Open Archive for Miscellaneous Data User Guide V1.3

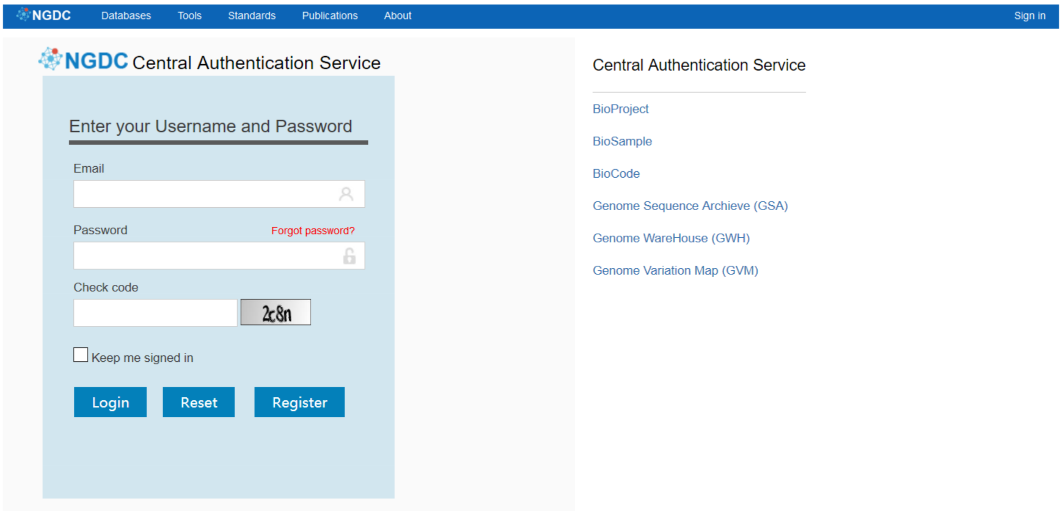
Last Updated: June 27, 2025

Introduction

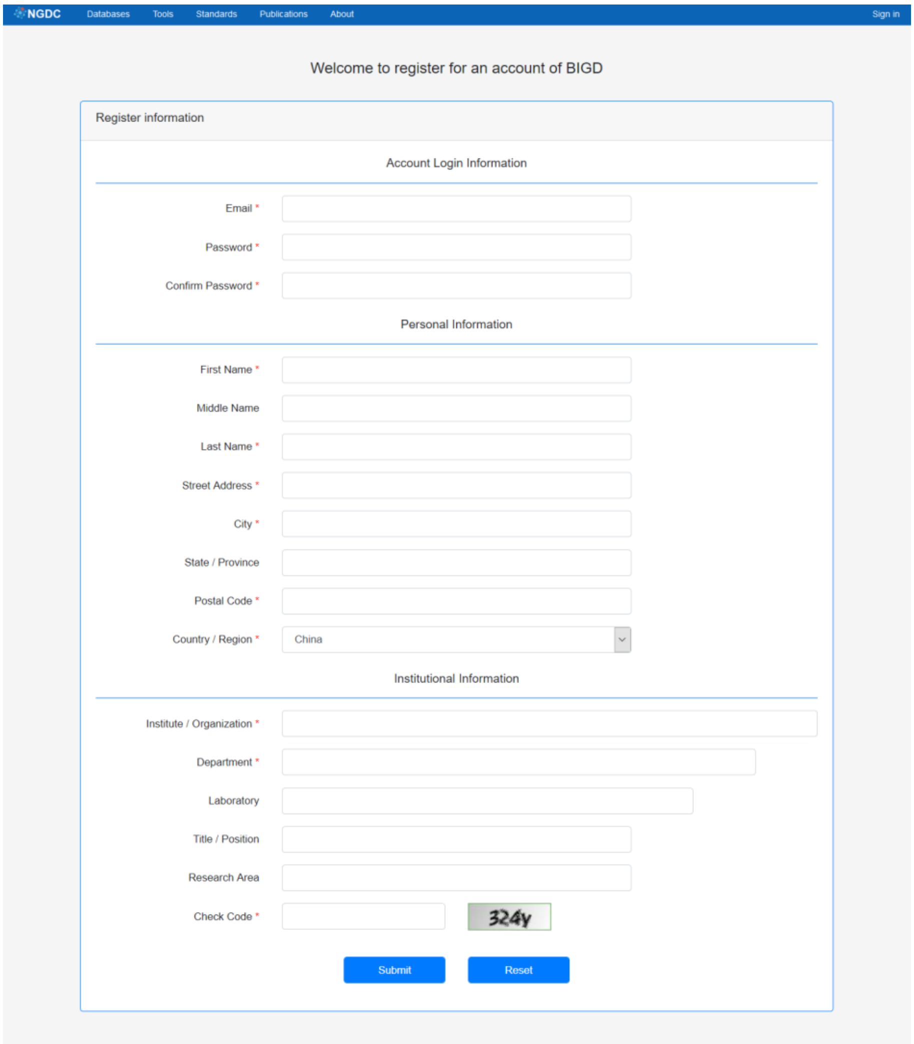
To meet the diverse data submission needs of users, the GSA team at the National Genomics Data Center (NGDC) has developed the Open Archive for Miscellaneous Data (OMIX). OMIX accepts not only processed raw sequencing and microarray data but also multi-omics and related data, including metabolomics (such as lipidomics) and proteomics. In terms of human genetic resources management, the OMIX database strictly adheres to the Regulations of the People’s Republic of China on the Administration of Human Genetic Resources, offering two modes of data access: open access and controlled access. Controlled access data requires authorization from the data submitter before it can be downloaded or used.

User Registration

All users can register an account through the BIG Single Sign-On (BIG SSO) system. However, the OMIX database recommends that data submissions be made by project leaders or principal investigators of research groups. You can complete account registration via the BIG Single Sign-On system. If you encounter any issues during registration or usage, please contact gsa@big.ac.cn.



⇩



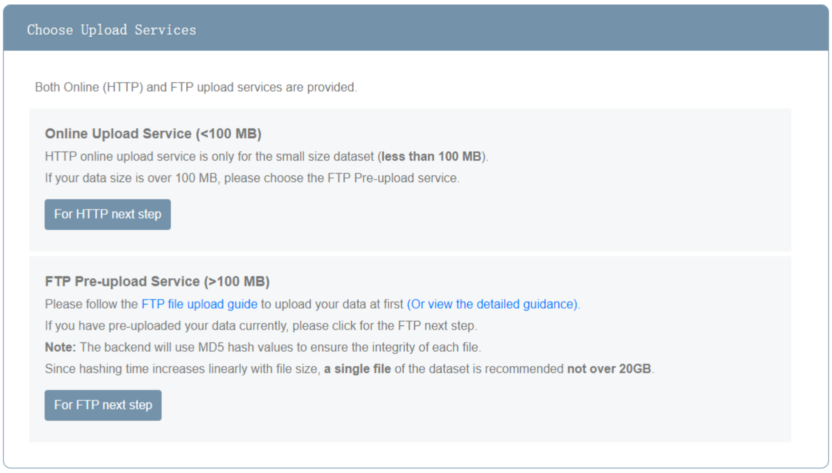
Welcome to register for an account of BIGD

Country / Region * China

Data Submission Guidelines

The OMIX database provides two methods for data submission:

1. Online Upload Service — submission via the web interface.
2. FTP Pre-upload Service — submission via FTP.



Both Online (HTTP) and FTP upload services are provided.

Online Upload Service (<100 MB)

HTTP online upload service is only for the small size dataset (**less than 100 MB**).

If your data size is over 100 MB, please choose the FTP Pre-upload service.

FTP Pre-upload Service (>100 MB)

Please follow the FTP file upload guide to upload your data at first (Or view the detailed guidance).

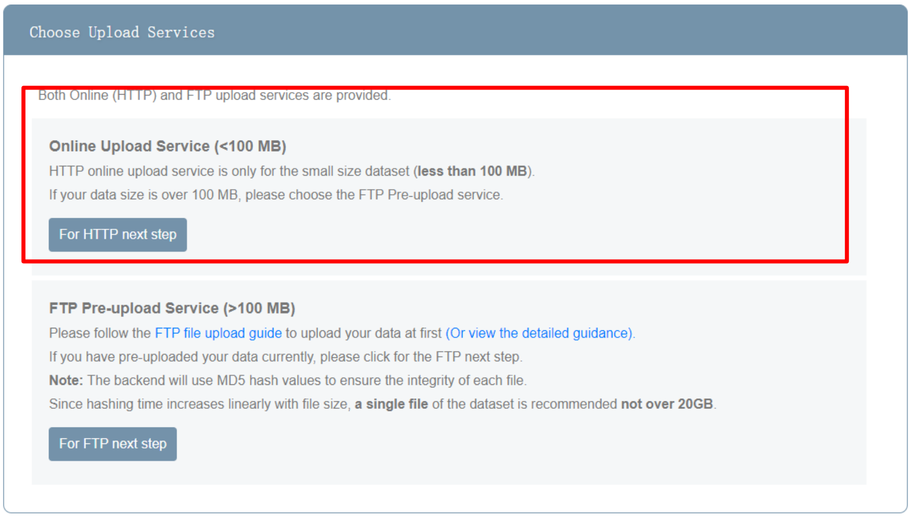
If you have pre-uploaded your data currently, please click for the FTP next step.

Note: The backend will use MD5 hash values to ensure the integrity of each file.

Since hashing time increases linearly with file size, **a single file** of the dataset is recommended **not over 20GB**.

Online Upload Service

For data submissions with a total size of less than 100 MB, it is recommended to use the Online Upload Service. This method allows users to complete both metadata entry and data file upload directly through the web interface.



Both Online (HTTP) and FTP upload services are provided.

Online Upload Service (<100 MB)

HTTP online upload service is only for the small size dataset (**less than 100 MB**).

If your data size is over 100 MB, please choose the FTP Pre-upload service.

FTP Pre-upload Service (>100 MB)

Please follow the FTP file upload guide to upload your data at first (Or view the detailed guidance).

If you have pre-uploaded your data currently, please click for the FTP next step.

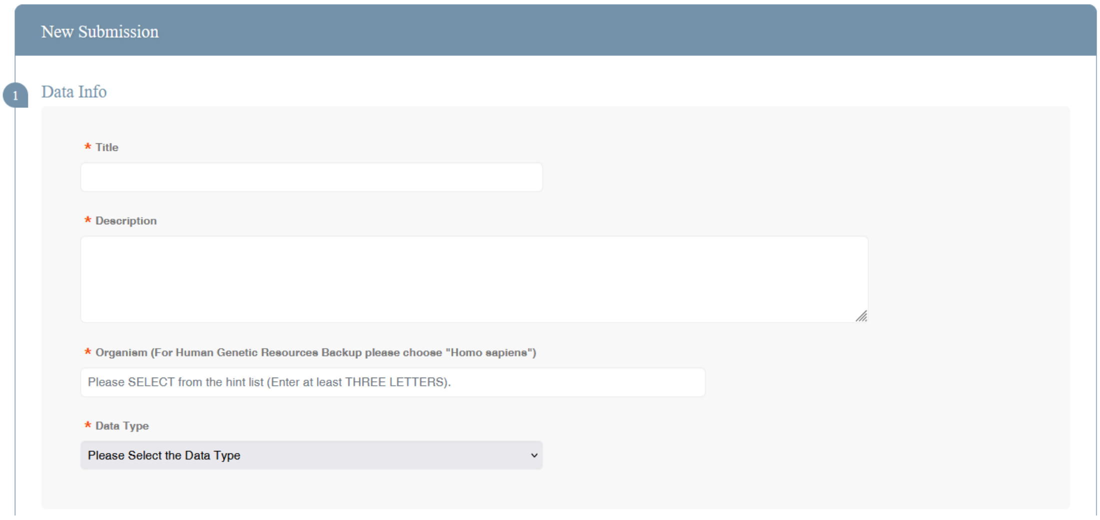
Note: The backend will use MD5 hash values to ensure the integrity of each file.

Since hashing time increases linearly with file size, **a single file** of the dataset is recommended **not over 20GB**.

Basic Information (Data Info)

This section is used to collect the dataset name, a brief description, and species information. If a file involves multiple species, please indicate it as “mixed sample”.

Note: For backup data of human genetic resources, the species must be specified as Human (Homo sapiens). The system provides two access modes for human genetic resources data: controlled access and open access.



New Submission

1 Data Info

* Title

* Description

* Organism (For Human Genetic Resources Backup please choose "Homo sapiens")

Please SELECT from the hint list (Enter at least THREE LETTERS).

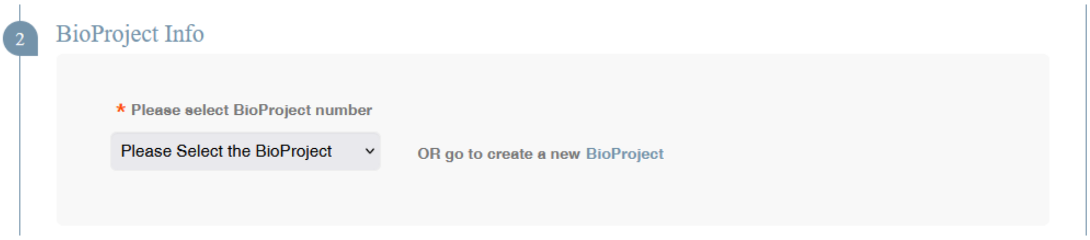
* Data Type

Please Select the Data Type

- If you set the data to Open Access, it will be automatically released on the specified release date or upon system detection of the related publication. Once released, users can browse the dataset information online and download the data.
- If you set the data to Controlled Access, users will still need your authorization to download the data even after the dataset is released. After Controlled Access release, downloads are not supported unless authorized by you.

BioProject Information (BioProject Info):

This section is used to collect the BioProject information associated with the dataset. If you have not yet created a BioProject, you can click the link to create one in the BioProject database. The review status of the BioProject does not affect the subsequent data submission process in OMIX.



2 BioProject Info

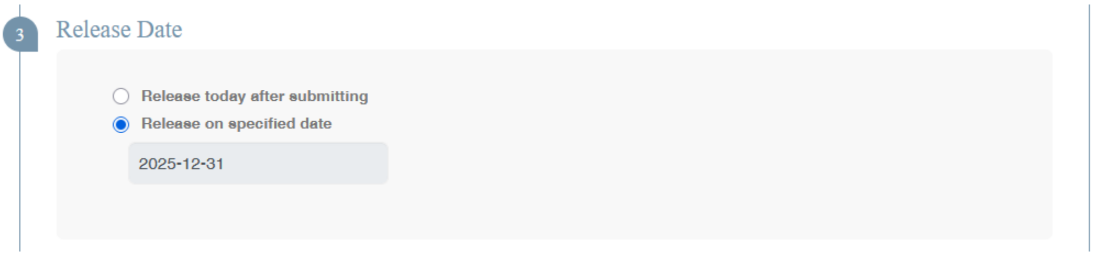
* Please select BioProject number

Please Select the BioProject

OR go to create a new BioProject

Release Date:

This section is used to collect the release date of the dataset.



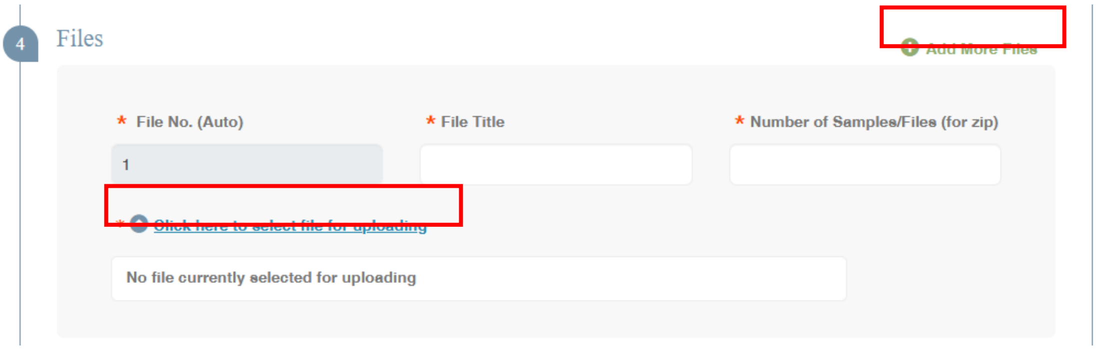
○ Release today after submitting
● Release on specified date

2025-12-31

Files:

- This section is used for file submission and collecting file information. You can add multiple files in each submission. Click “Add More File” as needed and fill in the corresponding metadata.

