

User Guide - Apus

Apu Pickup

INTRODUCTION

Version: ApuPickup_v202507

Release Date: 2025.07.15

OS: Windows 10 (x64) or Windows 11 (x64)

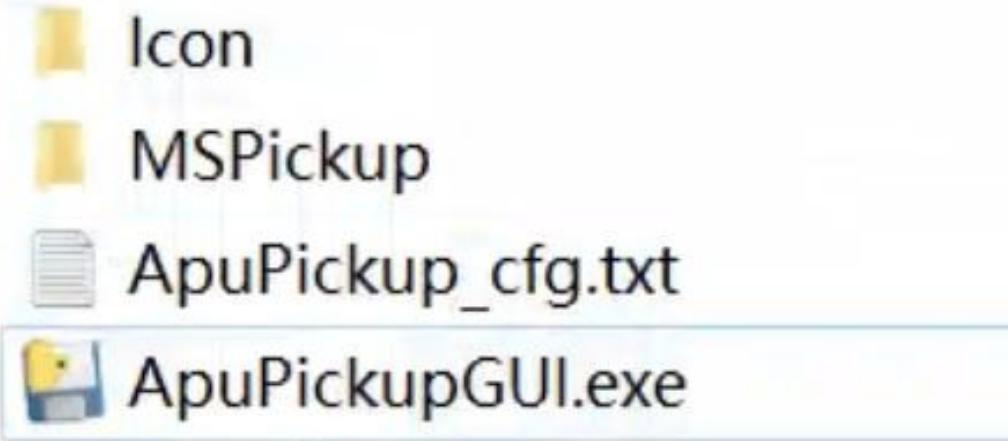
ApuPickup is a local client system within the Apus solution, responsible for the **automated and centralized management of raw data**. Deployed on the control computer of the LC-MS/MS instrument, its primary functions are:

1. **Real-time monitoring**: To continuously scan for and identify newly generated raw data files (e.g., .raw files).
2. **Automated data aggregation**: To securely transfer these new files to a central Network-Attached Storage (NAS) server via the local area network (LAN).

STEP-1

1. Download the ApuPickup installation package from github (<https://github.com/BUAA-LiuLab/Apus>) to the **control computer of the LC-MS/MS instrument**. (If you are only testing the software, it can be installed on any Windows computer that supports FTP).
2. Navigate to the “/ApuPickup_v202507/” directory.
3. Execute **ApuPickupGUI.exe** to launch the application. This will open the main software interface.

②



③



STEP-2

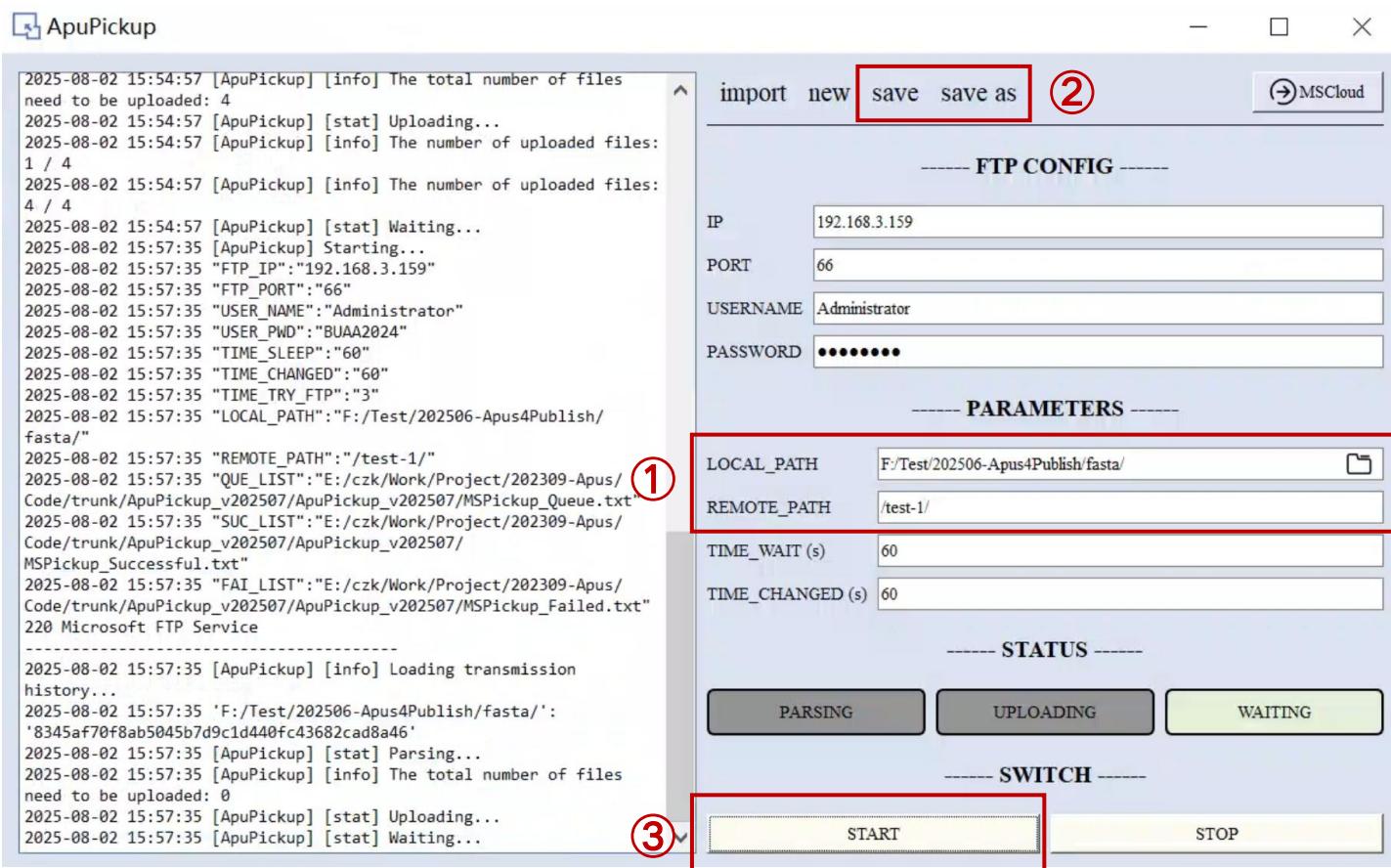
1. Use **IIS (Internet Information Services)** to establish an **FTP connection** between the local machine and the central NAS server.
2. Input these **FTP configuration settings** into the software's interface. (The following figure is an example).

----- FTP CONFIG -----

IP	192.168.3.159
PORT	66
USERNAME	Administrator
PASSWORD	••••••••

STEP-3

1. In the **LOCAL_PATH** field, specify the directory path where the mass spectrometry data files are generated.
2. In the **REMOTE_PATH** field, specify the target directory path on the central NAS server designated for data storage.
3. Click **Save or Save As** to store the configuration settings locally as a file.
4. Click **START** to initiate the process. The software will then cycle through three operational states.



ATTENTIONS

1. ApuPickup operates based on the **File Transfer Protocol (FTP)**. Therefore, it is essential to ensure that the local machine and the central NAS server are on the same **local area network (LAN)** and that a **stable FTP connection** has been established between them.
2. ApuPickup features a **breakpoint resume mechanism** to handle network fluctuations, ensuring the integrity and stability of large-scale data transfers. A history of all transfer activities is recorded and saved as .pkl files within the **“/ApuPickup_v202507/MSPickup/Uploaded Files/”** directory. Files that have already been successfully transferred will not be re-uploaded. To re-transmit any files, the **Uploaded Files** folder must be deleted.

ApuPioneer

INTRODUCTION

ApuPioneer is an automated and parallel system for identification, built for large-cohort studies. By default, its DDA mode uses the pFind open-search engine.

1. **Automation & parallelization:** The system automates the protein identification process. It runs tasks in parallel on multiple compute nodes.
2. **Real-time analysis:** ApuPioneer monitors a folder and automatically starts analyzing new raw files as they are transferred. You don't need to wait for all samples to be acquired.
3. **Simple to use:** The system uses optimized default parameters. Users only need to provide basic info, like the data path and database file, to start an analysis.

ApuPioneer's output follows the standard format of the search engine, including peptide-spectrum matches and protein inference results. These results can be used directly by the ApuHorizon system for quantification.

INTRODUCTION

Version: ApuPioneer-DDA_v202507

Release Date: 2025.07.15

Computer configuration

CPU: Intel or AMD processor with 64-bit support, 2.3 GHz or faster processor with at least 2 cores is recommended

