

# ProteomeXchange 上传数据说明

## 1. 介绍:

---

在进行蛋白质组学相关领域的期刊投稿时，通常会要求作者在第三方平台上提交蛋白质组学的原始数据。其中 ProteomeXchange（简称 PX）是应用较为广泛的蛋白质组学质谱数据的存储平台（<http://www.proteomexchange.org>），研究者可以根据需要选择通过 ProteomeXchange 客户端直接上传质谱的原始数据。

## 详细教程

1. 进入 Proteomexchange 网站界面，地址：<http://www.proteomexchange.org/>



2. 下拉页面，点击Submit Data

### Public Data



Public PXD datasets can be browsed over at [ProteomeCentral](#). An [RSS feed](#) is also available.

### Data Submission

ProteomeXchange fully supports both MS/MS proteomics and SRM data submission. Submissions of other types of proteomics data is also possible using the Partial Submission mechanism.



### Subscribe

Subscribe to receive all new ProteomeXchange announcements!



3. 点击 Download 获得数据上传小工具



## Data submission

ProteomeXchange supports submission of experiments coming from all proteomics data workflows.

### PRIDE - PRoteomics IDentifications Database



The GUI based PX Submission Tool can be **downloaded** to start data upload.

下载上传工具

### PeptideAtlas - PASSEL



SRM/MRM data can be submitted to [PASSEL](#).

### MassIVE

Shotgun proteomics data can be submitted to [MassIVE](#).

下载后的数据上传工具需要保证电脑安装有 java【安装包路径：<https://www.java.com/zh-CN/download/>】

解压缩下载工具px-submission-tool.zip，双击 px-submission-tool-x.x.x.jar 打开数据上传工具。

4. 打开上传工具后，首次使用需要进行注册，注册完成后，密码会发送到注册的邮箱中，之后在上传工具中使用该账号密码进行登陆即可。有账号的可以直接填入信息，登陆。

The screenshot shows the ProteomeXchange login and registration interface. At the top, it says "Login" and "Login to your PRIDE account (version 2.5.2)". Below this is a form with two input fields: "Email\*" and "Password\*". To the right of the "Password\*" field is a link that says "Register New User". A red arrow points to this link with the text "首次使用点击进行注册". At the bottom right, there is a red arrow pointing to the "Register New User" link with the text "已经有账号的填★邮箱密码，点击登陆". The bottom of the screen has a navigation bar with a question mark icon, a "C..." button, a "Back" button, and a "L..." button.

5. 登陆后会显示选择是完全上传还是部分上传，虽然网站显示的更为推荐的选项为全部上传（Complete Submission），但一般来讲只需要上传原始数据的话选择部分上传（Partial Submission）就可以；

PRIDE Submission Tool Version - 2.5.2

Step 1: Submission Type (1/11)  
Selection of Submission Type

ProteomeXchange

Choose submission option below ☐ Controlled access data

**Complete Submission (recommended)**

☒ Use this option if you can provide your identification results in either mzIdentML or mzTab. It will then be possible to fully integrate the results in PRIDE and visualise them (e.g. as required by MCP).

In addition to a PXD identifier, a permanent Digital Object Identifier (DOI) will be provided to uniquely identify the dataset.

**Partial Submission**

☐ You should only choose this option if your search results cannot be converted to mzIdentML or mzTab. Identifications will not be integrated in PRIDE. However, files will be made available to download (and may be visualised with other external tools).

A PXD identifier will be provided to uniquely identify the dataset, but not a DOI.

[Resubmission](#) [Bulk submission](#) [Submission guidelines](#)

? C... < Back Next >

选择部分上传后会出现点击 Next 的界面，直接确认继续即可；继续后有相应的部分上传（Partial Submission）中所需要的数据提示，这一步也直接点击 Next 即可

PRIDE Submission Tool Version - 2.5.2

ProteomeXchange

Prerequisites

Before we begin, please check you have the following information

Search Engine Results

Protein/Peptide Identificat...

	✓
	✓

Raw Data

MS instrument raw output

.O

101010100

100001001

010101001

011100001

PRIDE Login

PRIDE user credentials

Experiment Details

Sample, protocols, species, t...