Note: For compressed packages, please specify the Number of Samples/Files as the number of files contained within the package. For non-compressed files, please specify the number of samples involved in the file.



* File No. (Auto): 1

* File Title:

* Number of Samples/Files (for zip):

No file currently selected for uploading

- Please click “Click here to select file for uploading” to upload your data files.

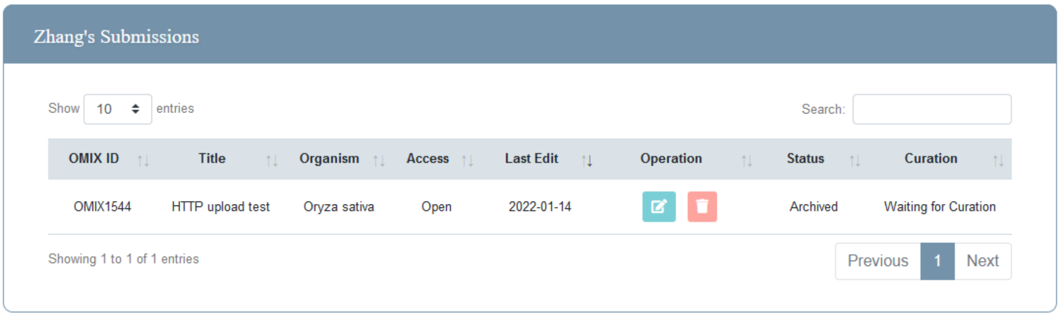
Ministry of Science and Technology Filing Number (Accessions in the Ministry of Science and Technology).

This field is **optional for controlled-access datasets**, but recommended for datasets involving **human genetic resources**.

- For details, refer to the Human Genetic Resource Information Filing Process.
- Only genomic data derived from samples collected from populations within China are required to undergo information filing.

Submission:

After completing and verifying all information, please click “Submit” to submit the dataset. Once submitted, please wait for the system administrator to review it. After approval, the dataset’s Status/Review in the list below will display as “🔗 ”. Clicking it will generate a data sharing link. For details, please refer to “How to Generate a Dataset Sharing Link.”

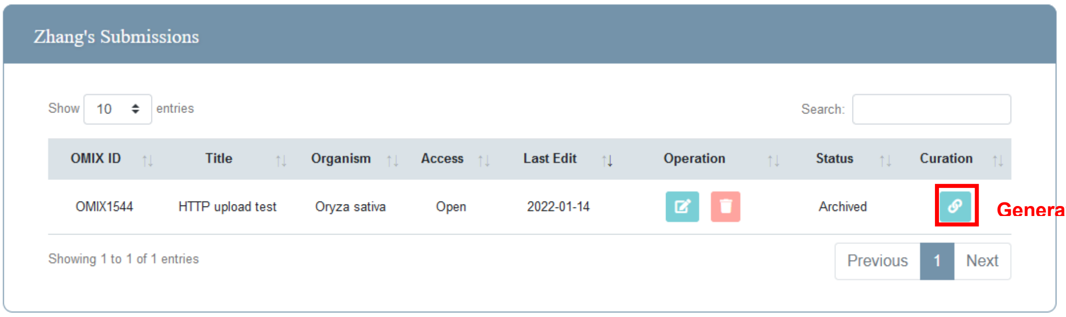


Zhang's Submissions

Show 10 entries Search:

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14	✎ 🗑	Archived	Waiting for Curation

Showing 1 to 1 of 1 entries



Zhang's Submissions

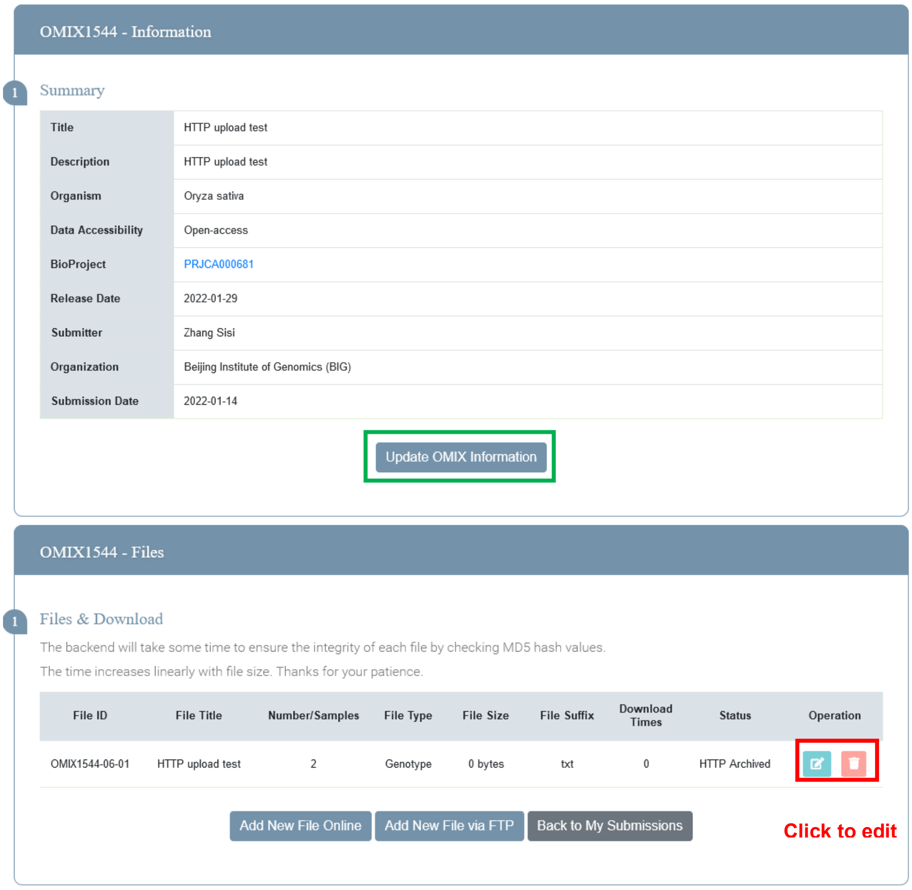
Show 10 entries Search:

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14	✎ 🗑	Archived	🔗

Showing 1 to 1 of 1 entries

Generate Sharing

Before the dataset’s release date, you can modify the basic information and data file details in OMIX at any time. However, data files cannot be replaced directly. If replacement is needed, please first click “Add New File Online” or “Add New File via FTP” to upload the new data, and then delete the original incorrect files.



OMIX1544 - Information

1 Summary

Title	HTTP upload test
Description	HTTP upload test
Organism	Oryza sativa
Data Accessibility	Open-access
BioProject	PRJCA000681
Release Date	2022-01-29
Submitter	Zhang Sisi
Organization	Beijing Institute of Genomics (BIG)
Submission Date	2022-01-14

OMIX1544 - Files

1 Files & Download

The backend will take some time to ensure the integrity of each file by checking MD5 hash values.
The time increases linearly with file size. Thanks for your patience.

File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Status	Operation
OMIX1544-06-01	HTTP upload test	2	Genotype	0 bytes	txt	0	HTTP Archived	

FTP Submission Method

If your total data size is **greater than 100 MB**, it is recommended to use this method for submitting data files and related information.

Step 1: Prepare for Data Pre-upload

① **File Preparation**

- If the number of data files is small, you can upload them directly.
- If you have many files, consider categorizing and compressing them into .tar or .zip packages.
- It is recommended that each compressed file package not exceed **10 GB**.
- **File names must not contain spaces**, as this will affect data archiving.

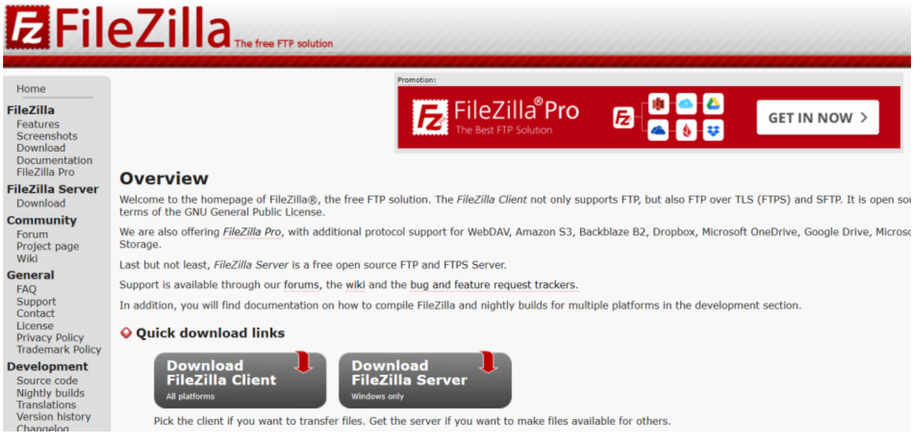
② **Calculate MD5 Checksums**

- The **MD5 checksum** is a 32-character alphanumeric string. Use command-line tools or software to generate the checksum for each compressed package.
  - **Mac/Linux** users can generate it with the command md5sum (Linux) or md5 (Mac).
  - **Windows** users need third-party software. Recommended tools include **WinMD5Free** or **Haozip 2345** (which supports batch MD5 generation).

③ **Data Upload Procedure**

Step 1: Use an FTP client (e.g., **FileZilla Client**) to log in to the FTP server to upload your data. FileZilla is used here as an example.

Step 2: Download FileZilla from the official site: https://filezilla-project.org/. On the download page, click the red box labeled **“Download FileZilla Client”**, and follow the instructions to install.



Overview

Welcome to the homepage of FileZilla®, the free FTP solution. The *FileZilla Client* not only supports FTP, but also FTP over TLS (FTPS) and SFTP. It is open so
terms of the GNU General Public License.

We are also offering *FileZilla Pro*, with additional protocol support for WebDAV, Amazon S3, Backblaze B2, Dropbox, Microsoft OneDrive, Google Drive, Microsc
Storage.

Last but not least, *FileZilla Server* is a free open source FTP and FTPS Server.

Support is available through our forums, the wiki and the bug and feature request trackers.

In addition, you will find documentation on how to compile FileZilla and nightly builds for multiple platforms in the development section.

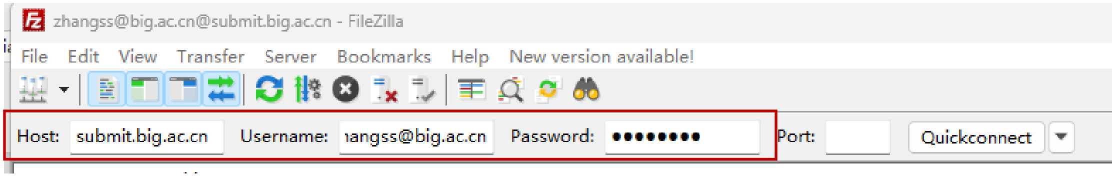
Quick download links

Pick the client if you want to transfer files. Get the server if you want to make files available for others.

Step 3: Open FileZilla. In the interface, enter the following:

- **Host**: submit.big.ac.cn
- **Username**: your email address
- **Password**: your password

  Click **"Quick connect"**. If the status bar shows a successful connection, you're in. For troubleshooting, see "Common FTP Upload Issues".

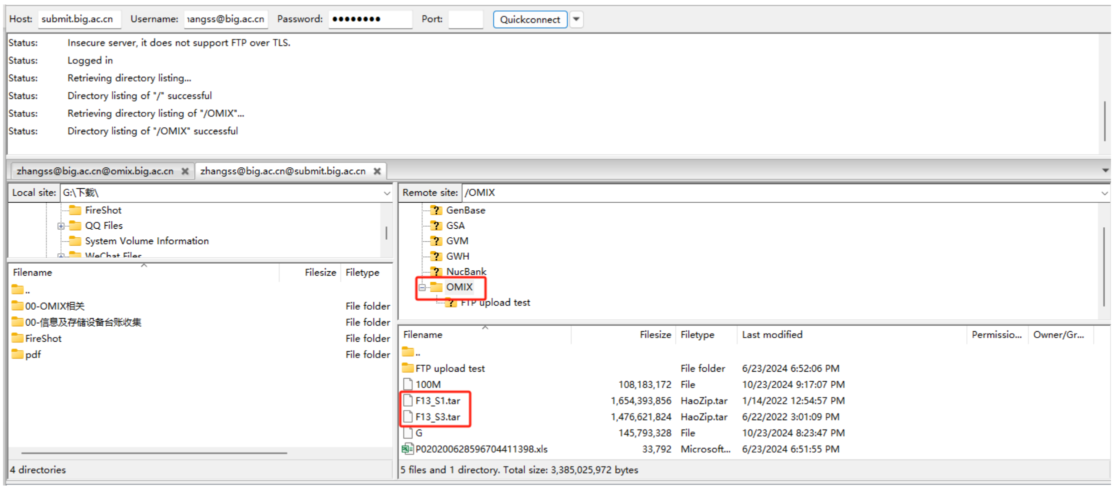


Host: submit.big.ac.cn Username: ıangss@big.ac.cn Password: •••••••• Port:

Step 4:

- In **"Local site"**, navigate to the local path of your data.
- In **"Remote site"**, locate the OMIX folder.

  ⚠ **Note**: Ensure that neither file names nor the **remote path** contain spaces or special characters.



Host: submit.big.ac.cn Username: ıangss@big.ac.cn Password: ••••••••• Port:

```
Status:      Insecure server, it does not support FTP over TLS.
Status:      Logged in
Status:      Retrieving directory listing...
Status:      Directory listing of "/" successful
Status:      Retrieving directory listing of "/OMIX"...
Status:      Directory listing of "/OMIX" successful
```

Local site: G:\下载\

Filename	Filesize	Filetype
..		
00-OMIX相关		File folder
00-信息及存储设备台账收集		File folder
FireShot		File folder
pdf		File folder

4 directories

Remote site: /OMIX

Filename	Filesize	Filetype	Last modified	Permissio...	Owner/Gr...
..					
FTP upload test		File folder	6/23/2024 6:52:06 PM		
100M	108,183,172	File	10/23/2024 9:17:07 PM		
F13_S1.tar	1,654,393,856	HaoZip.tar	1/14/2022 12:54:57 PM		
F13_S3.tar	1,476,621,824	HaoZip.tar	6/22/2022 3:01:09 PM		
G	145,793,328	File	10/23/2024 8:23:47 PM		
P020200628596704411398.xls	33,792	Microsoft...	6/23/2024 6:51:55 PM		

5 files and 1 directory. Total size: 3,385,025,972 bytes

Step 5:

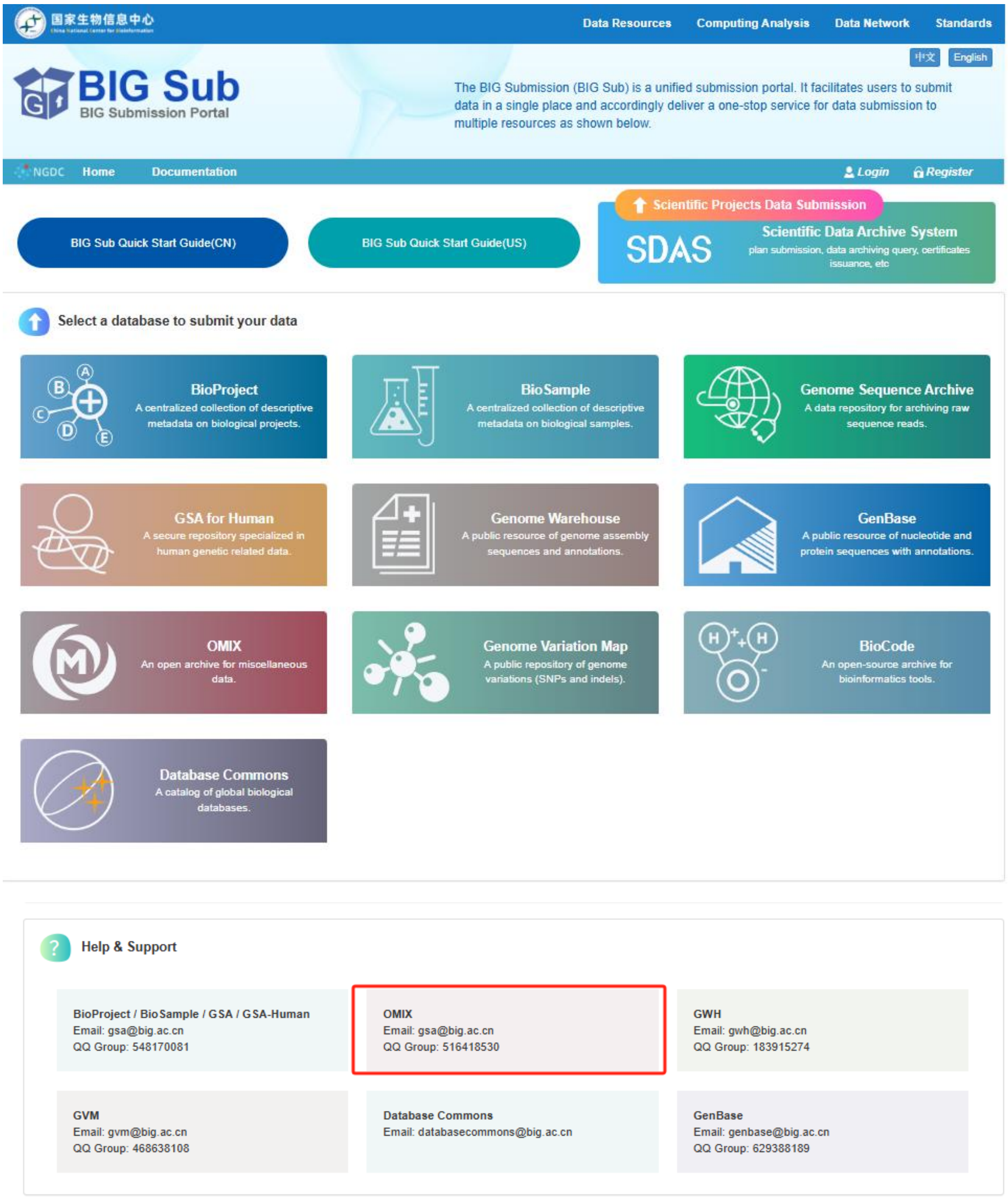
- In the **"Local site"**, right-click the data file or folder to upload and select **"Upload"**, or simply drag it to the **"Remote site"**.