RAM: 16G or higher is recommended

ROM: for one raw data (1-2 G) or higher is recommended

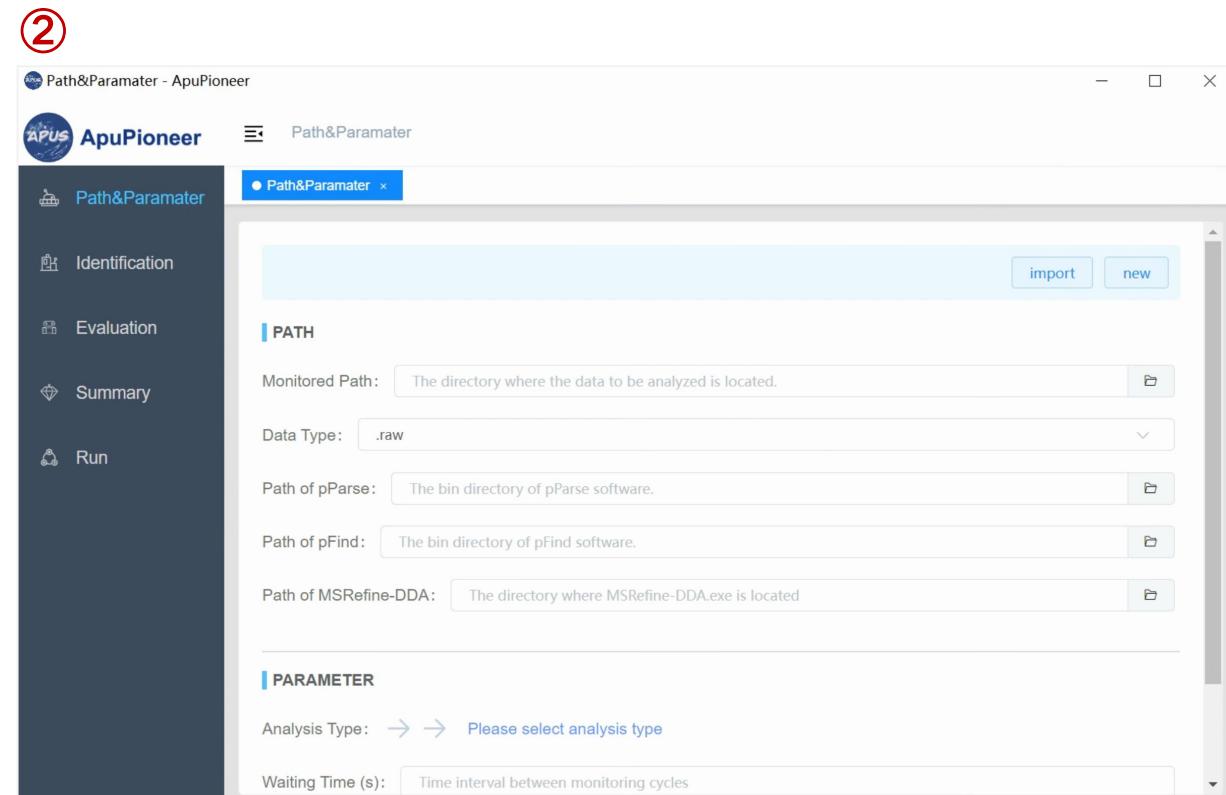
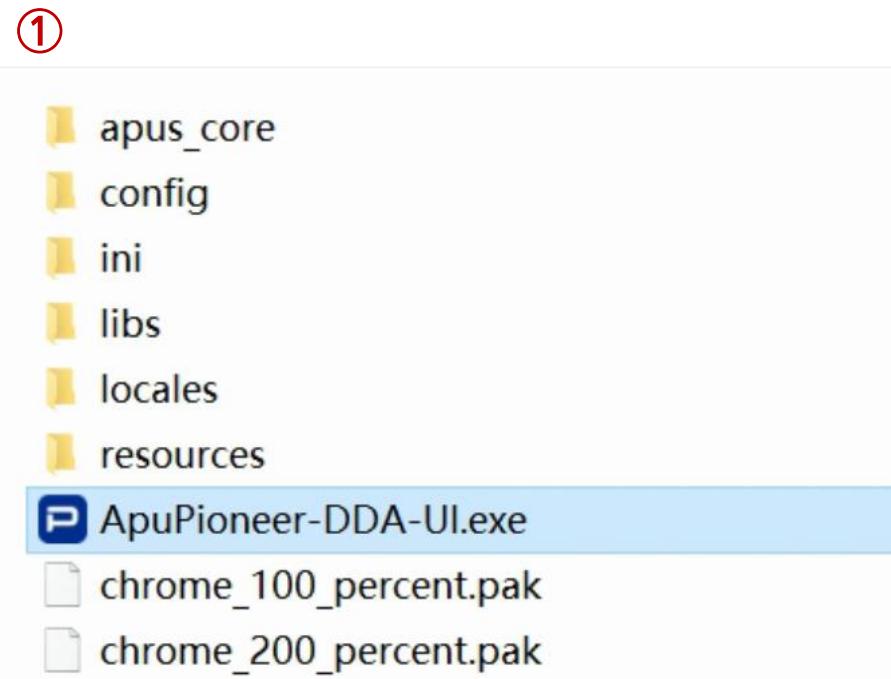
OS: Windows 10 (x64) or Windows 11 (x64)

STEP-1 Preparation

1. **Install software:** Deploy ApuPioneer-DDA, pFind, and MSRefine-DDA on **every compute node** of your cluster.
 - Download ApuPioneer-DDA and MSRefine-DDA from [*https://github.com/BUAA-LiuLab/Apus*](https://github.com/BUAA-LiuLab/Apus).
 - Download pFind from [*https://pfind.ict.ac.cn/se/pfind/#Downloads*](https://pfind.ict.ac.cn/se/pfind/#Downloads).
2. **Get licenses:** Obtain the **licenses** for pFind and MSRefine-DDA by following the instructions in their manuals:
 - /pFind3/pFind 3 UserGuide.pdf
 - /MSRefine-DDA_v202407/Manual.for.MSRefine.pdf
3. **Mount storage:** Use the **SMB** protocol to mount your data storage system to the local filesystem of the compute nodes, so it can be accessed like a local disk.

STEP-2

1. Navigate to the “/ApuPioneer-DDA_v202507/” directory.
2. Execute **ApuPioneer-DDA-UI.exe** to launch the application. This will open the main software interface.



STEP-3 Path&Paramater

1. Fill in **data information**:

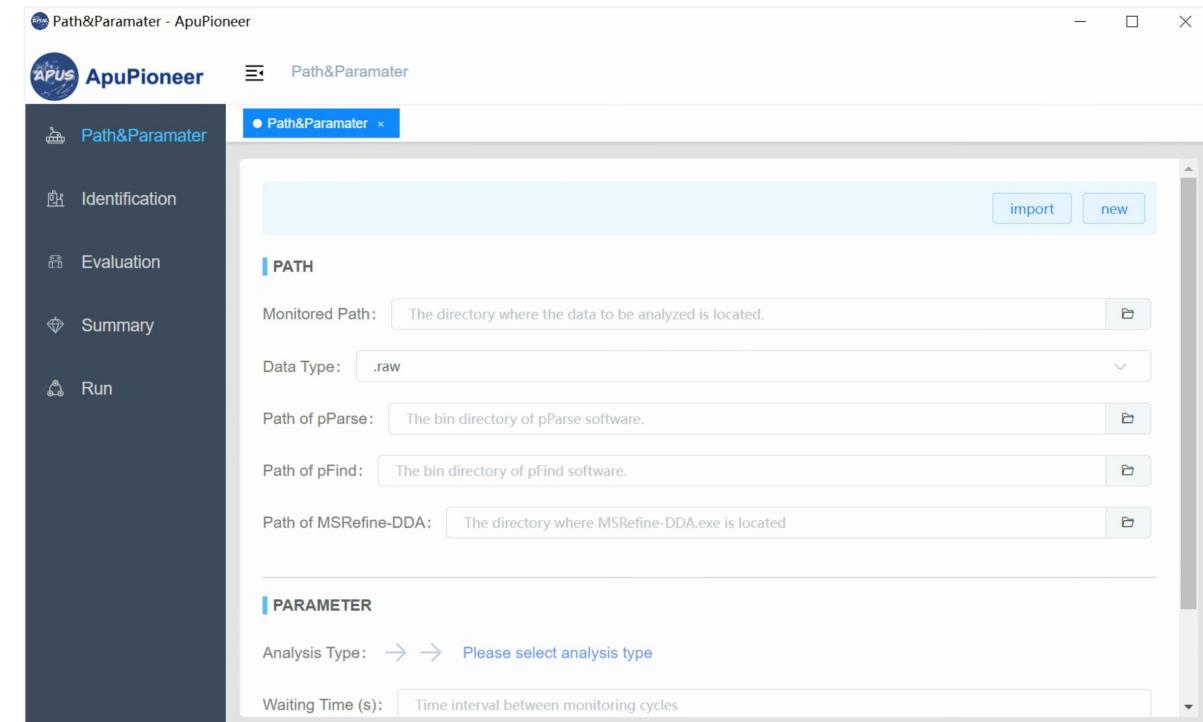
1. Monitored path: The folder path where your mass spectrometry data is stored.
2. Data type: The format of your data files (e.g., .raw, .d, .wiff).

2. Fill in **software paths**:

Enter the installation directory paths for pParse, pFind, and MSRefine-DDA. This should be the folder containing the main executable file (e.g., MSRefine-DDA.exe).

3. Fill in **analysis parameters**:

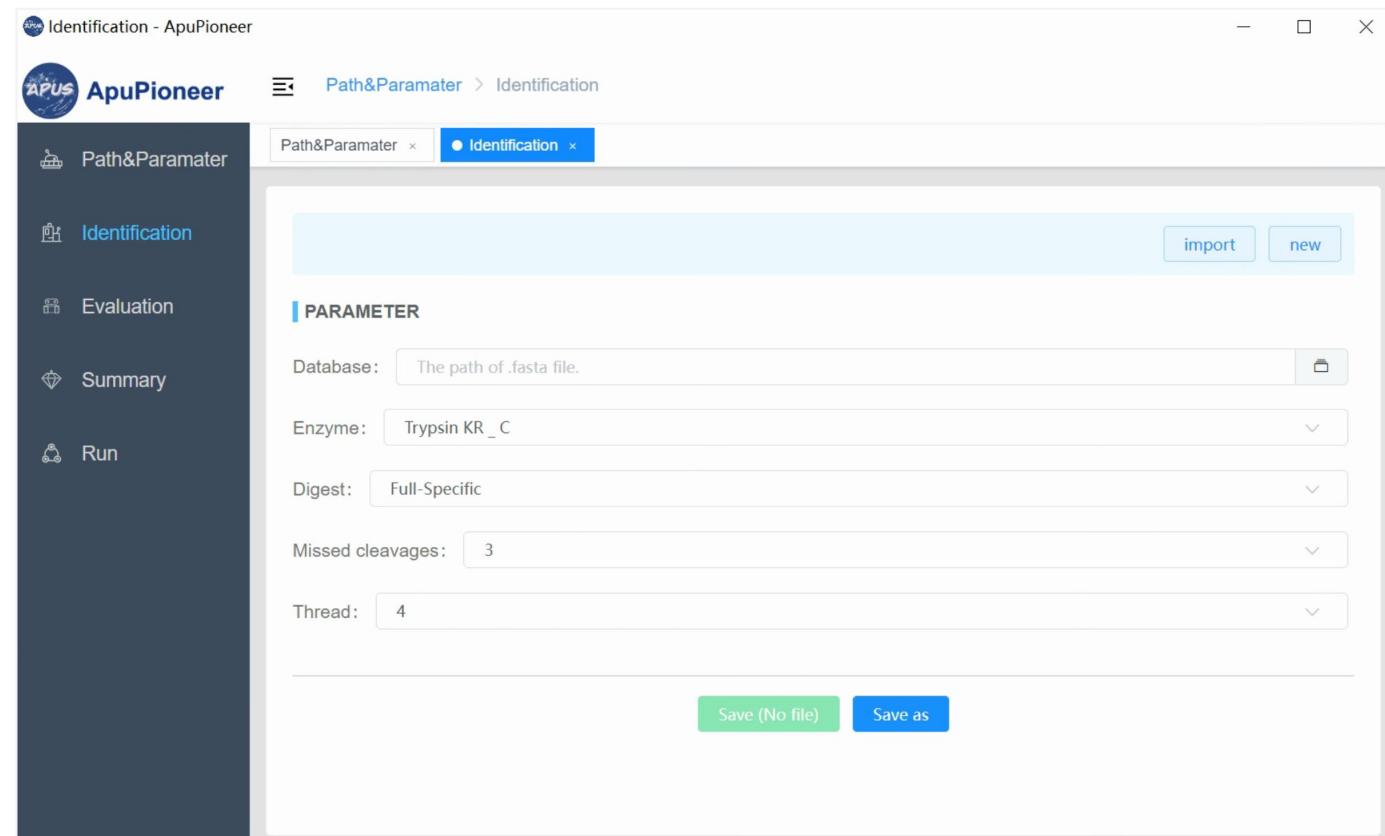
1. Analysis type: "database search only" or "database search + quality control".
2. Waiting time: The pause time between cycles of checking for new files.