Lab Head

Name, email, and affiliation

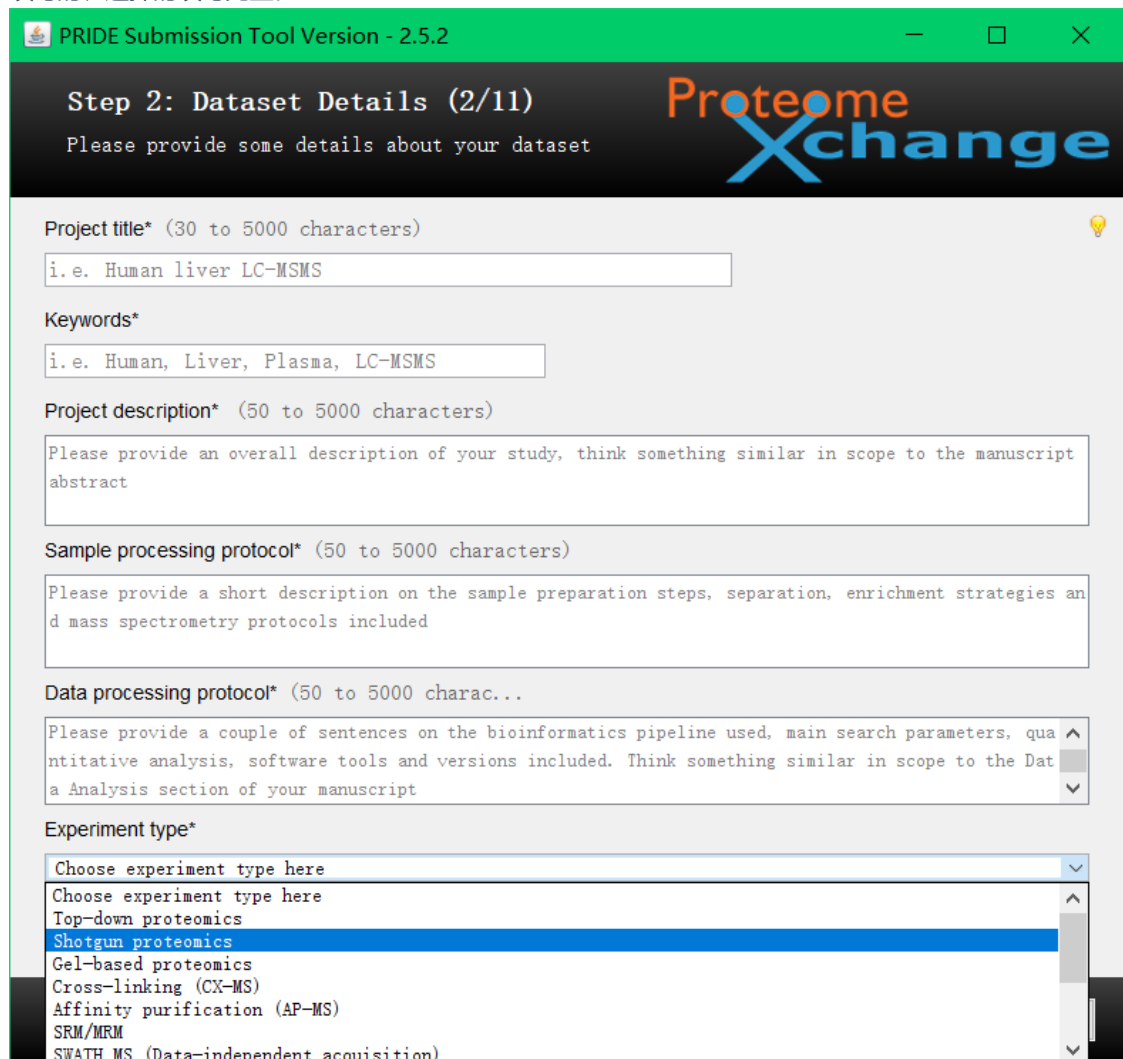
?

✕C...