Step 6:

- Uploaded files will appear in the **Queued files** panel and be uploaded in order.
- Successfully uploaded files will move to **Successful transfers**.
- Failed uploads will go to **Failed transfers** — you can retry them using the **Resume** function.

Step 2: Complete Metadata Information

Visit the unified data submission portal of the National Genomics Data Center: https://ngdc.cncb.ac.cn/gsub/



The BIG Submission (BIG Sub) is a unified submission portal. It facilitates users to submit data in a single place and accordingly deliver a one-stop service for data submission to multiple resources as shown below.

plan submission, data archiving query, certificates issuance, etc

Select a database to submit your data

BioProject
A centralized collection of descriptive metadata on biological projects.

BioSample
A centralized collection of descriptive metadata on biological samples.

Genome Sequence Archive
A data repository for archiving raw sequence reads.

GSA for Human
A secure repository specialized in human genetic related data.

Genome Warehouse
A public resource of genome assembly sequences and annotations.

GenBase
A public resource of nucleotide and protein sequences with annotations.

OMIX
An open archive for miscellaneous data.

Genome Variation Map
A public repository of genome variations (SNPs and indels).

BioCode
An open-source archive for bioinformatics tools.

Database Commons
A catalog of global biological databases.

Help & Support

BioProject / BioSample / GSA / GSA-Human
Email: gsa@big.ac.cn
QQ Group: 548170081

OMIX
Email: gsa@big.ac.cn
QQ Group: 516418530

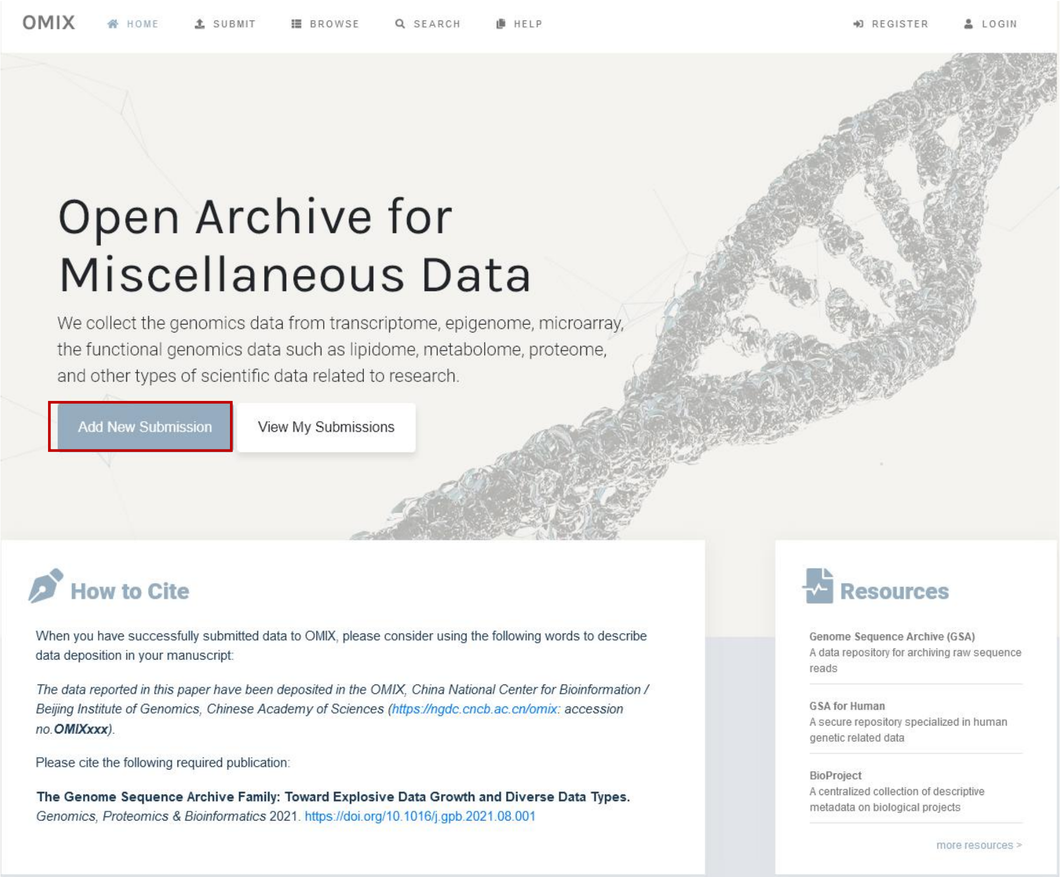
GWH
Email: gwh@big.ac.cn
QQ Group: 183915274

GVM
Email: gvm@big.ac.cn
QQ Group: 468638108

Database Commons
Email: databasecommons@big.ac.cn

GenBase
Email: genbase@big.ac.cn
QQ Group: 629388189

⇩



Open Archive for Miscellaneous Data

We collect the genomics data from transcriptome, epigenome, microarray, the functional genomics data such as lipidome, metabolome, proteome, and other types of scientific data related to research.

How to Cite

When you have successfully submitted data to OMIX, please consider using the following words to describe data deposition in your manuscript:

The data reported in this paper have been deposited in the OMIX, China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (https://ngdc.cncb.ac.cn/omix: accession no. ***OMIXxxx****).*

Please cite the following required publication:

The Genome Sequence Archive Family: Toward Explosive Data Growth and Diverse Data Types. *Genomics, Proteomics & Bioinformatics* 2021. https://doi.org/10.1016/j.gpb.2021.08.001

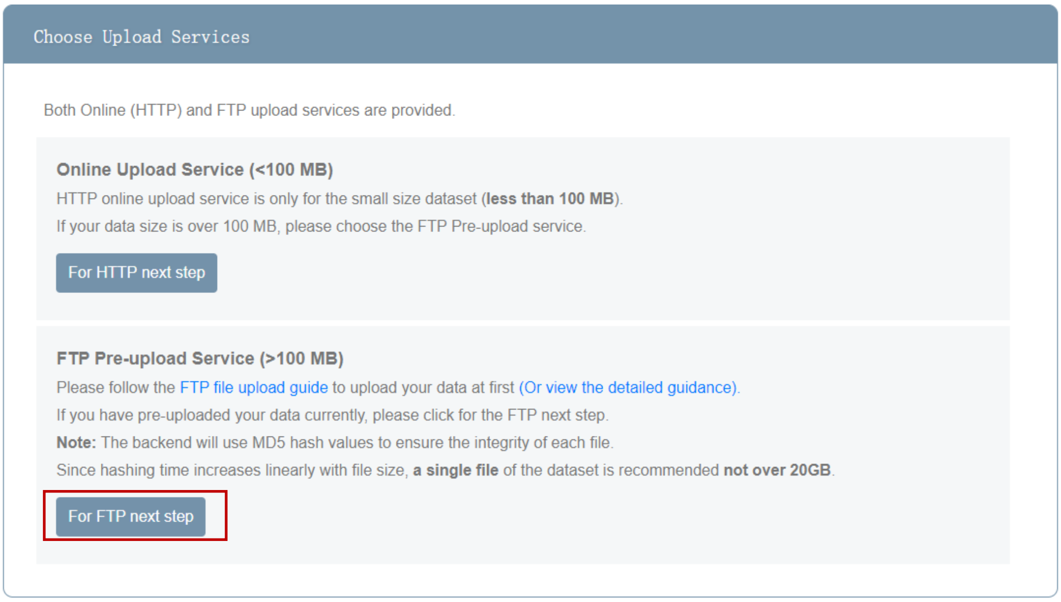
Resources

Genome Sequence Archive (GSA)
A data repository for archiving raw sequence reads

GSA for Human
A secure repository specialized in human genetic related data

BioProject
A centralized collection of descriptive metadata on biological projects

more resources >



Both Online (HTTP) and FTP upload services are provided.

Online Upload Service (<100 MB)

HTTP online upload service is only for the small size dataset (**less than 100 MB**).

If your data size is over 100 MB, please choose the FTP Pre-upload service.

FTP Pre-upload Service (>100 MB)

Please follow the FTP file upload guide to upload your data at first (Or view the detailed guidance).

If you have pre-uploaded your data currently, please click for the FTP next step.

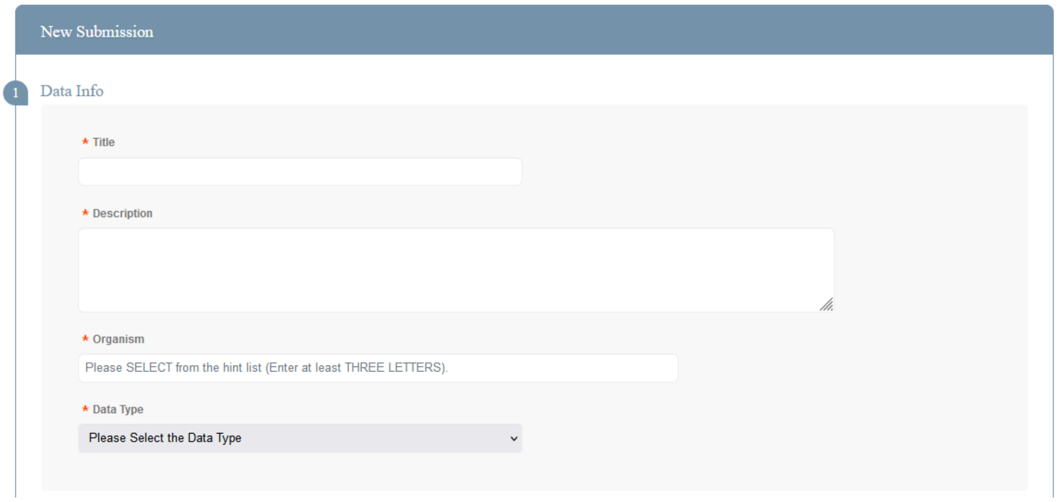
Note: The backend will use MD5 hash values to ensure the integrity of each file.

Since hashing time increases linearly with file size, **a single file** of the dataset is recommended **not over 20GB**.

① Basic Information (Data Info)

Used to collect the dataset name, brief description, and organism information.

- If a file involves multiple organisms, mark as **mixed sample**.


New Submission

1 Data Info

* Title

* Description

* Organism

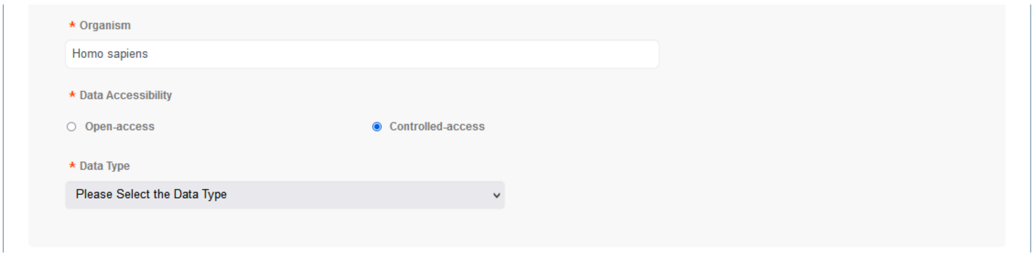
Please SELECT from the hint list (Enter at least THREE LETTERS).

* Data Type

Please Select the Data Type

You can choose two types of data access for human genetic resources:

- **Open-access**: Data will be automatically released upon the publication date or article detection.
- **Controlled-access**: Users must obtain your authorization before downloading, even after the dataset is published.


* Organism

Homo sapiens

* Data Accessibility

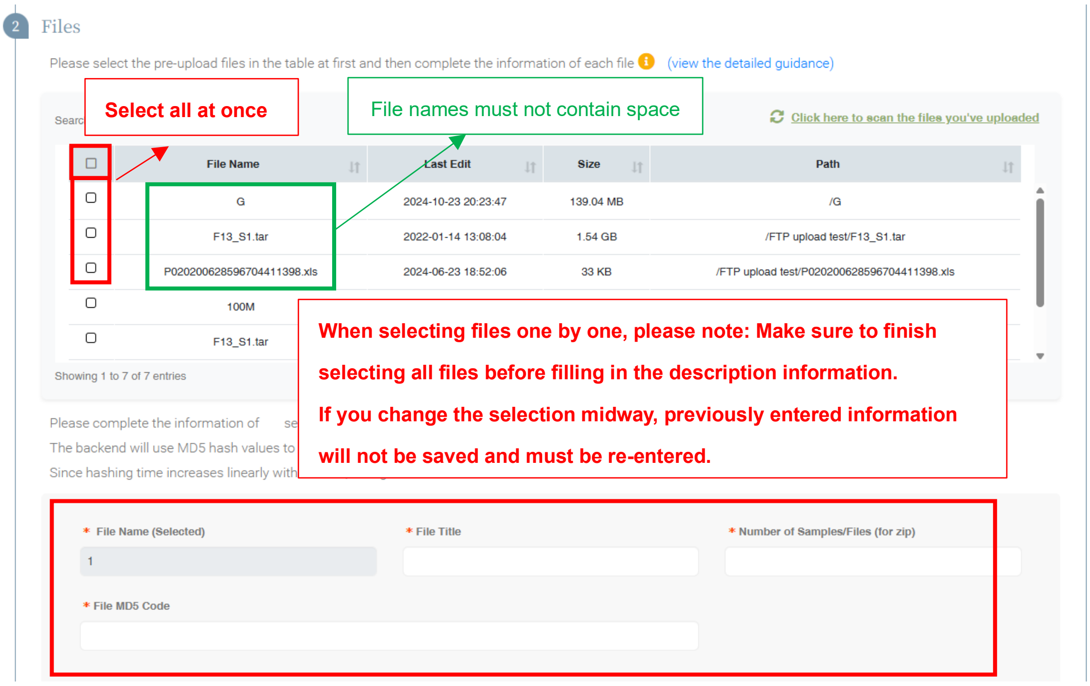
○ Open-access ● Controlled-access

* Data Type

Please Select the Data Type

② File Information (Files)

- Select uploaded files from the list, and then fill in their descriptive information.
- ⚠️ Make sure to **select all relevant files first**, then start filling in their descriptions. If you change the selection midway, all previously entered descriptions will be lost.
- ⚠️ Ensure no file names contain spaces.



2 Files

Please select the pre-upload files in the table at first and then complete the information of each file (view the detailed guidance)

Select all at once

File names must not contain space

Searc

Click here to scan the files you've uploaded

☐	File Name	Last Edit	Size	Path
☐	G	2024-10-23 20:23:47	139.04 MB	/G
☐	F13_S1.tar	2022-01-14 13:08:04	1.54 GB	/FTP upload test/F13_S1.tar
☐	P02020062859670441 1398.xls	2024-06-23 18:52:06	33 KB	/FTP upload test/P02020062859670441 1398.xls
☐	100M			
☐	F13_S1.tar			

Showing 1 to 7 of 7 entries

When selecting files one by one, please note: Make sure to finish selecting all files before filling in the description information.