STEP-4 Identification

Fill in the following pFind search parameters for DDA analysis. You can customize: Database, Enzyme, Digest, Missed cleavages, Thread

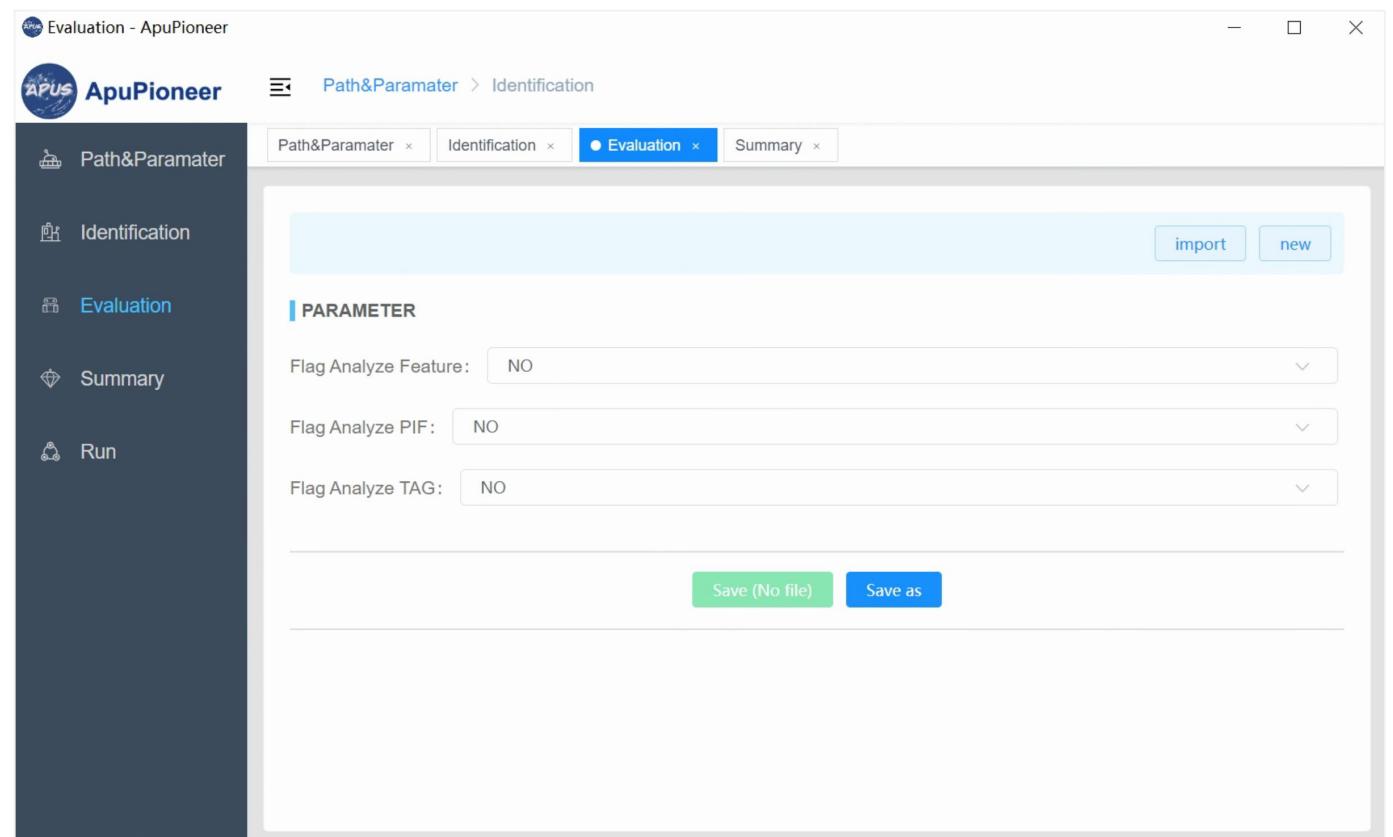
More advanced pFind settings are already set to default. You can view or edit them in the file: "/ApuPioneer-DDA_v202507/config/pFind_config.txt"



STEP-5 Evaluation

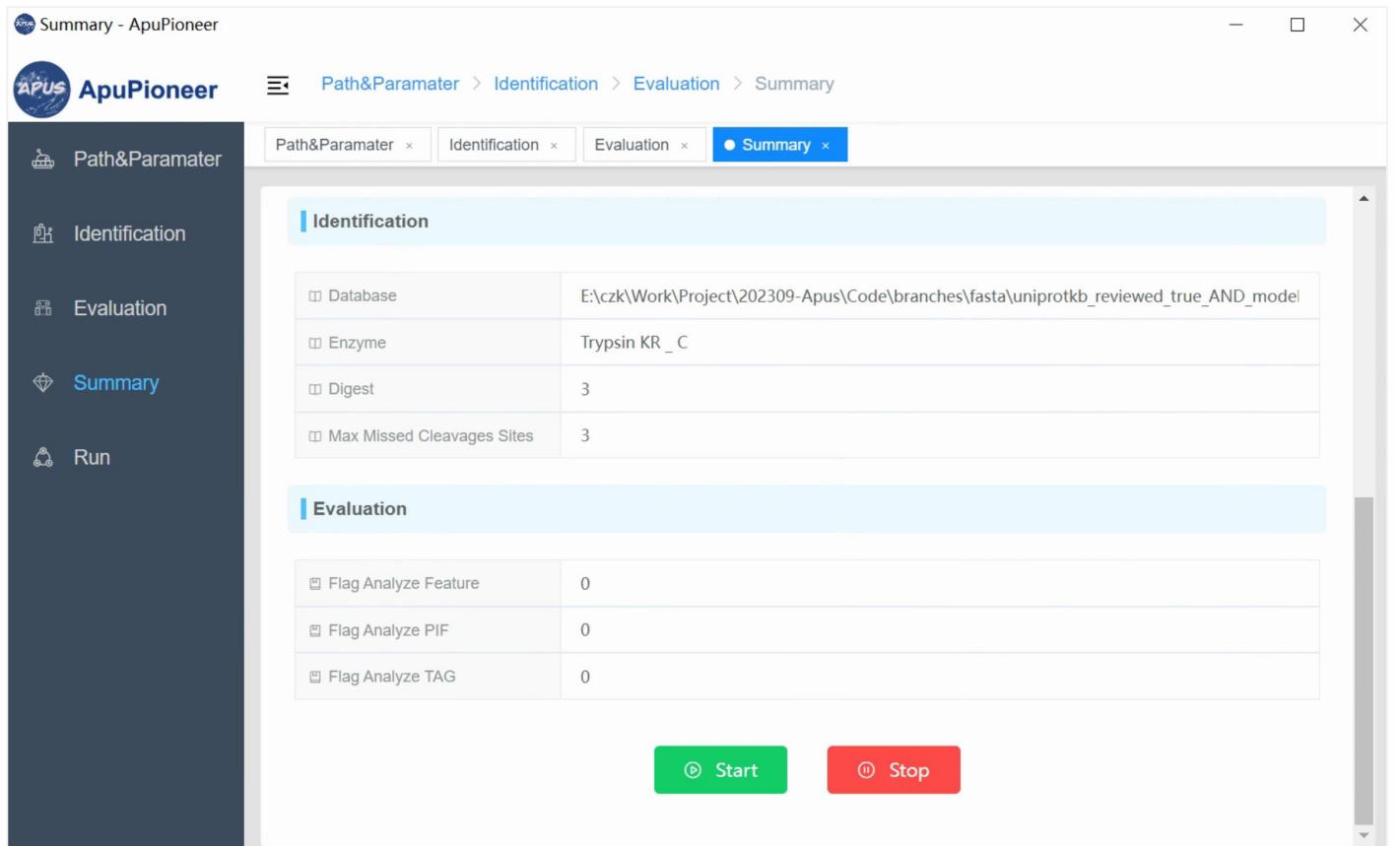
Fill in the following **MSRefine-DDA parameters** for quality control in DDA analysis. You can configure: Flag Analyze Feature, Flag Analyze PIF, Flag Analyze TAG

More advanced **MSRefine-DDA settings are already set to default**. You can view or edit them in the file: “/ApuPioneer-DDA_v202507/config/MSRefine-DDA_config.txt”.



STEP-6 Summary

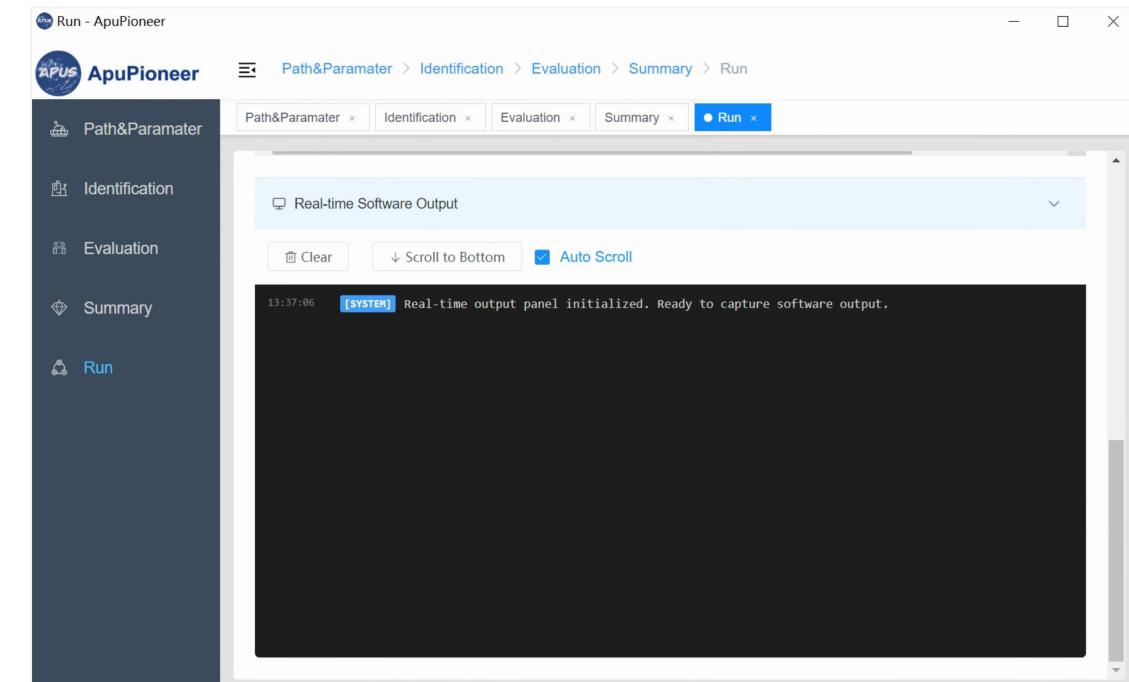
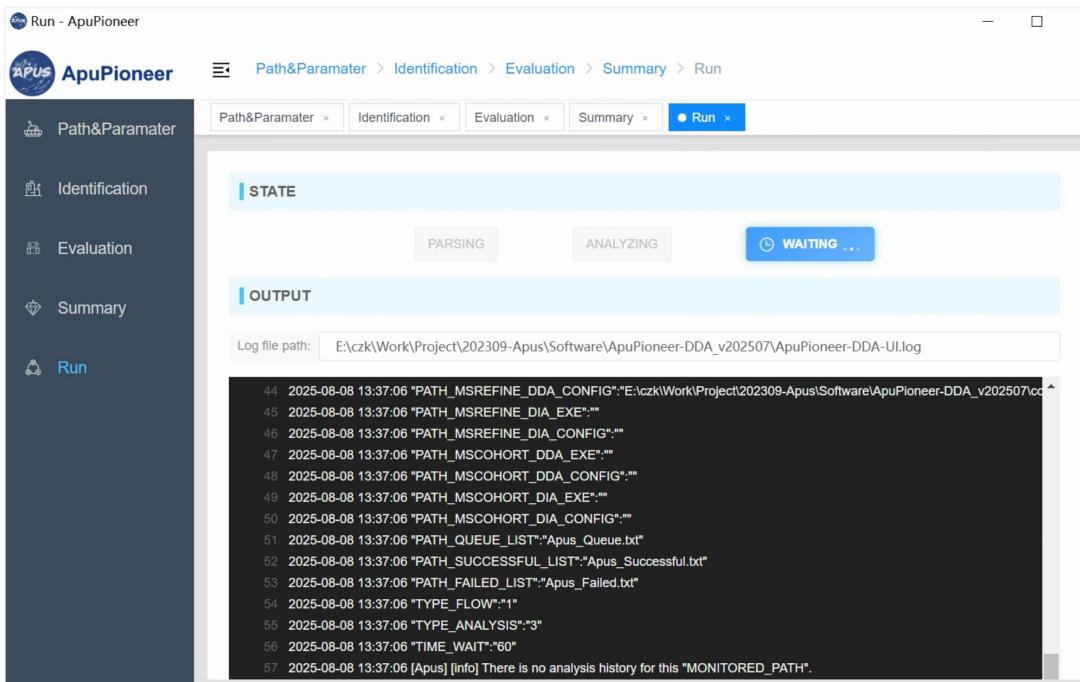
- 1. Review Configuration:** Navigate to the Summary page to review all previously configured parameters.
- 2. Start Process:** After verifying that all settings are correct, click the **Start** button to launch the application.