< Back

Next >

6. 按照以下空格处的示例，填写具体的项目信息，此处注意好每一栏的字数限制，超出限制或者不达标的情况下，是无法进行提交的；最下一栏中有下拉框需要选择具体的组学方法，按照项目类型选择即可。把\*要填写的、选择的填写完整。



PRIDE Submission Tool Version - 2.5.2

## Step 2: Dataset Details (2/11)

Please provide some details about your dataset

**Project title\*** (30 to 5000 characters)

i.e. Human liver LC-MSMS

**Keywords\***

i.e. Human, Liver, Plasma, LC-MSMS

**Project description\*** (50 to 5000 characters)

Please provide an overall description of your study, think something similar in scope to the manuscript abstract

**Sample processing protocol\*** (50 to 5000 characters)

Please provide a short description on the sample preparation steps, separation, enrichment strategies and mass spectrometry protocols included

**Data processing protocol\*** (50 to 5000 characters)

Please provide a couple of sentences on the bioinformatics pipeline used, main search parameters, quantitative analysis, software tools and versions included. Think something similar in scope to the Data Analysis section of your manuscript

**Experiment type\***

Choose experiment type here

- Choose experiment type here
- Top-down proteomics
- Shotgun proteomics**
- Gel-based proteomics
- Cross-linking (CX-MS)
- Affinity purification (AP-MS)
- SRM/MRM
- SWATH MS (Data-independent acquisition)

7. 填写相关信息后，就是上传原始数据的界面了。

点击“Add Files”，一般 RAW 文件指蛋白组学下机的原始数据（格式为.RAW），我们交付的Raw-文件夹；Search 为搜库数据，我们交付的Match\_result文件夹，上传后的 File Type 选择 Search 即可；大多数情况下，我们选择的都是部分上传（Partial Submission），这部分 RAW 和 Search 为必须文件。

PRIDE Submission Tool Version - 2.5.2

Step 3: Add Files (3/11)

Add the files you want to submit

ProteomeXchange

Add Files
Remove...

? Which are the requir...

File Name	PATH / URL	...	File Type	Remove
checksum.txt	C:\Users\dec\Desktop\px-submission-tool\px...	15	OTHER	✗
HFX6_FPEP210001...	C:\Users\dec\Desktop\HFX6_FPEP210001977-1A...	6...	RAW	✗
Match_result-X1...	C:\Users\dec\Desktop\Match_result-X101SC21...	1...	SEARCH	✗