If you change the selection midway, previously entered information will not be saved and must be re-entered.

Please complete the information of se

The backend will use MD5 hash values to

Since hashing time increases linearly with

* File Name (Selected): 1

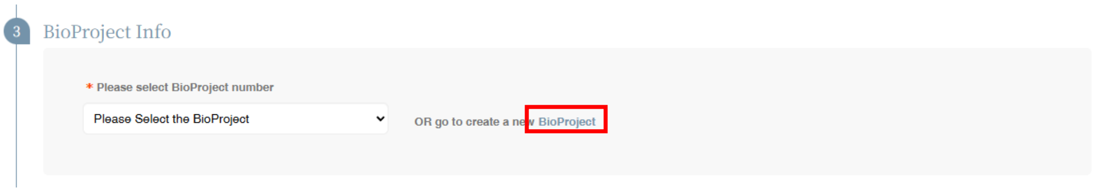
* File Title:

* Number of Samples/Files (for zip):

* File MD5 Code:

③ Project Information (BioProject Info)

- Used to collect information about the BioProject associated with the dataset.
- If you haven't created one yet, you can create it via the BioProject database.
- Note: **BioProject approval status does not affect OMIX data submission.**



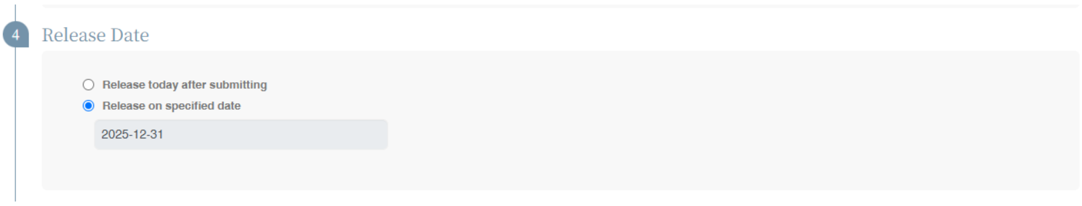
3 BioProject Info

* Please select BioProject number: Please Select the BioProject

OR go to create a new BioProject

④ Dataset Release Date (Release Date)

- Specifies the intended release date of the dataset.



4 Release Date

○ Release today after submitting

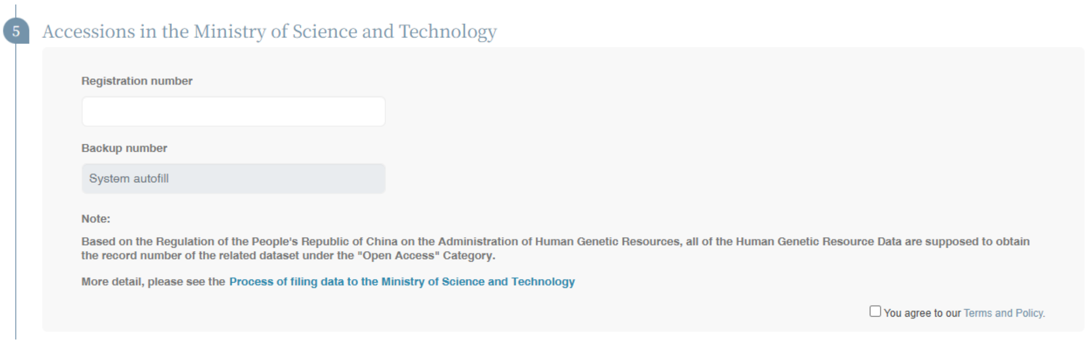
● Release on specified date

2025-12-31

⑤ Ministry of Science and Technology Filing Number (Accessions in the Ministry of Science and Technology).

This field is **optional for controlled-access datasets**, but recommended for datasets involving **human genetic resources**.

- For details, refer to the **Human Genetic Resource Information Filing Process**.
- Only genomic data derived from samples collected from populations within China are required to undergo information filing.



Note:

Based on the Regulation of the People's Republic of China on the Administration of Human Genetic Resources, all of the Human Genetic Resource Data are supposed to obtain the record number of the related dataset under the "Open Access" Category.

More detail, please see the Process of filing data to the Ministry of Science and Technology

☐ You agree to our Terms and Policy.

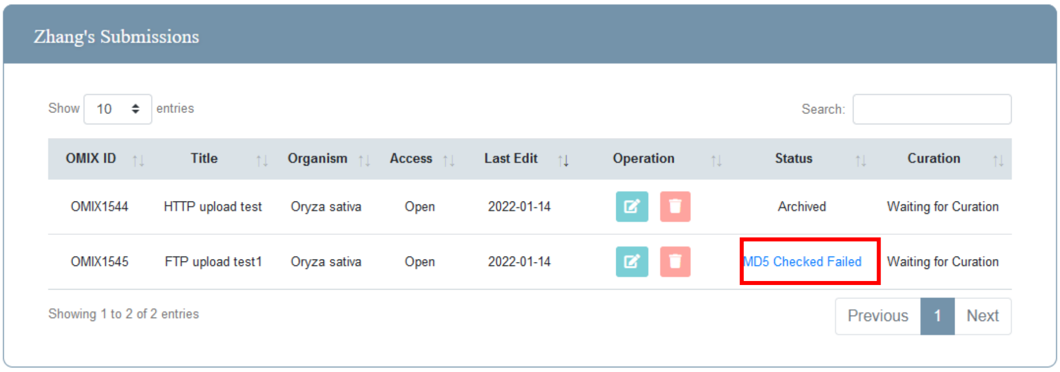
⑥ **Submit**

- After filling and reviewing all information, click **“Submit”** to complete the dataset submission.
- Once submitted, wait for the system administrator to review. Upon approval, the **"Status/Review"** column in the data list will indicate approval.
- You can then generate a **data sharing link**, explained in the guide “How to Generate a Dataset Sharing Link”.

Common Issues and Solutions for Data Submission

Issue 1: File Status Displays "MD5 Checked Failed" — How to Resolve?

If a file in your dataset fails the MD5 checksum validation, the file status under “Status/Review” will show **"MD5 Checked Failed"** (the system will also send a failure notification email). Please follow these steps to resolve the issue:

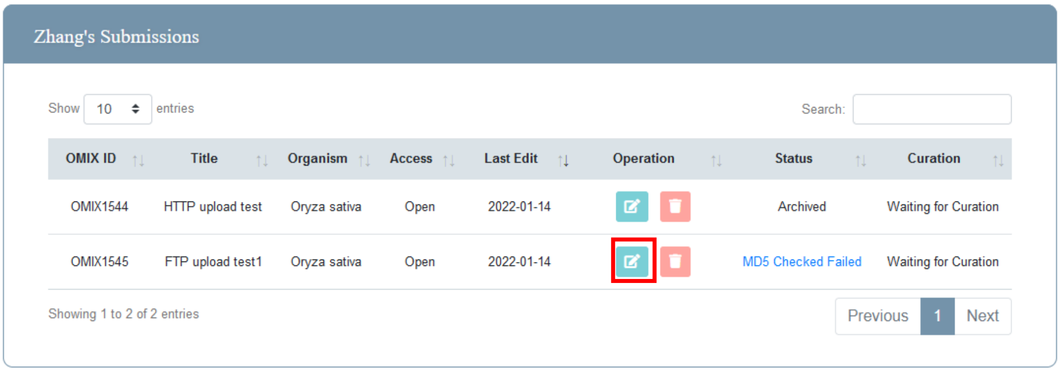


Zhang's Submissions

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14		Archived	Waiting for Curation
OMIX1545	FTP upload test1	Oryza sativa	Open	2022-01-14		MD5 Checked Failed	Waiting for Curation

Showing 1 to 2 of 2 entries

Step 1: Click the edit icon “✎” to open the file with the failed MD5 check.

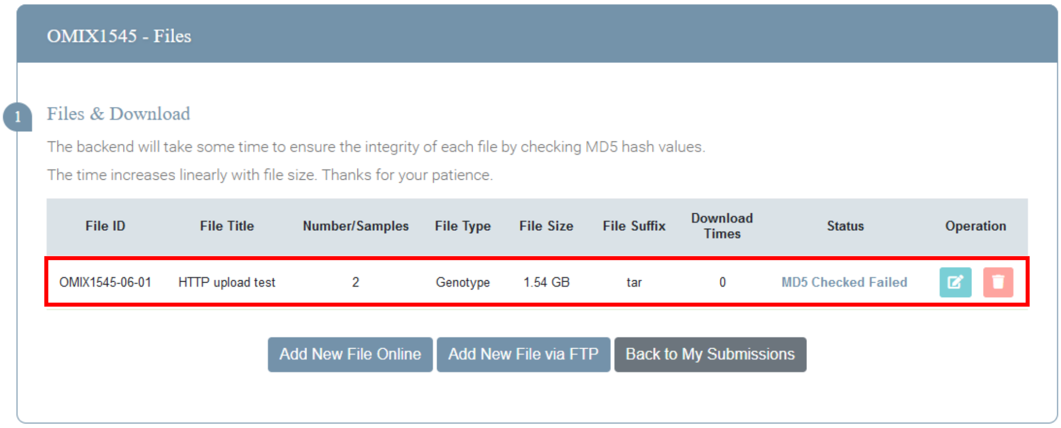


Zhang's Submissions

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14		Archived	Waiting for Curation
OMIX1545	FTP upload test1	Oryza sativa	Open	2022-01-14		MD5 Checked Failed	Waiting for Curation

Showing 1 to 2 of 2 entries

⇩



OMIX1545 - Files

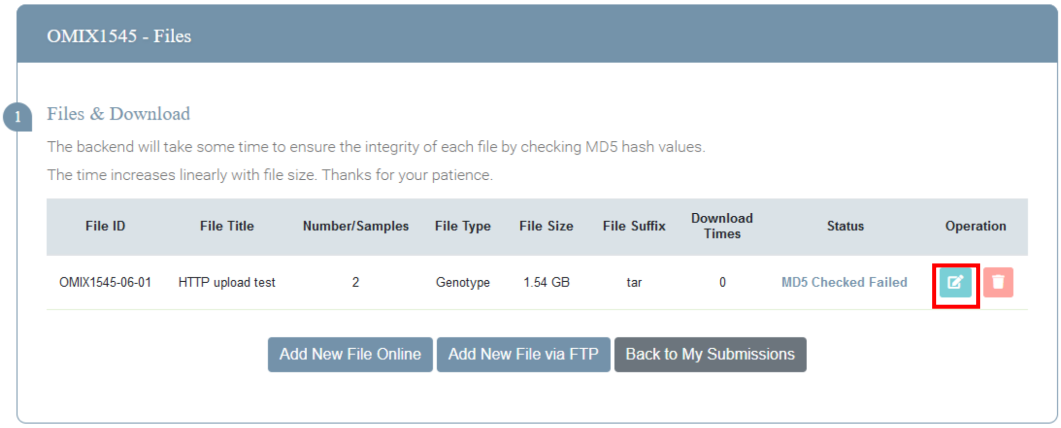
1 Files & Download

The backend will take some time to ensure the integrity of each file by checking MD5 hash values.
The time increases linearly with file size. Thanks for your patience.

File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Status	Operation
OMIX1545-06-01	HTTP upload test	2	Genotype	1.54 GB	tar	0	MD5 Checked Failed	

Step 2: Troubleshoot the MD5 failure cause:

- **Incorrect MD5 entry**: Recalculate the MD5 checksum of the original file and compare it with the submitted value. If they differ, it is a submission error. Please enter the correct MD5 value in the file editing interface and resubmit.

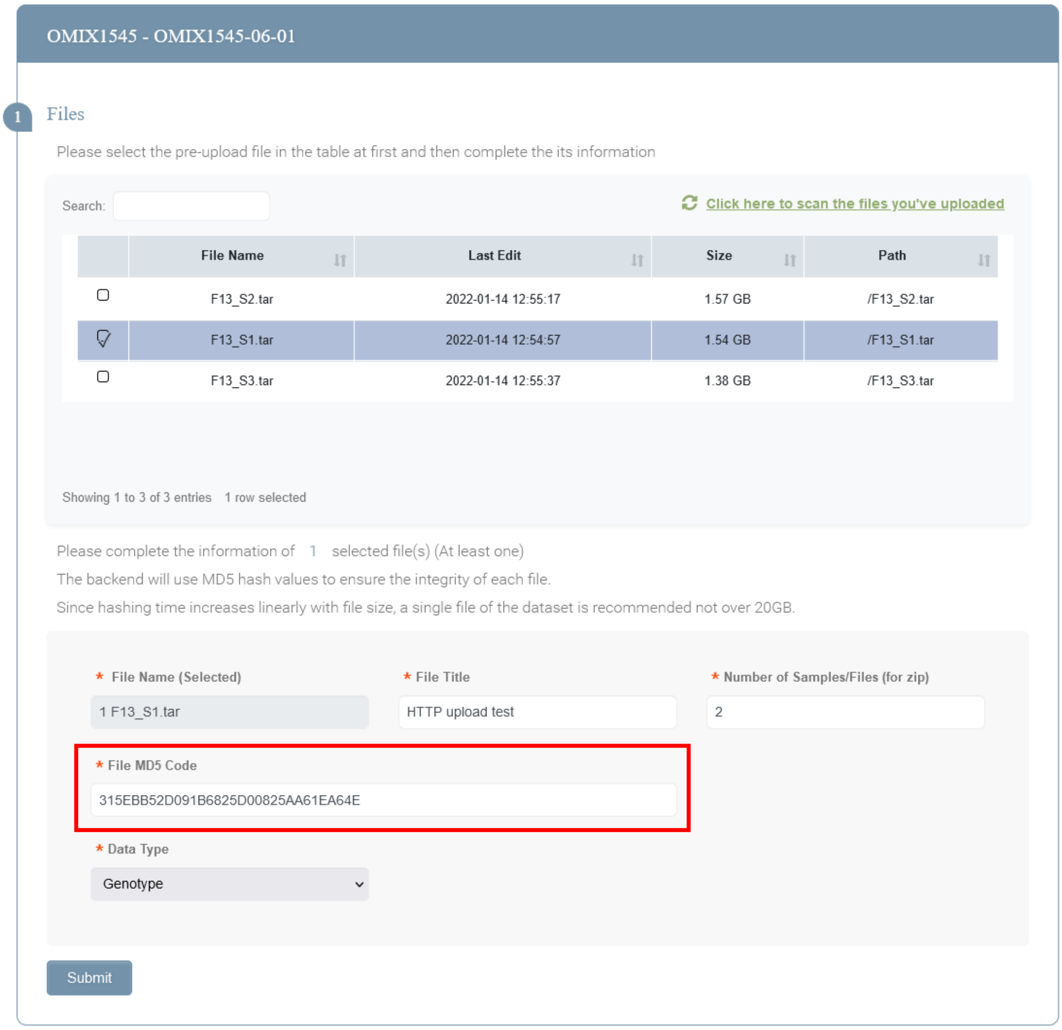


OMIX1545 - Files

1 Files & Download

The backend will take some time to ensure the integrity of each file by checking MD5 hash values.
The time increases linearly with file size. Thanks for your patience.

File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Status	Operation
OMIX1545-06-01	HTTP upload test	2	Genotype	1.54 GB	tar	0	MD5 Checked Failed	



OMIX1545 - OMIX1545-06-01

1 Files

Please select the pre-upload file in the table at first and then complete the its information

Search:

Click here to scan the files you've uploaded

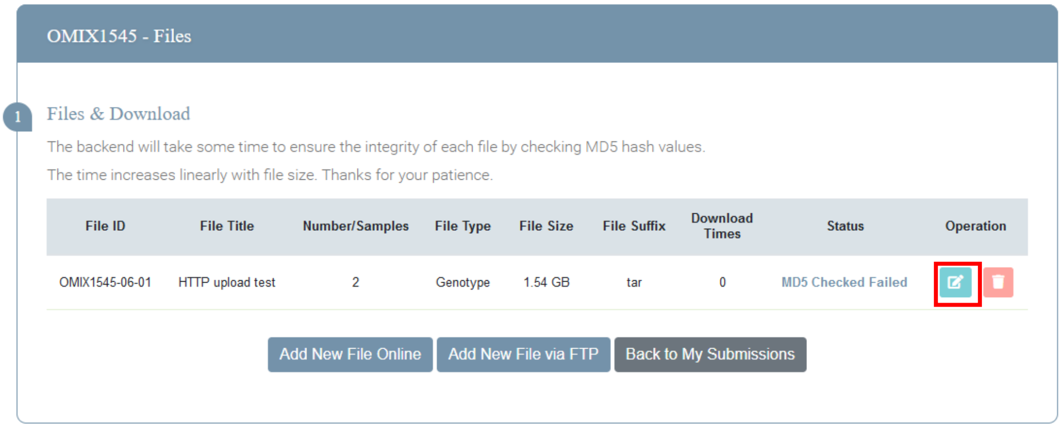
	File Name	Last Edit	Size	Path
☐	F13_S2.tar	2022-01-14 12:55:17	1.57 GB	/F13_S2.tar
☑	F13_S1.tar	2022-01-14 12:54:57	1.54 GB	/F13_S1.tar
☐	F13_S3.tar	2022-01-14 12:55:37	1.38 GB	/F13_S3.tar

Showing 1 to 3 of 3 entries 1 row selected

Please complete the information of 1 selected file(s) (At least one)
The backend will use MD5 hash values to ensure the integrity of each file.
Since hashing time increases linearly with file size, a single file of the dataset is recommended not over 20GB.

