCap Database:

Sequence data and contextual data/metadata (refer to protocol "[Establishing processes to capture standardized contextual data](#)" for details) was entered into the REDCap database of the previous project. Figure 3 shows the landing page after entering the login into the REDCap database, which is based on the data access group entered from the last login. Individual sample sequence data and metadata entries may be uploaded manually through the webpage GUI (through the green button "Add new record" and manual entry in individual sub-entry pages). Figure 4 shows the case metadata sub-entry page. Figure 5 shows the analysis metadata sub-entry page. Another approach to uploading data is through the "Data Import Tool" (Figure 6). Sequence data may also be uploaded through the use of the provided custom script "redcap_import_virusname_consensus.py" (Figure 2).

Data can be "scraped", or collected, from the same database by using the provided custom script "gisaidprep.py", which outputs a spreadsheet that conforms to GISAID submission requirements (Figure 7).

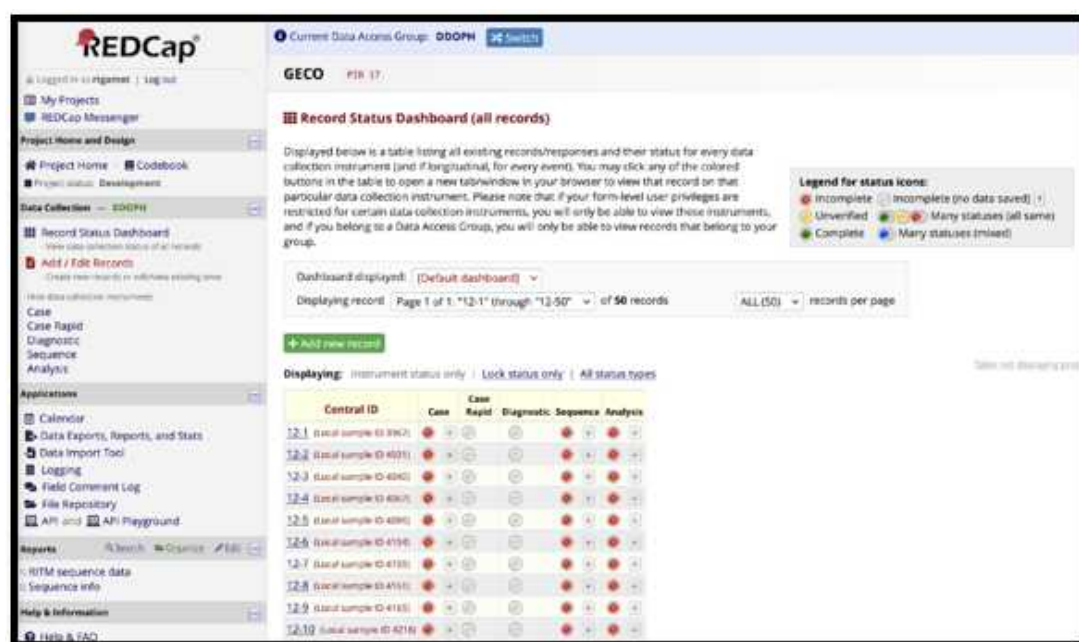


Figure 3. Main landing page after entering login of the REDCap database

Case

Current instance: 1 - 2021-05-26

Editing existing Central ID 12-1 (Local sample ID 3967)

Central ID	12-1
Local sample ID	3967
gvsid name	PH-DDOPH-12-1
city/municipality	
province	Davao del Oro
region	Region XI (Davao Region)
country	Philippines
date sample collected	2021-05-26
The date the sample was received.	2021-05-26
One of serum, oropharyngeal/nasopharyngeal swab, other	<input type="radio"/> serum <input checked="" type="radio"/> oropharyngeal/nasopharyngeal swab <input type="radio"/> other
Age of the patient in years	21
Sex of the patient. One of: F M Other	<input checked="" type="radio"/> Female <input type="radio"/> Male

Figure 4. Case metadata sub-entry page of a REDCap database sample entry

Analysis

Current instance: 1 - 2021-05-26

Editing existing Central ID 12-1 (Local sample ID 3967)