?
✗ C...
Back
Next

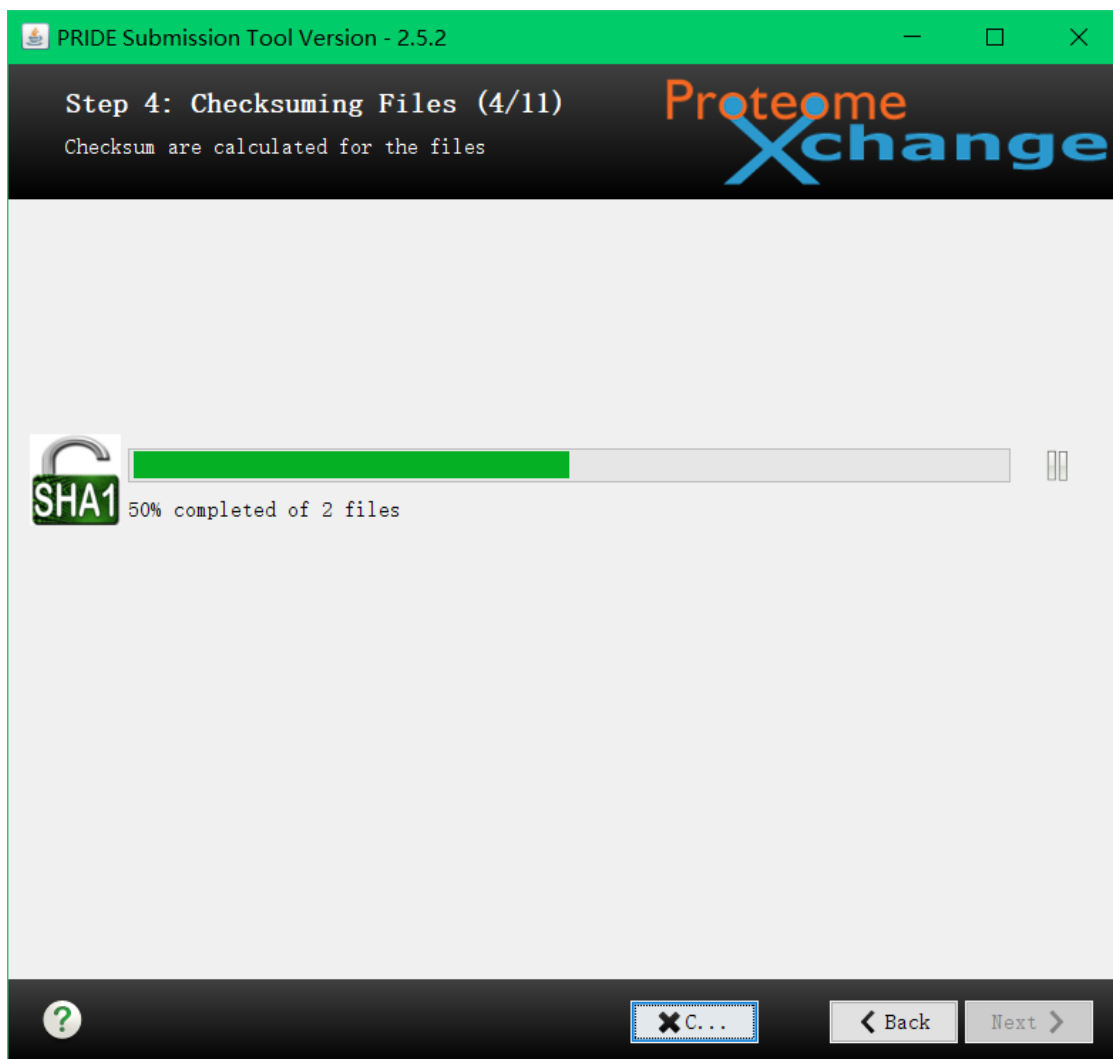
## File Type

[Previous](#) | [Next](#)

The File Selection step asks the user to assign correct file types to the files to be submitted. The submission tool does the best to assign file types automatically, however, it is worth double-checking the file types.