* File Name (Selected): 1 F13_S1.tar
* File Title: HTTP upload test
* Number of Samples/Files (for zip): 2
* File MD5 Code: 315EBB52D091B6825D00825AA61EA64E
* Data Type: Genotype

- **File transfer error**: If the recalculated MD5 matches the originally submitted value, the issue may stem from an upload error. Re-upload the data file. For detailed instructions, refer to the “Pre-upload Data” section. After confirming the upload is complete, follow the modification steps again.

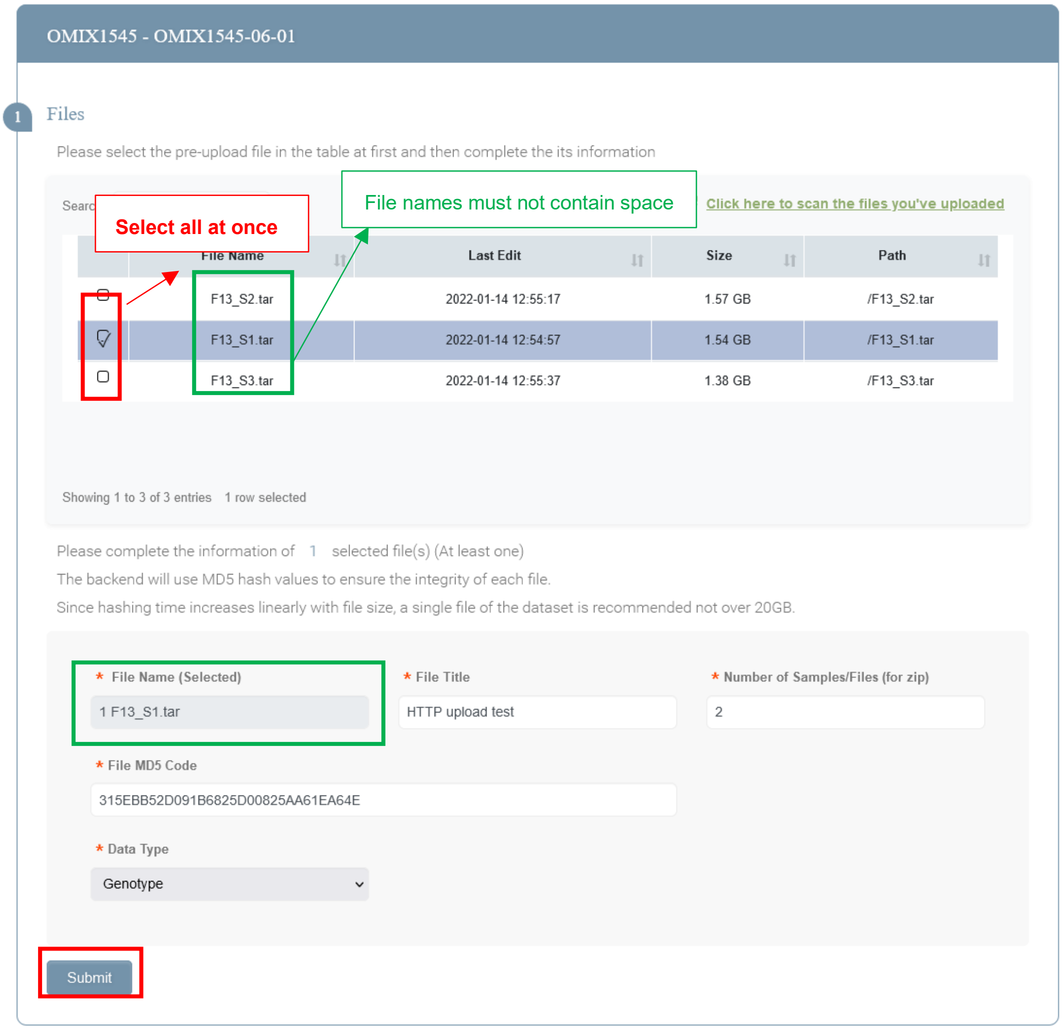


OMIX1545 - Files

1 Files & Download

The backend will take some time to ensure the integrity of each file by checking MD5 hash values.
The time increases linearly with file size. Thanks for your patience.

File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Status	Operation
OMIX1545-06-01	HTTP upload test	2	Genotype	1.54 GB	tar	0	MD5 Checked Failed	



OMIX1545 - OMIX1545-06-01

1 Files

Please select the pre-upload file in the table at first and then complete the its information

Select all at once

File names must not contain space

	File Name	Last Edit	Size	Path
☐	F13_S2.tar	2022-01-14 12:55:17	1.57 GB	/F13_S2.tar
☑	F13_S1.tar	2022-01-14 12:54:57	1.54 GB	/F13_S1.tar
☐	F13_S3.tar	2022-01-14 12:55:37	1.38 GB	/F13_S3.tar

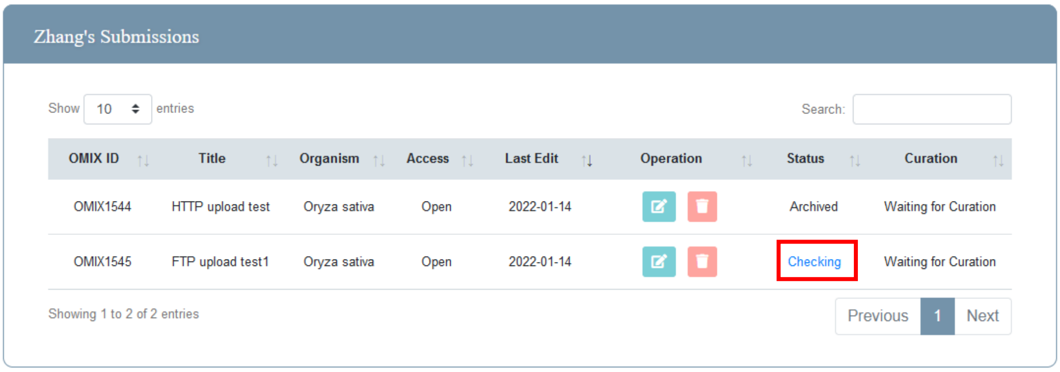
Showing 1 to 3 of 3 entries 1 row selected

Please complete the information of 1 selected file(s) (At least one)
The backend will use MD5 hash values to ensure the integrity of each file.
Since hashing time increases linearly with file size, a single file of the dataset is recommended not over 20GB.

* File Name (Selected): 1 F13_S1.tar
* File Title: HTTP upload test
* Number of Samples/Files (for zip): 2
* File MD5 Code: 315EBB52D091B6825D00825AA61EA64E
* Data Type: Genotype

Issue 2: File Status Remains "Checking" — What to Do?

When a data file is undergoing MD5 verification, its status will show **"Checking"** under "Status/Review." Verification time increases with file size — it is recommended that individual files remain under 20 GB.



Zhang's Submissions

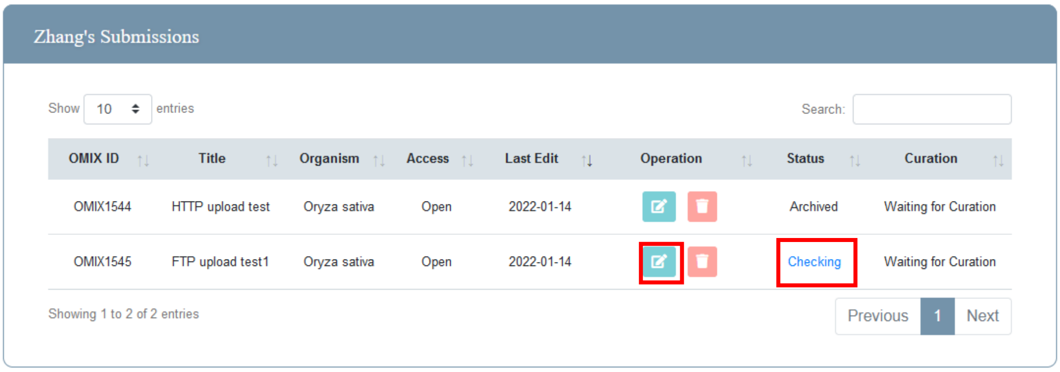
Show 10 entries Search:

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14		Archived	Waiting for Curation
OMIX1545	FTP upload test1	Oryza sativa	Open	2022-01-14		Checking	Waiting for Curation

Showing 1 to 2 of 2 entries

However, if a file smaller than 20 GB stays in the "Checking" state for over 24 hours, consider the following:

- **File upload failed**: Click the edit icon to verify whether the file was uploaded correctly. If not, please delete it and re-submit.

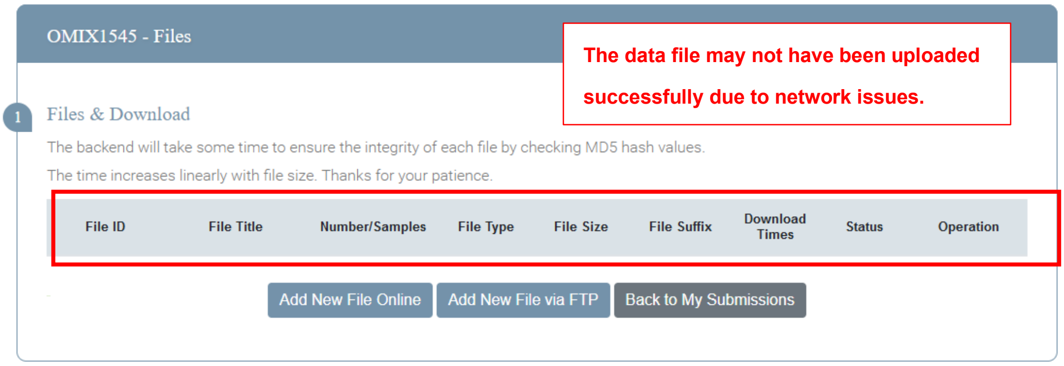


Zhang's Submissions

Show 10 entries Search:

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14		Archived	Waiting for Curation
OMIX1545	FTP upload test1	Oryza sativa	Open	2022-01-14		Checking	Waiting for Curation

Showing 1 to 2 of 2 entries



OMIX1545 - Files

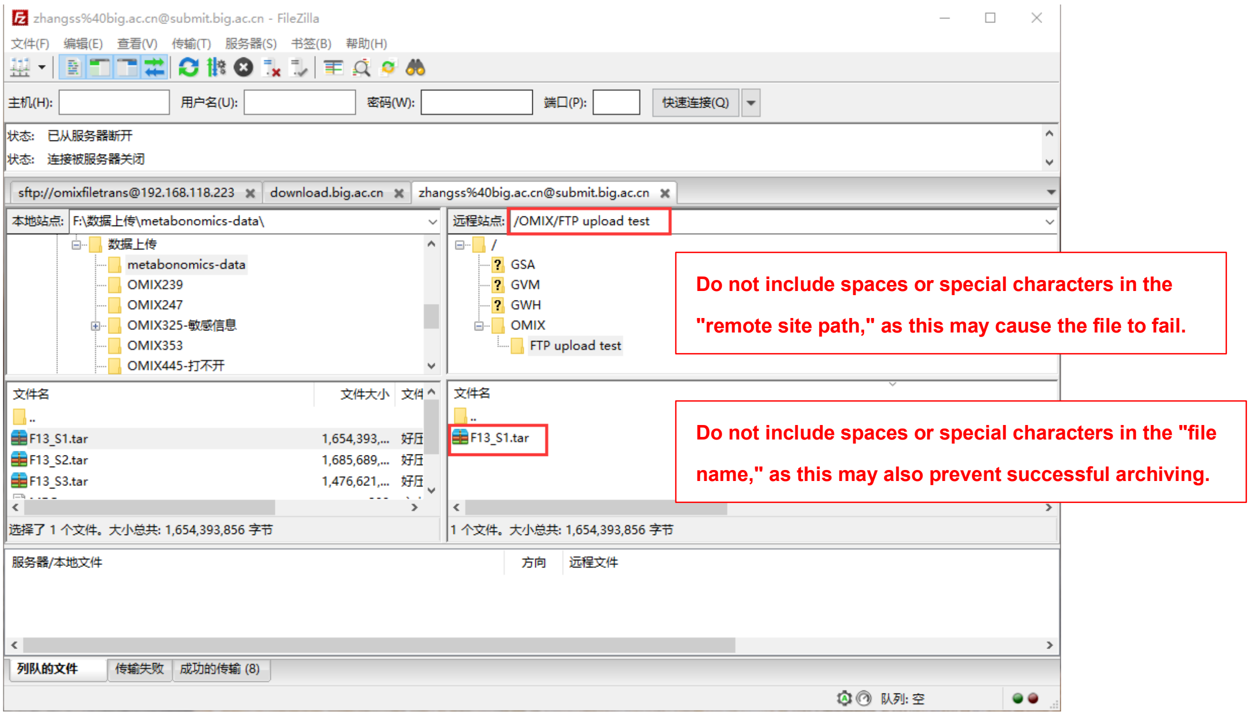
The data file may not have been uploaded successfully due to network issues.

1 Files & Download

The backend will take some time to ensure the integrity of each file by checking MD5 hash values.
The time increases linearly with file size. Thanks for your patience.

File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Status	Operation

- **Archiving error**: Check whether the file name or remote site path includes spaces or special characters. These may interfere with archiving. If none of these apply, contact the database support team for troubleshooting.



zhangss%40big.ac.cn@submit.big.ac.cn - FileZilla

状态: 已从服务器断开
状态: 连接被服务器关闭

本地站点: F:\数据上传\metabonomics-data\

远程站点: /OMIX/FTP upload test

文件名	文件大小
..	
F13_S1.tar	1,654,393,...
F13_S2.tar	1,685,689,...
F13_S3.tar	1,476,621,...

文件名
..
F13_S1.tar

选择了 1 个文件。大小总共: 1,654,393,856 字节

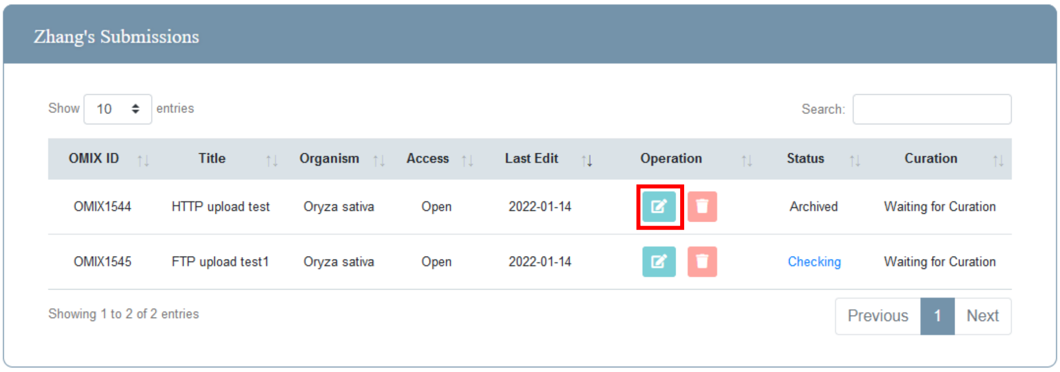
1 个文件。大小总共: 1,654,393,856 字节

Do not include spaces or special characters in the "remote site path," as this may cause the file to fail.

Do not include spaces or special characters in the "file name," as this may also prevent successful archiving.

Issue 3: How to Edit Dataset Basic Information?

Before a dataset is published on OMIX, you can edit the metadata by clicking the edit icon "✎" on the dataset list page to enter the **Update OMIX Information** interface.