Central ID	12-1
Local sample ID	3967
Combined runs	PH-DDOPH-12-1.fasta (0.03 MB)
Consensus sequence	152.42
GISAID Accession Number	29773
GISAID Authors List	86.41
Assembly method	3942
Consensus average depth	B.1.351
Sequence length	2021-07-28
Missing coverage	
Number of sites with gaps	
Pangolin lineage	
Lineage support	
Pangolin version	

Figure 5. Analysis metadata sub-entry page of a REDCap database sample entry

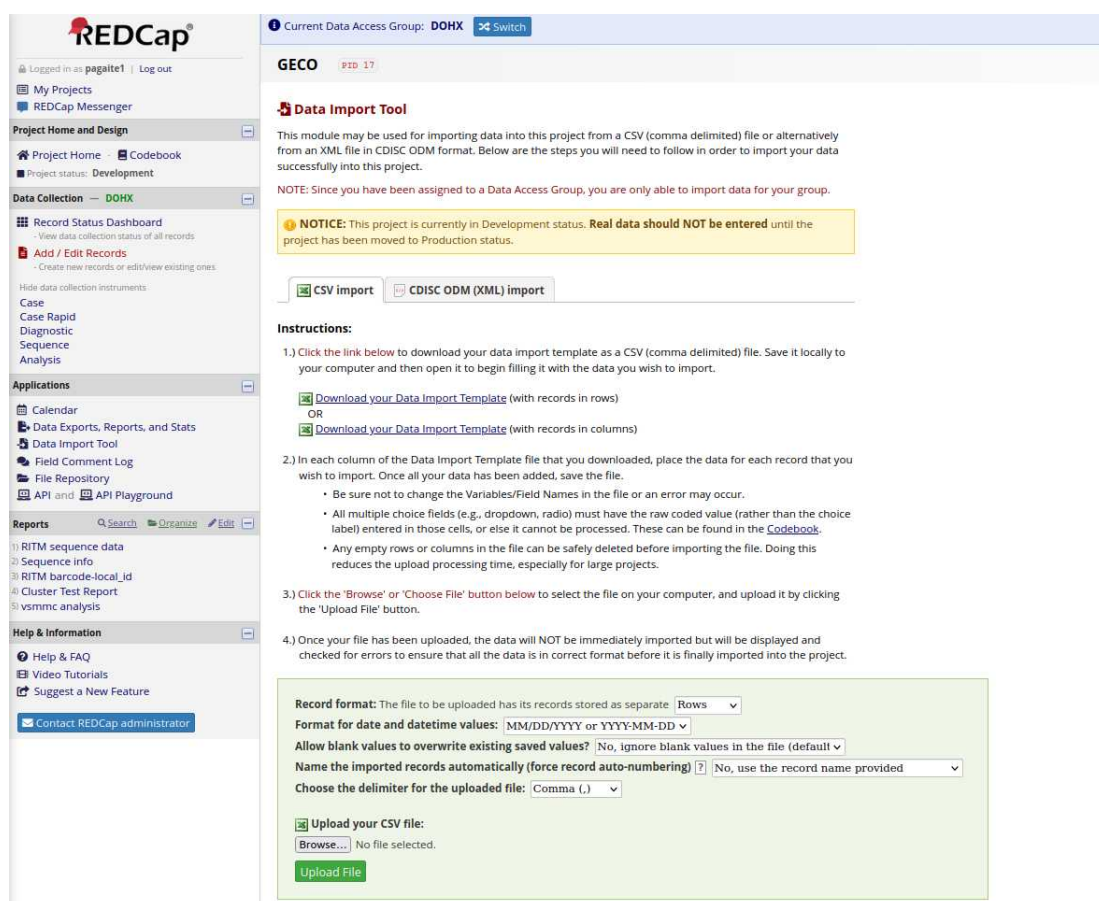


Figure 6. Data import tool of REDCap

2.3 Submission to GISAID database:

The spreadsheet outputted by "gisaidprep.py" (containing the sequence metadata) from the previous section, along with the final formatted sequence FASTA file, will be used for uploading the sequence data and metadata to GISAID. Figure 7 shows the spreadsheet outputted by "gisaidprep.py" metadata that will be submitted to GISAID in conjunction with the sequence data. Standard GISAID submission procedure was performed, wherein the recommended approach is through the use of the command-line tool CLI2. This process requires a live client ID for authentication, which may be requested by emailing the support team for this tool at GISAID (clisupport@gisaid.org). Figure 8 shows a number of the commands that can be issued by the CLI2 command-line tool.