STEP-7 Run

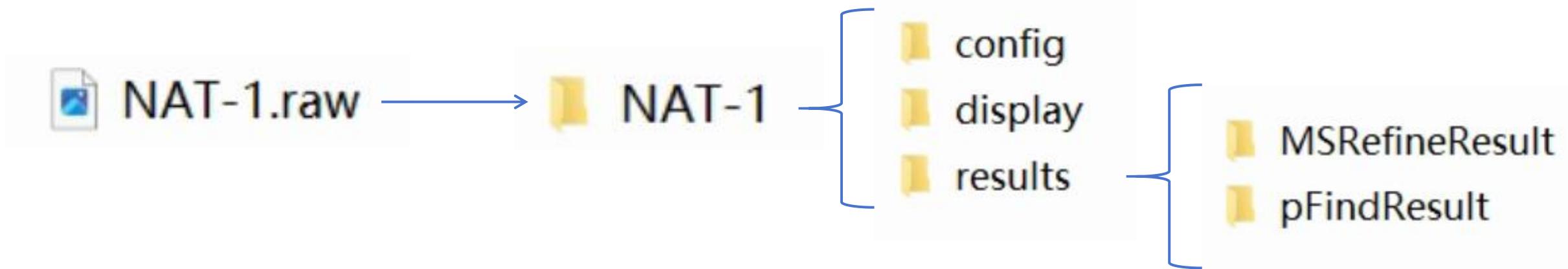
Go to the Run page to watch the program's progress.

- **STATE:** Shows the current status of the program's loop.
- **OUTPUT:** Here you can see live logs:
 1. Logs from ApuPioneer-DDA.
 2. More detailed logs from the tools it calls (like pFind).



STEP-8 View analysis results

Enter "Monitored Path" and view the data analysis results in the **folder with the same name as the mass spectrometry data file**.



ATTENTIONS

1. Always click **Save** or **Save As** before leaving a page. If you don't, your changes will be lost.
2. ApuPioneer keeps a record of processed files to run efficiently and prevent re-running completed tasks if the analysis is interrupted. It saves two types of history:
 1. **Local History**: Stored in the “`/ApuPioneer-DDA_v202507/AP_LocalHistory/`” folder (as .pkl files).
 2. **Global History**: Stored in the “`/Monitored Path/AP_GlobalHistory.txt`” file.
3. **To re-analyze all data in a folder from the beginning, you must delete the old history files first:**
 1. Delete the AP_LocalHistory folder.
 2. Delete the AP_GlobalHistory.txt file.

INTRODUCTION

ApuPioneer is an automated and parallel system for identification, built for large-cohort studies. Its DIA mode defaults to **DIA-NN**.

1. **Automation & parallelization**: The system automates the protein identification process. It runs tasks in parallel on multiple compute nodes.
2. **Real-time analysis**: ApuPioneer monitors a folder and automatically starts analyzing new raw files as they are transferred. You don't need to wait for all samples to be acquired.
3. **Simple to use**: The system uses optimized default parameters. Users only need to provide basic info, like the data path and database file, to start an analysis.

ApuPioneer's output follows the standard format of the search engine, including peptide-spectrum matches and protein inference results. These results can be used directly by the ApuHorizon system for quantification.

INTRODUCTION

Version: ApuPioneer-DIA_v202507

Release Date: 2025.07.15

Computer configuration

CPU: Intel or AMD processor with 64-bit support, 2.3 GHz or faster processor with at least 2 cores is recommended

RAM: 16G or higher is recommended

ROM: for one raw data (1-2 G) or higher is recommended

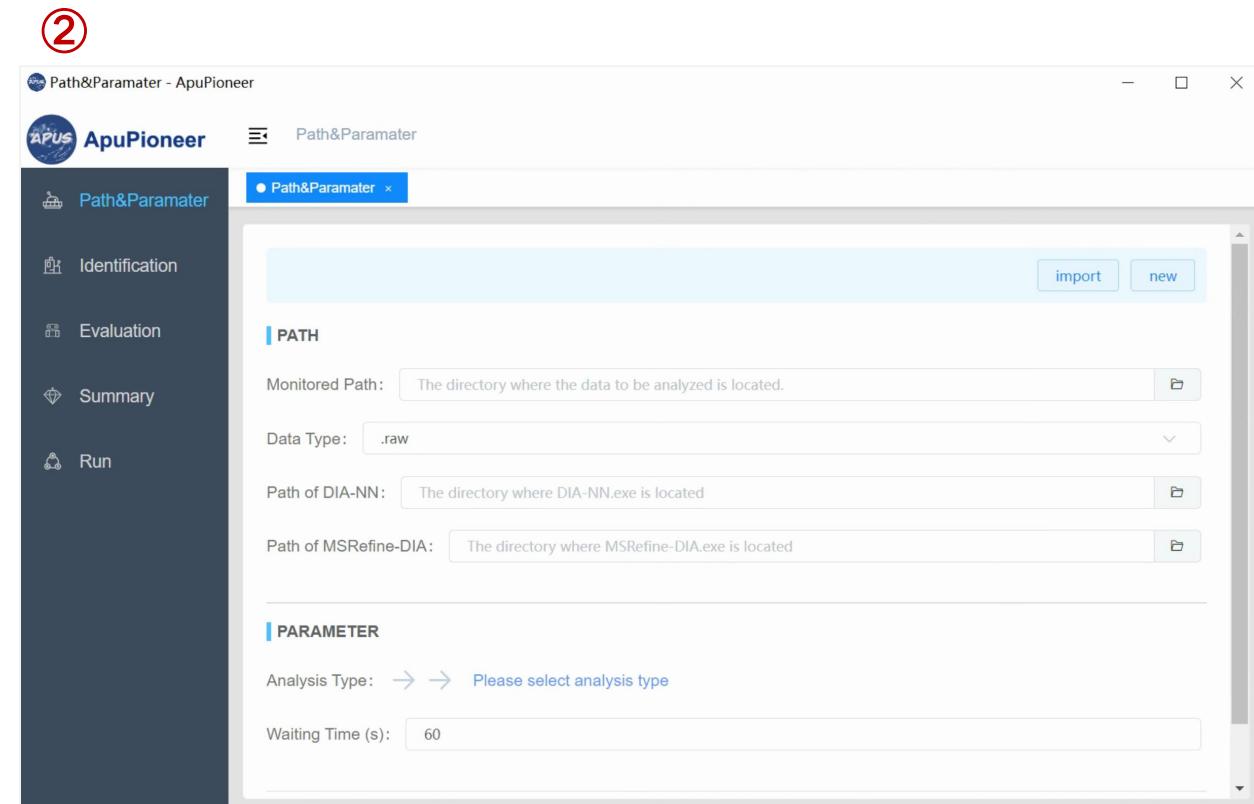
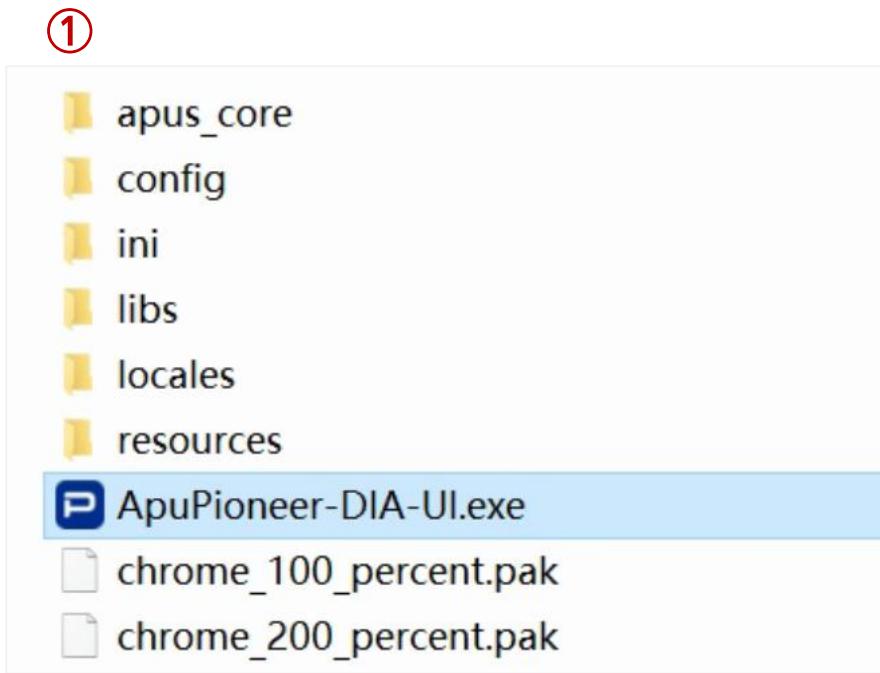
OS: Windows 10 (x64) or Windows 11 (x64)

STEP-1 Preparation

1. **Install software:** Deploy ApuPioneer-DIA, DIA-NN, and MSRefine-DIA on all **compute nodes** of your cluster.
 - Download ApuPioneer-DIA and MSRefine-DDA from <https://github.com/BUAA-LiuLab/Apus>.
 - Download DIA-NN from <https://github.com/vdemichev/DiaNN>.
2. **Mount storage:** Use the **SMB** protocol to mount your data storage system to the local filesystem of the compute nodes, so it can be accessed like a local disk.