Zhang's Submissions

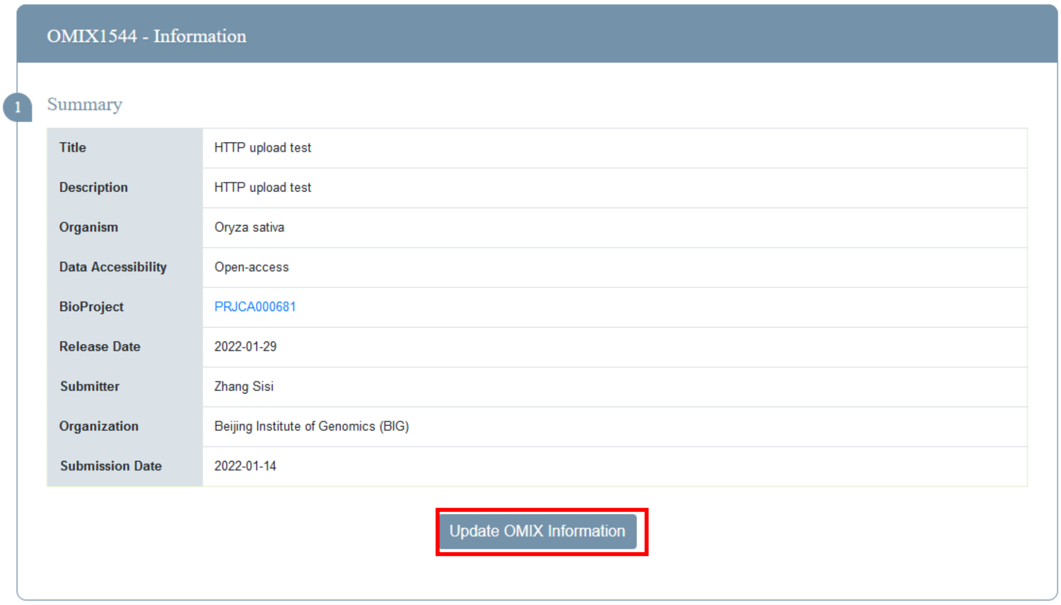
Show 10 entries

Search:

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14	✎ 🗑	Archived	Waiting for Curation
OMIX1545	FTP upload test1	Oryza sativa	Open	2022-01-14	✎ 🗑	Checking	Waiting for Curation

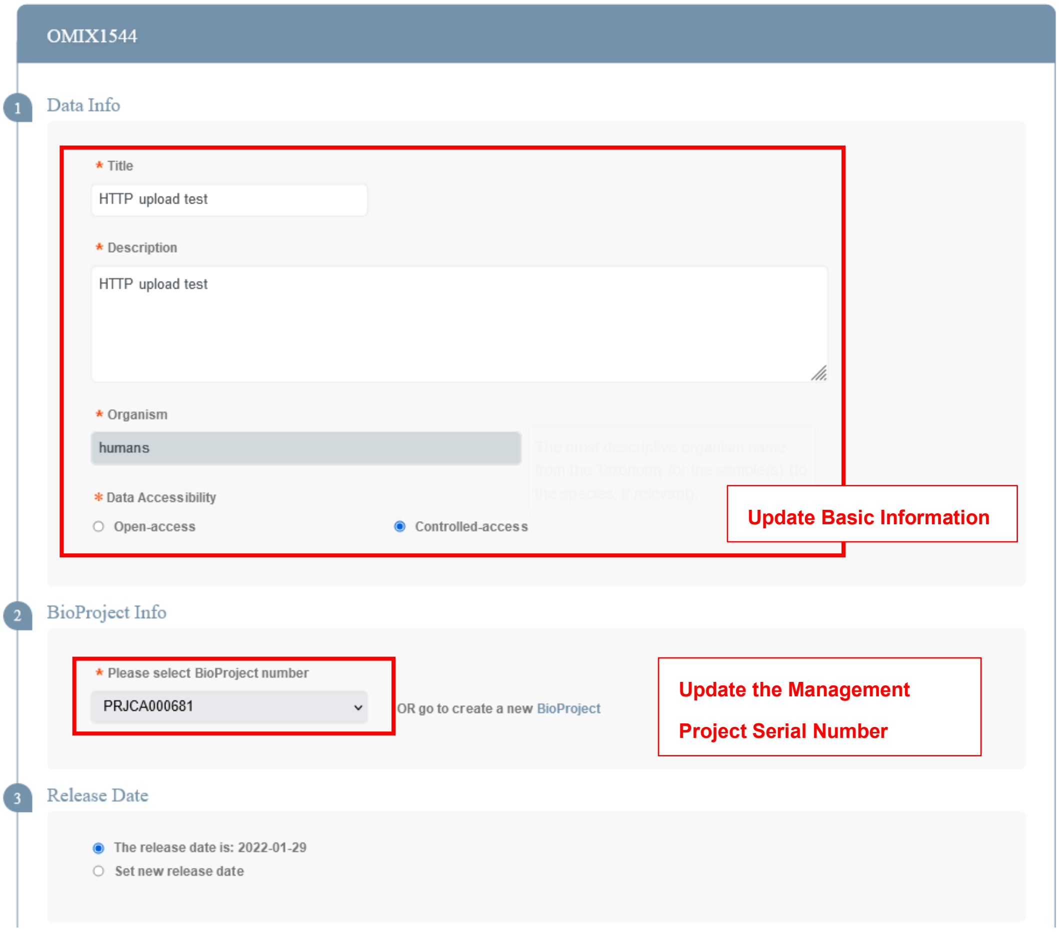
Showing 1 to 2 of 2 entries

⇩


OMIX1544 - Information

1 Summary

Title	HTTP upload test
Description	HTTP upload test
Organism	Oryza sativa
Data Accessibility	Open-access
BioProject	PRJCA000681
Release Date	2022-01-29
Submitter	Zhang Sisi
Organization	Beijing Institute of Genomics (BIG)
Submission Date	2022-01-14


OMIX1544

1 Data Info

* Title: HTTP upload test

* Description: HTTP upload test

* Organism: humans

The most descriptive organism name from the Taxonomy for the sample(s) (to the species, if relevant).

* Data Accessibility: ○ Open-access ● Controlled-access

2 BioProject Info

* Please select BioProject number: PRJCA000681

OR go to create a new BioProject

3 Release Date

● The release date is: 2022-01-29

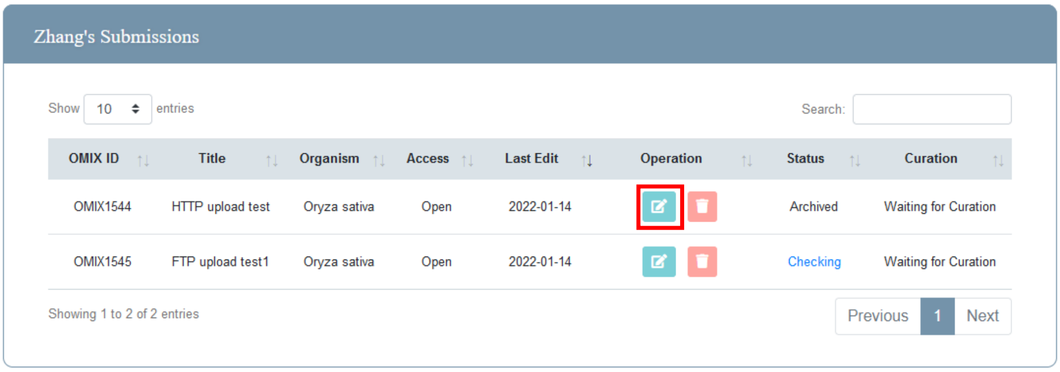
○ Set new release date

Issue 4: How to Release a Dataset Early?

Before publication, you can update the release time by clicking the edit icon “✎ ”

on the dataset list page and accessing the **Update OMIX Information** section.

⚠ Note: The system may take some time to process the release depending on data volume. Please be patient.

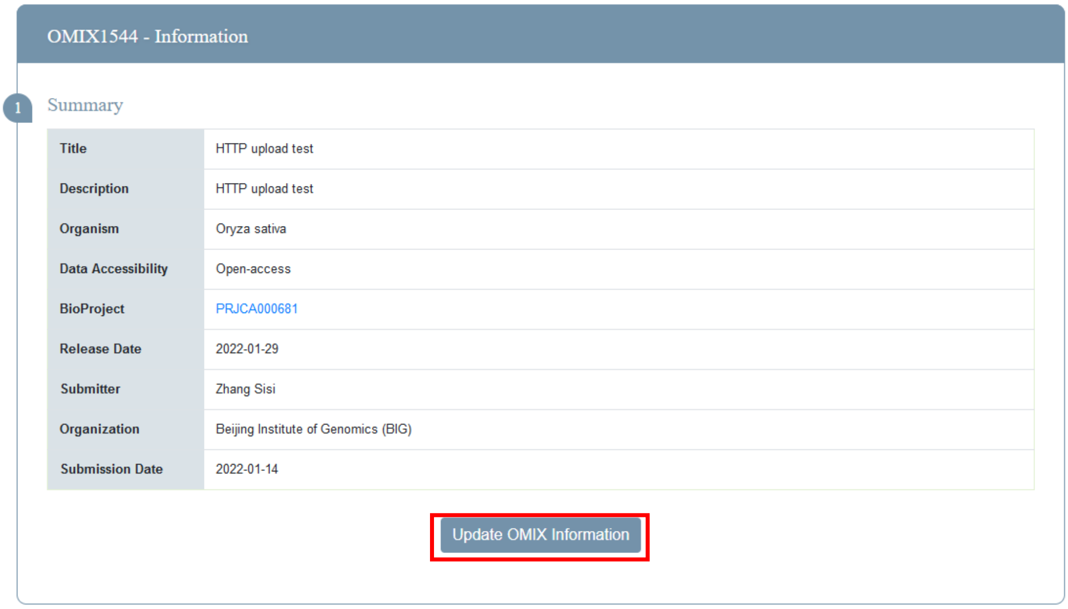


Zhang's Submissions

Show 10 entries Search:

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14	✎ 🗑	Archived	Waiting for Curation
OMIX1545	FTP upload test1	Oryza sativa	Open	2022-01-14	✎ 🗑	Checking	Waiting for Curation

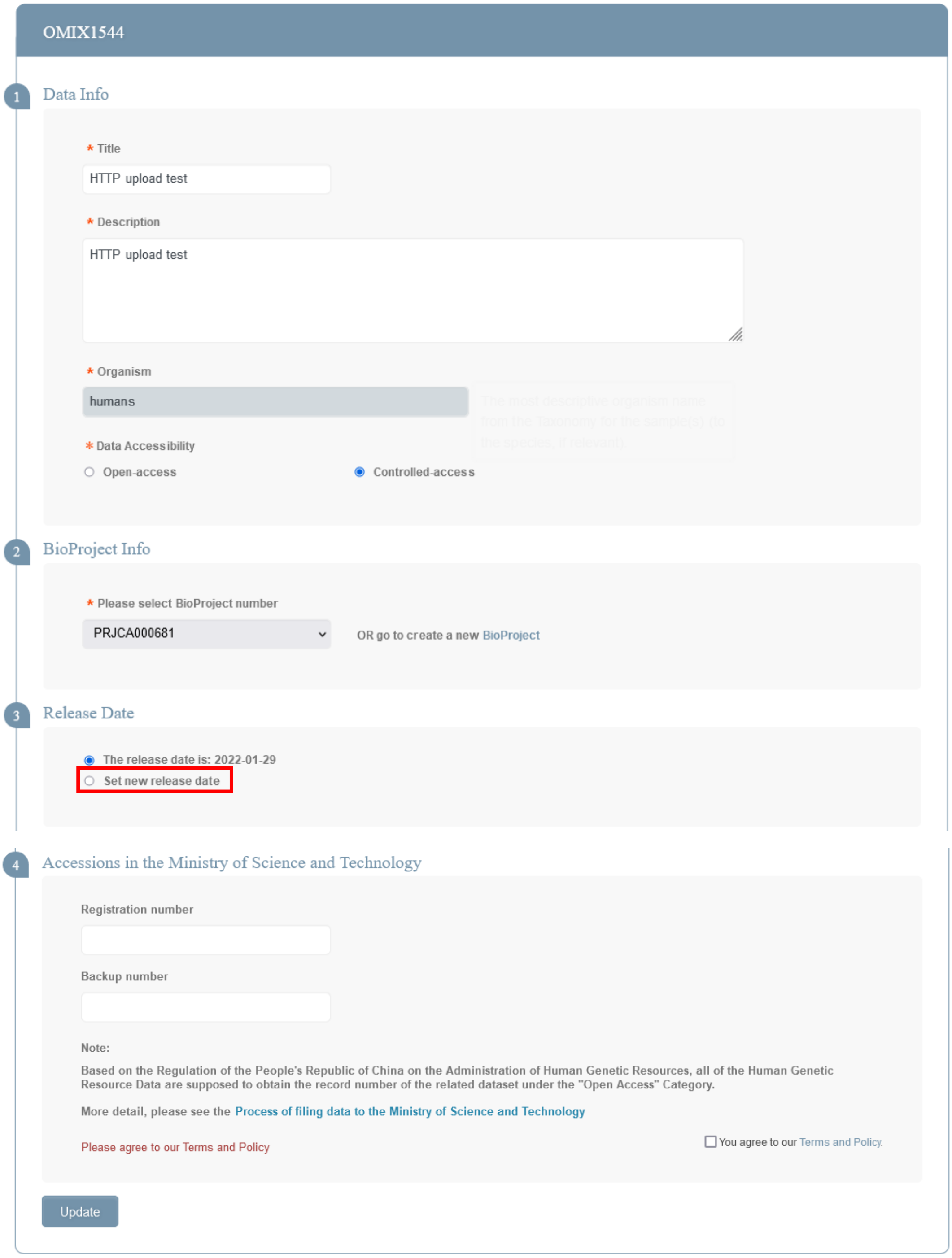
Showing 1 to 2 of 2 entries



OMIX1544 - Information

1 Summary

Title	HTTP upload test
Description	HTTP upload test
Organism	Oryza sativa
Data Accessibility	Open-access
BioProject	PRJCA000681
Release Date	2022-01-29
Submitter	Zhang Sisi
Organization	Beijing Institute of Genomics (BIG)
Submission Date	2022-01-14



OMIX1544

1 Data Info

* Title: HTTP upload test

* Description: HTTP upload test

* Organism: humans

The most descriptive organism name from the Taxonomy for the sample(s) (to the species, if relevant).

* Data Accessibility: ○ Open-access ● Controlled-access

2 BioProject Info

* Please select BioProject number: PRJCA000681

OR go to create a new BioProject

3 Release Date

● The release date is: 2022-01-29

○ Set new release date

4 Accessions in the Ministry of Science and Technology

Registration number:

Backup number:

Note:

Based on the Regulation of the People's Republic of China on the Administration of Human Genetic Resources, all of the Human Genetic Resource Data are supposed to obtain the record number of the related dataset under the "Open Access" Category.

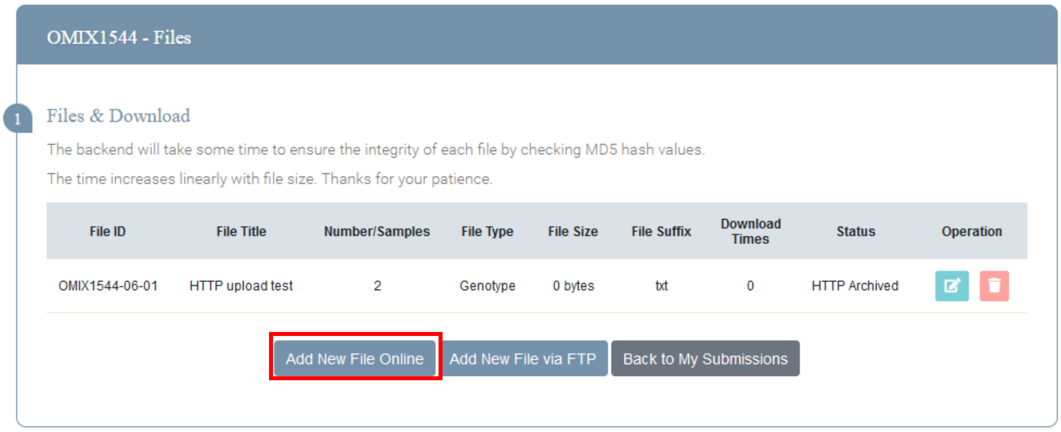
More detail, please see the Process of filing data to the Ministry of Science and Technology

Please agree to our Terms and Policy

☐ You agree to our Terms and Policy.

