多元数据归档库（OMIX）使用说明 V1.3

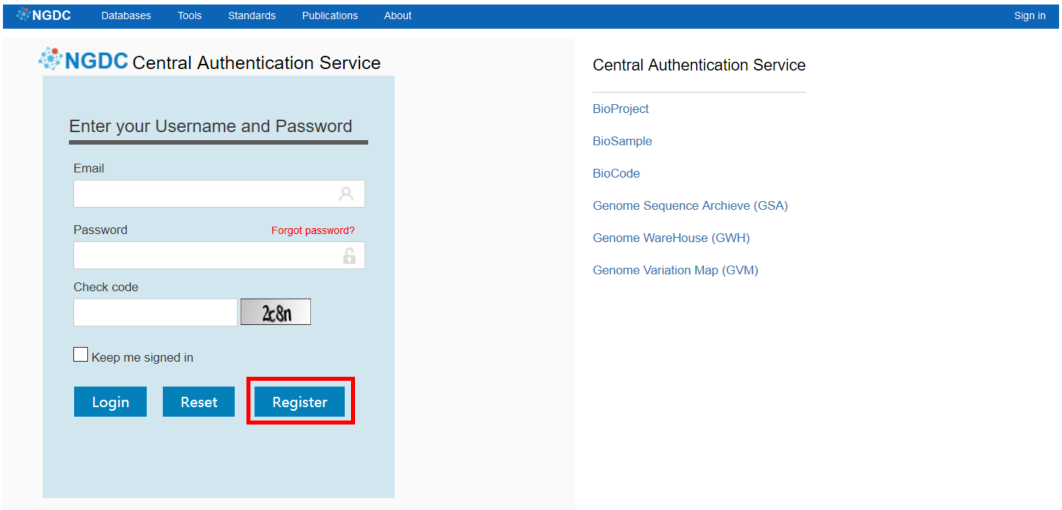
更新时间：2024 年 7 月 12 日

系统简介

为了满足用户多样的数据汇交需求，国家基因组科学数据中心（National Genomics Data Center，NGDC）GSA 团队开发了多元数据归档库（Open Archive for Miscellaneous Data，OMIX）。OMIX 不但接收组学原始测序分析后和芯片数据，还接收代谢组（包括脂质组）、蛋白组等多元组学及其相关数据。在人类遗传资源管理方面，OMIX 数据库严格遵循《中华人民共和国人类遗传资源管理条例》规定，为用户提供两种数据访问方式：公开访问和受控访问。受控访问的数据，需要通过数据提交方授权，才能获得下载和使用权限。

用户注册

所有用户都可通过中心单点登录系统（BIG Single Sign-On，BIG SSO）完成**账号注册**。但 OMIX 数据库建议使用研究组的项目负责人级别人员提交数据。您可通过中心单点登录系统完成**账号注册**。如果您在账号注册或使用过程中遇到任何问题，请联系 gsa@big.ac.cn。



NGDC Databases Tools Standards Publications About Sign in

NGDC Central Authentication Service

Enter your Username and Password

Email

Password Forgot password?

Check code 2c8n

☐ Keep me signed in

Login Reset Register

Central Authentication Service

BioProject

BioSample

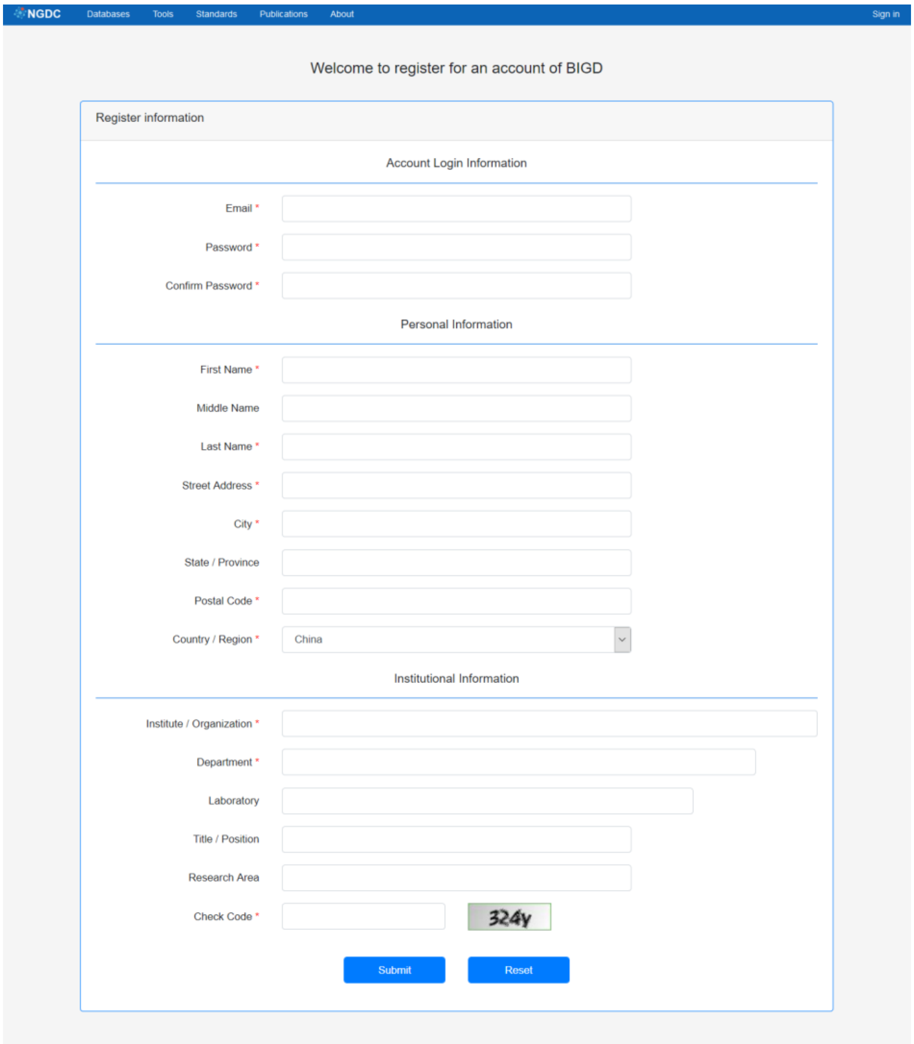
BioCode

Genome Sequence Archieve (GSA)

Genome WareHouse (GWH)

Genome Variation Map (GVM)

⇩



Welcome to register for an account of BIGD

Register information

Account Login Information

Email *

Password *

Confirm Password *

Personal Information

First Name *

Middle Name

Last Name *

Street Address *

City *

State / Province

Postal Code *

Country / Region * China

Institutional Information

Institute / Organization *

Department *

Laboratory

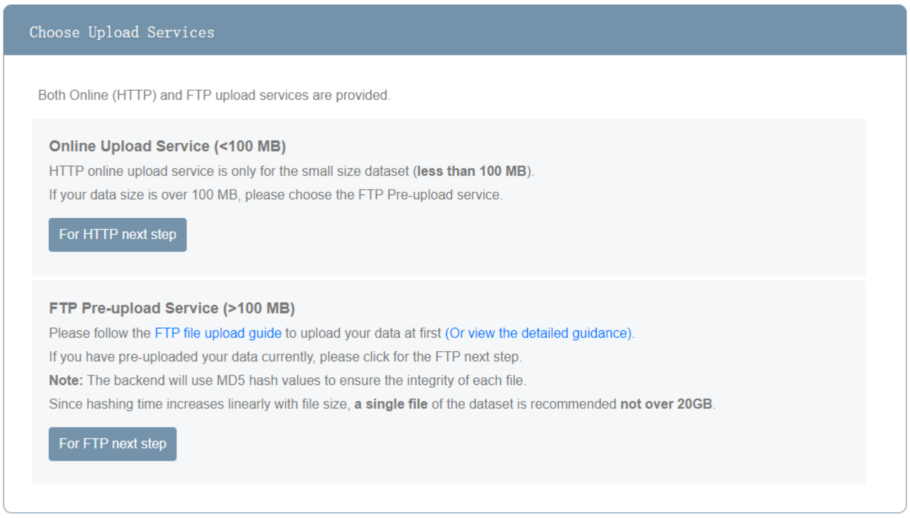
Title / Position

Research Area

Check Code *

数据递交指导

OMIX 数据库提供数据提交方式，分别页面提交方式（Online Upload Service）和 FTP 提交方式（FTP Pre-upload Service）。



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).

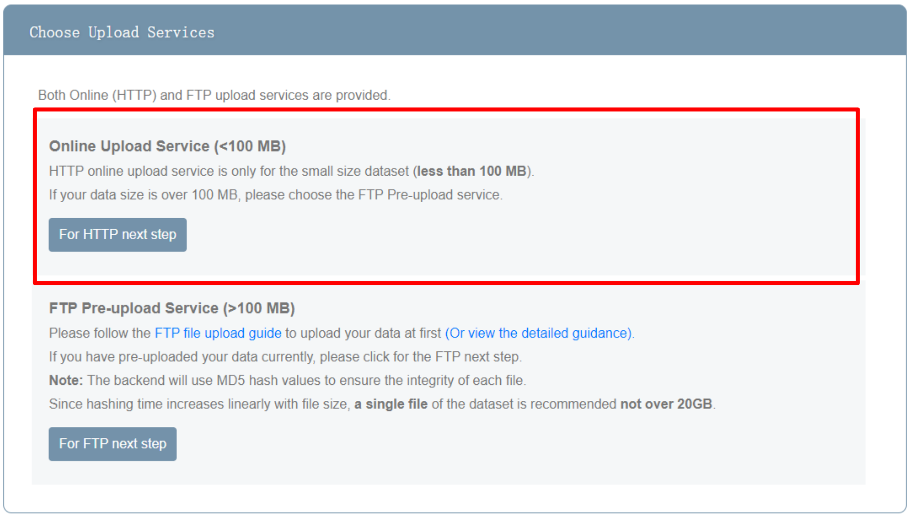
If you have pre-uploaded your data currently, please click for the FTP next step.

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**.

页面提交方式

如果一次性上传数据量＜100 MB，推荐使用页面提交方式（Online Upload Service），并完成元数据信息和数据文件提交。



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 **基本信息（Data Info）**用于收集数据集名称、简单描述和物种信息。如果一个文件涉及多个物种，可填写 mixed sample。

请注意：如果是人遗备份数据的物种务必填写——人（*Homo sapiens*）。系统为人类遗传资源数据提供两种数据访问方式，受控访问和公开访问。

简体中文 English

可自由切换中英文

OMIX

欢迎使用OMIX，请选择您预上传的数据并填写表中信息！

新建提交

1 数据集信息

* 标题

* 描述信息

* 物种名称

请输入至少三个字母，然后从提示列表中选出所需的物种名称。

* 数据类型

请选择数据类型

人遗备份数据，物种请填写 *Homo sapiens*

- 如果您将数据设定为公开访问（Open-access），数据将在到达发布日期或（系统检测到）文章发表后自动释放。数据发布后，数据使用者即可在线浏览信息并下载该数据
- 如果您将数据设定为受控访问（Controlled-access），即使数据集发布，使用者也需获得您方许可才能下载。受控访问发布后，暂时没法支持下载，除非是联系您们授权。

* 物种名称

Homo sapiens

* 数据访问权限

○ 公开访问 ● 受控访问

* 数据类型

请选择数据类型

② **项目信息（BioProject Info）：**用于收集数据集相关联的 BioProject 信息。如果还未创建 BioProject，可点击前往 BioProject 数据库进行创建，BioProject 审核状态不影响 OMIX 数据后续填写提交。

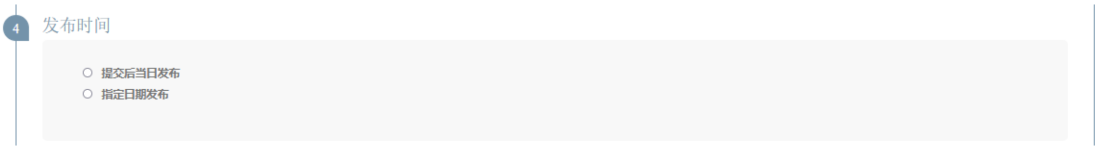
3 项目信息

* 请选择项目编号

请选择项目编号

或者去创建新 BioProject

③ **数据集释放时间（Release Date）：用于收集数据集发布时间。**



4 发布时间

○ 提交后当日发布
○ 指定日期发布

④ 文件信息（Files）：用于提交文件和收集文件信息。

- 每次提交可以添加多个文件。根据情况点击“增加文件（add more file）”并填写相应的元数据信息。请注意：对于压缩包，样本数/文件个数（Number of Samples/Files）请填写压缩包内包含文件数。非压缩包，请填写文件涉及样本数。

4 文件信息

添加文件

* 文件序号（自动）: 1

* 文件标题:

* 样本数/文件个数 (对于压缩文件):

* 单击此处选择上传文件

当前暂未上传文件

点击 add more file
可添加多个文件集

点击上传文件

- 请点击“**单击此处选择上传文件（Click here to select file for uploading）**”上传数据文件。

⑤ 科技部备案备份编号（Accessions in the Ministry of Science and Technology）：此项受控访问数据集为选填项。依照目前规定所有人相关数据都为人遗数据申报填写，具体情况可参考“人类遗传资源信息备案备份流程”。

5 科技处备案信息

备案编号

备份编号

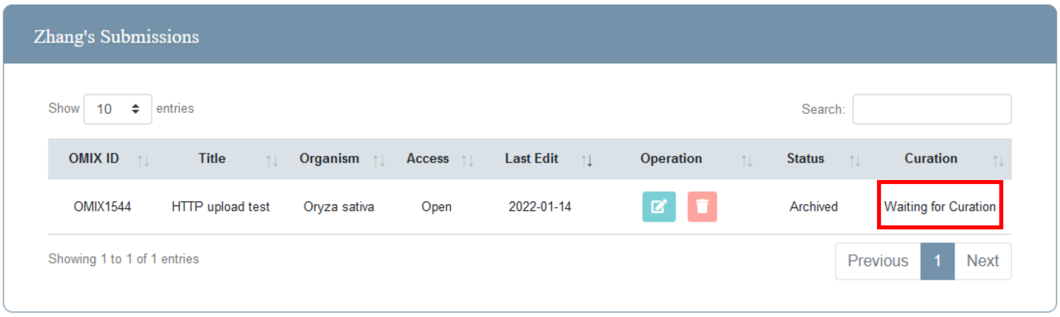
注意：根据国家法律规定，所有人类遗传资源相关数据都需要在科技部完成信息备份和备案工作

具体流程详见 《人类遗传资源信息备案流程》

☐ You agree to our Terms and Policy.