STEP-2

1. Navigate to the “/ApuPioneer-DIA_v202507/” directory.
2. Execute **ApuPioneer-DIA-UI.exe** to launch the application. This will open the main software interface.



STEP-3 Path&Paramater

1. Fill in **data information**:

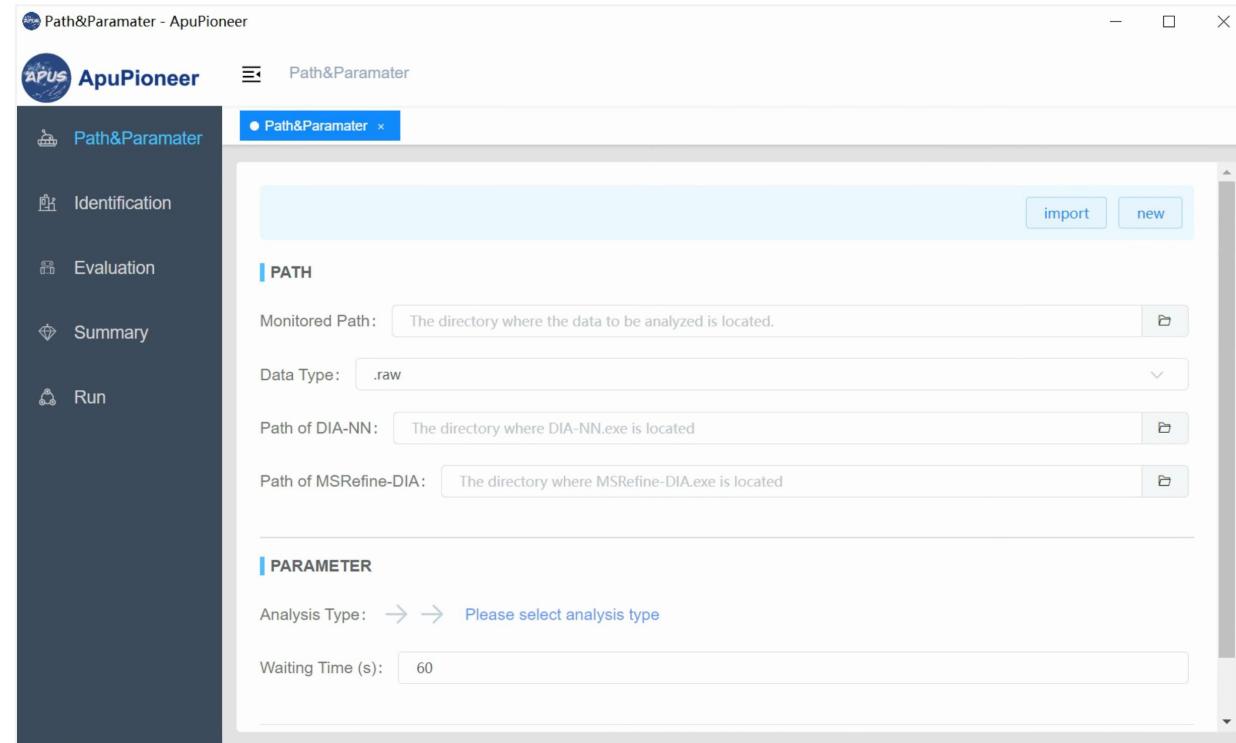
1. Monitored path: The folder path where your mass spectrometry data is stored.
2. Data type: The format of your data files (e.g., .raw, .d, .wiff).

2. Fill in **software paths**:

Enter the installation directory paths for DIA-NN and MSRefine-DIA. This should be the folder containing the main executable file (e.g., MSRefine-DIA.exe).

3. Fill in **analysis parameters**:

1. Analysis type: "database search only" or "database search + quality control".
2. Waiting time: The pause time between cycles of checking for new files.



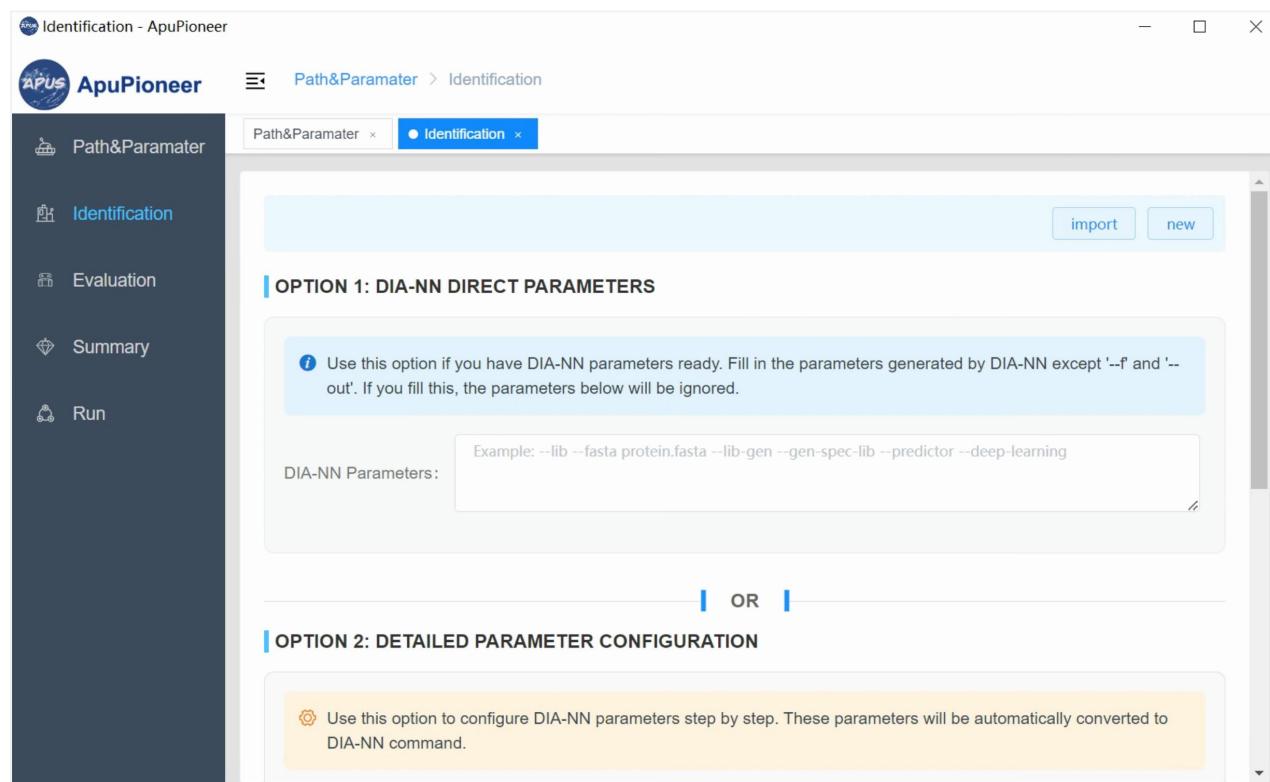
STEP-4 Identification

Configure DIA-NN parameters for ApuPioneer-DIA.

1. Many search parameters are preset with optimized default values. Users only need to specify essential parameters, such as the FASTA Database and Library-free Search options, to run the analysis.
2. Notably, for DIA mode, besides setting common DIA-NN parameters via the GUI, ApuPioneer allows users to input a full command-line argument string.

This string should be generated from the native DIA-NN interface, excluding the `--f` and `--out` arguments.

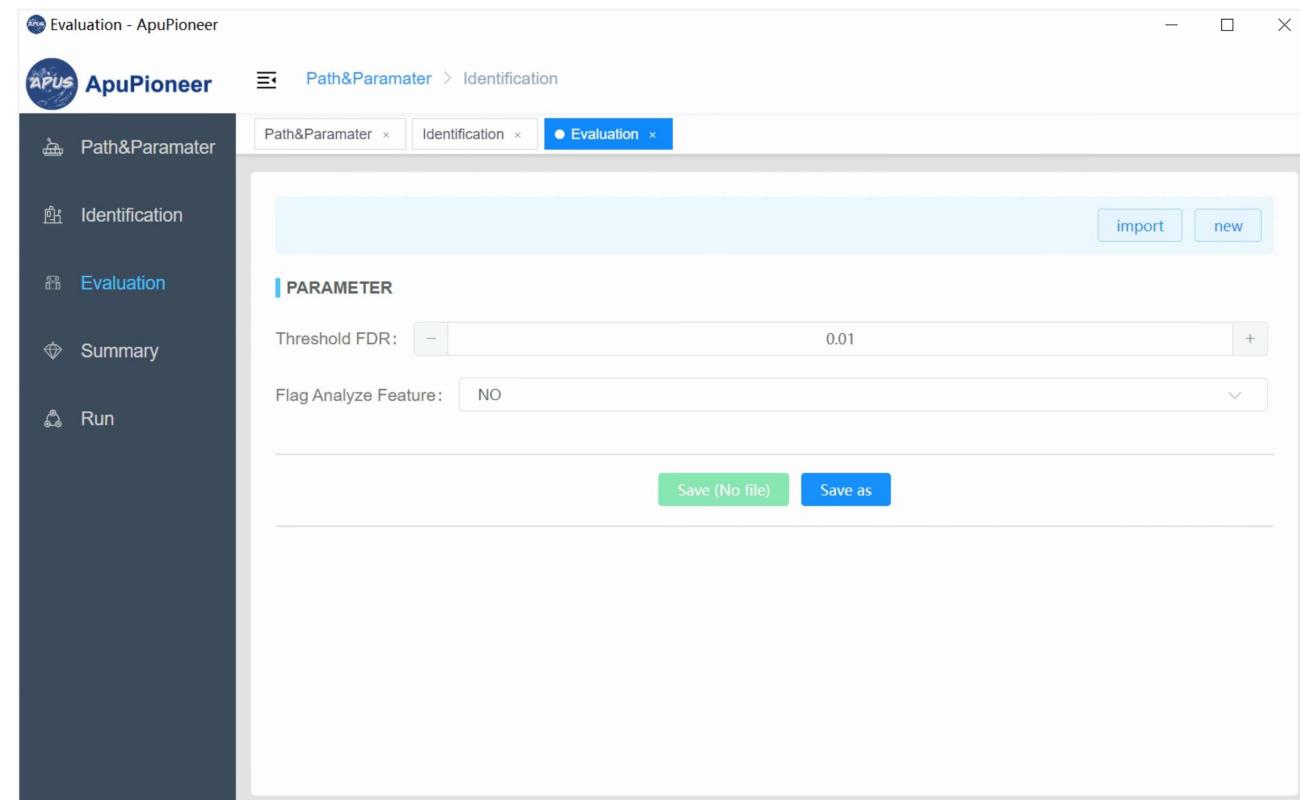
This feature provides fine-grained control over the identification process.



