CNSA Handbook (simple version in English) V1.5

Updated: March 21, 2023

Register/Login

Please use the email or mobile number to register/login on the registration page and fill in the Submitter's information.

Enter the submission portal

Click "Submit" on the CNSA homepage or click "Submission portal" on the homepage navigation bar to enter the Submission portal page.

Submit project

1. Enter the submission process

   Click "Project" on the Submission portal page to enter the submission process.

2. Submit project information

   Select data management form -> fill in the genernal information -> fill in the details -> overview -> submit

Notes:

1. The first step in project submission requires the choice of data management forms. If you choose "Public" or "Controlled", the release date can refer to the date of article publication, and the release date should be earlier than the date of article publication.

2. The publication information of project also can be supplemented after the article is published.

3. After the project is submitted successfully, you can get the CNSA assigned project accession (prefixed with CNP) in "My submission-Project".

Submit sample

1. Enter the submission process

   Click "Sample" on the Submission portal page to enter the submission process.

2. Submit sample information

   If you submit only one sample at a time, we recommend that you choose a single submission method. If you submit multiple samples at a time, we recommend that you choose the batch submission method.

   （1）Single submission: Select "Submit a single sample" -> select sample model -> fill in sample attributes -> Fields pass check-> overview -> submit

   （2）Batch submission: Select "Submit batch samples" -> Select sample model -> Download template -> Upload completed template -> Template pass check -> Submit

Notes:

1. Please select the sample model correctly and you can't modify it by yourself after submitting.

2. The sample name cannot be duplicated.

3. When filling out the batch template file, please view the related description and field comments first. If some required fields are missing, you can fill in 'not collected', 'not applicable' or 'missing'. If the taxonomy ID or scientific name of the organism is unclear, you can enter the single submission process to search and ensure that the information is correct.

4. Collection date supports 4 data formats, YYYY, YYYY-MM, YYYY-MM-DD, YYYY-YYYY.

5. The number of uploaded file lines cannot exceed 2000. If it exceeds 2000, please submit in multiple processes.

6. After the sample is submitted successfully, you can get the sample accession assigned by CNSA (prefixed by CNS) in "My submission-Sample".

Submit experiment/run

1. Enter the submission process

   Click "Experiment/run" on the Submission portal page to enter the submission process.

2. Submit data files and metadata

   If you submit only one experiment/run at a time, we recommend that you choose a single submission method. If you submit multiple experiments/runs at a time, we recommend that you choose the batch submission method.

   (1) Single submission: Select submission type (Submit a single experiment/run) -> Fill in basic information -> Fill in metadata -> Metadata pass check -> Submit data files -> Data files pass check -> Overview ->Submit

   (2) Batch submission: Select submission type (Submit batch experiments/runs) ->Upload data files->Download metadata template->Upload completed metadata template->Metadata pass check ->Data files pass check -> submit

Notes:

1. It is recommended to upload the data files first. All users can upload data via FTP or mail the hard drive.

   (1) You just finished uploading after the data is uploaded to the FTP personal directory.

   (2) FTP server, username and password can be viewed in the submission process or "My service". Each user has a unique FTP account.

2. When filling out the batch template file, please view the related instructions and field comments first. One line represents a run. If a sample is associated with multiple data files, please submit them in multiple lines to ensure that the experiment information is consistent and the library name is unique. The file name and MD5 value of each data file are unique.

3. The number of metadata file lines cannot exceed 2000. If it exceeds 2000, please submit in multiple processes.

4. After the experiment/run is submitted successfully, you can obtain the accession assigned by CNSA in "My submission- Experiment/run" (Experiment: prefixed with CNX; Run: prefixed with CNR)

Submit assembly

1. Enter the submission process

   Click "Assembly" on the Submission portal page to enter the submission process.

2. Submit data files and metadata

(1) Single submission: Select submission type (Submit a single assembly) -> Fill in basic information -> Fill in metadata -> Metadata pass check -> Submit data files -> Data files pass check -> Overview ->Submit

(2) Batch submission: Select submission type (Submit batch assemblies)->Upload data files->Download metadata template->Upload completed metadata template->Metadata pass check ->Data files pass check -> submit

Notes:

1. It is recommended to upload the data files first (currently only supports FASTA format). All users can upload data via FTP or mail the hard disk.