Metadata_DOHX_27Oct2021 .XLSX ☆ ☰ ☰ Last edit was seconds ago

File Edit View Insert Format Data Tools Help

100% 10 Arial

Submitter	FASTA filename	Virus name	Type	Passage details/history	Collection date	Location
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-10/2021	betacoronaviru	Original		2021-05-08	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-11/2021	betacoronaviru	Original		2021-06-25	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-12/2021	betacoronaviru	Original		2021-06-25	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-13/2021	betacoronaviru	Original		2021-06-22	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-15/2021	betacoronaviru	Original		2021-06-22	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-16/2021	betacoronaviru	Original		2021-06-22	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-17/2021	betacoronaviru	Original		2021-06-22	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-18/2021	betacoronaviru	Original		2021-06-22	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-19/2021	betacoronaviru	Original		2021-06-23	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-20/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-21/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-22/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-23/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-24/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-25/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-26/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-27/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-28/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-29/2021	betacoronaviru	Original		2021-06-23	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-30/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-31/2021	betacoronaviru	Original		2021-06-21	Asia /
rtchiemaega all_sequences fasta	hCoV-19/Philippines/PH-DOHX-14-32/2021	betacoronaviru	Original		2021-06-21	Asia /

Instructions Submissions Count: 30 Explore

Figure 7. Spreadsheet outputted by "gisaidprep.py" script containing sequence metadata conforming to GISAID submission requirements


```

To get help on the upload sub-command, do cli2 upload -h:
1 usage: cli2 upload [-h] [--database {EpiCoV,EpiFlu,EpiRSV}] [--token TOKEN] --metadata
  ↪ METADATA --fasta FASTA
2               [--frameshift {catch_all,catch_novel,catch_none}] [--failed FAILED]
  ↪ [--proxy PROXY] [--debug] [--log LOG]
3
4 Perform upload of sequences and metadata to GISAID's curation zone.
5
6 optional arguments:
7 -h, --help            show this help message and exit
8 --database {EpiCoV,EpiFlu,EpiRSV}
9                       Target GISAID database. (default: EpiCoV)
10 --token TOKEN         Authentication token. (default: ./gisaid.authtoken)
11 --metadata METADATA   The csv-formatted metadata file. (default: None)

```

3

```

12 --fasta FASTA         The fasta-formatted nucleotide sequences file. (default: None)
13 --frameshift {catch_all,catch_novel,catch_none}
14                       'catch_none': catch none of the frameshifts and release
  ↪ immediately; 'catch_all': catch all frameshifts and
  ↪ require email
15                       confirmation; 'catch_novel': catch novel frameshifts and require
  ↪ email confirmation. (default: catch_all)
16 --failed FAILED      Name of CSV output to contain failed records. (default: ./failed.
  ↪ out)
17 --proxy PROXY        Proxy-configuration for HTTPS-Request in the form: http(s)://
  ↪ username:password@proxy:port. (default: None)
18 --debug              Switch off debugging information (dev purposes only). (default:
  ↪ True)
19 --log LOG            All output logged here. (default: ./upload.log)

```

Figure 8. A number of the commands that can be issued by the CLI2 command-line tool

After submission of the sequence data and metadata to GISAID through the CLI2 tool, the data will be reviewed by curators at GISAID. The uploader will be notified by email of successfully-released samples (together with their corresponding individual accession IDs) from submit@gisaid.org, and the samples made immediately available to registered GISAID users. On the other hand, the uploader will be notified by email of unsuccessful sample/s from hcov-19@gisaid.org with the reason for the non-release of each of these sample/s. The uploader may resubmit with the corrected sequence/s or inform GISAID that the sequence/s is correct and is supported by the raw sequencing data.