需勾选协议

⑥ **提交：**所有信息填写完并核对后请点击“Submit”提交数据集。数据提交后请等待系统管理员进行审核，审核通过后数据下图的列表中“**数据状态和共享链接（Status/Review）**”显示为“🔗”。点击即可生成数据共享链接，具体详见“如何生成数据集共享链接”；

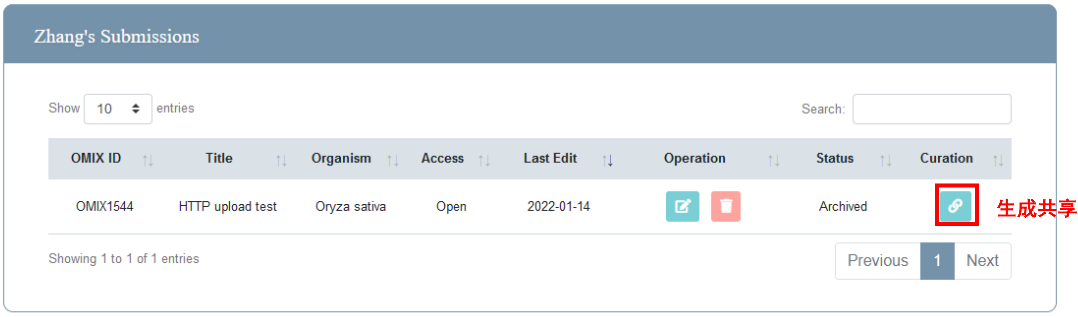


Zhang's Submissions

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

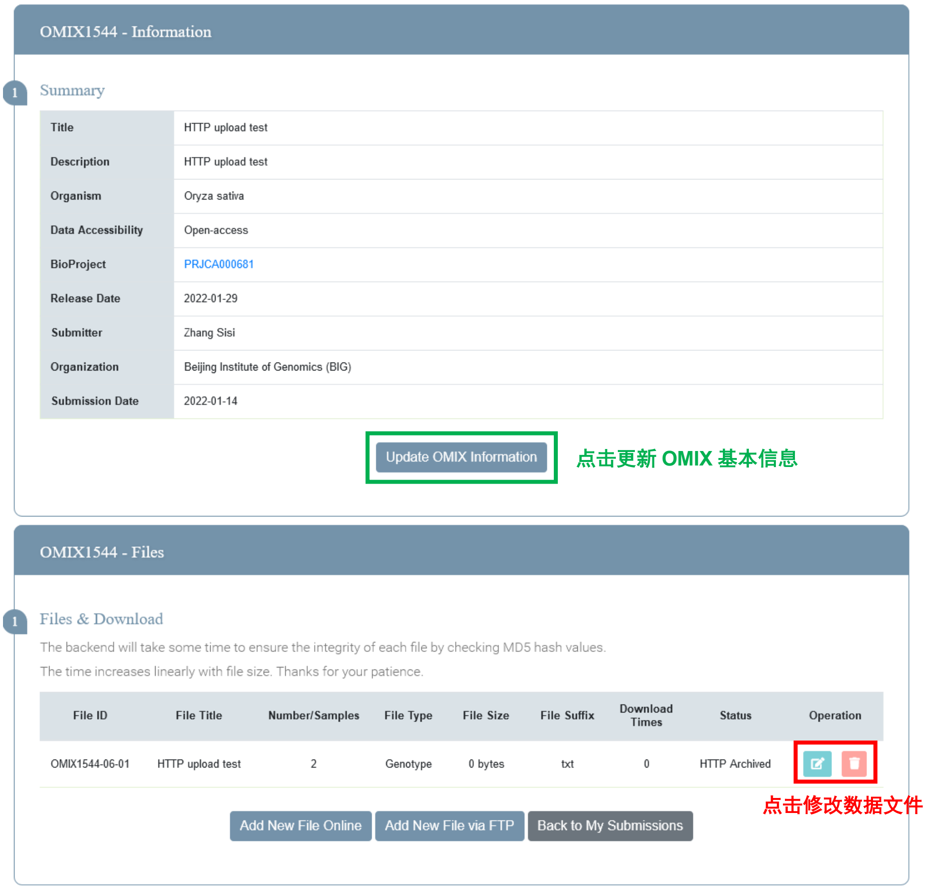
Show 10 entries | Search:

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

生成共享链

数据未到达发布时间前，您可随时修改 OMIX 基本和数据文件信息。但无法替换数据文件。如需替换，可先点击“Add New File Online/Add New File via FTP”添加一条新数据，然后再删除原有错误文件。



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

点击更新 OMIX 基本信息

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	

点击修改数据文件

FTP 提交方式

如果一次性上传数据量＞100 MB，建议使用本方法完成数据文件和信息提交。

完成数据预上传

① 文件准备

如果您的数据文件个数比较少，您可以直接上传。如果您的文件个数比较多，也可以考虑分类压缩为 tar 或 zip 的包，建议每个文件压缩包的大小不宜超过 10GB。文件名称不可包含空格，否则会影响数据归档

② 计算 MD5 码：

MD5 码是一个 32 个字符的字母数字字符串。请通过**命令行或软件计算**获得每一个文件压缩包的 MD5 码。

- Mac 和 Linux 系统的用户，可在本地运行命令行"md5sum" (Linux) 或"md5"(Mac OX)便可生成 MD5 码。
- Windows 用户需使用第三方程序，才能生成 MD5 码。推荐软件：winmd5free，“好压 2345”（可支持批量生成 md5 码）

③ 数据上传流程

第1步 用户需要使用 FTP 客户端软件（比如 FileZilla Client）登录 FTP 服务器上传数据，文档中以 FileZilla 作为实例。

第2步 下载客户端软件（https://filezilla-project.org/ ），下载页面如下图所示，点击红色框中的“Download FileZilla Client”，并按照提示安装软件；

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, Microso 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.

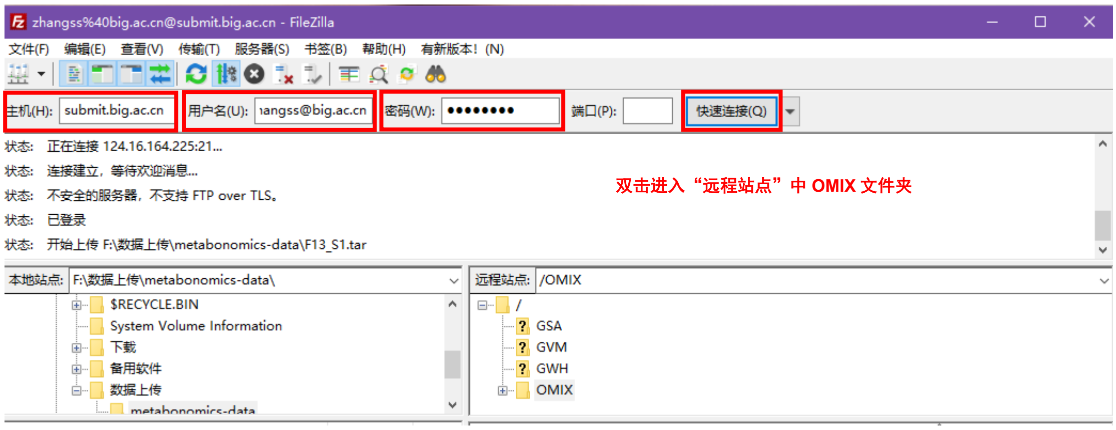
In addition, you will find documentation on how to compile FileZilla and nightly builds for multiple platforms in the development section.

Quick download links

下载 FileZilla Client 软件

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

第3步 打开软件，界面下图所示，填写主机信息为“submit.big.ac.cn”，用户登陆账号邮箱和密码，然后点击“快速连接”，状态栏显示登陆成功。常见问题详见“FTP 上传常见问题与对策”；


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

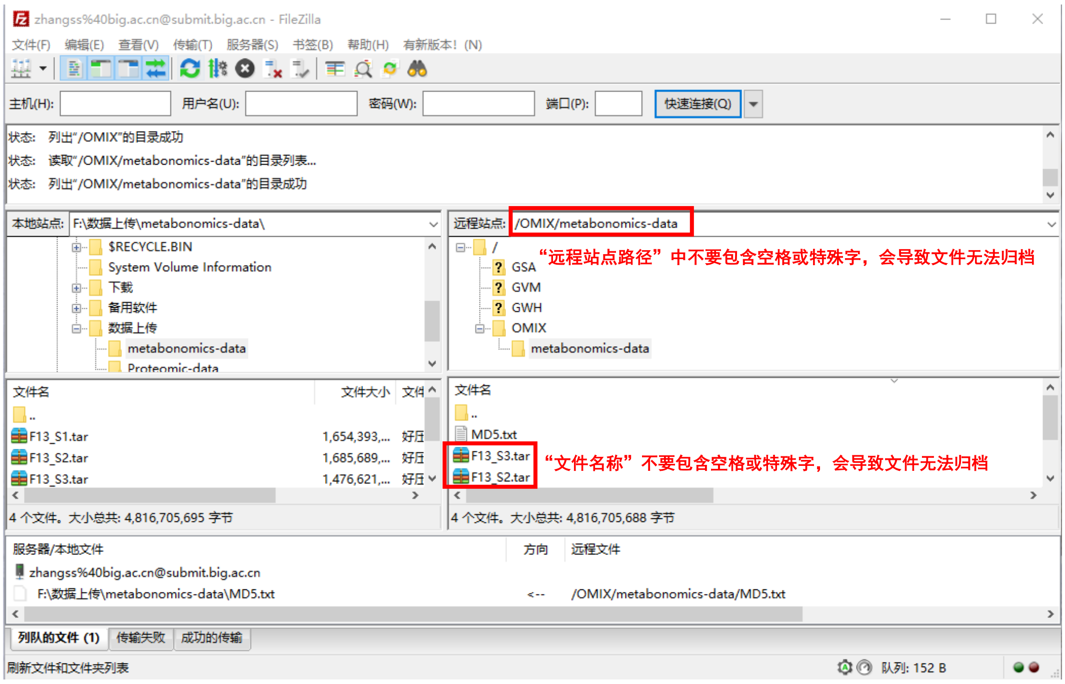
主机(H): submit.big.ac.cn
用户名(U): ıangss@big.ac.cn
密码(W): ••••••••
端口(P):

```
状态:	正在连接 124.16.164.225:21...
状态:	连接建立，等待欢迎消息...
状态:	不安全的服务器，不支持 FTP over TLS。
状态:	已登录
状态:	开始上传 F:\数据上传\metabonomics-data\F13_S1.tar
```

双击进入“远程站点”中 OMIX 文件夹

本地站点: F:\数据上传\metabonomics-data\
远程站点: /OMIX

第4步 登陆成功后，“本地站点”选择需要上传数据的本地数据路径，“远程站点”中的 OMIX 文件中。 注意：上传文件名称及其所在“远程站点路径”都不要包含空格或特殊字符。


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

```
状态:	列出“/OMIX”的目录成功
状态:	读取“/OMIX/metabonomics-data”的目录列表...
状态:	列出“/OMIX/metabonomics-data”的目录成功
```

本地站点: F:\数据上传\metabonomics-data\
远程站点: /OMIX/metabonomics-data

“远程站点路径”中不要包含空格或特殊字符，会导致文件无法归档

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

4 个文件。大小总共: 4,816,705,695 字节

文件名
..
MD5.txt
F13_S3.tar
F13_S2.tar

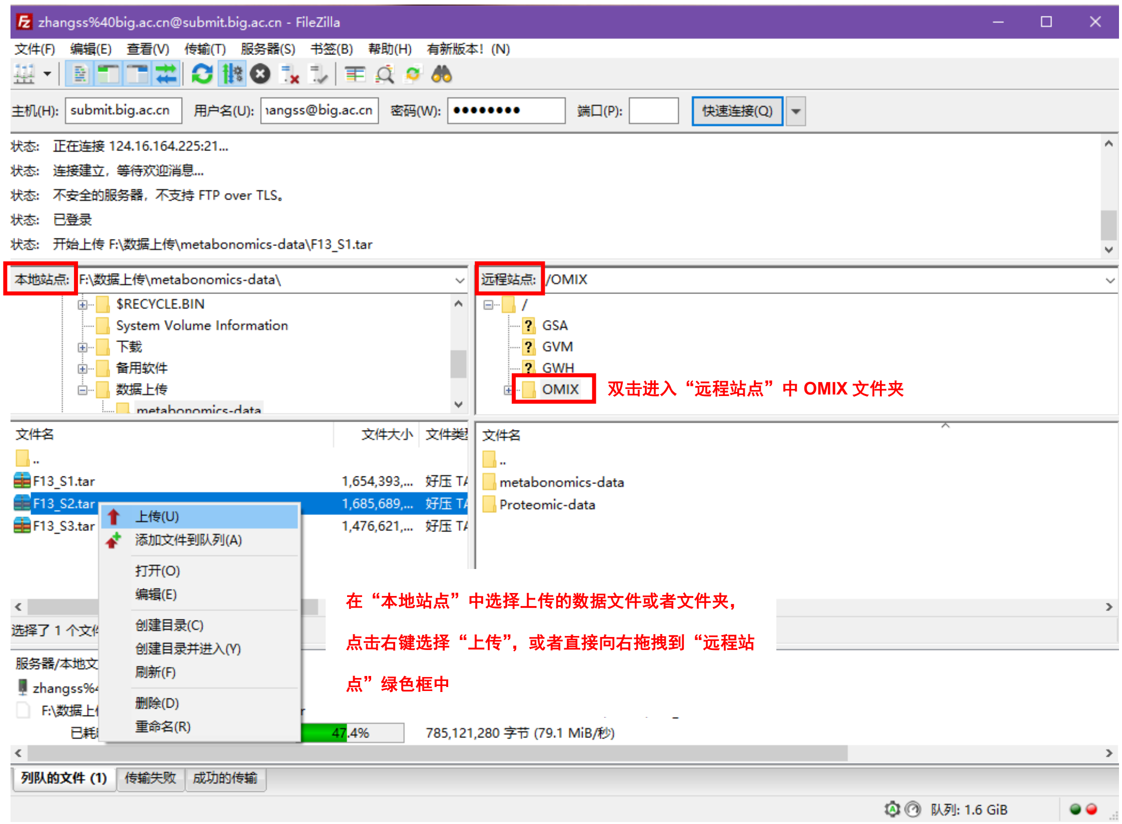
“文件名称”不要包含空格或特殊字符，会导致文件无法归档

4 个文件。大小总共: 4,816,705,688 字节

服务器/本地文件	方向	远程文件
zhangss%40big.ac.cn@submit.big.ac.cn		
F:\数据上传\metabonomics-data\MD5.txt	<--	/OMIX/metabonomics-data/MD5.txt

刷新文件和文件夹列表
队列: 152 B

第5步 在“本地站点”中选择上传的数据文件或者文件夹，点击右键，选择“上传”，或者直接拖拽到“远程站点”，如下图所示。



```
状态:	正在连接 124.16.164.225:21...
状态:	连接建立，等待欢迎消息...
状态:	不安全的服务器，不支持 FTP over TLS。
状态:	已登录
状态:	开始上传 F:\数据上传\metabonomics-data\F13_S1.tar
```

双击进入“远程站点”中 OMIX 文件夹

在“本地站点”中选择上传的数据文件或者文件夹，点击右键选择“上传”，或者直接向右拖拽到“远程站点”绿色框中

第6步 上传的所有数据会进入“队列的文件”，排队上传，上传成功后数据信息会转移到“成功的传输”中，如果上传不成功会转移到“传输失败”，需要重新上传，可以选择“断点续传”。

完成元数据信息填写

国家基因组科学数据中心数据提交统一入口：https://ngdc.cncb.ac.cn/gsub/

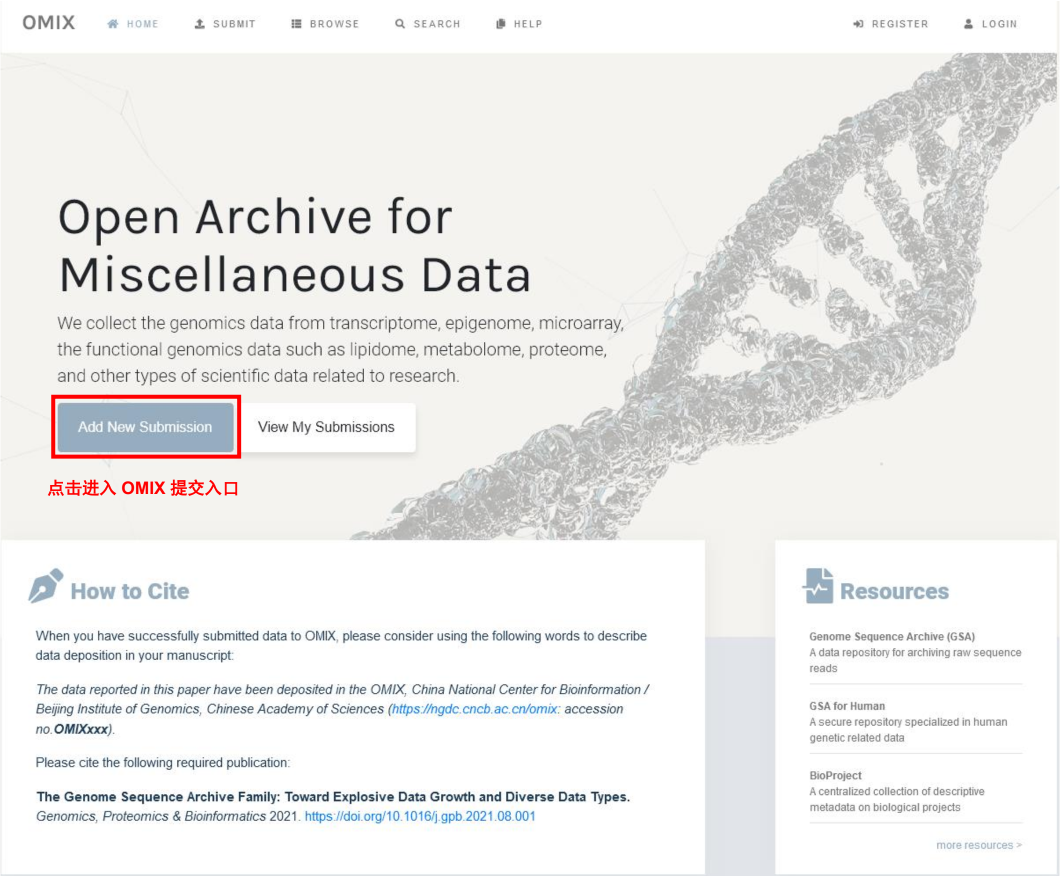
请选择您要提交的数据类型

生物数据递交系统(BIG Submission, BIG Sub)是国家基因组科学数据中心生物数据统一汇交入口，为用户提供一站式数据递交服务。

专项数据汇交

点击进入 OMIX 数据库

⇩



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.