   (1) You just finished uploading after the data is uploaded to the FTP personal directory.
   (2) FTP server, username and password can be viewed in the submission process or "My service". Each user has a unique FTP account.

2. When filling out the batch template file, please view the related instructions and field comments first. One line represents an assembly, ensuring that the assembly name of each assembly is unique, and the file name and MD5 value of each data file are unique.

3. The number of metadata file lines cannot exceed 2000. If it exceeds 2000, please submit in multiple processes.

4. After the assembly is submitted successfully, you can get the CNSA assigned assembly accession (prefixed with CNA) in "My submission-Assembly".

Submit variation

1. Enter the submission process

   Click "Variation" on the Submission portal page to enter the submission process.

2. Submit data files and metadata

   (1) Submit SNP: Select variation type (SNP) -> Upload data files to ftp-> Download metadata template -> Upload completed metadata template -> Metadata pass check -> Data files pass check -> Submit

   (2) Submit SV: Select variant type (SV) -> Upload data files to ftp (optional) -> Download template -> Upload completed template -> Pass check -> Data files pass check (if submitted) ->Submit

   (3) Submit CAHV: Select variation type (CAHV)->Download template->Upload completed template-> Pass check->Submit

Notes:

1. If the selected variation type is SNP, you are advised to first upload the data files (currently only supports VCF format). All users can upload data via FTP or mail the hard disk.

   (1) You just finished uploading after the data is uploaded to the FTP personal directory.
   (2) FTP server, username and password can be viewed in the submission process or "My service". Each user has a unique FTP account.

2. When filling out the batch template file, please view the description and field comments first. If you need to upload VCF files, make sure that the file name and MD5 value of each data file are unique.

3. After the data has been reviewed, you can get the variation accessions (prefixed by varc) assigned by CNSA in "My submission-Variation".

Submit metabolism

1. Enter the submission process

Click "Metabolism" on the [Submission portal page](#) to enter the submission process.

2. Submit data files and metadata

Upload data files-> Download "Descriptions" template -> Upload completed "Descriptions" template -> Add "Assay" ->Download the template of the created "Assay" -> upload the completed "Assay" template ->Metadata pass check -> Data files pass check -> Submit

Notes:

1. You are advised to first upload the data files. All users can upload data via FTP or mail the hard disk.

   (1) You just finished uploading after the data is uploaded to the FTP personal directory.

   (2) FTP server, username and password can be viewed in the submission process or "[My service](#)". Each user has a unique FTP account.

2. The "descriptions" information is required, and at least one assay needs to be created. A new assay can be added after the previous assay is uploaded.

3. The number of lines in each sheet of the metadata file cannot exceed 2000. If it exceeds 2000, please submit in multiple processes.

4. When filling out the template file, please view the description and field comments first, and make sure that the file name and MD5 value of each data file are unique.

5. After the data has been reviewed, you can get the metabolism accession (prefixed by METM) assigned by CNSA in "[My submission- Metabolism](#)".

Submit single cell

1. Enter the submission process

   Click "Single cell" on the [Submission portal page](#) to enter the submission process.

2. Select the associated project accession and fill in the number of cells.

3. Submit data files and metadata

   Upload data files-> Download metadata template -> Upload completed metadata template -> Metadata pass check -> Data files pass check -> Submit

Notes:

1. You are advised to first upload the data files. All users can upload data via FTP or mail the hard disk.

   (1) You just finished uploading after the data is uploaded to the FTP personal directory.

   (2) FTP server, username and password can be viewed in the submission process or "[My service](#)". Each user has a unique FTP account.

2. The expression matrix file is required, and other file types are optional. If your gene expression files, metadata files, and cluster files need to be grouped, please add the group name to the file name.

3. The number of lines in each sheet of the metadata file cannot exceed 2000. If it exceeds 2000, please submit in multiple processes.

4. When filling out the template file, please view the description and field comments first, and make sure that the file name and MD5 value of each data file are unique.

5. After the data has been reviewed, you can get the metabolism accession (prefixed by CSE) assigned by CNSA in "[My submission-Single cell](#)".

Submit virus sequence

1. Enter the submission process

Click "Virus sequence" on the Submission portal page to enter the submission process.