STEP-5 Evaluation

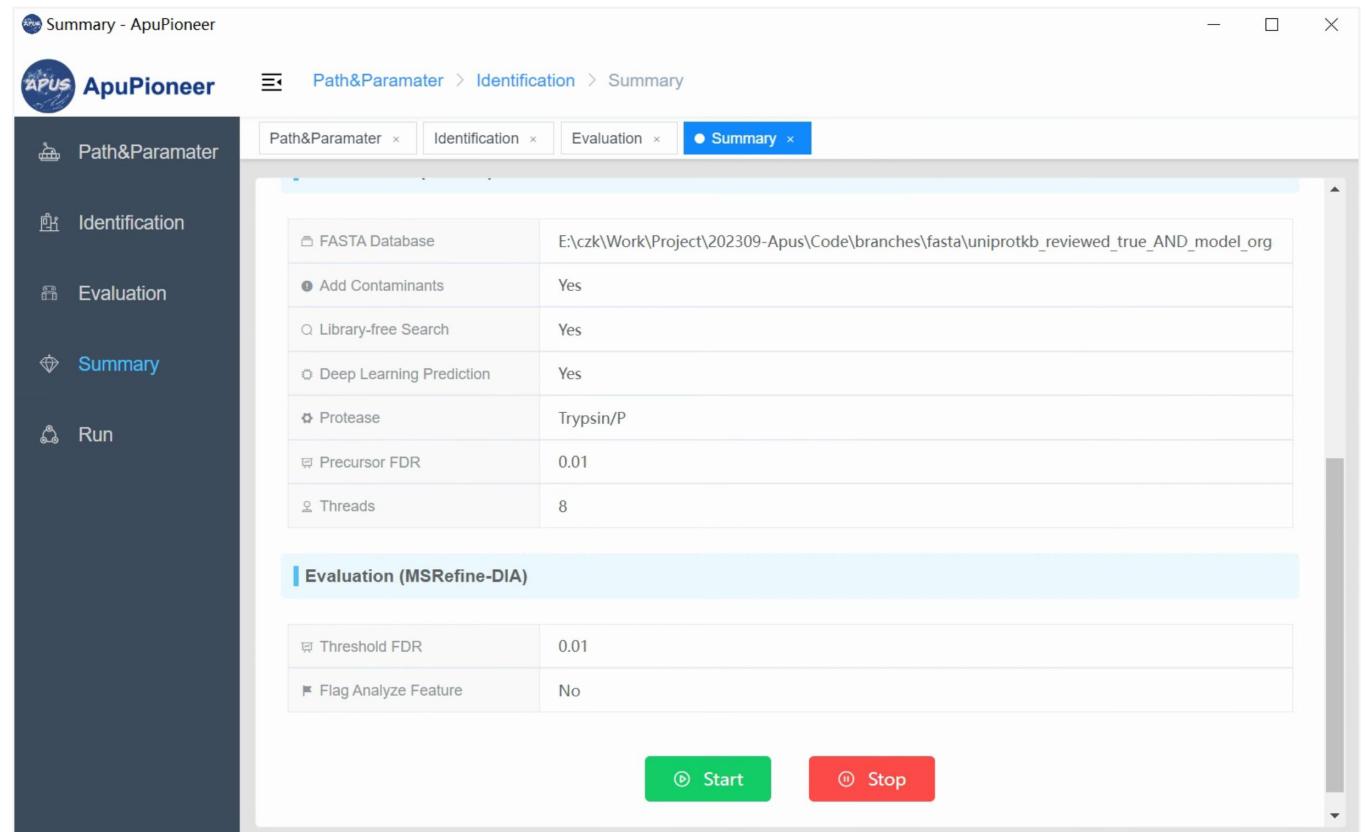
Fill in the following **MSRefine-DIA parameters** for quality control in DIA analysis. You can configure:
Threshold FDR, Flag Analyze Feature.

More advanced MSRefine-DIA settings are already set to default. You can view or edit them in the file: “/ApuPioneer-DIA_v202507/config/MSRefine-DIA_config.txt”.



STEP-6 Summary

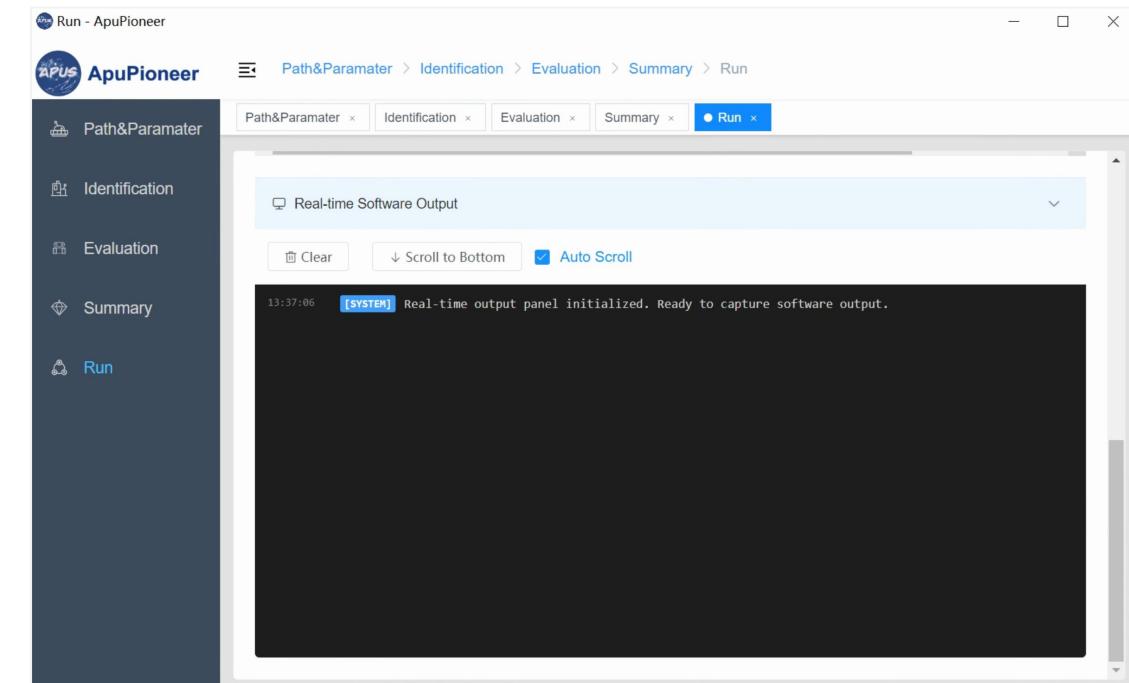
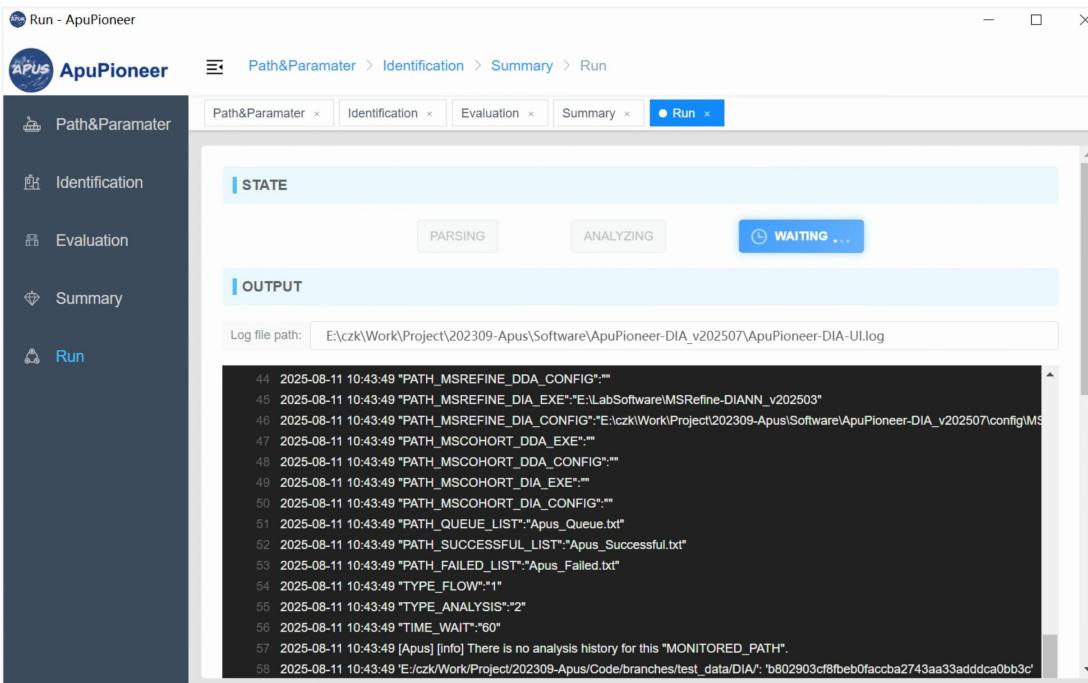
- Review configuration:** Navigate to the Summary page to review all previously configured parameters.
- Start process:** After verifying that all settings are correct, click the **Start** button to launch the application.



STEP-7 Run

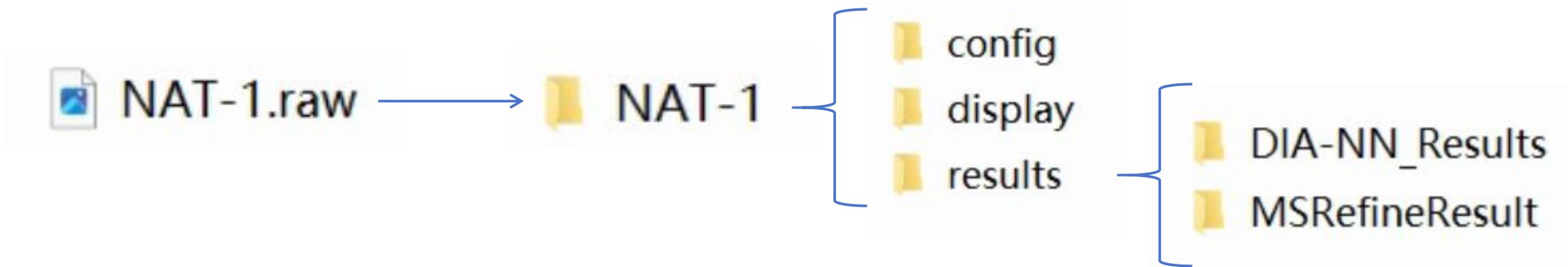
Go to the Run page to watch the program's progress.

- **STATE:** Shows the current status of the program's loop.
- **OUTPUT:** Here you can see live logs:
 1. Logs from ApuPioneer-DIA.
 2. More detailed logs from the tools it calls (like DIA-NN)



STEP-8 View analysis results

Enter "Monitored Path" and view the data analysis results in the **folder with the same name as the mass spectrometry data file**.



ATTENTIONS

1. Always click **Save** or **Save As** before leaving a page. If you don't, your changes will be lost.
2. ApuPioneer keeps a record of processed files to run efficiently and prevent re-running completed tasks if the analysis is interrupted. It saves two types of history:
 1. **Local History**: Stored in the “`/ApuPioneer-DIA_v202507/AP_LocalHistory/`” folder (as .pkl files).
 2. **Global History**: Stored in the “`/Monitored Path/AP_GlobalHistory.txt`” file.
3. **To re-analyze all data in a folder from the beginning, you must delete the old history files first:**
 1. Delete the AP_LocalHistory folder.
 2. Delete the AP_GlobalHistory.txt file.