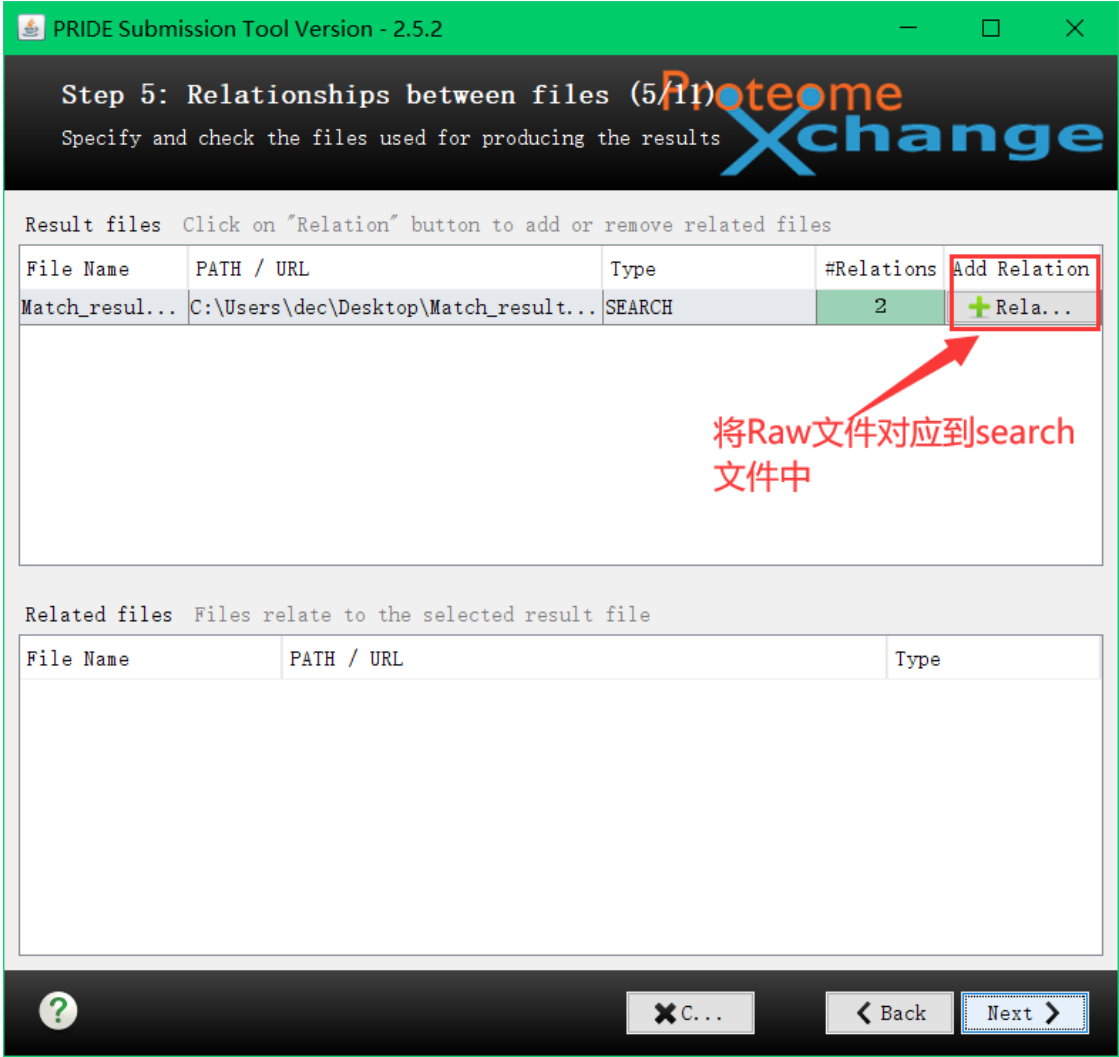
### Supported File Types

File Type	Requirement	Description
RESULT	Required (for <b>Complete Submissions</b> ), not applicable for Partial Submissions.	Results of peptide/protein identification. Usually produced and exported by the Search Engine. mzIdentML and mzTab are the main supported file formats. A legacy file format, PRIDE XML is also accepted but no longer actively supported.
RAW	Required for <b>both Complete and Partial Submissions</b>	MS instrument raw output, for example Thermo .raw or Bruker .baf files. Non-processed mzML or mzXML files are also accepted. In cases where instrument output is contained in directories (e.g. Waters Synapt instruments), those should be zipped.
SEARCH	Required (for <b>Partial Submissions</b> ), Optional for Complete Submissions (recommended to include).	Search Engine or analysis pipeline output. The files should contain peptide and protein identifications. Examples include MASCOT .dat files, or Proteome Discoverer .msf files.
PEAK	Required for <b>mzIdentML-based Complete submissions</b> . Recommended for Partial submissions.	Processed peak list files, such as: mgf, dta, ms2, pkl, or processed mzML/mzXML files.
QUANT	Optional	Quantification analysis output files, for example: list of protein fold changes, MaxQuant proteinGroups.txt file, mzQuantML.
GEL	Optional	Gel image file such as: TIF, JPG, or PNG files
FASTA	Optional	Sequence Database file which contains protein sequences used by the Search Engine
SPECTRUM_LIBRARY	Optional	When applicable, the files used to perform mass Spectral Search.
MS_IMAGE_DATA	Optional, but <b>required when there are MS image raw files</b>	MS image data, such as: imzML
OPTICAL_IMAGE	Optional	Optical images
OTHER	Optional	Any other files related to the submission (e.g. scripts)



8. 上传完成后，即可点击 Next，进入下一步；

这一步需要点击 Relate 按钮将 RAW 文件对应到 Search 文件中，这部分也比较好理解，就是如果是所有的 RAW 文件一起查库的话，就是对应同一个查库文件；有多个查库文件的情况下，需要将查库结果和原始数据一一对应进去。





9. 提交后需要填写项目的具体信息，再次提交后会弹出填写实验室基本信息的界面，这部分填写相应的信息即可；

PRIDE Submission Tool Version - 2.5.2

Step 7: Additional Details (7/11)

ProteomeXchange

Please give additional details about your submission

<b>Species*</b>		<b>Tissue*</b>	
Choose sample species here		Choose tissue here	
Homo sapiens (Human)	✖	Cell culture	✖
物种信息		样本的类型	
<b>Modification*</b>		<b>Instrument*</b>	
Choose modifications here		Choose MS instruments here	
Label:180(2)	✖	Q Exactive HF	✖
修饰		仪器型号	
<b>Cell type</b>		<b>Disease</b>	
Choose cell type here		Choose disease here	
细胞类型		疾病相关	
<b>Quantification method</b>			
Choose quantification method here			
定量方式			

?

✖ C...