2. Select the release date of this submission.
3. Submit data files and metadata

Upload data files to ftp-> Download metadata template -> Upload completed metadata template -> Metadata pass check -> Data files pass check -> Submit

Notes:

1. You are advised to first upload the data files (currently only supports FASTA format). All users can upload data via FTP or mail the hard disk.

(1) You just finished uploading after the data is uploaded to the FTP personal directory.
(2) FTP server, username and password can be viewed in the submission process or "My service". Each user has a unique FTP account.

2. When filling out the template file, please view the description and field comments first, and make sure that the file name and MD5 value of each data file are unique.

3. After the data has been reviewed, you can get the virus sequence accessions (prefixed by N_) assigned by CNSA in "My submission-Virus sequence".

Submit sequence

1. Enter the submission process

Click "Sequence" on the Submission portal page to enter the submission process.

2. Select the release date of this submission.
3. Submit data files and metadata

Upload data files-> Download file list template -> Upload completed file list template -> File list pass check -> Data files pass check -> Submit

Notes:

1. You are advised to first upload the data files (currently only supports GBFF format). All users can upload data via FTP or mail the hard disk.

(1) You just finished uploading after the data is uploaded to the FTP personal directory.
(2) FTP server, username and password can be viewed in the submission process or "My service". Each user has a unique FTP account.

2. When filling out the template file, please view the description and field comments first, and make sure that the file name and MD5 value of each data file are unique.

3. After the data has been reviewed, you can get the virus sequence accessions (prefixed by N_) assigned by CNSA in "My submission-Sequence".

Memo

1. Data release

The release date can be set in the "Data Management" in the project submission process. Only the date today or within two years after today can be selected. The data will not be made public until the submitted data is reviewed by the reviewer and reaches the release date set by the user. If the data is about to reach the release date, the system will send a reminder email 15 days in advance.

2. Modify and delete

On the "My Submission" page, you can click the "pencil icon" in the status column to modify. If the status column does not have a "pencil icon", please send an email to datasubs@cngb.org and indicate the submission ID or data accession and the reason for the modification.

Set and modify the release date.

（1）Modify release date

a) If the status of the project is "Unfinished", go to "My submission", and click "pencil icon" in the project status column to enter the modification process.

b) If the status of the project is "Processing" or "Processed", you can modify the release date by clicking the date in the release date column or the "pencil icon".

c) If the status of the project is "Public" or "Controlled", the release date cannot be modified by yourself. If you need to make changes, please send an email to datasubs@cngb.org and indicate the project accession and the reason for the change.

（2）Delete submission

If the status is "unfinished", click the "trash can icon" to delete the submission. If it is in other status, please send an email to datasubs@cngb.org and indicate the submission ID and the reason for the deletion.

3. Data association

The information of the sample will only be triggered after the experiment/run or assembly is submitted; only after the experiment/run or assembly is reviewed and the data is public, all the information associated with the project can be retrieved according to the project accession. Otherwise, the information of the project and sample can only be retrieved separately, and no association will occur.

4. MD5 check

Please fill in the file name and MD5 value of the uploaded data file, and then click "Check", there are maybe four statuses:

（1）Not detected: The data file is not uploaded or system has not detected the file yet. Please click "Check" a few mins late if you have uploaded the data file, it still shows "Not detected".

（2）Calculating: The data file has been uploaded, but the MD5 value of the file has not been calculated or is being calculated. Please click "Check" later.

（3）MD5 mismatch: The MD5 value calculated by the system is inconsistent with the MD5 value you filled in. If the data file is only uploaded a part and in check, the md5 value calculated by the system will be inconsistent with the one you filled in. Please click "Check" later. If you click "Check" after a long time (such as half an hour), the status still shows "MD5 mismatch", please recalculate and fill in the MD5 of the data file. If the status still shows "MD5 mismatch", please contact datasubs@cngb.org and indicate the file name in the email.

（4）Check finished: The data file has been uploaded and passed the check.

5. View accessions and submitted metadata

On "My submission" page, you can directly view the accession of a single submission, download the batch-submitted attribute files with accessions in the status column, or click on the completed submission ID to view the details.

6. Contact us

If you have any questions, please contact the administrator at datasubs@cngb.org or 0755-33945586.