点击进入 OMIX 提交入口

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

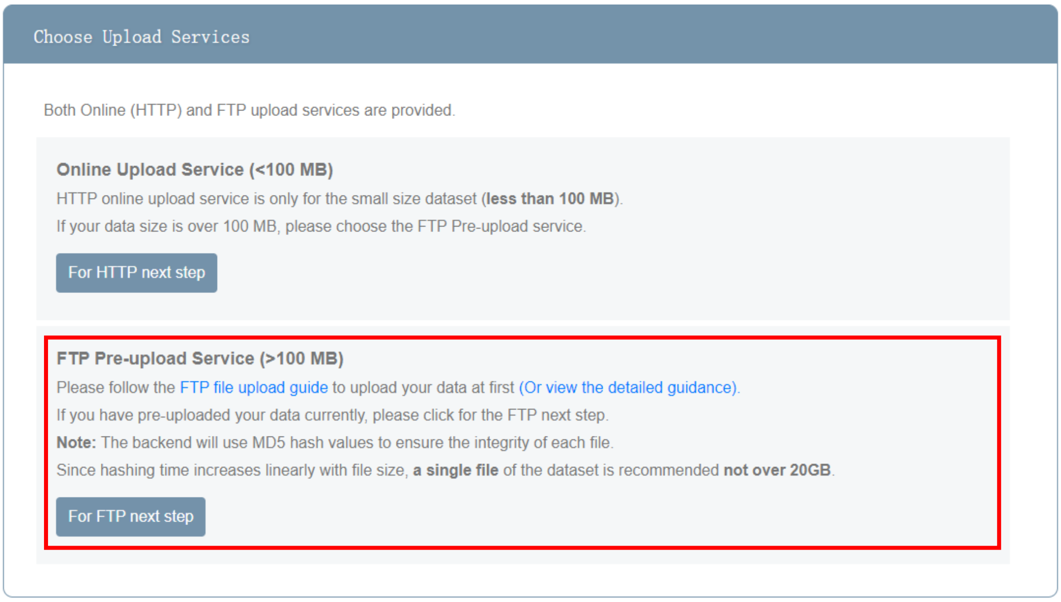
Genome Sequence Archive (GSA)
A data repository for archiving raw sequence reads

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

BioProject
A centralized collection of descriptive metadata on biological projects

more resources >

⇩



Choose Upload Services

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).

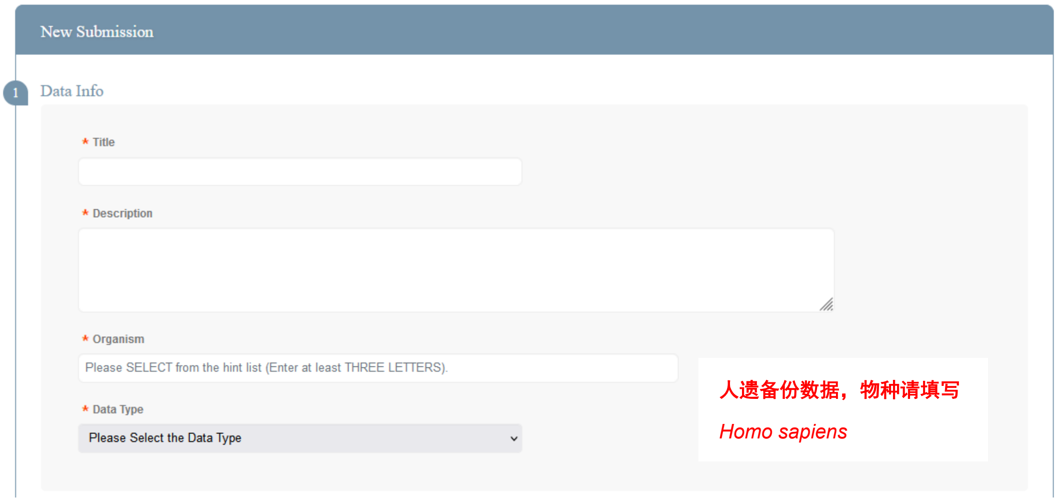
If you have pre-uploaded your data currently, please click for the FTP next step.

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 **基本信息（Data Info）**：用于收集数据集名称、简单描述和物种信息。如果一个文件涉及多个物种，可填写 mixed sample。

请注意：如果是人遗备份数据的物种务必填写——人（*Homo sapiens*）。系统为人类遗传资源数据提供两种数据访问方式，受控访问和公开访问。


New Submission

1 Data Info

* Title

* Description

* Organism

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

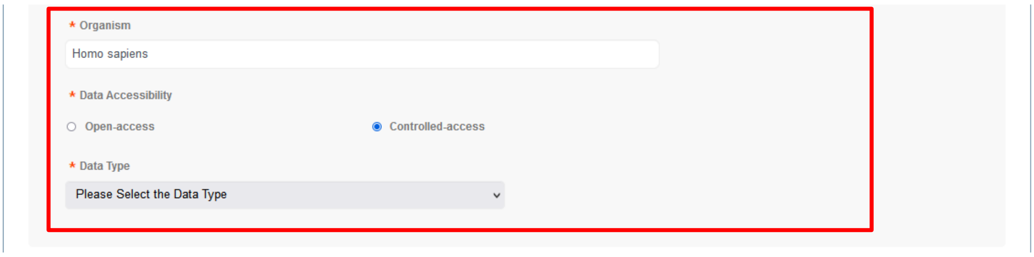
* Data Type

Please Select the Data Type

人遗备份数据，物种请填写

Homo sapiens

- 如果您将数据设定为公开访问（Open-access），数据将在到达发布日期或（系统检测到）文章发表后自动释放。数据发布后，数据使用者即可在线浏览信息并下载该数据
- 如果您将数据设定为受控访问（Controlled-access），即使数据集发布，使用者也需获得您方许可才能下载。受控访问发布后，暂时没法支持下载，除非是联系您们授权。


* Organism

Homo sapiens

* Data Accessibility

○ Open-access ● Controlled-access

* Data Type

Please Select the Data Type

② **文件信息（Files）：**您可从列表中勾选出已上传的数据文件在填写对应的数据描述信息。数据文件上传方法详见“完成数据预上传”

- 请务必**一次性“勾选”完毕后**，再开始**填写描述信息**。请勿一边**“勾选”**一边**填写**。因为如果**中途更改“勾选”项**，之前所填写的**信息将不会被保留**，全部需要重新填写
- 请务必确保已上传文件的文件名中**不允许有空格**

2 文件信息

请首先在下表中选择您预上传的文件，然后为对应的每个文件补全相关信息 (查看详细使用说明文档)

Search:

一次性全选

文件名中不允许有空格

单击此处扫描获取您已上传的文件列表

□	File Name	Last Edit	Size	Path
□	F13_S1.tar	2022-01-14 13:08:04	1.54 GB	/FTP upload test/F13_S1.tar
□	F13_S1.tar	2022-01-14 12:54:57	1.54 GB	/F13_S1.tar
□	F13_S3.tar	2022-06-22 15:01:09	1.38 GB	/F13_S3.tar

Showing 1 to 3 of 3 entries

逐个勾选时，请注意：务必全部勾选完毕后，再填写描述信息，如中途更改勾选项，原填写信息将不会保留，全部需要重新填写

请补全所选的　个文件的信息（至少选择一个）

后台会校验MD5码以确保每个文件的完整性

由于校验时间会随文件大小增长，因此建议数据集中单个文件小于20GB

* 文件名称（已勾选）：1

* 文件标题：

* 样本数/文件个数 (对于压缩文件)：

* 文件MD5码：

③ **项目信息（BioProject Info）：**用于收集数据集相关联的 BioProject 信息。如果还未创建 BioProject，可点击前往 BioProject 数据库进行创建，BioProject 审核状态不影响 OMIX 数据后续填写提交。

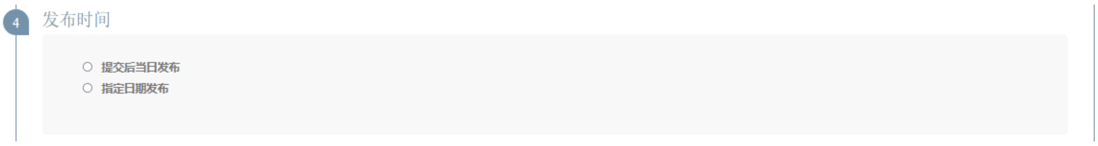
3 项目信息

* 请选择项目编号

请选择项目编号

或者去创建新BioProject

④ **数据集释放时间（Release Date）：用于收集数据集发布时间。**



4 发布时间

○ 提交后当日发布

○ 指定日期发布

⑤ **科技部备案备份编号（Accessions in the Ministry of Science and Technology）：此项受控访问数据集为选填项。建议人遗数据申报填写，**具体情况可参考“人类遗传资源信息备案流程”。

注意：根据国家法律规定，所有人类遗传资源相关数据都需要在科技部完成信息备份和备案工作

具体流程详见 《人类遗传资源信息备案流程》

需勾选协议

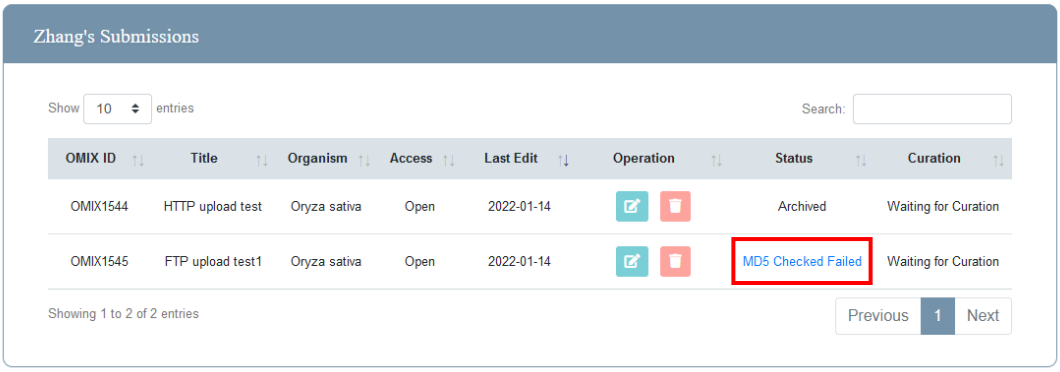
☐ You agree to our Terms and Policy.

⑥ **提交：**所有信息填写完并核对后请点击“Submit”提交数据集。数据提交后请等待系统管理员进行审核，审核通过后数据下图的列表中“**数据状态和共享链接（Status/Review）**”显示为“🔗”。点击即可生成数据共享链接，具体详见“如何生成数据集共享链接”；

数据递交常见问题与对策

问题 1：文件状态（Status）显示为 MD5 Checked Failed，如何处理？

通常情况下，数据集内如果有文件 **MD5 码校验不通过**，下图中**数据状态和共享链接（Status/Review）显示为“MD5 Checked Failed”**（系统会发送数据归档失败通知邮件）。修改步骤如下：



Zhang's Submissions

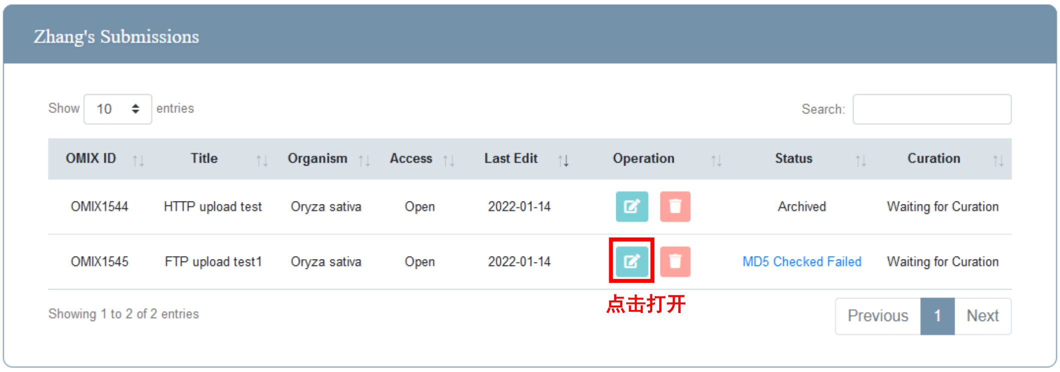
Show 10 entries

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	✎ 🗑	MD5 Checked Failed	Waiting for Curation

Showing 1 to 2 of 2 entries

第一步 请点击编辑按钮“✎”打开找到 **MD5 码校验不通过**的数据文件



Zhang's Submissions

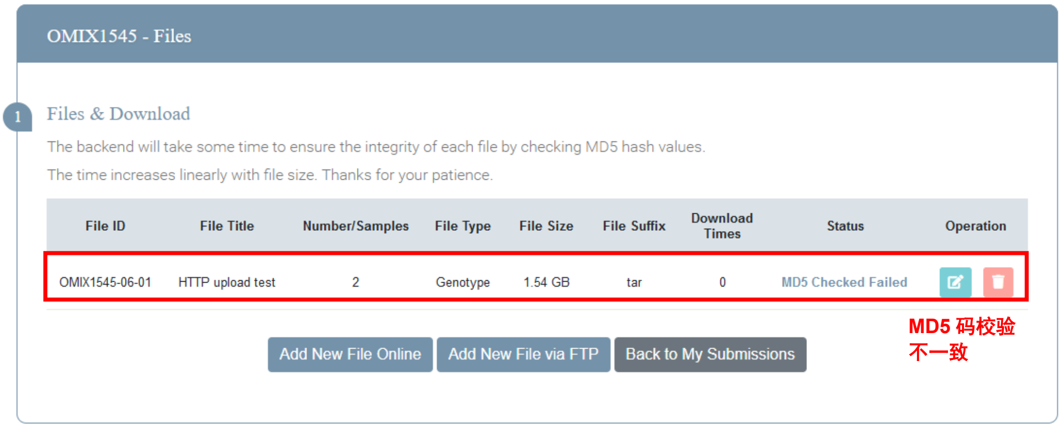
Show 10 entries

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	✎ 🗑	MD5 Checked Failed	Waiting for Curation

Showing 1 to 2 of 2 entries

点击打开



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.