3 PHA4GE contextual data package

The contextual data template spreadsheet from the package can be used to standardize submission to repositories. See protocol on "[Establishing processes to capture standardized contextual data](#)" for details.

4 NCBI GenBank submission process

BugSeq has conducted an orientation session on submission of sequence and contextual data to INSDC repositories, particularly NCBI GenBank. Sequences previously-generated by PGC Mindanao were submitted in actual during the orientation. This section outlines this process.

Figure 9 shows the landing page or submission portal for sequences in GenBank. The process is done entirely on webpage GUI at the online NCBI GenBank page. The page also shows the overview of previous sequence submissions.

Figure 10 shows the first step to submission, asking the uploader the type of sequences to be uploaded and other details (e.g. if it is SARS-CoV-2).

Figure 11 shows the page for the second step, which asks for details regarding the submitter.

Figure 12 shows the third step of the process, which asks for details regarding the sequencing technology used (e.g. sequencing platform and method).

Figure 13 shows the fourth step, which is the upload page for the sequence file and also asks when the sequences are to be released.

Figure 14 presents the fifth step, which is the sequence processing page and also where the uploader is given an option to automatically remove failed sequences.

Figure 15 shows the sixth step, which is the source information page and where the uploader is asked about the details on the sequence IDs.

Figures 16 and 17 show the seventh step of the process wherein source modifiers, or sample/sequence metadata, are provided. The uploader is given an option to provide the metadata by filling out the editable table within the page or uploading a tab-delimited table file containing the metadata. Figure 17 shows the editable table where the source modifiers are provided in the submission process in this case.

Figure 18 shows the eighth step asking details on references, such as sequence author information and status of the publication linked to the sequences.

Figure 19 shows the ninth step, which is the review of the sequence submission before uploading and submitting the sequence data and corresponding metadata to the NCBI GenBank database.