Issue 5: How to Add or Replace Data Files?

Via Web Interface: Use the “Add or Replace File” function (for replacement, delete the original file first). Enter the file submission page to complete both metadata and file upload.



The backend will take some time to ensure the integrity of each file by checking MD5 hash values.
The time increases linearly with file size. Thanks for your patience.

File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Status	Operation
OMIX1544-06-01	HTTP upload test	2	Genotype	0 bytes	txt	0	HTTP Archived	

Note: For **“Number of Samples/Files”**, input the number of samples represented. If the file is a compressed archive or sample count is unavailable, you may enter the number of files.

Via FTP: Use the “Add or Replace File” function (delete original files before replacement). Go to the file submission page, select the already uploaded files from the list, and fill in metadata. For upload instructions, see **“FTP Pre-upload”**.

☑ Tips:

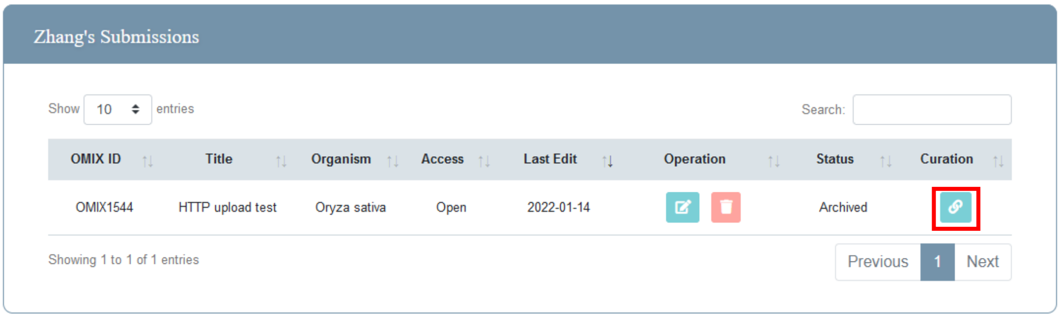
- Always complete all selections before filling in metadata.
- Avoid changing selections mid-process, or all filled data will be lost.
- Do **not** use spaces in file names.

Issue 6: How to Generate a Dataset Sharing Link?

After all files in the dataset are archived, the **Status/Review** column will indicate availability for sharing. You can then click “🔗” to generate an online sharing link.

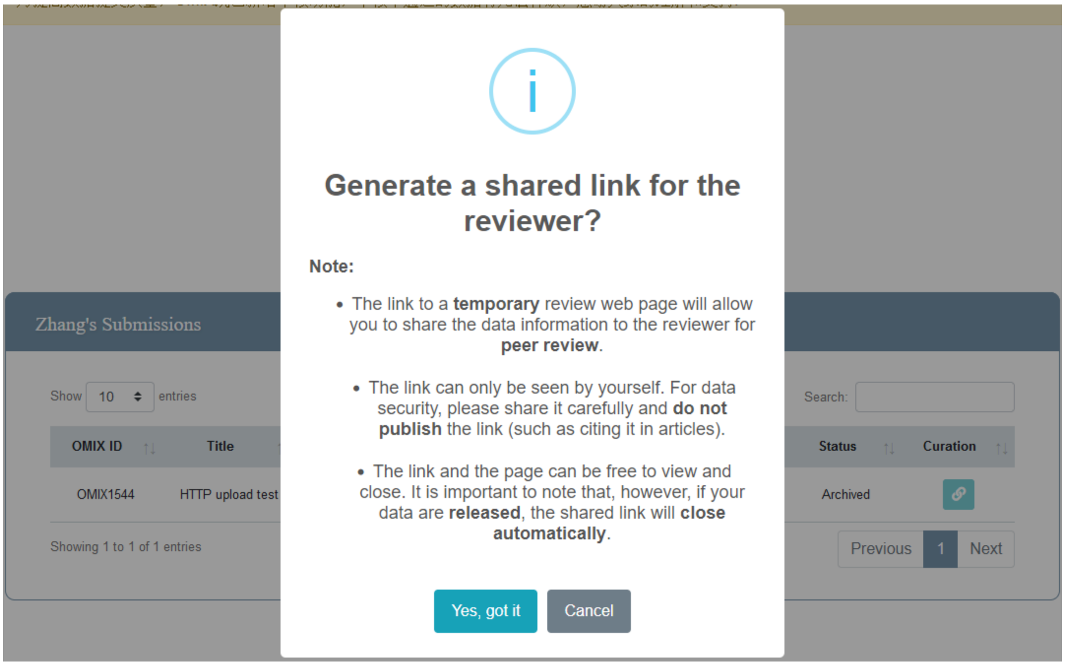
⚠ Notes:

- For data security, shared links **do not support file downloads**.
- **Do not include links in publications** — they will become invalid after the dataset is published.


Zhang's Submissions

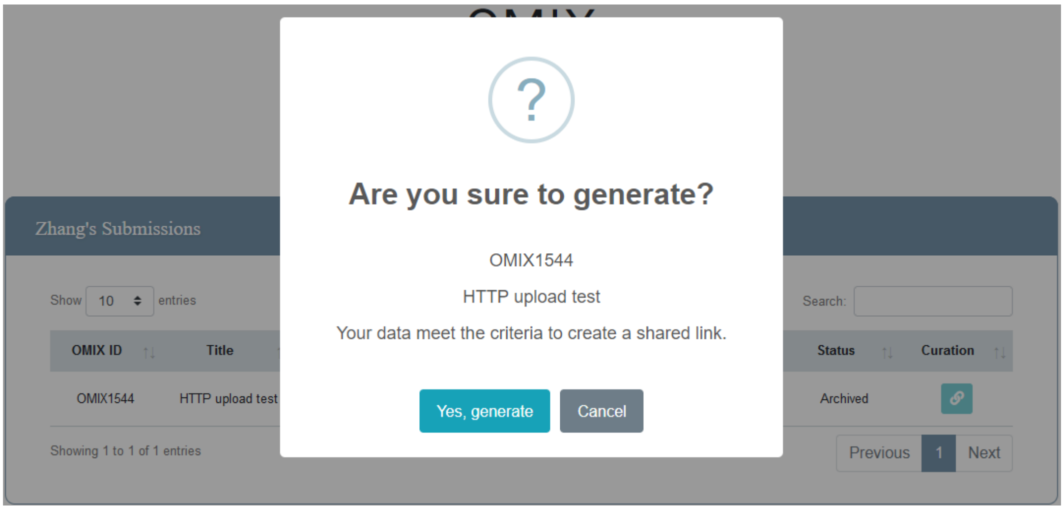
OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14		Archived	

Showing 1 to 1 of 1 entries


Generate a shared link for the reviewer?

Note:

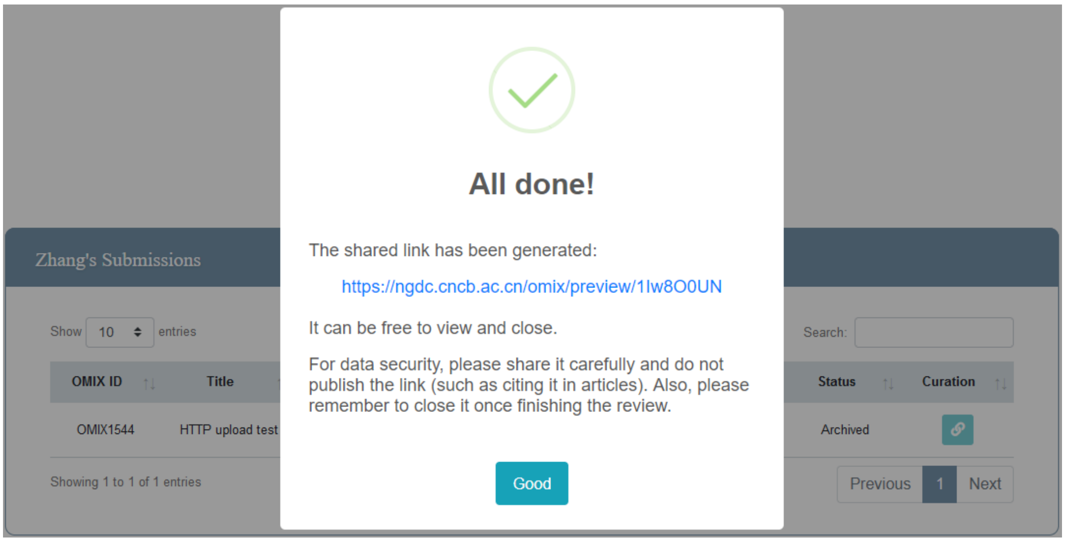
- The link to a **temporary** review web page will allow you to share the data information to the reviewer for **peer review**.
- The link can only be seen by yourself. For data security, please share it carefully and **do not publish** the link (such as citing it in articles).
- The link and the page can be free to view and close. It is important to note that, however, if your data are **released**, the shared link will **close automatically**.


Are you sure to generate?

OMIX1544

HTTP upload test

Your data meet the criteria to create a shared link.



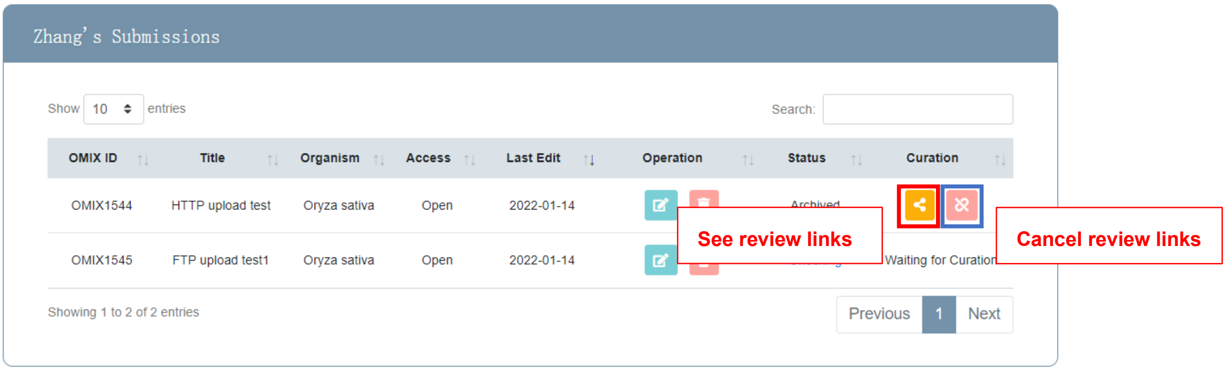
All done!

The shared link has been generated:

https://ngdc.cncb.ac.cn/omix/preview/1Iw8O0UN

It can be free to view and close.

For data security, please share it carefully and do not publish the link (such as citing it in articles). Also, please remember to close it once finishing the review.



Zhang's Submissions

Show 10 entries

Search:

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX1544	HTTP upload test	Oryza sativa	Open	2022-01-14		Archived	
OMIX1545	FTP upload test1	Oryza sativa	Open	2022-01-14			Waiting for Curation

Showing 1 to 2 of 2 entries

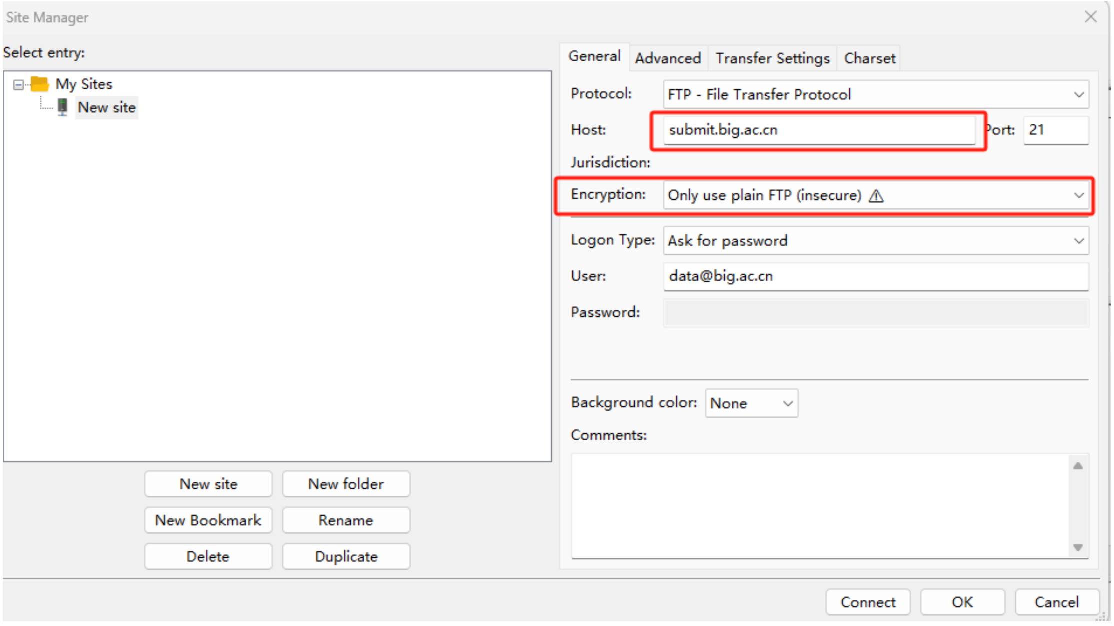
Common FTP Upload Issues and Solutions

Issue 1: Error Message "AUTH SSL" Appears During Login

主机(H): submit.big.ac.cn 用户名(U): chentt@big.ac.cn 密码(W): ●●●●●●●●● 端口(P):

```
响应:  220 Service ready for new user.
命令:  AUTH TLS
响应:  431 Service is unavailable.
命令:  AUTH SSL
响应:  431 Service is unavailable.
```

If the status bar displays an **"AUTH SSL"** error during FTP login, open the **Site Manager** by clicking “File” on the menu bar. Then, change the **Encryption** setting to **“Only use plain FTP”** or a similar option, and make sure to fill in the correct **Host** address: *submit.big.ac.cn,* along with your **username and password**. Finally, click **Connect**.



Protocol: FTP - File Transfer Protocol

Host: submit.big.ac.cn

Port: 21

Encryption: Only use plain FTP (insecure) ⚠

Logon Type: Ask for password

User: data@big.ac.cn

Background color: None

Issue 2: Error Message "MLSD" Appears During Login

主机(H): submit.big.ac.cn 用户名(U): atanan@163.com 密码(W): ●●●●●●●●●●●● 端口(P):

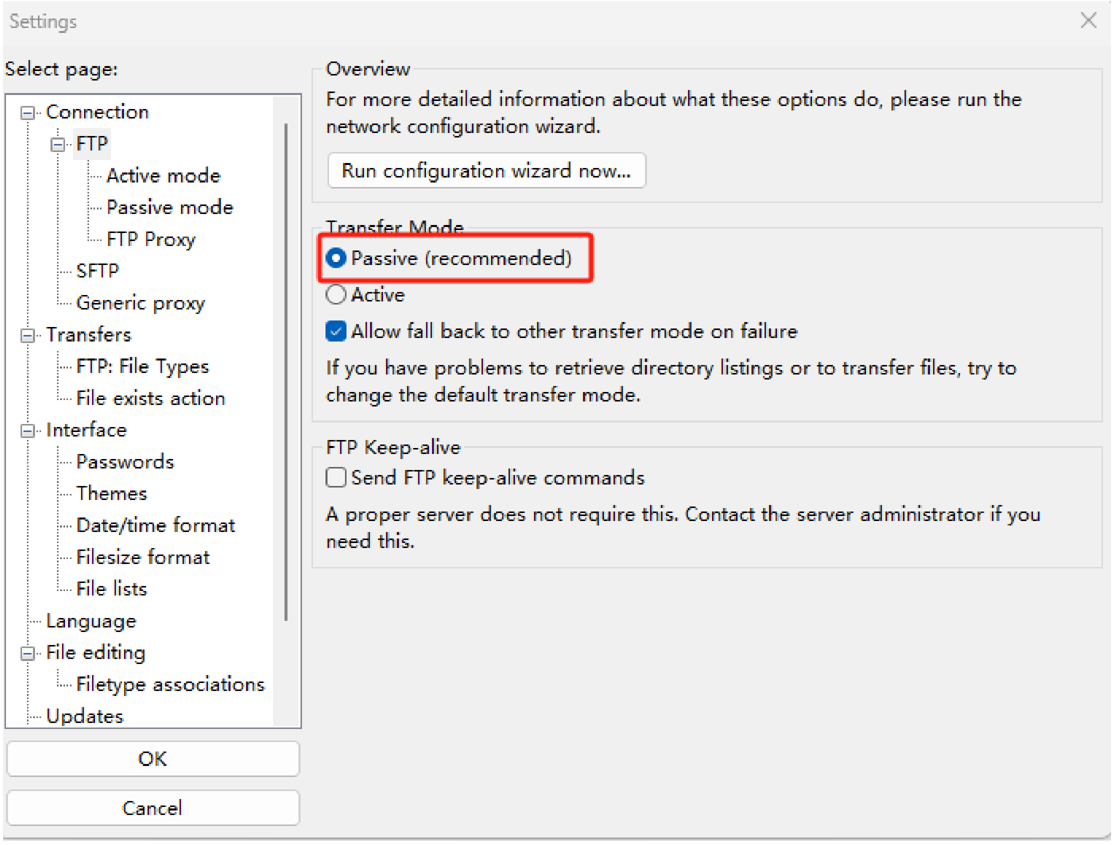
```
状态:	读取目录列表...
命令:	PWD
响应:	257 "/" is current directory.
命令:	TYPE I
响应:	200 Command TYPE okay.
命令:	PORT 192,168,1,2,235,88
响应:	200 Command PORT okay.
命令:	MLSD
响应:	150 File status okay; about to open data connection.
错误:	20 秒后无活动，连接超时
错误:	读取目录列表失败
```

本地站点: C:\Users\wanglin\ 远程站点:

Solution:

If the status bar displays an **"MLSD"** error and shows **"Failed to retrieve directory listing"**, open FileZilla and go to:

Edit → **Settings** → **Transfer Mode**, and switch the mode to **Passive** (as shown below).