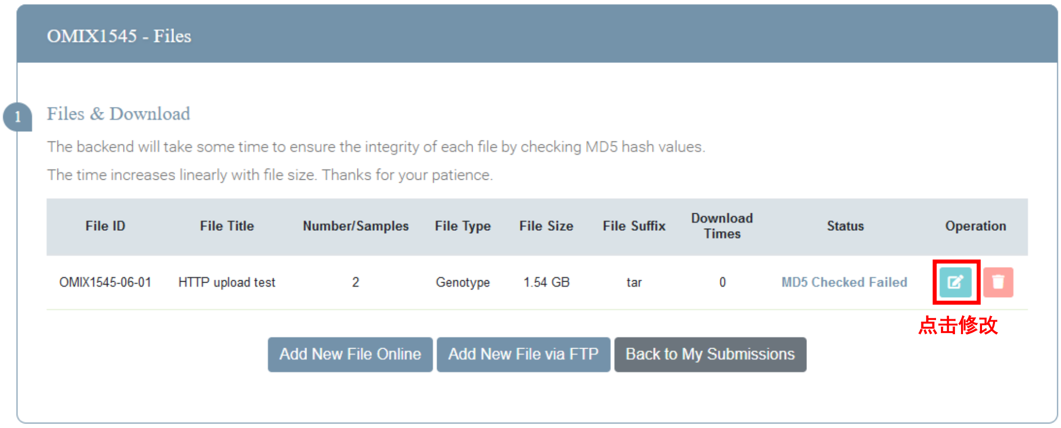
File ID	File Title	Number/Samples	File Type	File Size	File Suffix	Download Times	Status	Operation
OMIX1545-06-01	HTTP upload test	2	Genotype	1.54 GB	tar	0	MD5 Checked Failed	✎ 🗑

MD5 码校验不一致

第二步 排除 MD5 码校验失败原因：

- **MD5 码填写问题：**请重新计算原始文件的 MD5 码，将其与提交的 MD5 码

进行对比，如果二者不同，即为填写问题，请在文件编辑界面填入正确的 md5 码再次提交，具体修改步骤如下：

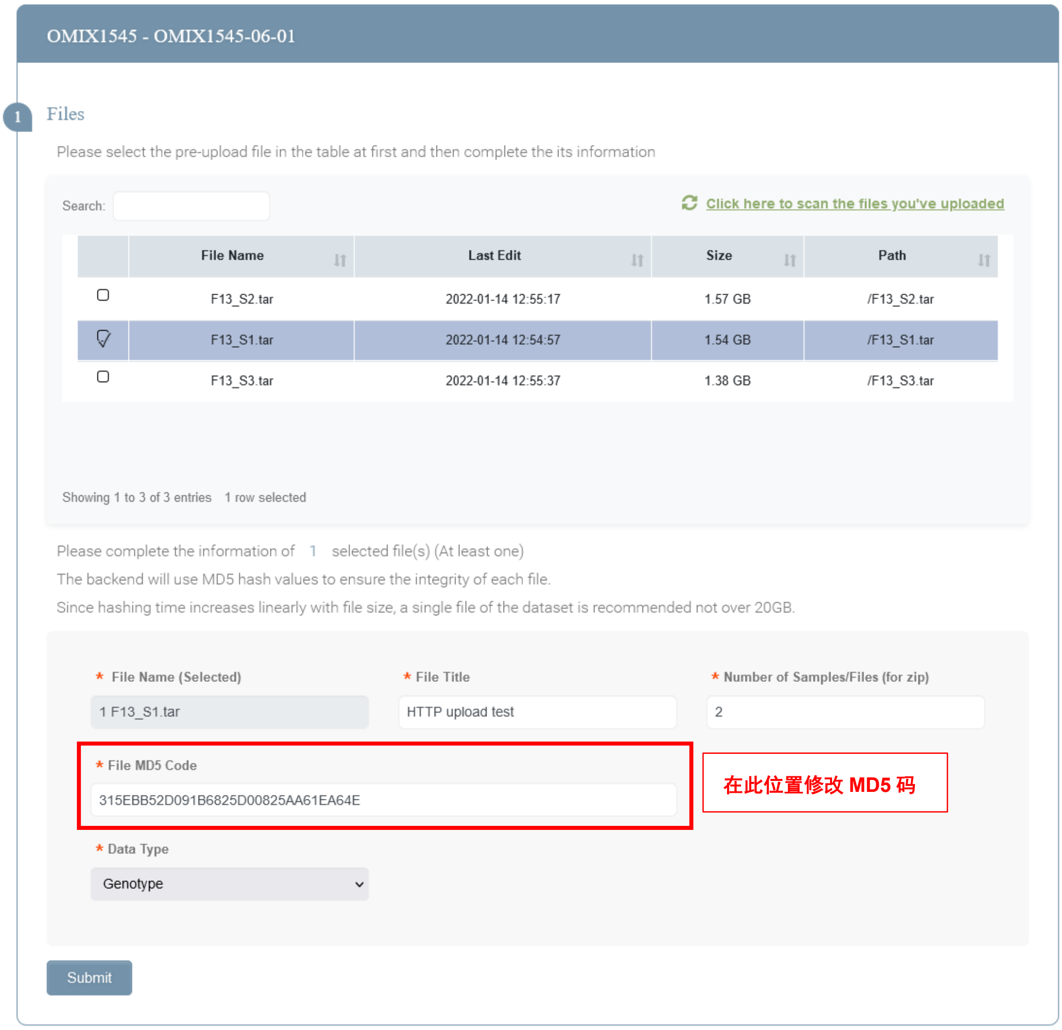


OMIX1545 - Files

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

	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

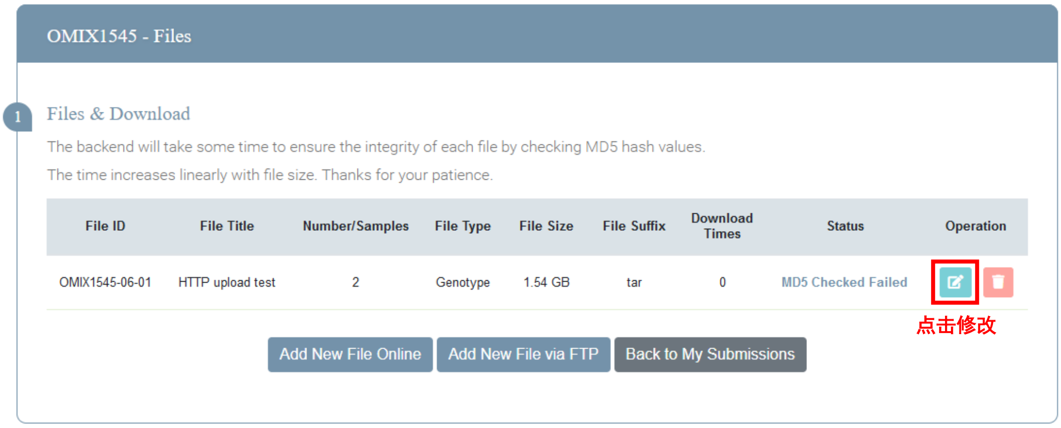
Showing 1 to 3 of 3 entries 1 row selected

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

- **数据文件传输问题**：请重新计算原始文件的 MD5 码，将其与提交的 MD5 码进行对比，如果二者相同，可能由于文件上传出错，导致校验失败。您需要

重新提交数据文件，数据预上传方法详见“完成数据预上传”，确定数据文件上传完成后，根据以下步骤完成修改：



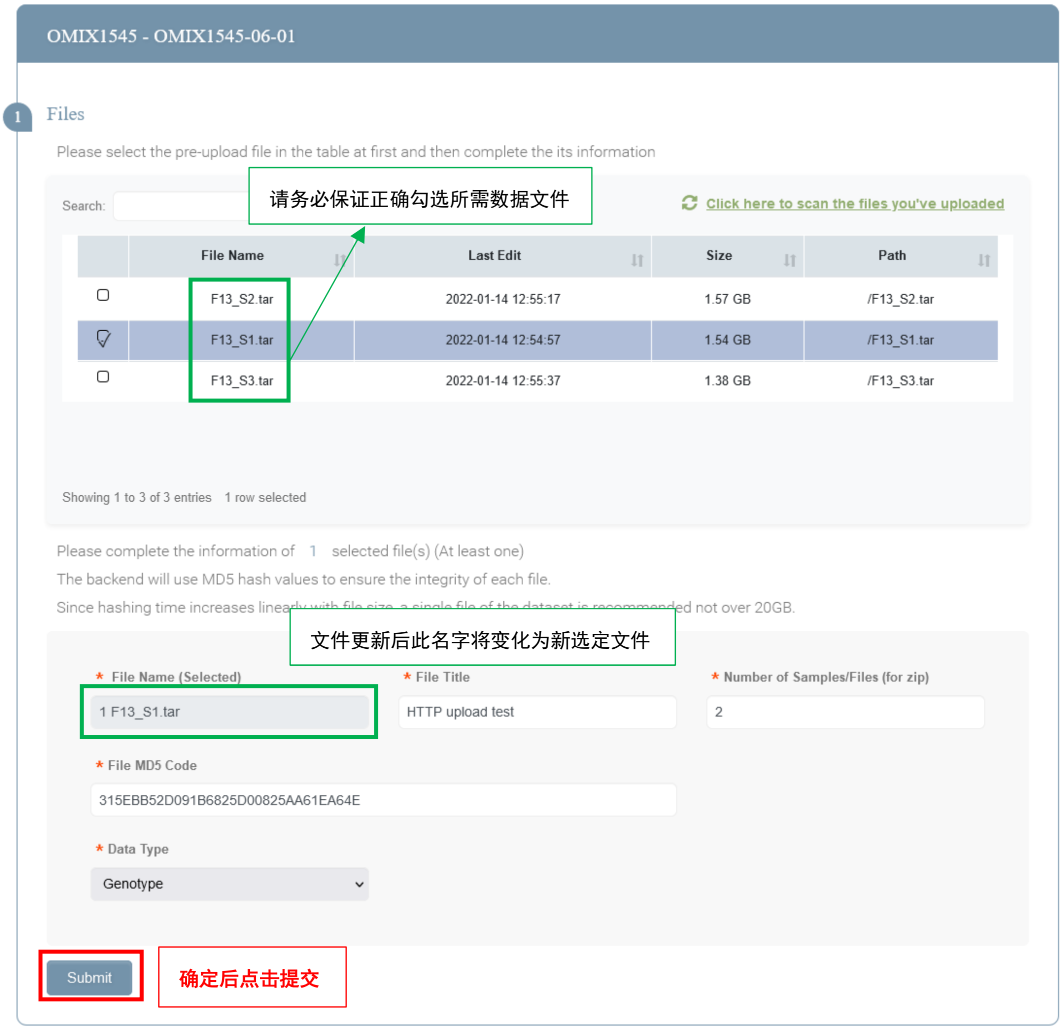
OMIX1545 - Files

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

	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

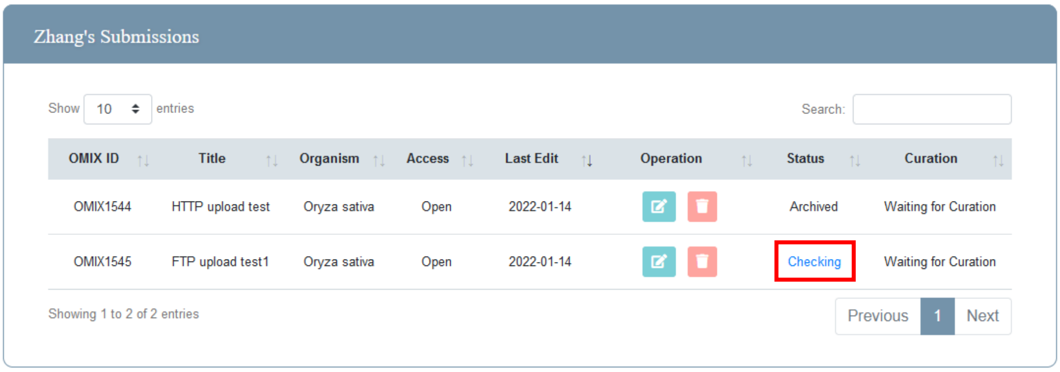
* File MD5 Code: 315EBB52D091B6825D00825AA61EA64E

* Data Type: Genotype

问题 2：文件状态（Status）一直显示为 Checking，如何处理？

通常情况下，当数据文件正在进行 MD5 码时，下图中 Status/Review 显示

为“Checking”。校验时间会随文件大小增长，因此建议数据集中单个文件小于20GB。



Zhang's Submissions

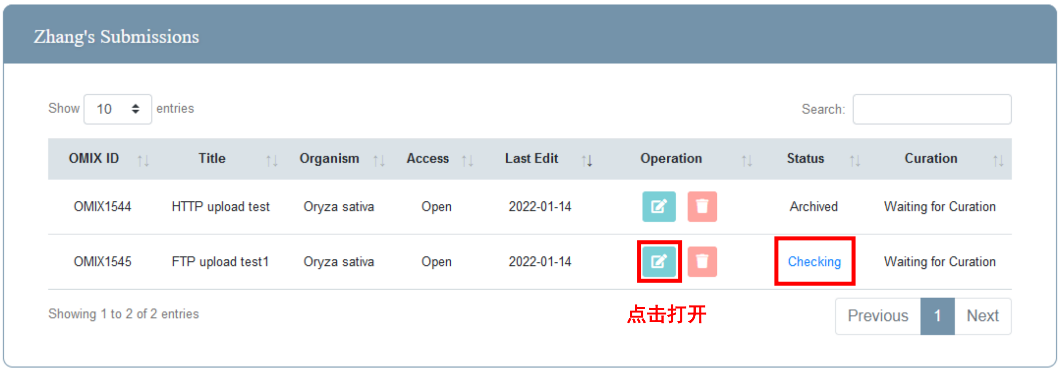
Show 10 entries Search:

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

但如果单个文件<20 GB，**维持此状态**超过 24 小时无变化，有可能为以下原因：

- **数据文件上传失败**：请点击“✎”打开确认一下文件是否正确上传，如果没有请删除后，重新提交一遍。



Zhang's Submissions

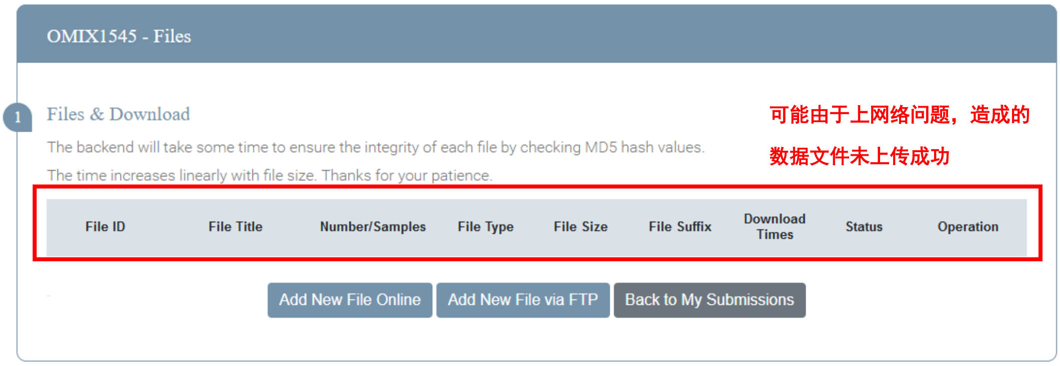
Show 10 entries Search:

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

点击打开

⇩



OMIX1545 - Files

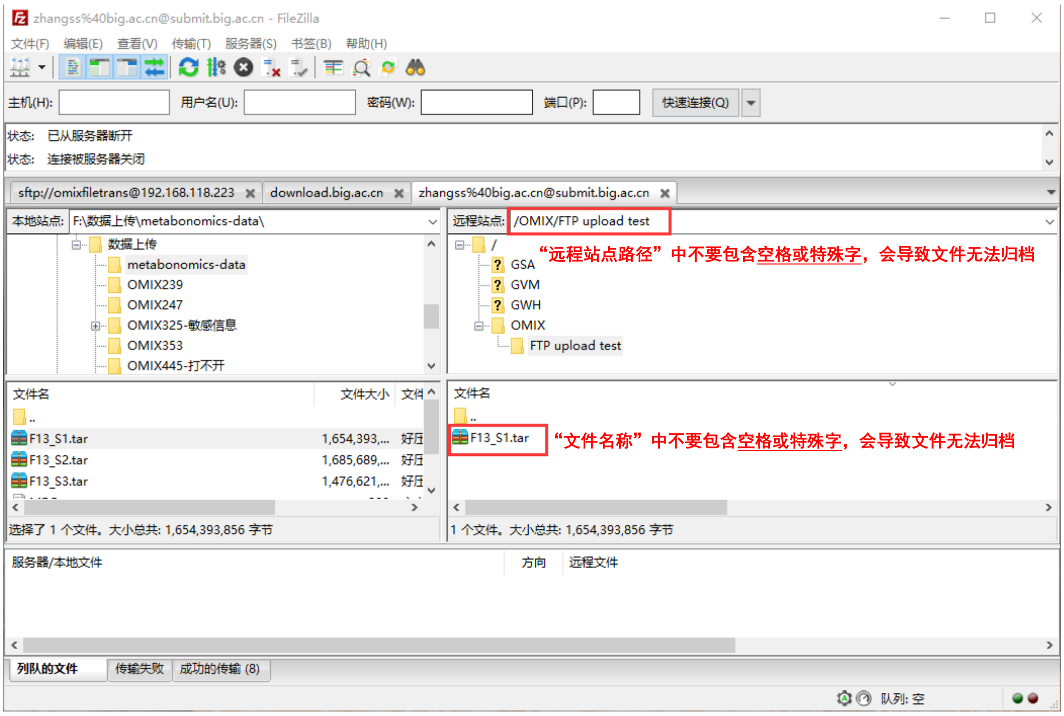
1 Files & Download

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

可能由于上网络问题，造成的数据文件未上传成功

- **数据文件归档异常**：请排查上传文件名称及其所在“远程站点路径”是否包含空格或特殊字符，如下图所示。如果包含空格或特殊字符，会影响数据归档。如果非以上情况，请及时联系数据库进行问题排查。



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	文件
..		
F13_S1.tar	1,654,393,...	好压
F13_S2.tar	1,685,689,...	好压
F13_S3.tar	1,476,621,...	好压

选择了 1 个文件。大小总共: 1,654,393,856 字节

文件名
..
F13_S1.tar

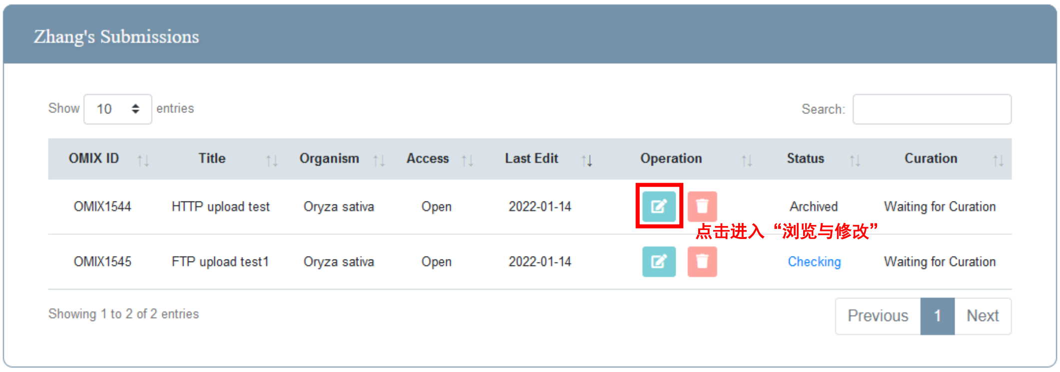
1 个文件。大小总共: 1,654,393,856 字节

“远程站点路径”中不要包含空格或特殊字，会导致文件无法归档

“文件名称”中不要包含空格或特殊字，会导致文件无法归档

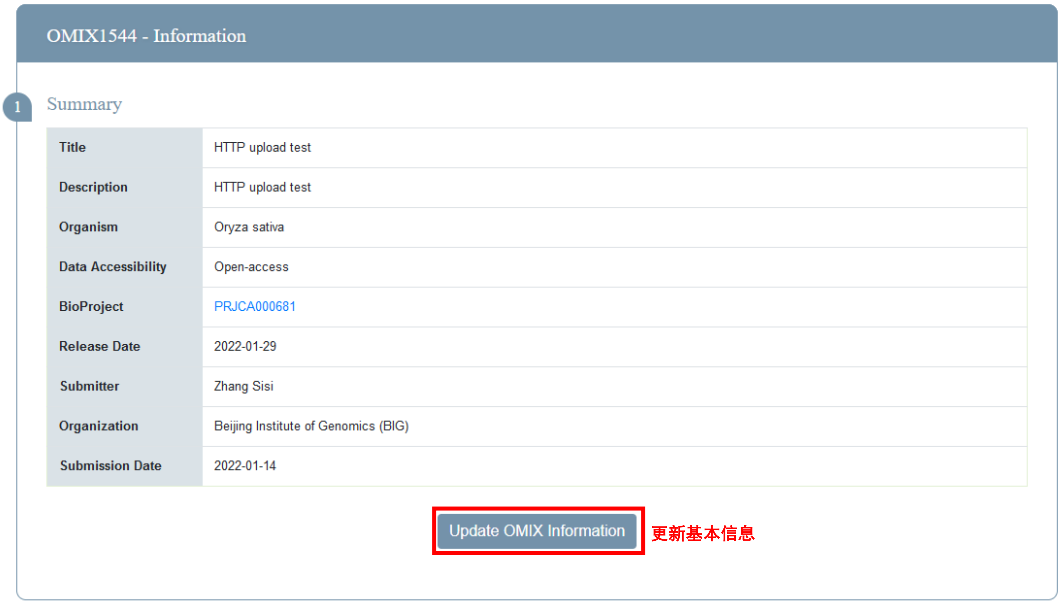
问题 3：如何修改数据集的基本信息？

在 OMIX 数据集未发布之前，用户可通过点击列表页的“✎”进入浏览与修改界面，更新基本信息（Update OMIX Information）。



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

点击进入“浏览与修改”

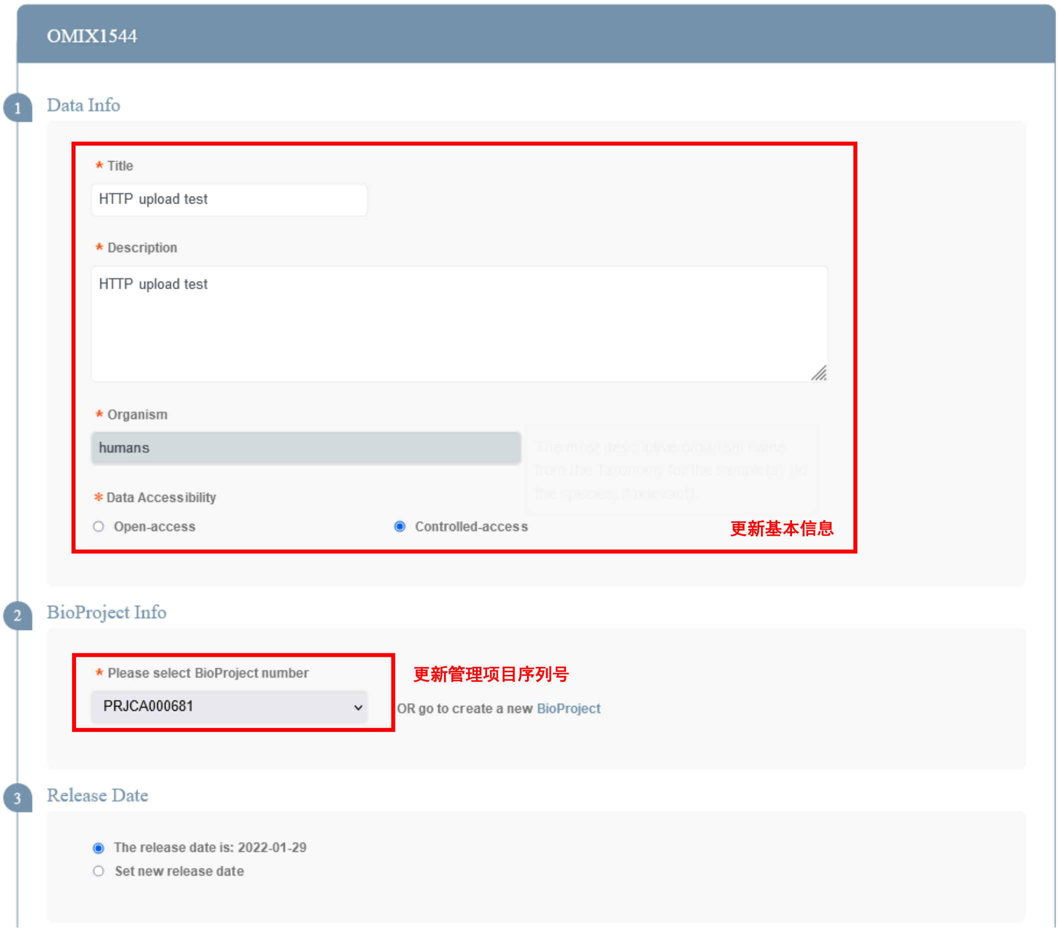


OMIX1544 - Information

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

更新基本信息



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

更新基本信息

2 BioProject Info

* Please select BioProject number

PRJCA000681

OR go to create a new BioProject

更新管理项目序列号

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

添加备案备份编号

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

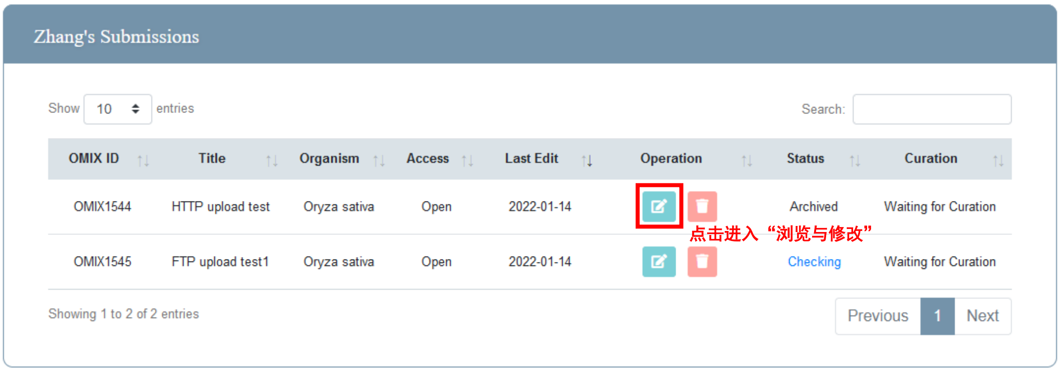
Please agree to our Terms and Policy

☐ You agree to our Terms and Policy.

问题 4：如何提前释放数据集？

在 OMIX 数据集未发布之前，用户可通过点击列表页的“✎”进入浏览与修改界面，更新基本信息（Update OMIX Information），重新设置发布时间。

注意：系统需要一段时间完成数据集释放（与数据量大小相关），请您耐心等待。



Zhang's Submissions

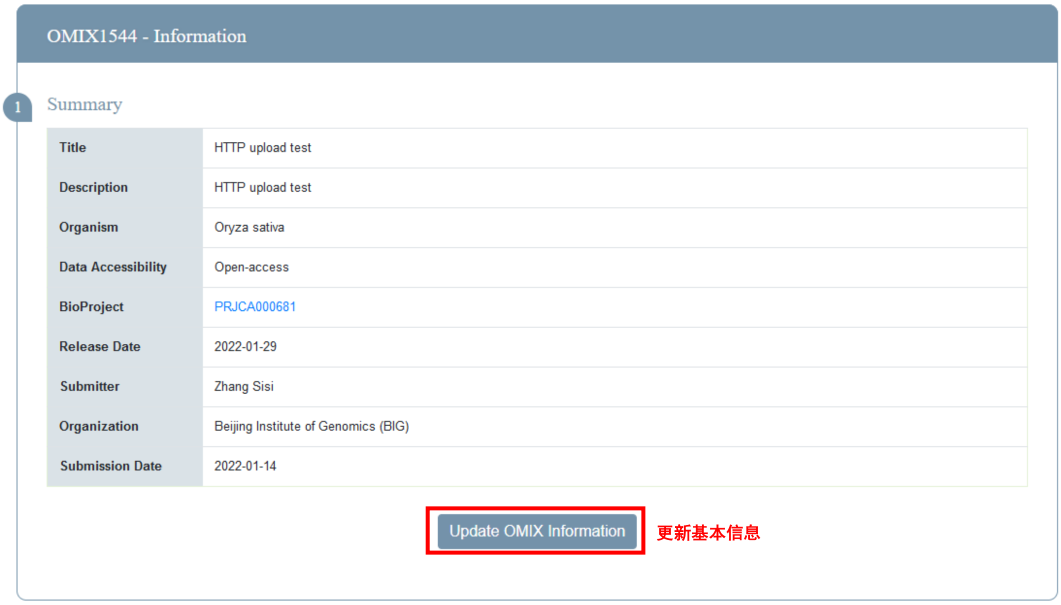
Show 10 entries　Search:

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

点击进入“浏览与修改”

Showing 1 to 2 of 2 entries

⇩



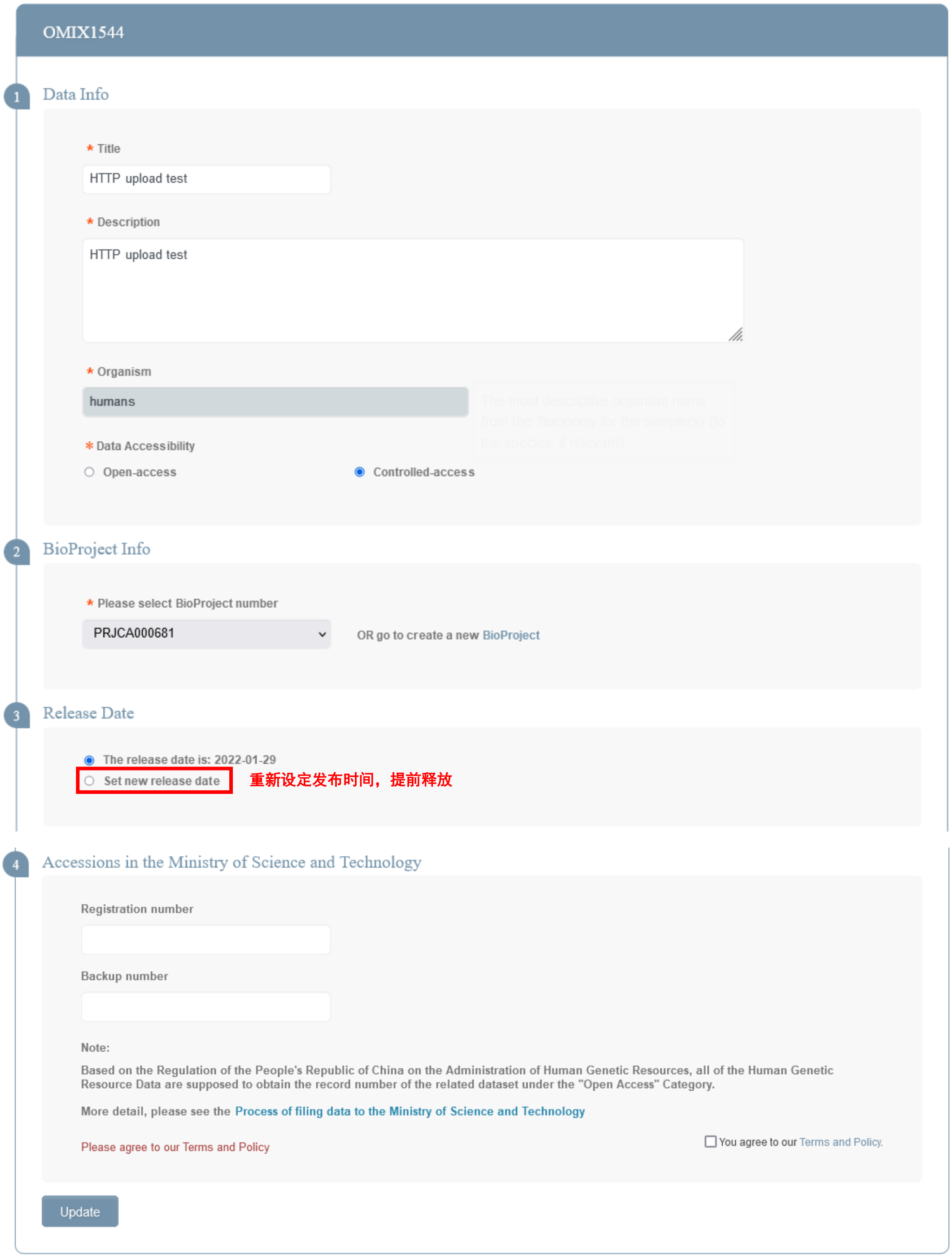
OMIX1544 - Information

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

更新基本信息

⇩



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 重新设定发布时间，提前释放

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.