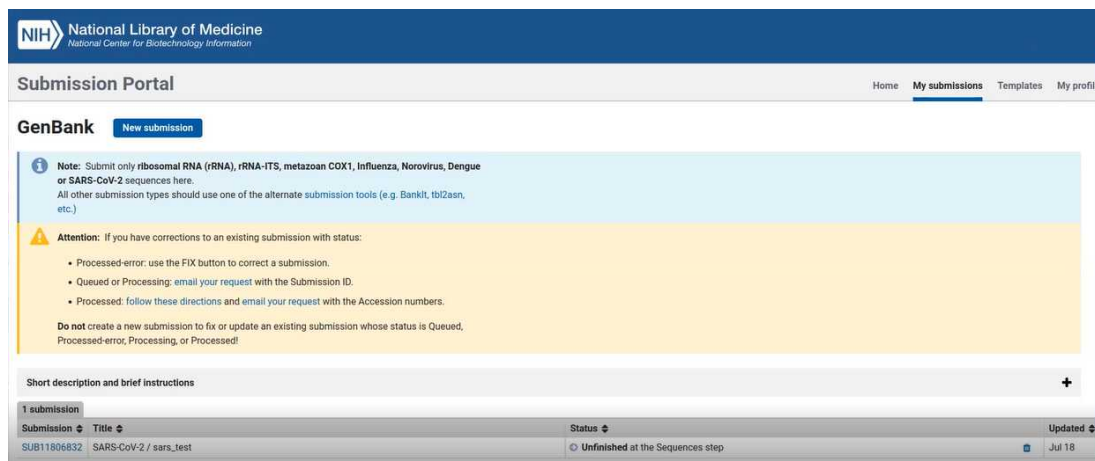


Figure 9. Landing page of the submission portal of NCBI GenBank

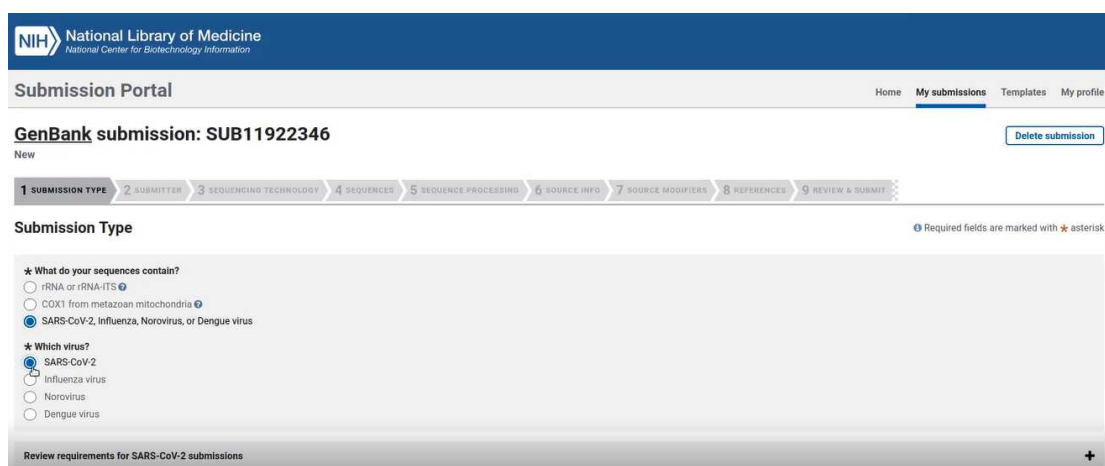


Figure 10. Submission type page

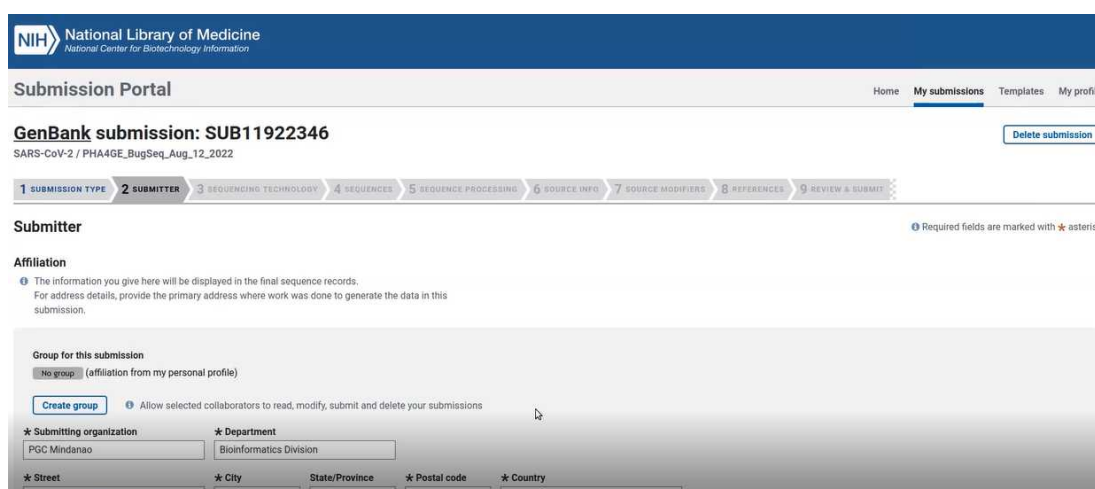


Figure 11. Submitter information page

NIH National Library of Medicine
National Center for Biotechnology Information

Submission Portal Home **My submissions** Templates My profile

GenBank submission: SUB11922346 [Delete submission](#)

SARS-CoV-2 / PHA4GE_BugSeq_Aug_12_2022

1 SUBMISSION TYPE 2 SUBMITTER **3 SEQUENCING TECHNOLOGY** 4 SEQUENCES 5 SEQUENCE PROCESSING 6 SOURCE INFO 7 SOURCE MODIFIERS 8 REFERENCES 9 REVIEW & SUBMIT

Sequencing Technology Required fields are marked with * asterisk

Method

* What methods were used to obtain these sequences? ?