ApuHorizon

INTRODUCTION

ApuHorizon: The protein quantification system of Apus. It is deployed on a dedicated server with substantial memory to handle the high-throughput quantification of a large cohort comprising hundreds to thousands of samples.

- **Input:** Receives identification results from ApuPioneer.
- **Process:** Performs LFQ using a proven engine (defaults to **pQuant** for DDA mode).
- **Output:** Generates a result file detailing analyte abundances per sample.
- **Integration:** Output is ready for immediate use in ApuInsight (for bioinformatics).

INTRODUCTION

Version: ApuHorizon-DDA_v202507

Release Date: 2025.07.15

Computer configuration

CPU: Intel or AMD processor, 32 cores or higher is recommended

RAM: 256GB or higher is recommended

ROM: 1TB or higher high-speed NVMe SSD is recommended

OS: Windows 10 (x64) or Windows 11 (x64)

STEP-1 Preparation

1. **Install software:** Deploy ApuHorizon-DDA, pQuant, and MSCohort-DDA on the quantification server with substantial memory. (Download them from <https://github.com/BUAA-LiuLab/Apus>.)
2. **Get licenses:** Refer to “/pQuant-pGlycoQuant_v202407/Manual for pQuant.pdf” to obtain the pQuant license.
3. **Mount storage:** Use the **SMB** protocol to mount your data storage system to the local filesystem of the compute nodes, so it can be accessed like a local disk.

STEP-2

1. Navigate to the “/ApuHorizon-DDA_v202507/” directory.
2. Execute **ApuHorizon-DDA-UI.exe** to launch the application. This will open the main software interface.

①

The screenshot shows a Windows File Explorer window. The path is 'C:\Program Files\ApuHorizon-DDA_v202507'. The contents of the folder are:

- apus_core
- config
- ini
- libs
- locales
- resources
- ApuHorizon-DDA-UI.exe** (highlighted in blue)
- chrome_100_percent.pak
- chrome_200_percent.pak

②

The screenshot shows the ApuHorizon software interface. The title bar says 'Path&Parameter - ApuHorizon'. The left sidebar has the following options:

- Path&Parameterer (selected)
- Quantification
- Evaluation
- Summary
- Run

The main panel is titled 'Path&Parameterer'. It contains the following fields:

- PATH**
 - Monitored Path: The directory where the data to be analyzed is located. (Input field with a browse button)
 - Data Type: .raw (Input field)
 - Path of pQuant: The directory where Quant.exe is located (Input field with a browse button)
 - Path of MSCohort-DDA: The directory where MSCohort-DDA.exe is located (Input field with a browse button)
- PARAMETER**
 - Analysis Type: → → No analysis selected (Input field)

At the bottom right are 'Save (No file)' and 'Save as' buttons.

STEP-3 Path&Paramater

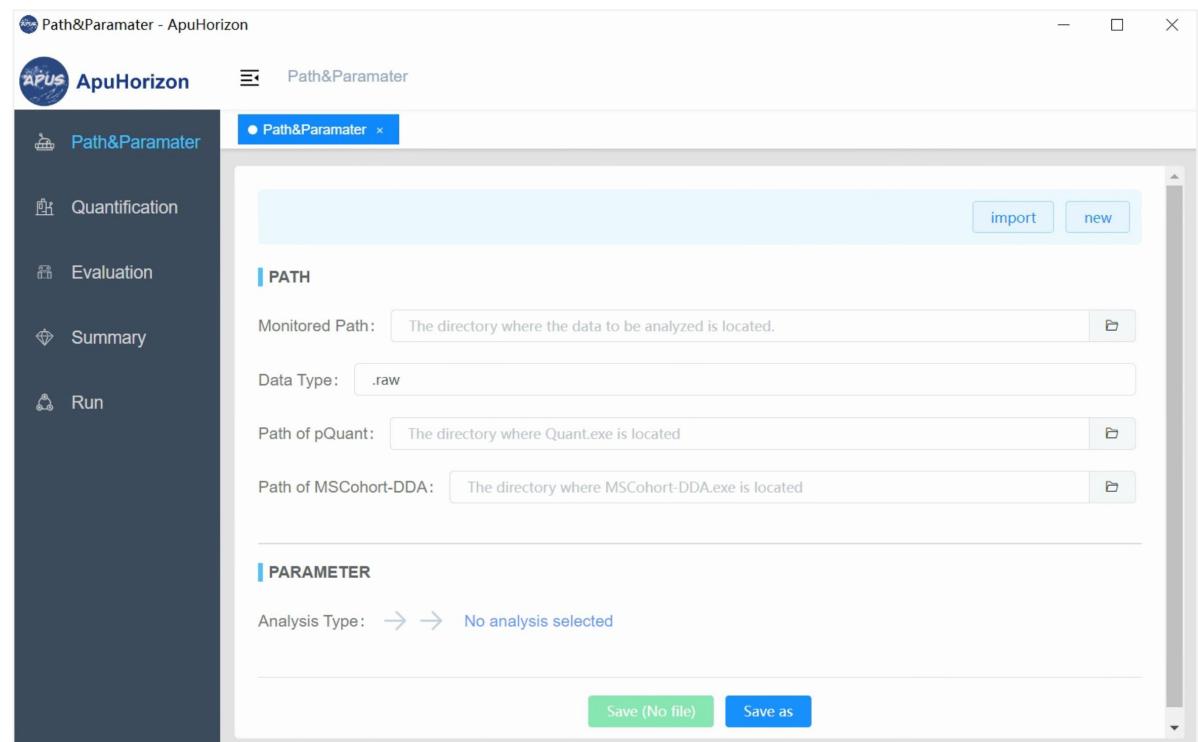
1. Fill in **data information**:

1. Monitored path: The folder path where your mass spectrometry data is stored.
2. Data type: The format of your data files (e.g., .raw, .d, .wiff).

2. Fill in **software paths**:

Enter the installation directory paths for pQuant and MSCohort-DDA. This should be the folder containing the main executable file (e.g., MSCohort-DDA.exe).

3. Fill in **analysis type**: "quantification only" or "quantification + quality control".

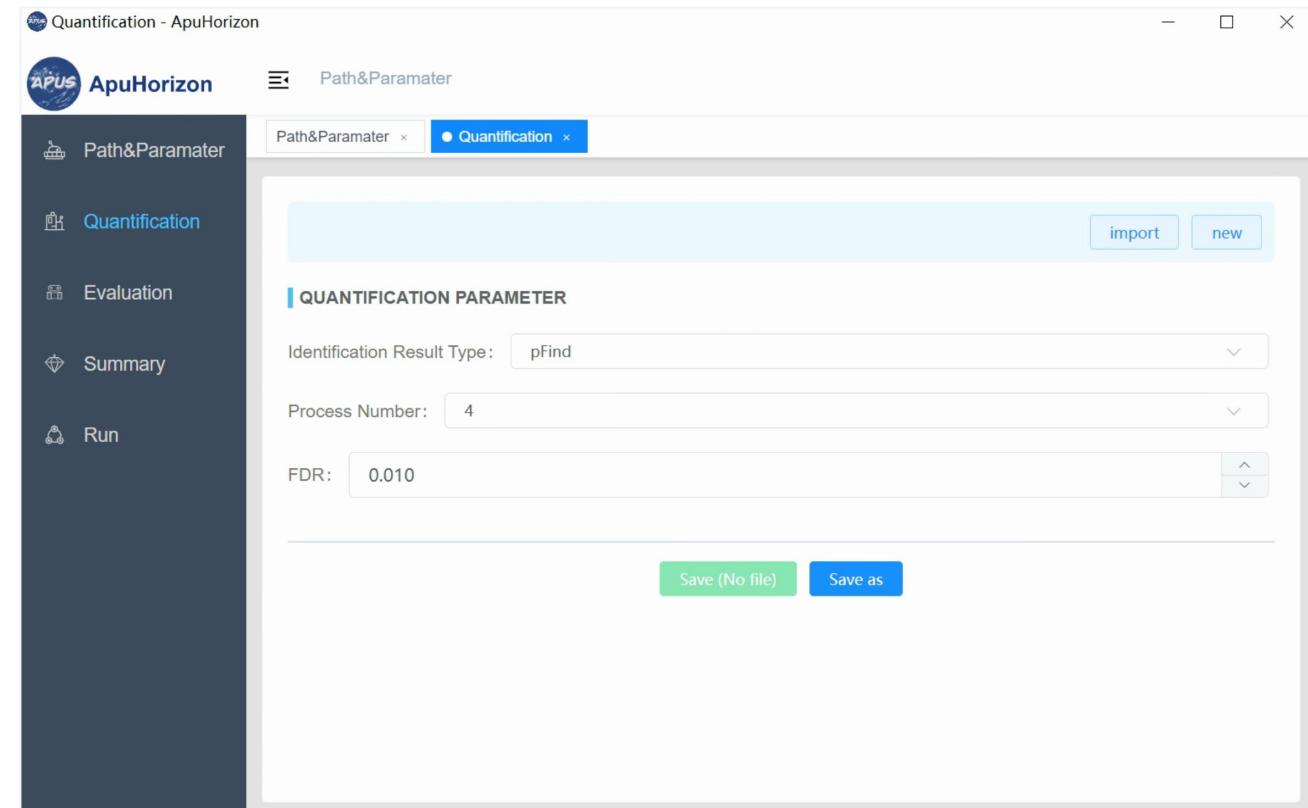


STEP-4 Quantification

Configure the **quantification parameters** for pQuant.

The system is pre-loaded with numerous default parameters and automatically locates the identification result files.