Overview

For more detailed information about what these options do, please run the network configuration wizard.

Transfer Mode

● Passive (recommended)

○ Active

☑ Allow fall back to other transfer mode on failure

If you have problems to retrieve directory listings or to transfer files, try to change the default transfer mode.

FTP Keep-alive

☐ Send FTP keep-alive commands

A proper server does not require this. Contact the server administrator if you need this.

Common Issues and Solutions for Data Download

Issue 1: How to Download Data via HTTP?

Method 1:

Click the **Download** button "⬇", select **"HTTP Download"**, and enter the download list page. Then, click the file name under the **Name** column to download the desired data.


Released Data : 2555 records

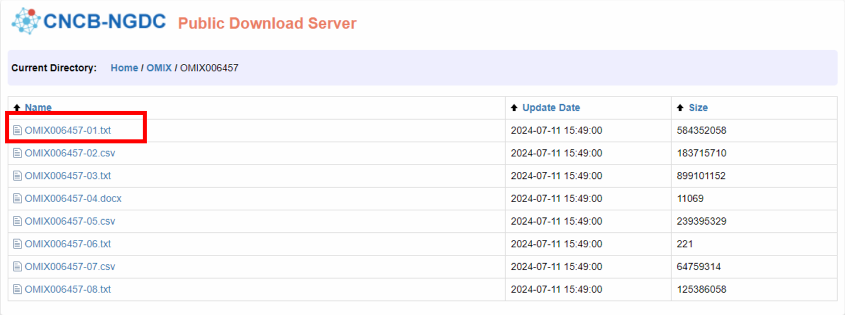
⚠ If the data cannot be downloaded/requested, as it has not yet been registered in the Human Genetic Resource Management Platform of MOST.

Show 10 entries

OMIX ID	Project	Title	Organism	Access Type	Release Time	Operation
OMIX006457	PRJCA006118	CMAISE-V1.0_RNA-seq_expressionMatrix	Homo sapiens	Open	2024-07-11	

Showing 1 to 1 of 1 entries (filtered from 2,555 total entries)

⇩



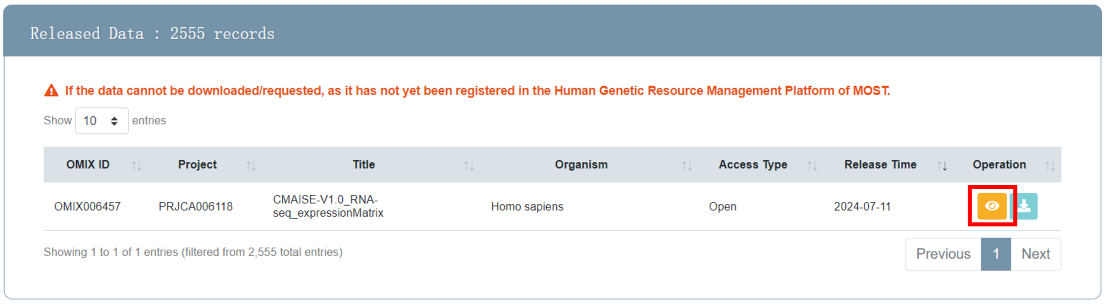
CNCB-NGDC Public Download Server

Current Directory: Home / OMIX / OMIX006457

Name	Update Date	Size
OMIX006457-01.txt	2024-07-11 15:49:00	584352058
OMIX006457-02.csv	2024-07-11 15:49:00	183715710
OMIX006457-03.txt	2024-07-11 15:49:00	899101152
OMIX006457-04.docx	2024-07-11 15:49:00	11069
OMIX006457-05.csv	2024-07-11 15:49:00	239395329
OMIX006457-06.txt	2024-07-11 15:49:00	221
OMIX006457-07.csv	2024-07-11 15:49:00	64759314
OMIX006457-08.txt	2024-07-11 15:49:00	125386058

Method 2:

Click the **View** button “👁” to access the dataset details page. In the **Files & Download** section, click the “HTTPs” button next to each file to download them individually.



Released Data : 2555 records

⚠ If the data cannot be downloaded/requested, as it has not yet been registered in the Human Genetic Resource Management Platform of MOST.

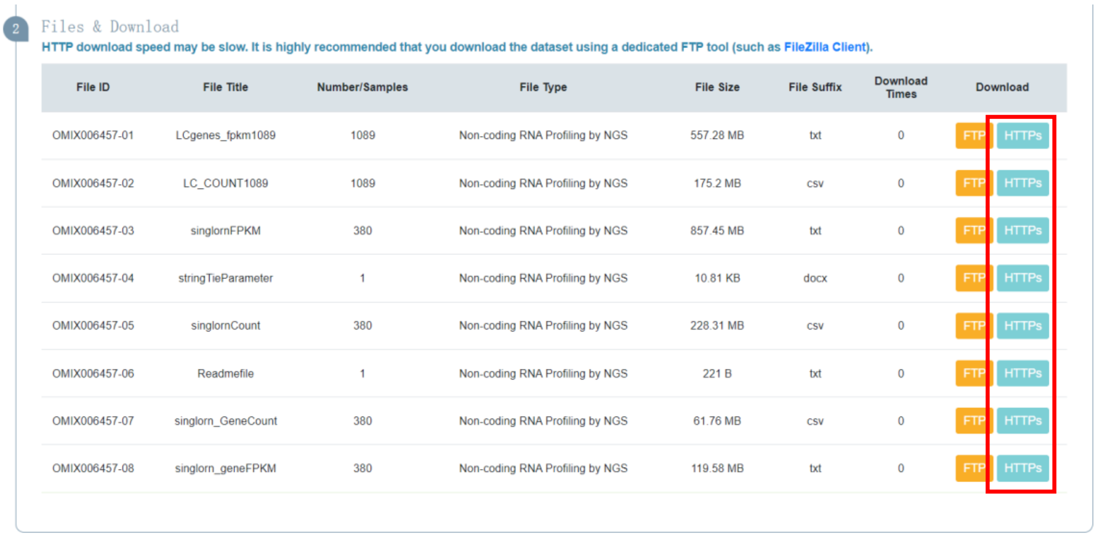
Show 10 entries

OMIX ID	Project	Title	Organism	Access Type	Release Time	Operation
OMIX006457	PRJCA006118	CMAISE-V1.0_RNA-seq_expressionMatrix	Homo sapiens	Open	2024-07-11	👁 ⤓

Showing 1 to 1 of 1 entries (filtered from 2,555 total entries)

Previous 1 Next

⇩



2 Files & Download

HTTP download speed may be slow. It is highly recommended that you download the dataset using a dedicated FTP tool (such as FileZilla Client).

File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Download
OMIX006457-01	LCgenes_fpkm1089	1089	Non-coding RNA Profiling by NGS	557.28 MB	txt	0	FTP HTTPs
OMIX006457-02	LC_COUNT1089	1089	Non-coding RNA Profiling by NGS	175.2 MB	csv	0	FTP HTTPs
OMIX006457-03	singlornFPKM	380	Non-coding RNA Profiling by NGS	857.45 MB	txt	0	FTP HTTPs
OMIX006457-04	stringTieParameter	1	Non-coding RNA Profiling by NGS	10.81 KB	docx	0	FTP HTTPs
OMIX006457-05	singlornCount	380	Non-coding RNA Profiling by NGS	228.31 MB	csv	0	FTP HTTPs
OMIX006457-06	Readmefile	1	Non-coding RNA Profiling by NGS	221 B	txt	0	FTP HTTPs
OMIX006457-07	singlorn_GeneCount	380	Non-coding RNA Profiling by NGS	61.76 MB	csv	0	FTP HTTPs
OMIX006457-08	singlorn_geneFPKM	380	Non-coding RNA Profiling by NGS	119.58 MB	txt	0	FTP HTTPs

Issue 2: How to Download Data via FTP?

Many modern browsers (such as **Chrome** and **Firefox**) have gradually removed support for FTP due to security concerns — the FTP protocol lacks encryption and is vulnerable to attacks. As a result, users cannot directly access files on FTP servers through these browsers.

Solution:

Use dedicated FTP client software such as **FileZilla** or **WinSCP**. Below is an example using **FileZilla**:

Step 1:

Download and install the FileZilla client from https://filezilla-project.org/. On the download page, click the red button labeled **“Download FileZilla Client”**, and follow the installation instructions.

Overview

Welcome to the homepage of FileZilla®, the free FTP solution. The *FileZilla Client* not only supports FTP, but also FTP over TLS (FTPS) and SFTP. It is open sou terms of the GNU General Public License.

We are also offering *FileZilla Pro*, with additional protocol support for WebDAV, Amazon S3, Backblaze B2, Dropbox, Microsoft OneDrive, Google Drive, Microsc Storage.

Last but not least, *FileZilla Server* is a free open source FTP and FTPS Server.

Support is available through our forums, the wiki and the bug and feature request trackers.

In addition, you will find documentation on how to compile FileZilla and nightly builds for multiple platforms in the development section.

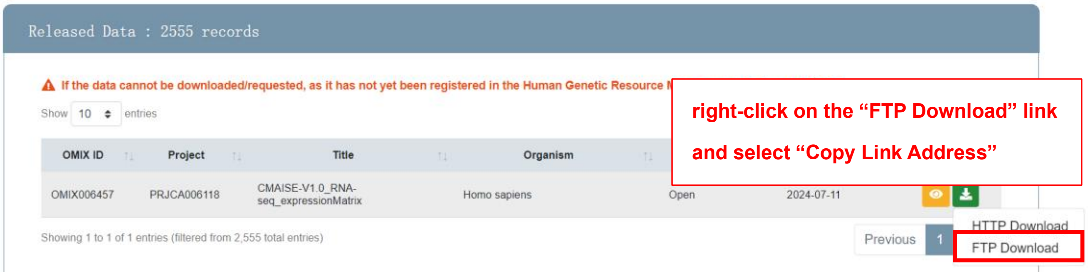
Quick download links

Pick the client if you want to transfer files. Get the server if you want to make files available for others.

Step 2:

On the dataset page, click the **Download** button “⬇”.

⚠️ **Do NOT directly click** “FTP Download”! Instead, **right-click** on the **“FTP Download”** link and select **“Copy Link Address”**.



⚠ If the data cannot be downloaded/requested, as it has not yet been registered in the Human Genetic Resource

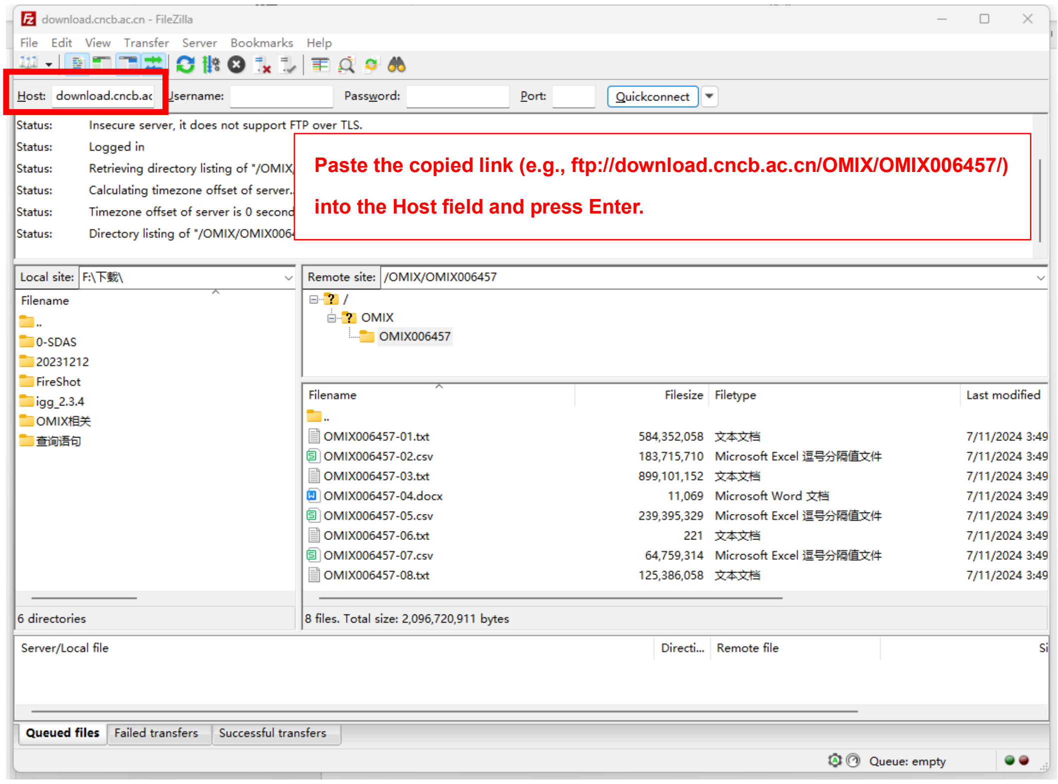
OMIX ID	Project	Title	Organism		
OMIX006457	PRJCA006118	CMAISE-V1.0_RNA-seq_expressionMatrix	Homo sapiens	Open	2024-07-11

Showing 1 to 1 of 1 entries (filtered from 2,555 total entries)

right-click on the “FTP Download” link and select “Copy Link Address”

Step 3:

Open the **FileZilla Client** application. Paste the copied FTP link (e.g., *ftp://download.cncb.ac.cn/OMIX/OMIX006457/*) into the **Host** field and press **Enter**. You will then see the list of files available for download.



Host: download.cncb.ac

```
Status:     Insecure server, it does not support FTP over TLS.
Status:     Logged in
Status:     Retrieving directory listing of "/OMIX
Status:     Calculating timezone offset of server.
Status:     Timezone offset of server is 0 second
Status:     Directory listing of "/OMIX/OMIX006
```

Paste the copied link (e.g., ftp://download.cncb.ac.cn/OMIX/OMIX006457/) into the Host field and press Enter.

Local site: F:\下载\

Filename
..
0-SDAS
20231212
FireShot
igg_2.3.4
OMIX相关
查询语句

6 directories

Remote site: /OMIX/OMIX006457

- /
  - OMIX
    - OMIX006457

Filename	Filesize	Filetype	Last modified
..			
OMIX006457-01.txt	584,352,058	文本文档	7/11/2024 3:49
OMIX006457-02.csv	183,715,710	Microsoft Excel 逗号分隔值文件	7/11/2024 3:49
OMIX006457-03.txt	899,101,152	文本文档	7/11/2024 3:49
OMIX006457-04.docx	11,069	Microsoft Word 文档	7/11/2024 3:49
OMIX006457-05.csv	239,395,329	Microsoft Excel 逗号分隔值文件	7/11/2024 3:49
OMIX006457-06.txt	221	文本文档	7/11/2024 3:49
OMIX006457-07.csv	64,759,314	Microsoft Excel 逗号分隔值文件	7/11/2024 3:49
OMIX006457-08.txt	125,386,058	文本文档	7/11/2024 3:49

8 files. Total size: 2,096,720,911 bytes

Queue: empty