☐ Sanger dideoxy sequencing
☐ 454
☐ Helicos
☐ Illumina
☐ IonTorrent
☐ Pacific Biosciences
☐ SOLID
☒ Other

* Method

Figure 12. Sequencing technology page

NIH National Library of Medicine
National Center for Biotechnology Information

Submission Portal Home **My submissions** Templates My profile

GenBank submission: SUB11922346 [Delete submission](#)

SARS-CoV-2 / PHA4GE_BugSeq_Aug_12_2022

1 SUBMISSION TYPE 2 SUBMITTER 3 SEQUENCING TECHNOLOGY **4 SEQUENCES** 5 SEQUENCE PROCESSING 6 SOURCE INFO 7 SOURCE MODIFIERS 8 REFERENCES 9 REVIEW & SUBMIT

Sequences Required fields are marked with * asterisk

Release date

Note: Release of BioProject or BioSample is also triggered by the release of linked data.

* When should this submission be released to the public?

☐ Release immediately following processing
☐ Release on specified date or upon publication, whichever is first


Sequences

* Upload a nucleotide FASTA formatted file.

[Choose file](#) or drag and drop it here

If you have multiple sequences, all of your sequences need to be in one file. [Help on FASTA file.](#)

Figure 13. Sequence upload and information page


National Library of Medicine
National Center for Biotechnology Information

Submission Portal

Home
My submissions
Templates
My profile

GenBank submission: SUB11922346

Delete submission

1 SUBMISSION TYPE
2 SUBMITTER
3 SEQUENCING TECHNOLOGY
4 SEQUENCES
5 SEQUENCE PROCESSING
6 SOURCE INFO
7 SOURCE MODIFIERS
8 REFERENCES
9 REVIEW & SUBMIT

Sequence Processing

Required fields are marked with * asterisk

Option to automatically remove failed sequences

If errors are found on sequences during [processing](#), they will be removed from this submission and the successful sequences accessioned. You will receive a detailed report on these errors.


* During processing, should NCBI remove sequences with errors and process the rest?

☐ Yes

☐ No

Continue

Figure 14. Sequence processing page


National Library of Medicine
National Center for Biotechnology Information

Submission Portal

Home
My submissions
Templates
My profile

GenBank submission: SUB11922346

Delete submission

1 SUBMISSION TYPE
2 SUBMITTER
3 SEQUENCING TECHNOLOGY
4 SEQUENCES
5 SEQUENCE PROCESSING
6 SOURCE INFO
7 SOURCE MODIFIERS
8 REFERENCES
9 REVIEW & SUBMIT

Source Information

Required fields are marked with * asterisk

The first few sequence IDs that we found are:

4974937e_barcode02

4974937e_barcode03

4974937e_barcode04

4974937e_barcode05

4974937e_barcode06

...

* Do your sequence IDs represent one of these?

☐ Isolate

☐ NONE of these

Values for these are typically alpha-numeric sample codes used in your laboratory to track individual samples. Select 'NONE of these' if it does not describe your sequence IDs or the sequence IDs contain more information than the [descriptions of these fields](#).

Continue

Figure 15. Source information page

National Library of Medicine
National Center for Biotechnology Information

Submission Portal

Home
My submissions
Templates
My profile

GenBank submission: SUB11922346

Delete submission

1 SUBMISSION TYPE
2 SUBMITTER
3 SEQUENCING TECHNOLOGY
4 SEQUENCES
5 SEQUENCE PROCESSING
6 SOURCE INFO
7 SOURCE MODIFIERS
8 REFERENCES
9 REVIEW & SUBMIT

Source Modifiers

Required fields are marked with * asterisk.
At least one of the fields marked with *,† or ‡ is required.

For each sequence, GenBank requires the following source information:

- collection-date,
- country,
- host, and
- isolate.

Current source modifiers - what you have provided so far

+

More help: [what is a source modifier](#), [description of each modifier](#), [how to provide source modifiers](#)

If you have already provided all the required information, you can press Continue to proceed.

* * How do you want to apply source modifiers?

☐ Use an editable table
☐ Upload a tab-delimited table (template file provided)

Figure 16. Source modifiers page

For each sequence, GenBank requires the following source information:

- collection-date,
- country,
- host, and
- isolate.

Current source modifiers - what you have provided so far

+

More help: [what is a source modifier](#), [description of each modifier](#), [how to provide source modifiers](#)

If you have already provided all the required information, you can press Continue to proceed.

* * How do you want to apply source modifiers?