< Back

Next >

如果蛋白定量方法下拉栏中无对应您的定量方法，可以选择最下方的 Others quantification method，点击后会弹出以下窗口，在 Term Name Search 一栏中输入对应词条点击 Search 按钮，系统会检索到对应条

目，选择该条目，点击右下角的 Use selected term 即可。

Ontology Lookup Service - (ols-dialog v3.4.16) ×

**Search Parameters**

Ontology PRIDE Controlled Vocabulary [PRIDE]

**Term Name Search** Term ID Search PSI-MOD Mass Search Browse Ontology

Term Name label free 13 Search

**Search Results**

Accession	CV Term
<a href="#">PRIDE:00004</a>	MS1 intensity based label-free quantification method
<a href="#">PRIDE:00003</a>	Label free
<a href="#">MS:1002038</a>	label free sample
<a href="#">PRIDE:00003</a>	Gel-free quantification method
<a href="#">CHMO:00010</a>	free-flow electrophoresis
<a href="#">CHMO:00015</a>	freeze drying
<a href="#">PRIDE:00005</a>	Label

**Selected Term** View Term Hierarchy

Name Value

Use Selected Term Cancel

10. 填写实验室信息，填写后点击next

PRIDE Submission Tool Version - 2.5.2 — □ ×

**Step 8: Lab Head (8/11)** ProteomeXchange

Please provide contact details of your lab head

**按照示例填写名字，邮箱联系方式即可（必填）**

**Name\***

Lab head's first name, middle name and last name, i.e. John Arthur Smith

**Email\***

Lab head's email address

**Affiliation\***

Lab head's affiliation, such as: department, lab, institute and country

NOTE: We are collecting this information for grouping submissions by lab and as a contact backup.

? ✕ C... < Back Next >

11. 填写基本信息后，系统会问你该提交的项目是否是属于其他大项目中的子项目，一般这一步都是可以直接下一步的；

**Step 9: Additional dataset details (9/11)** Step9全部为非必填项

Please provide additional details about your dataset

Parent project (optional) 如果该项目非大项目下的子项目，此项直接点击Next即可

If your project is part of a larger project, please select the parent project from the table below. If you would like to propose a new parent project, please contact us at: [pride-support@ebi.ac.uk](mailto:pride-support@ebi.ac.uk)

<input type="checkbox"/> Parent Project
<input type="checkbox"/> Antibodies (B/D-BFP)
<input type="checkbox"/> Bioinformatics Infrastructure for Life Sciences (BILS) network (Sweden)
<input type="checkbox"/> Biology/Disease-Driven Human Proteome Project (B/D-BFP)
<input type="checkbox"/> CPTAC Consortium
<input type="checkbox"/> Cancer (B/D-BFP)
<input type="checkbox"/> Cancer Research (B/D-BFP)

Published ID(s) (optional) 所属大项目已发表文章的PMID

Provide the PubMedID(s) if the dataset is associated with an existing publication (comma separated)

Reanalysis ProteomeXchange accession(s) (optional)

Only applicable if your results are based on the reprocessing of one or several previously submitted PX dataset(s)

Links to other 'Omics' datasets (optional)

Only applicable if proteomics results can be linked to other biological data submitted to other resources (e.g. ArrayExpress, GEO)

12. 确认相关项目信息以及上传文件数据无误；此处一定要记得勾选 PRIDE Archive dataset license 才能顺利的点击 submit 进入数据上传环节。