Update

问题 5：如何追加或替换数据文件？

如果选择使用**页面提交方式“追加或替换文件”**（替换文件用户，需删除原文件后，在进行以下操作），您可点击 Add New File Online 进入数据文件提交页面，分别完成数据描述信息和数据文件填写上传。注：对应“**样本数/文件个数（Number of Samples/Files）**”，请填写数据文件所涉及的样本个数，如果提交文件未压缩

包或无法给出样本个数的用户，可以填写文件数。

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	

通过“页面提交方式”追加或替换文件

OMIX546 - New File

1 Files

* File No. (Auto): 3

* File Title

* Number of Samples/Files (for zip)

* Data Type: Please Select the Data Type

* Click here to select file for uploading

No file currently selected for uploading

填写基本信息，如果无法提供样本数，特别是

选择文件数据类型

点击上传文件

如果通过 **FTP 提交方式“追加或替换文件”**（替换文件用户，需删除原文件后，在进行以下操作），您可点击进 Add New File via FTP 入数据文件提交页面。首先，可从列表中勾选出已上传的数据文件，再填写对应的数据描述信息。数据文件上传方法详见“FTP 数据预上传”。

- 请务必**一次性“勾选”完毕后**，再开始**填写描述信息**。请勿一边**“勾选”**一边**填写**。因为如果**中途更改“勾选”项**，之前所填写的**信息将不会被保留**，全部需要重新填写
- 请务必确保已上传文件的文件名中**不允许有空格**

OMIX1544 - Files

1 Files & Download

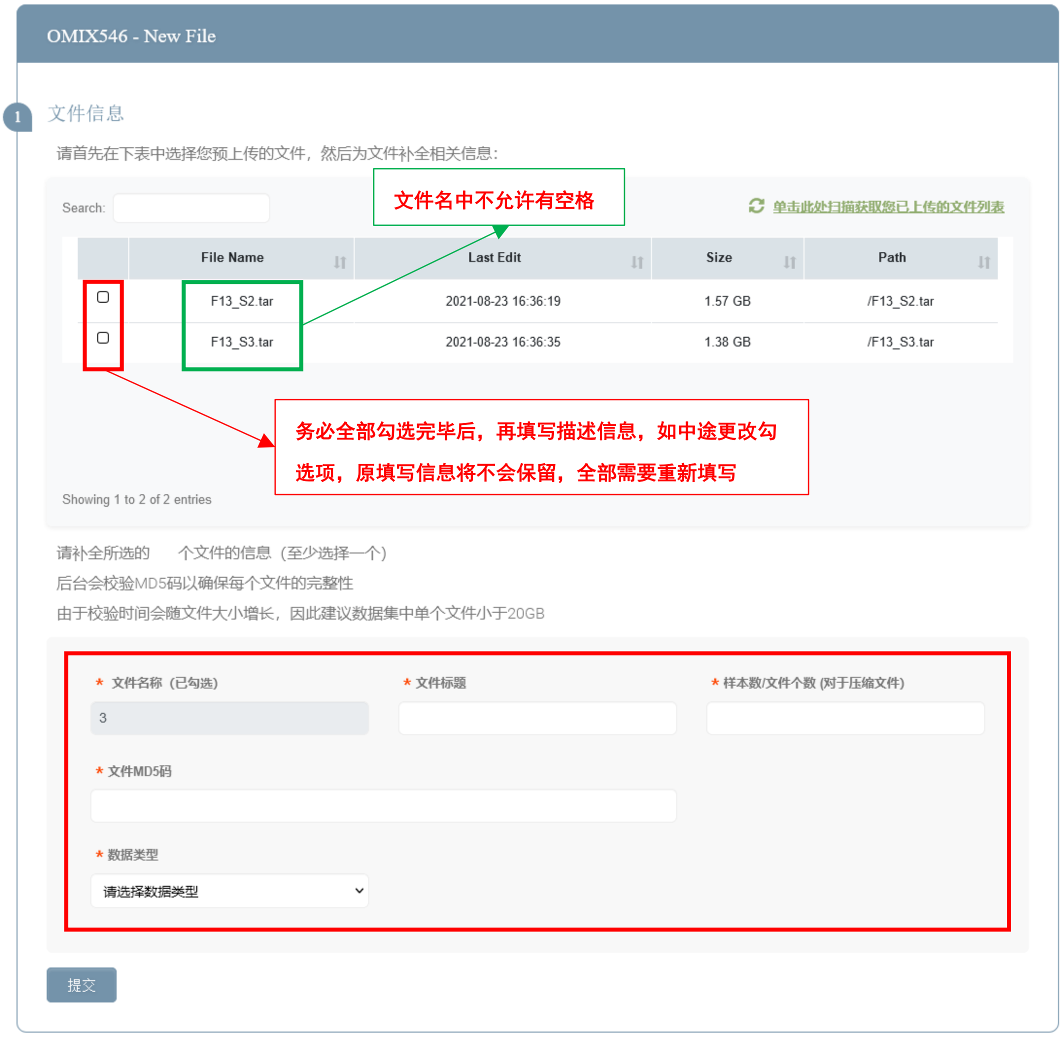
The backend will take some time to ensure the integrity of each file by checking MD5 hash values.

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	

通过“FTP 提交方式”追加文件

⇩



OMIX546 - New File

1 文件信息

请首先在下表中选择您预上传的文件，然后为文件补全相关信息：

文件名中不允许有空格

	File Name	Last Edit	Size	Path
□	F13_S2.tar	2021-08-23 16:36:19	1.57 GB	/F13_S2.tar
□	F13_S3.tar	2021-08-23 16:36:35	1.38 GB	/F13_S3.tar

务必全部勾选完毕后，再填写描述信息，如中途更改勾选项，原填写信息将不会保留，全部需要重新填写

Showing 1 to 2 of 2 entries

请补全所选的　　个文件的信息（至少选择一个）

后台会校验MD5码以确保每个文件的完整性

由于校验时间会随文件大小增长，因此建议数据集中单个文件小于20GB

* 文件名称（已勾选）: 3

* 文件标题

* 样本数/文件个数 (对于压缩文件)

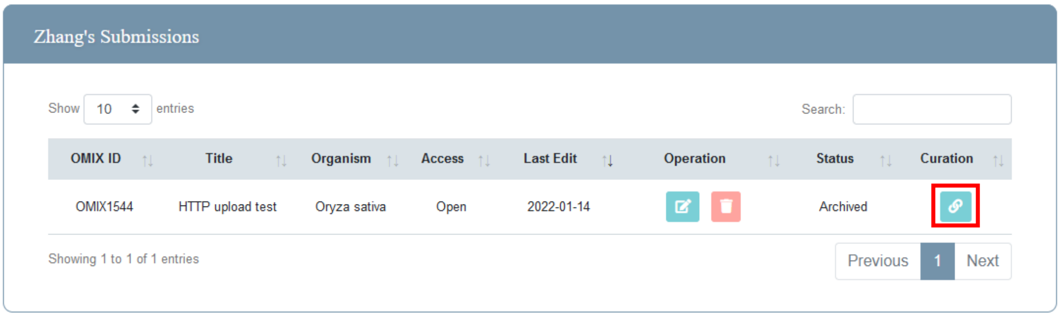
* 文件MD5码

* 数据类型: 请选择数据类型

问题 6：如何生成数据集共享链接？

数据集内所有文件归档后，“**数据状态和共享链接（Status/Review）**”显示为“🔗”。您可通过点击“🔗”在线生成共享链接。请注意以下问题。

- **为保护数据安全，共享链接暂不支持数据文件下载**
- **数据发布后共享链接将失效，请勿生成链接写在发表文章中！！！**



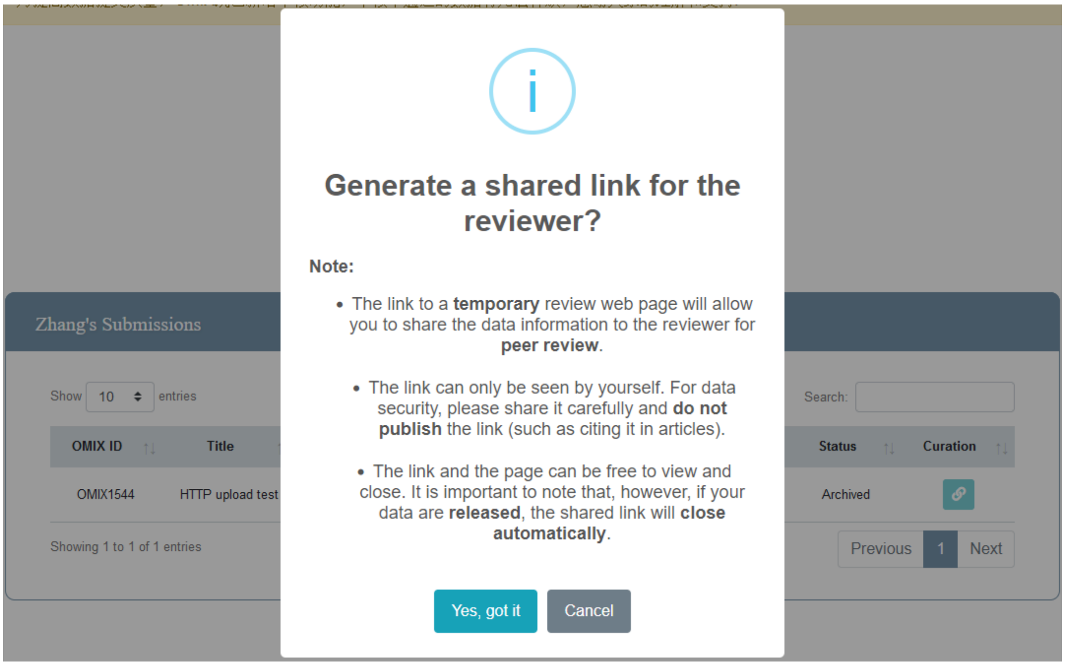
Zhang's Submissions

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	

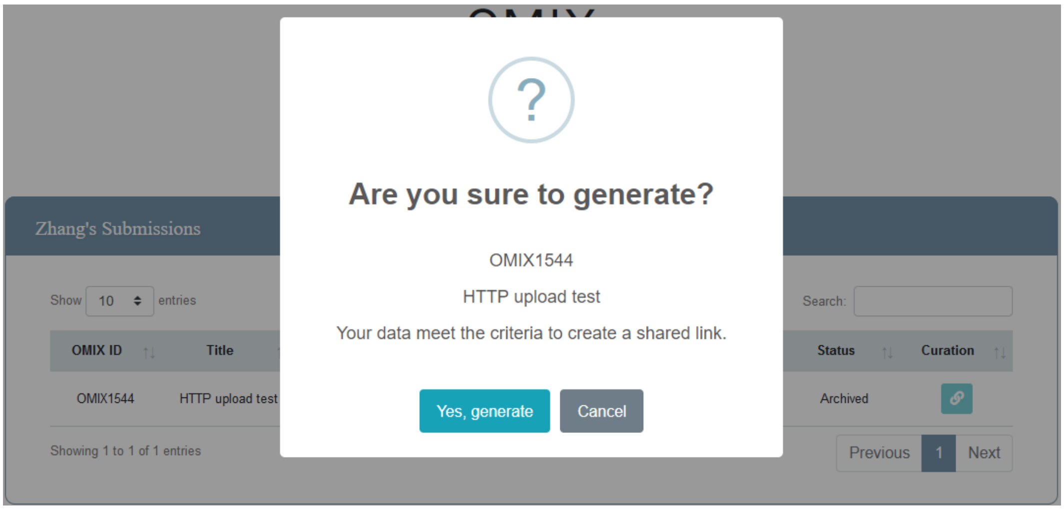
Showing 1 to 1 of 1 entries



Generate 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**.

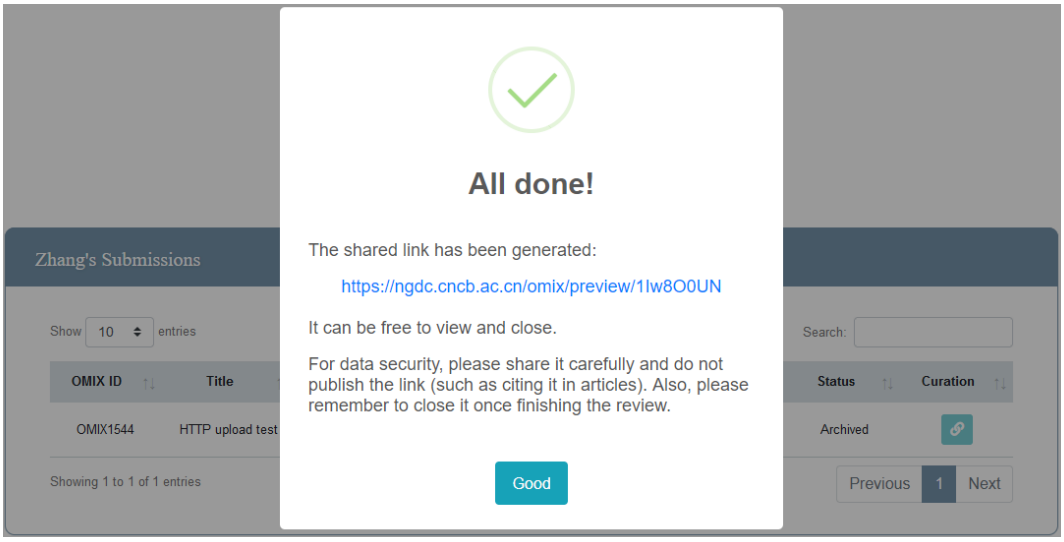


Are you sure to generate?