Apply source modifiers by editing a table

Example: USA, Washington, King county

* Sequence_ID	* country	* host	* isolate	* collection-date	isolation-source	Add column
1 4974937e_barcode02	Philippines					
2 4974937e_barcode03	Philippines					
3 4974937e_barcode04	Philippines					
4 4974937e_barcode05	Philippines					
5 4974937e_barcode06	Philippines					
6 4974937e_barcode07	Philippines					
7 4974937e_barcode08	Philippines					
8 4974937e_barcode09	Philippines					
9 4974937e_barcode10	Philippines					
10 4974937e_barcode11	Philippines					

Figure 17. Editable table for inputting source modifiers

GenBank submission: SUB11922346
SARS-CoV-2 / PHA4GE_BugSeq_Aug_12_2022

1 SUBMISSION TYPE 2 SUBMITTER 3 SEQUENCING TECHNOLOGY 4 SEQUENCES 5 SEQUENCE PROCESSING 6 SOURCE INFO 7 SOURCE MODIFIERS 8 REFERENCES 9 REVIEW & SUBMIT

References

Sequence authors
Who should be publicly credited as the submitter of this sequence data? Enter authors below. Drag and drop to reorder authors.

* First (given) name MI * Last (family) name Delete

Add another sequence author

Reference

* Publication status
☐ Unpublished ☐ Inpress ☐ Published

Figure 18. References page for the submission

GenBank submission: SUB11922346
SARS-CoV-2 / PHA4GE_BugSeq_Aug_12_2022

1 SUBMISSION TYPE 2 SUBMITTER 3 SEQUENCING TECHNOLOGY 4 SEQUENCES 5 SEQUENCE PROCESSING 6 SOURCE INFO 7 SOURCE MODIFIERS 8 REFERENCES 9 REVIEW & SUBMIT

Review & Submit

You have requested that your sequence data be released immediately following processing.

Submitter

Submitter Paul Lorenzo Gaithe
 Email: plgaithe@gmail.com, pagaithe1@up.edu.ph
 Institution PGC Mindanao
 Department Bioinformatics Division
 Street Maguindanao Street
 City Davao City
 State Davao del Sur
 Postal code 8022

GenBank Record Preview

REFERENCE: Nidovirales; Coronidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.
 1 (bases 1 to 29831)
 AUTHORS: Gaithe, P.L.A.
 TITLE: Direct Submission
 JOURNAL: Submitted (12-AUG-2022) Bioinformatics Division, PGC Mindanao, Maguindanao Street, Davao City, Davao del Sur 8022, Philippines
 COMMENT: ##Assembly-Data-START##
 Assembly Method: BugSeq v. JULY-2022
 Sequencing Technology: Oxford Nanopore Technology
 ##Assembly-Data-END##
 FEATURES: Location/Qualifiers
 source 1..29831
 /organism="Severe acute respiratory syndrome coronavirus 2"
 /mol_type="genomic DNA"
 /isolate="SARS-CoV-2/human/PHL/PH-CRMC-13-14/2021"
 /isolation_source="oronasopharynx"
 /host="Homo sapiens"

To proceed please review your submission, make any necessary changes using the tabs/steps above, then click on the Submit button below.

Figure 19. Review page for the submission

After submission of the sequence data and metadata to NCBI GenBank through its Submission Portal, the data will be reviewed by curators at NCBI GenBank. The uploader will be notified by email if the whole batch of sequences has been successfully-released or not, with the reason for the non-release of each unsuccessful sample. The uploader may resubmit with the corrected sequence/s through the Submission Portal or inform NCBI GenBank that the sequence/s is correct and is supported by the raw sequencing data.

5 Comparison of GISAID and GenBank submissions

Table 1 shows details of the sequences submitted to GISAID (from previous project) and GenBank (from this grant), such as if initial database submission was successful, step/s taken to

resolve initial unsuccessful submission, if resubmission was successful, and the corresponding accession ID given by the database (if applicable).