**PRIDE Submission Tool Version - 2.5.2**

**Step 10: Submission Summary (10/11)** ProteomeXchange

Please double-check before starting your submission

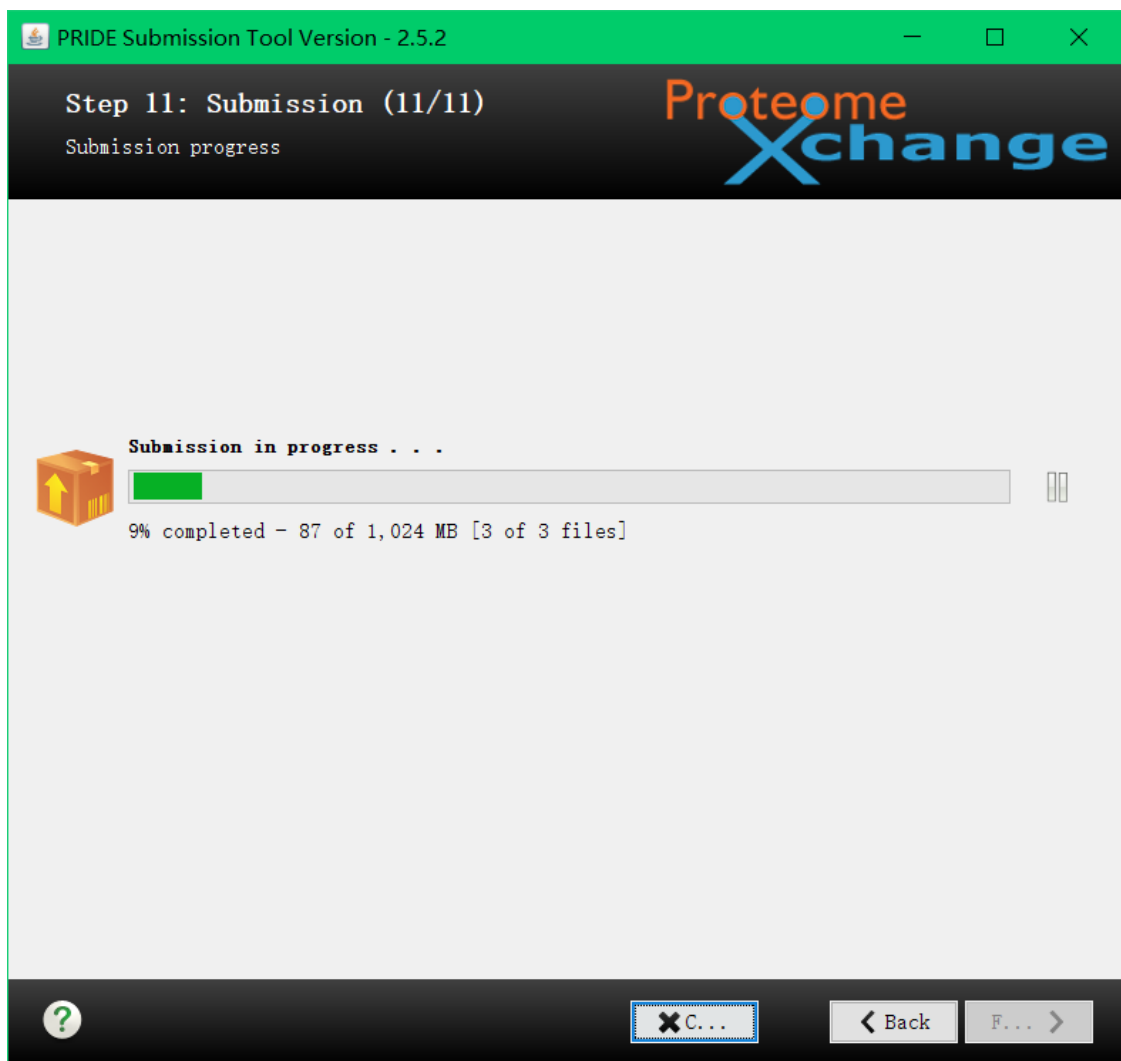
● Total file count: 3   ● Result files: 0   ● Raw files: 1   [Export submission.px](#)

● Peak files: 0   ● Search files: 1   ● Other files: 1   [Result File Validation](#)

File Name	PATH / URL	Type	Size (Mb)	#Mapped files
checksum.txt	C:\Users\dec\Des...	OTHER	0	0
Match_result...	C:\Users\dec\Des...	SEARCH	1024.613	1

☒ Please accept [PRIDE Archive dataset license](#)

[?](#) [C...](#) [Back](#) [S... >](#)




13. 显示 Submission in progress 后就可以耐心等待数据上传啦。一般我们建议数据上传过程中避免网络波动，同时如果这一步出现报错，可以尝试关闭电脑防火墙以及电脑管家等程序，再进行上传。上传数据成功后，系统会提供一个序列号，但是这个序列号是无法在文章中使用的。一般来讲，大概几个工作日后，数据审核完成会给注册的邮箱重新发送一个序列号，在文章中提供该邮件中提供的序列号就可以啦。

PRIDE Submission Tool Version - 2.5.2

Step 11: Submission (11/11)






Submission progress

ProteomeXchange




All your files have been uploaded successfully

100% completed - 1,024 of 1,024 MB [3 of 3 files]



Click here to provide more detailed feedback




Your submission reference:

1-20211020-32077

Please note: your submission reference is not an accession number.

We aim to process all submissions within five working days, however at certain times, it might take longer.



+ New

> F...