OMIX1544

HTTP upload test

Your data meet the criteria to create a shared link.



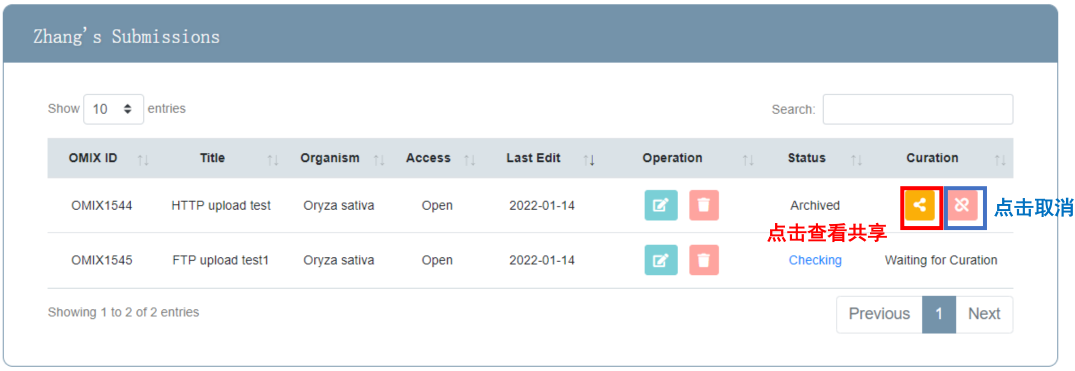
All done!

The shared link has been generated:

https://ngdc.cncb.ac.cn/omix/preview/1lw8O0UN

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.



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		Checking	Waiting for Curation

点击取消共享

点击查看共享

问题 7：如何补充人类遗传资源事先报告登记号？

数据集未发布前，您可随时补充事先报告登记号。首先在提交列表页找到目标数据集，点击"operation"列中的“✎”Edit 按键，进入信息编辑页面。

Zhang’s Submissions

Show 10 entries

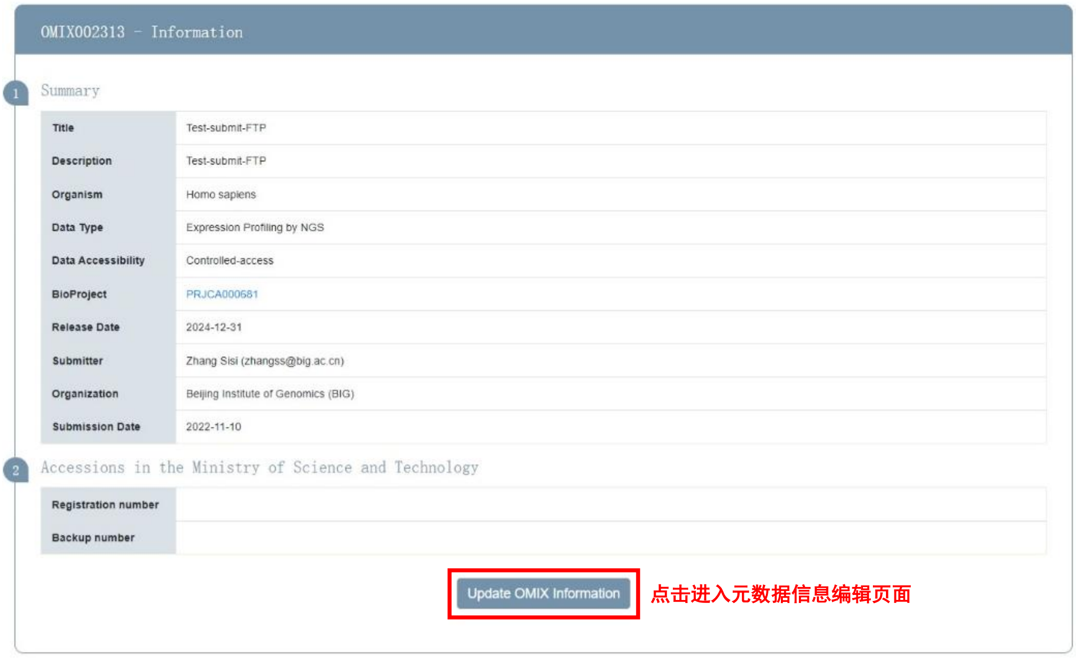
Search: OMIX002313

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX002313	Test-submit-FTP	Homo sapiens	Controlled	2022-11-10	✎ 🗑	Archived	📁

Showing 1 to 1 of 1 entries (filtered from 36 total entries)

点击进入数据集浏览页面

然后，点击“update OMIX Information”进入元数据信息编辑页面。



OMIX002313 - Information

1 Summary

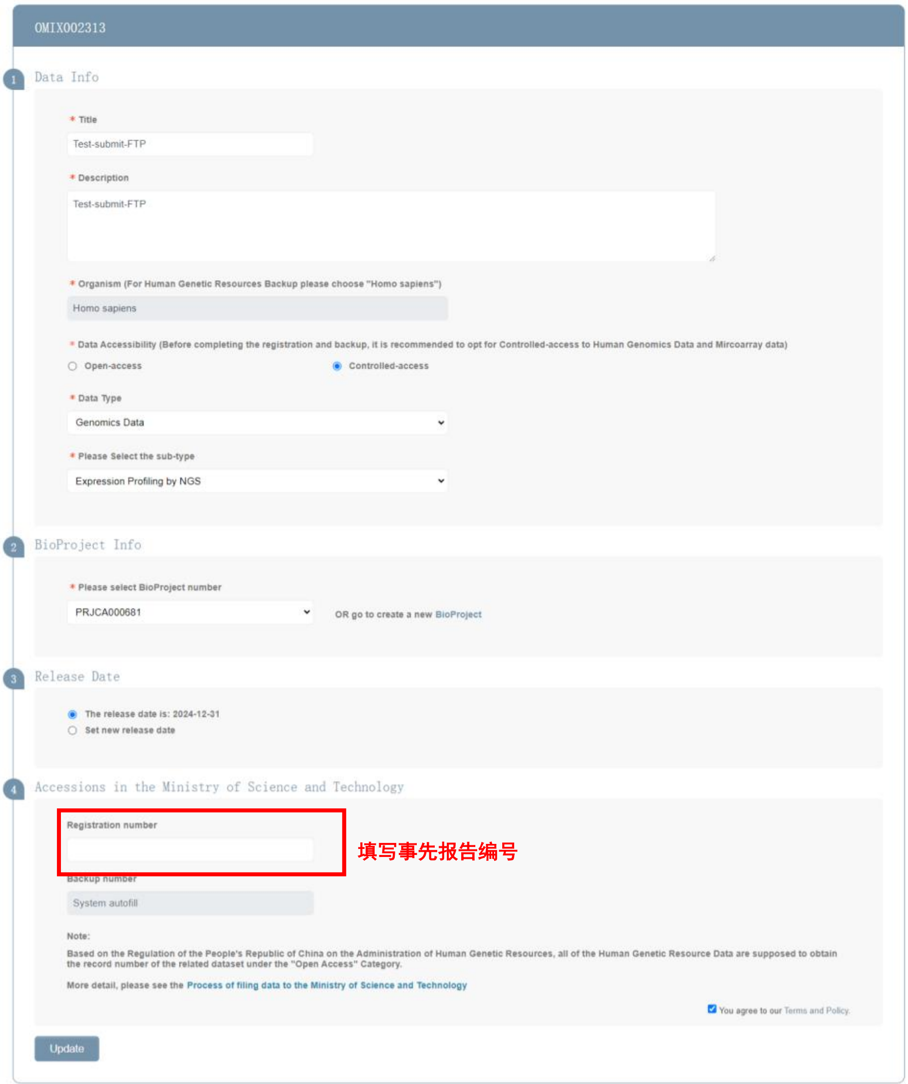
Title	Test-submit-FTP
Description	Test-submit-FTP
Organism	Homo sapiens
Data Type	Expression Profiling by NGS
Data Accessibility	Controlled-access
BioProject	PRJCA000681
Release Date	2024-12-31
Submitter	Zhang Sisi (zhangss@big.ac.cn)
Organization	Beijing Institute of Genomics (BIG)
Submission Date	2022-11-10

2 Accessions in the Ministry of Science and Technology

Registration number	
Backup number	

点击进入元数据信息编辑页面

在“Accessions in the Ministry of Science and Technology”的“Registration number”中完成事先报告登记号添加。



OMIX002313

① Data Info

* Title

Test-submit-FTP

* Description

Test-submit-FTP

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

Homo sapiens

* Data Accessibility (Before completing the registration and backup, it is recommended to opt for Controlled-access to Human Genomics Data and Mircoarray data)

○ Open-access ● Controlled-access

* Data Type

Genomics Data

* Please Select the sub-type

Expression Profiling by NGS

② BioProject Info

* Please select BioProject number

PRJCA000681 OR go to create a new BioProject

③ Release Date

● The release date is: 2024-12-31
○ Set new release date

④ Accessions in the Ministry of Science and Technology

Registration number

填写事先报告编号

Backup number

System autofill

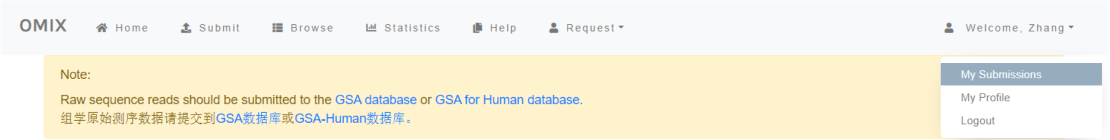
Note:

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.

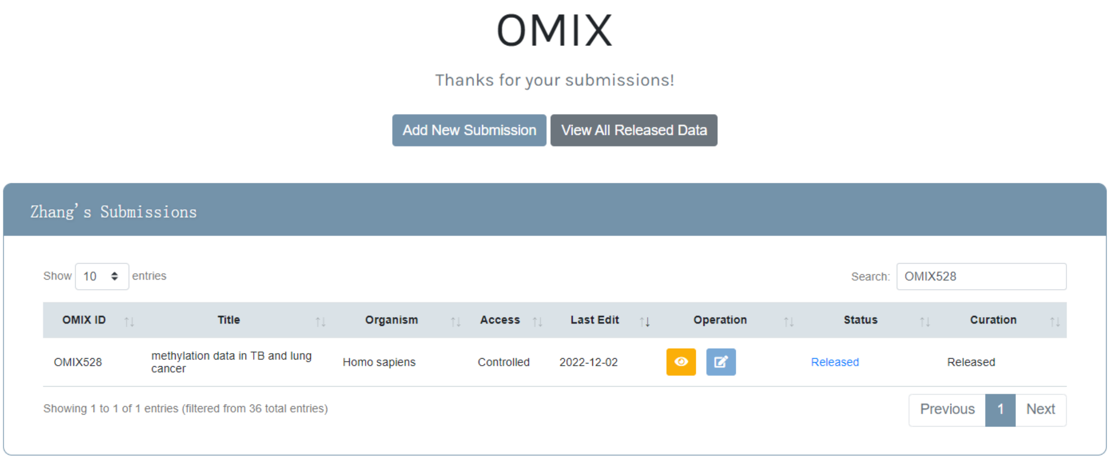
数据集发布后，也可自行补充事先报告登记号。首先登录 OMIX 系统，找到并打开数据提交列表。



Note:

Raw sequence reads should be submitted to the GSA database or GSA for Human database.
组学原始测序数据请提交到GSA数据库或GSA-Human数据库。

通过 search 功能找出目标数据集，点击"operation"列中的“✎”update 按键，进入信息编辑页面。



OMIX

Thanks for your submissions!

Zhang's Submissions

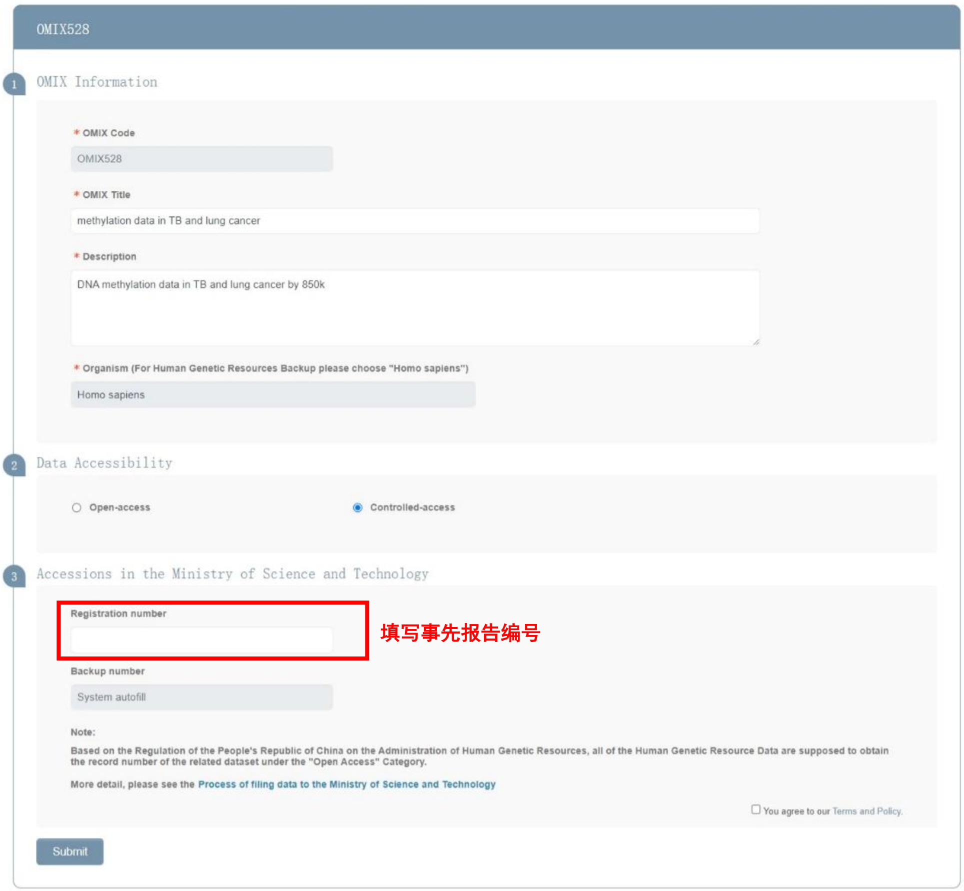
Show 10 entries

Search: OMIX528

OMIX ID	Title	Organism	Access	Last Edit	Operation	Status	Curation
OMIX528	methylation data in TB and lung cancer	Homo sapiens	Controlled	2022-12-02		Released	Released

Showing 1 to 1 of 1 entries (filtered from 36 total entries)

然后，在“Accessions in the Ministry of Science and Technology”的“Registration number”中完成事先报告登记号添加。



OMIX528

1 OMIX Information

* OMIX Code: OMIX528

* OMIX Title: methylation data in TB and lung cancer

* Description: DNA methylation data in TB and lung cancer by 850k

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

2 Data Accessibility

○ Open-access ● Controlled-access

3 Accessions in the Ministry of Science and Technology

Registration number:

填写事先报告编号

Backup number: System autofill

Note:

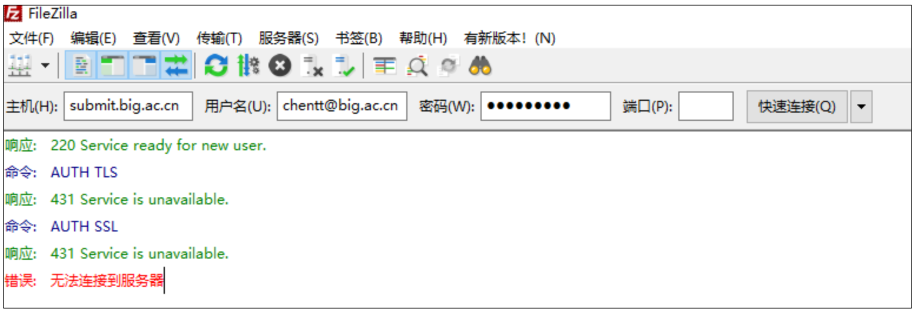
Based on the Regulation of the People's Republic of China on the Administration of Human Genetic Resources, all of the Human Genetic Resource Data are supposed to obtain the record number of the related dataset under the "Open Access" Category.