A	B	C	D	E	F	G	H	I	J	K
Sample ID	GISA ID submitted?	Initial GISAID submission successful?	Resolution if initial GISAID submission unsuccessful	Resubmission to GISAID successful?	GISAID accession ID	NCBI submitted?	Initial NCBI submission successful?	Resolution if initial NCBI submission unsuccessful	Resubmission to NCBI successful?	NCBI accession ID
hCoV-19/Philippines/PH-CRMC-13-14/2021	Yes	Yes	-	-	EPI_ISL_5934896	Yes	No	Corrected sequence	Yes	OP522426
hCoV-19/Philippines/PH-CRMC-13-15/2021	Yes	Yes	-	-	EPI_ISL_5934897	Yes	Yes	-	-	OP522427
hCoV-19/Philippines/PH-CRMC-13-16/2021	Yes	Yes	-	-	EPI_ISL_5934898	Yes	Yes	-	-	OP522428
hCoV-19/Philippines/PH-CRMC-13-17/2021	Yes	Yes	-	-	EPI_ISL_5934899	Yes	Yes	-	-	OP522429
hCoV-19/Philippines/PH-CRMC-13-18/2021	Yes	Yes	-	-	EPI_ISL_5934900	Yes	Yes	-	-	OP522430
hCoV-19/Philippines/PH-CRMC-13-19/2021	Yes	Yes	-	-	EPI_ISL_5934901	Yes	Yes	-	-	OP522431
hCoV-19/Philippines/PH-CRMC-13-20/2021	Yes	Yes	-	-	EPI_ISL_5934902	Yes	Yes	-	-	OP522432
hCoV-19/Philippines/PH-CRMC-13-21/2021	Yes	Yes	-	-	EPI_ISL_5934903	Yes	Yes	-	-	OP522433
hCoV-19/Philippines/PH-CRMC-13-22/2021	Yes	Yes	-	-	EPI_ISL_5934904	Yes	Yes	-	-	OP522434
hCoV-19/Philippines/PH-CRMC-13-23/2021	Yes	No	Corrected sequence	Yes	EPI_ISL_5934905	Yes	No	Corrected sequence	No	-
hCoV-19/Philippines/PH-DDOPH-12-1/2021	Yes	Yes	-	-	EPI_ISL_5934981	Yes	Yes	-	-	OP522435

A	B	C	D	E	F	G	H	I	J	K
hCoV-19/Philippines/PH-DDOPH-12-2/2021	Yes	Yes	-	-	EPI_ISL_5934985	Yes	Yes	-	-	OP522436
hCoV-19/Philippines/PH-DDOPH-12-3/2021	Yes	Yes	-	-	EPI_ISL_5934986	Yes	Yes	-	-	OP522437
hCoV-19/Philippines/PH-DDOPH-12-4/2021	Yes	Yes	-	-	EPI_ISL_5934987	Yes	Yes	-	-	OP522438
hCoV-19/Philippines/PH-DDOPH-12-5/2021	Yes	Yes	-	-	EPI_ISL_5934988	Yes	Yes	-	-	OP522439
hCoV-19/Philippines/PH-DDOPH-12-6/2021	Yes	Yes	-	-	EPI_ISL_5934989	Yes	Yes	-	-	OP522440
hCoV-19/Philippines/PH-DDOPH-12-7/2021	Yes	Yes	-	-	EPI_ISL_5934990	Yes	Yes	-	-	OP522441
hCoV-19/Philippines/PH-DDOPH-12-8/2021	Yes	Yes	-	-	EPI_ISL_5934991	Yes	No	Corrected sequence	Yes	OP522442
hCoV-19/Philippines/PH-DDOPH-12-9/2021	Yes	Yes	-	-	EPI_ISL_5934992	Yes	Yes	-	-	OP522443
hCoV-19/Philippines/PH-DDOPH-12-10/2021	Yes	Yes	-	-	EPI_ISL_5934982	Yes	Yes	-	-	OP522444
hCoV-19/Philippines/PH-DDOPH-12-11/2021	Yes	Yes	-	-	EPI_ISL_5934983	Yes	Yes	-	-	OP522445
hCoV-19/Philippines/PH-DDOPH-12-12/2021	Yes	Yes	-	-	EPI_ISL_5934984	Yes	Yes	-	-	OP522446

Table 1. Comparison of GISAID and NCBI submission details