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

☐ You agree to our Terms and Policy.

FTP 上传常见问题与对策

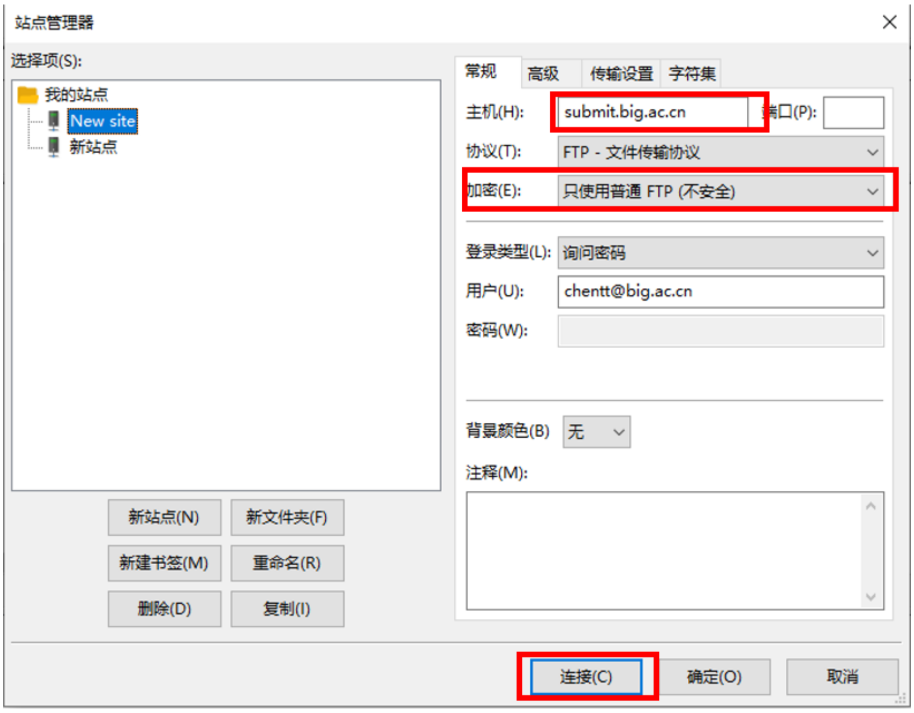
问题 1：登陆时，状态栏显示 AUTH SSL 的报错信息



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.
错误:  无法连接到服务器
```

解决方案： FTP 登陆时状态栏出现 AUTH SSL 的报错信息，请通过下图所示点击菜单栏“文件”中的“站点管理器”，修改“加密”选项为“只使用普通 FTP”或者“”，同时，填写正确的主机地址：submit.big.ac.cn，帐号和密码信息。最后点击“连接”即可。



主机(H): submit.big.ac.cn 端口(P):

协议(T): FTP - 文件传输协议

加密(E): 只使用普通 FTP (不安全)

登录类型(L): 询问密码

用户(U): chentt@big.ac.cn

密码(W):

背景颜色(B) 无

注释(M):

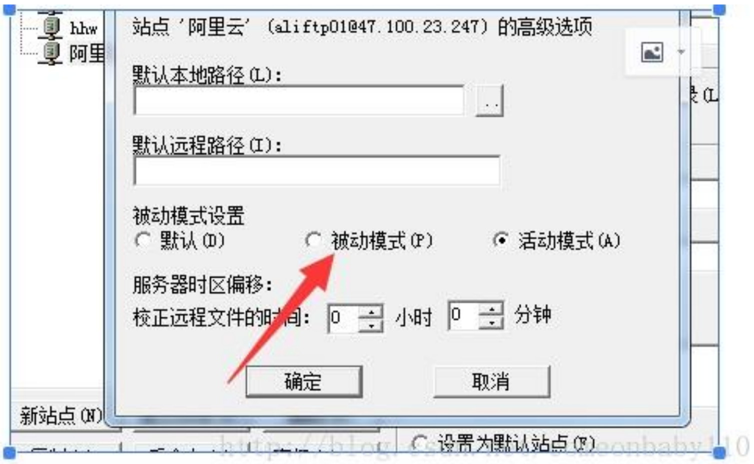
问题 2：登陆时，状态栏显示 MLSD 的报错信息

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\

解决方案：FTP 登陆时，状态栏出现 MLSD 的报错，并显示“读取目录列表失败”如何处理。Filezila –>编辑->设置中修改传输模式，改为被动模式（如下图所示）。



站点 ‘阿里云’ （aliftp01@47.100.23.247）的高级选项

默认本地路径(L):

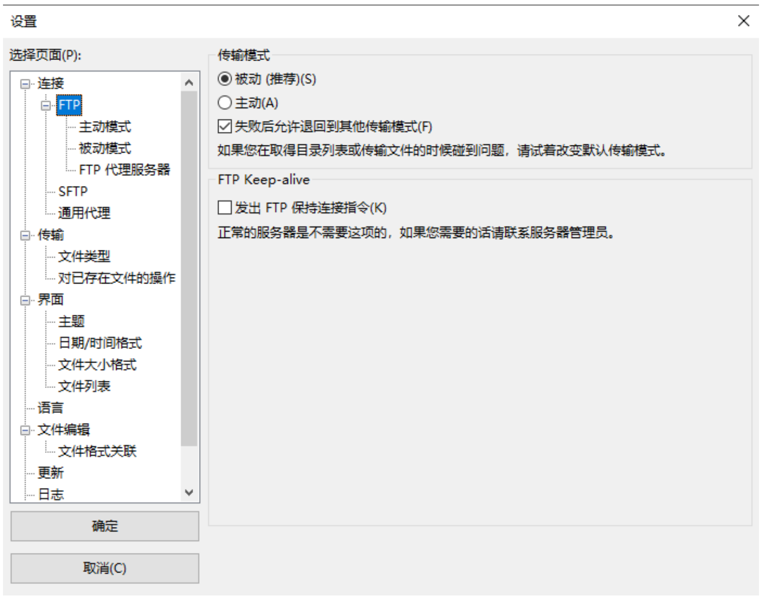
默认远程路径(I):

被动模式设置

○ 默认(D) ○ 被动模式(P) ● 活动模式(A)

服务器时区偏移：

校正远程文件的时间: 0 小时 0 分钟



设置

选择页面(P):

传输模式

● 被动 (推荐)(S)

○ 主动(A)

☑ 失败后允许退回到其他传输模式(F)

如果您在取得目录列表或传输文件的时候碰到问题，请试着改变默认传输模式。

FTP Keep-alive

☐ 发出 FTP 保持连接指令(K)

正常的服务器是不需要这项的，如果您需要的话请联系服务器管理员。

数据下载常见问题与对策

问题 1：如何通过 HTTP 方式下载数据？

方法一：点击下载按钮“⬇”，选中“HTTP Download”并点击进入下载列表页面，通过点击文件名称（Name）下载所需数据。

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)

⇩

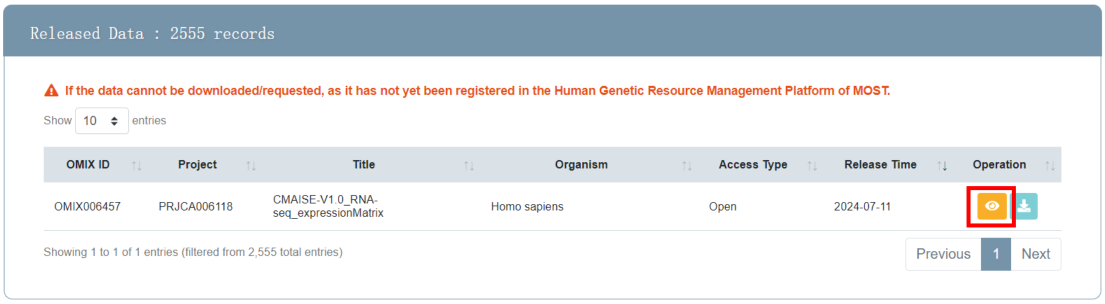
download.cncb.ac.cn/OMIX/OMIX006457/

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

点击名称（Name）下载数据

方法二：点击查看按钮“👁”进入数据集详细页面。通过点击 Files & Download 部分的“HTTPs”按钮逐个下载数据。



Released Data : 2555 records

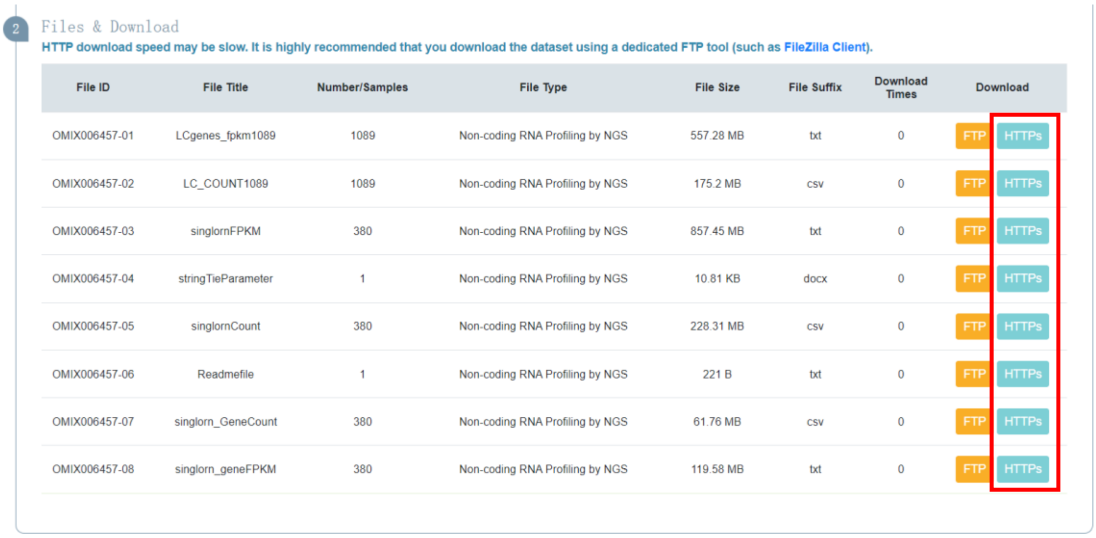
⚠ If the data cannot be downloaded/requested, as it has not yet been registered in the Human Genetic Resource Management Platform of MOST.

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)

⇩



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

问题 2：如何通过 FTP 方式下载数据？

许多浏览器（如 Chrome、Firefox）出于安全性考虑（FTP 协议缺乏加密，容易被攻击），已经逐步取消对 FTP 文件传输协议的支持。这意味着用户无法通过这些浏览器直接访问 FTP 服务器上的文件。

解决方案：请使用 FileZilla、WinSCP 等专用 FTP 客户端。以 FileZilla 为例，具体方法如下：

第 1 步 下载并安装 FTP 客户端（https://filezilla-project.org/ ），下载页面如下图所示，点击红色框中的“Download FileZilla Client”，并按照提示安装软件；

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

下载 FileZilla Client 软件

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

第 2 步 请点击下载按钮“⬇”，**不要直接点击“FTP Download”！！**相反，请右键点击“FTP Download”，然后选择“**复制链接地址**”。

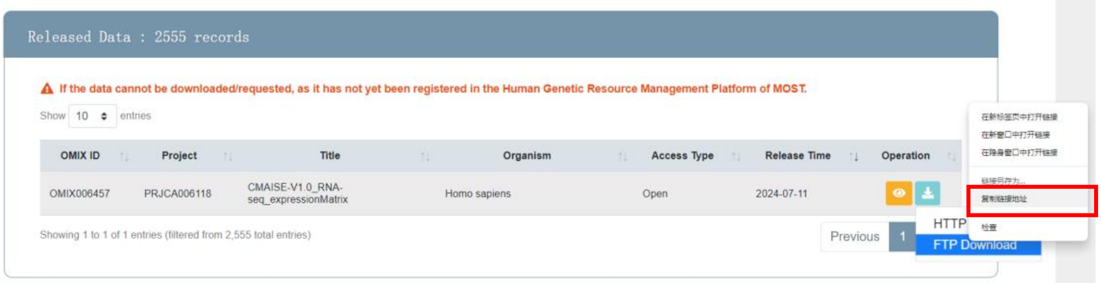
Released Data : 2555 records

⚠ 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)



Released Data : 2555 records

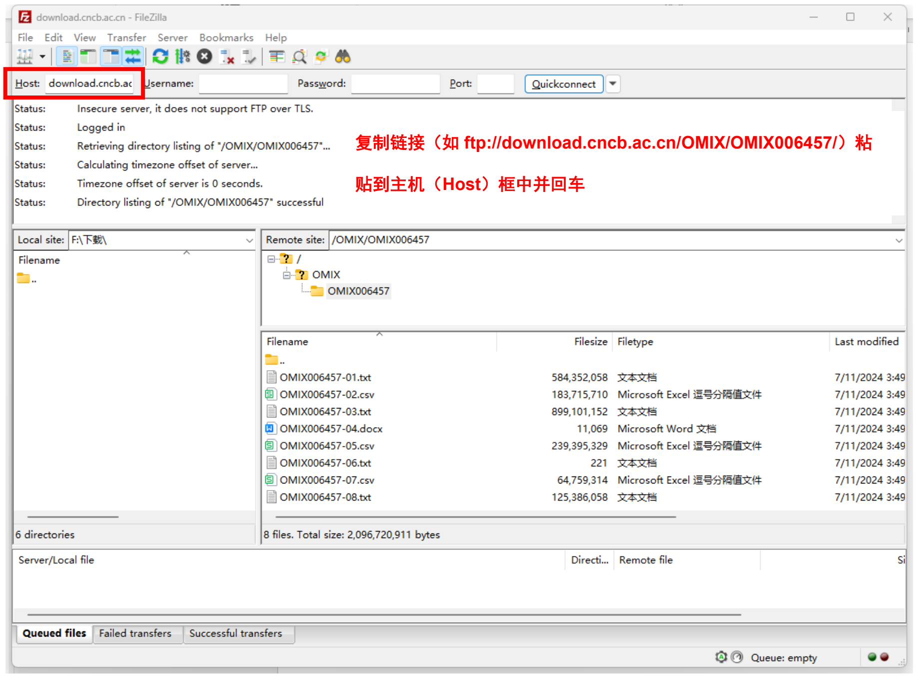
⚠ If the data cannot be downloaded/requested, as it has not yet been registered in the Human Genetic Resource Management Platform of MOST.

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)

第 3 步 打开“FileZilla Client”软件，将复制的链接地址（如ftp://download.cncb.ac.cn/OMIX/OMIX006457/）粘贴到主机（Host）框中并回车，便可看到所需数据。



Host: download.cncb.ac Username: Password: Port: Quickconnect

```
Status:     Insecure server, it does not support FTP over TLS.
Status:     Logged in
Status:     Retrieving directory listing of "/OMIX/OMIX006457"...
Status:     Calculating timezone offset of server...
Status:     Timezone offset of server is 0 seconds.
Status:     Directory listing of "/OMIX/OMIX006457" successful
```

复制链接（如 ftp://download.cncb.ac.cn/OMIX/OMIX006457/）粘贴到主机（Host）框中并回车

Local site: F:\下载\

Remote site: /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

6 directories

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