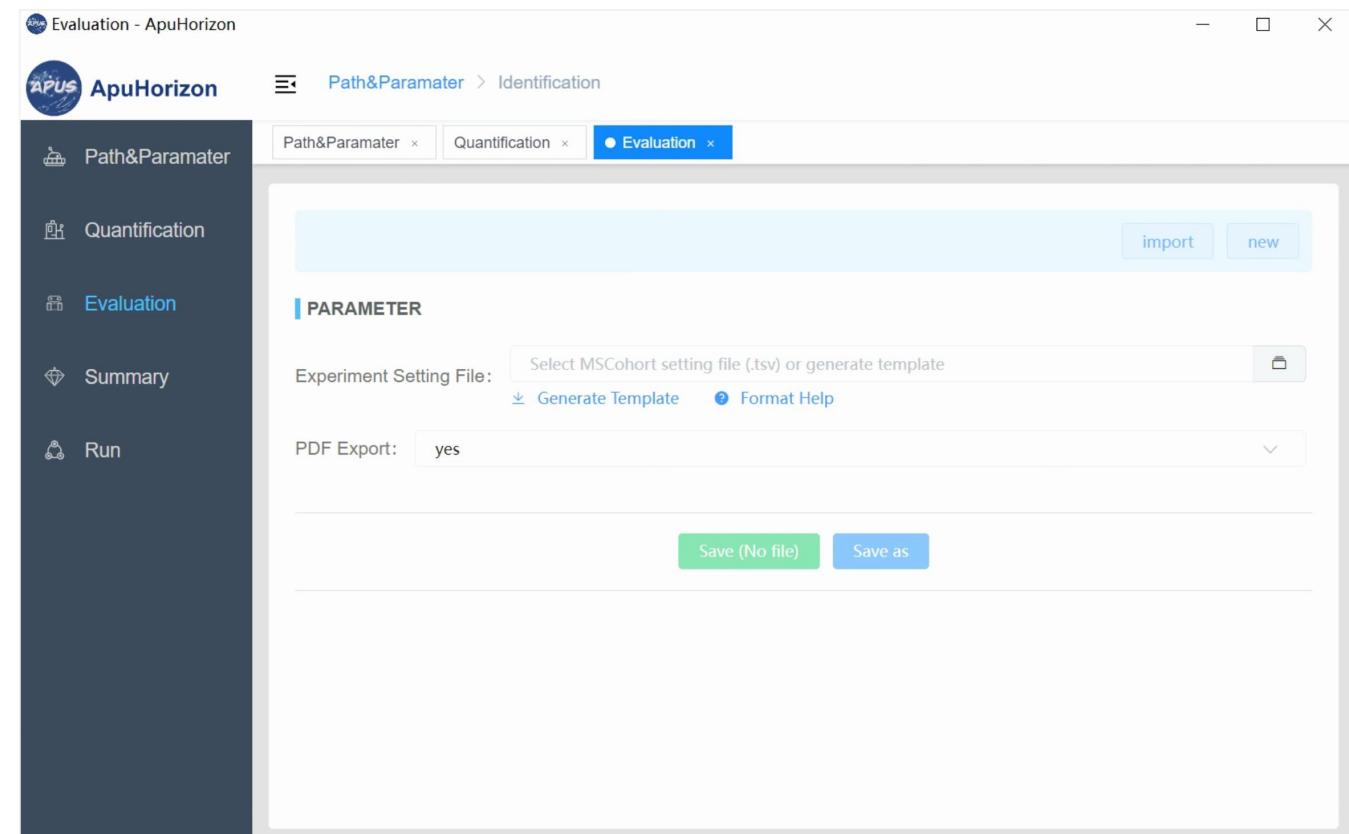
For more specific pQuant parameters, please refer to the configuration file at "/ApuHorizon-DDA_v202507/config/pQuant_config.txt".



STEP-5 Evaluation

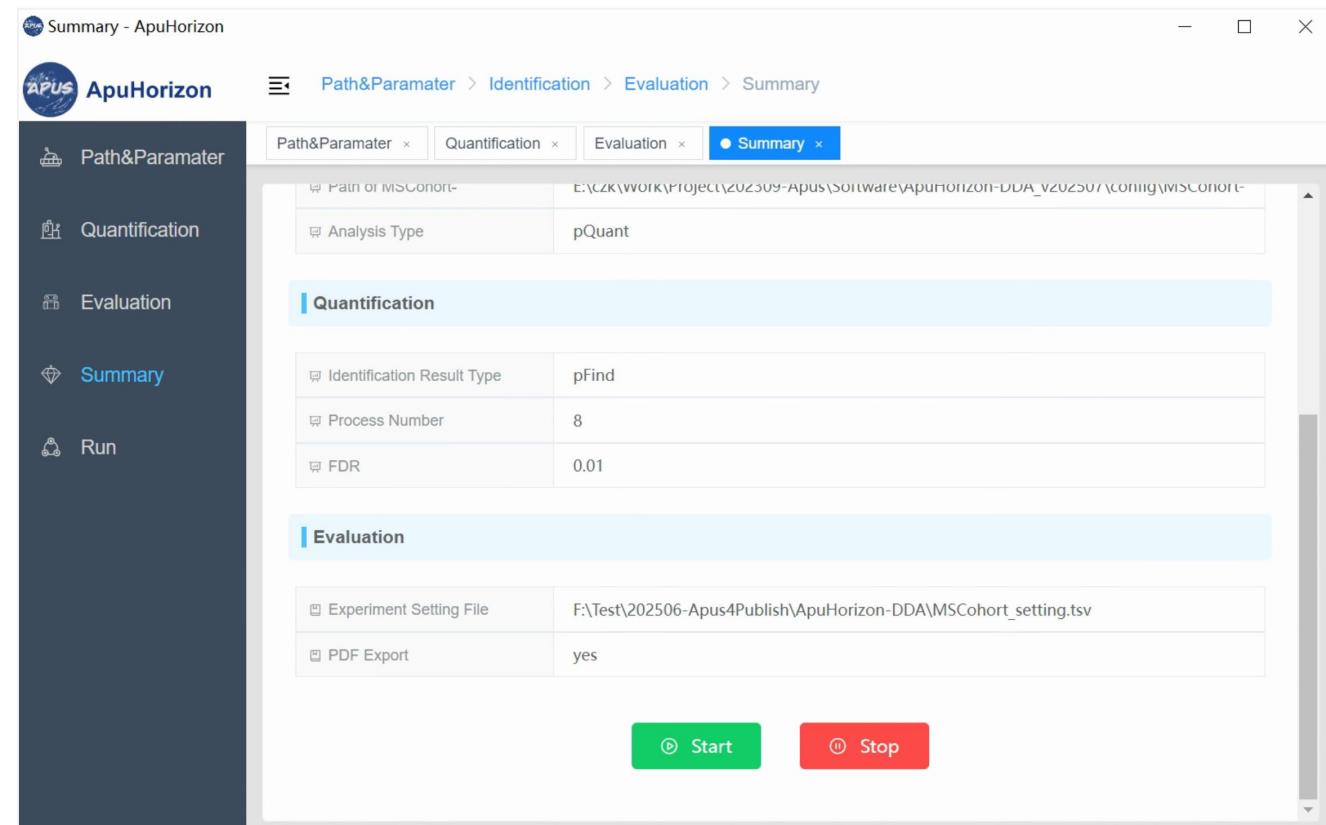
Fill in the following **MSCohort-DDA parameters** for quality control.

More advanced **MSCohort-DDA settings** are already set to default. You can view or edit them in the file: “/ApuHorizon-DDA_v202507/config/MSCohort-DDA_config.txt”.



STEP-6 Summary

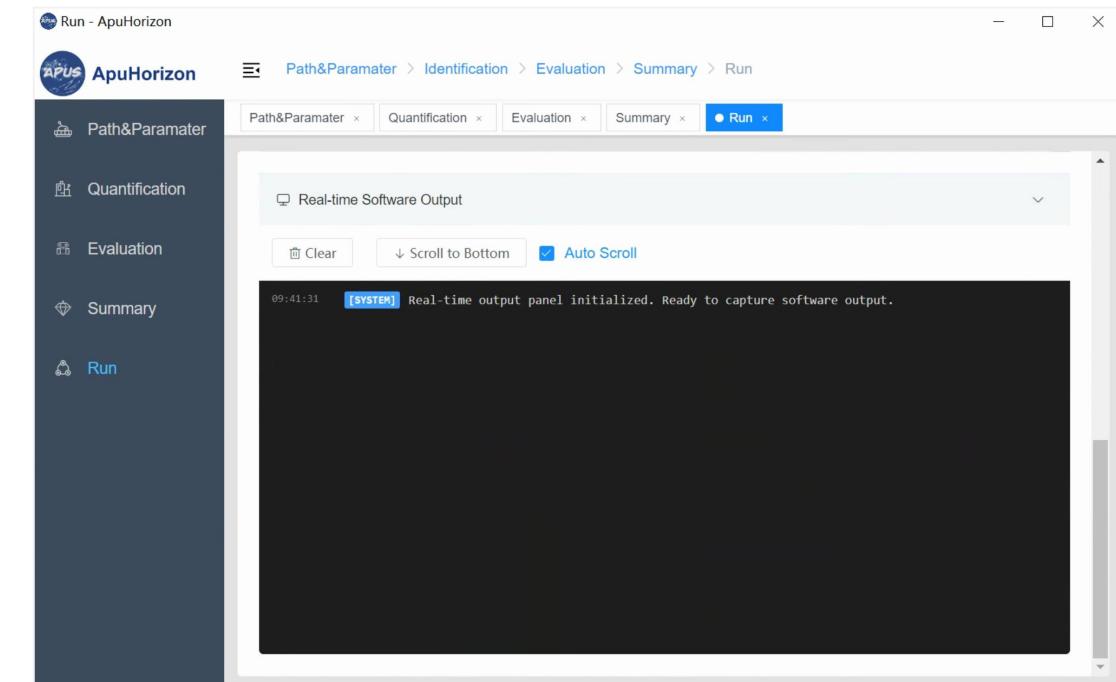
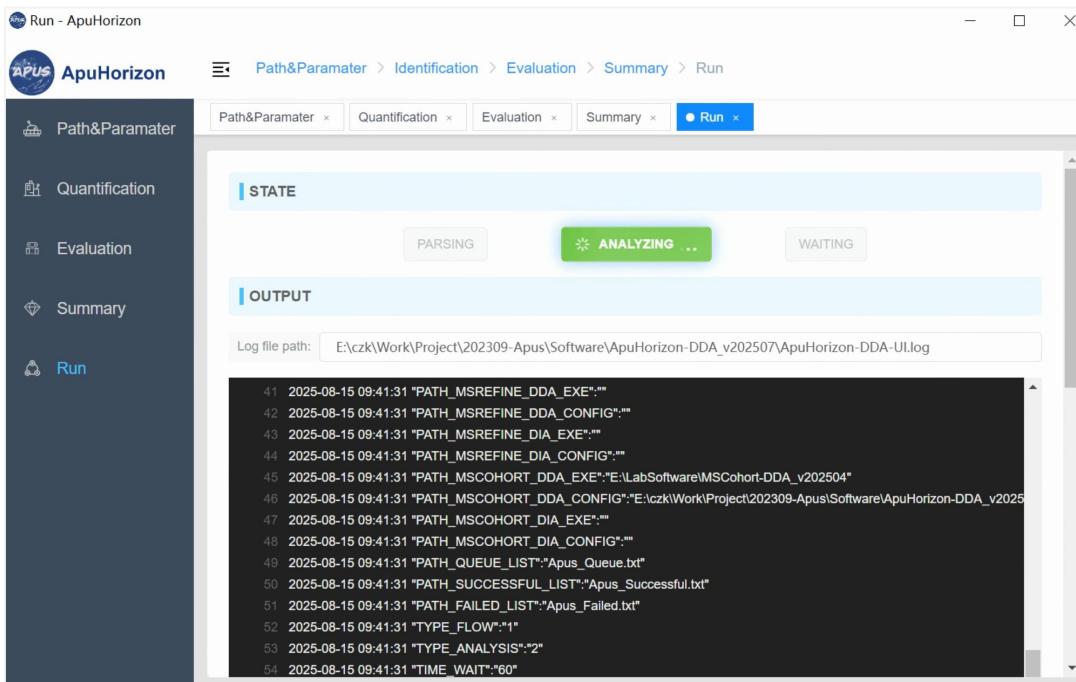
- Review configuration:** Navigate to the Summary page to review all previously configured parameters.
- Start process:** After verifying that all settings are correct, click the **Start** button to launch the application.



STEP-7 Run

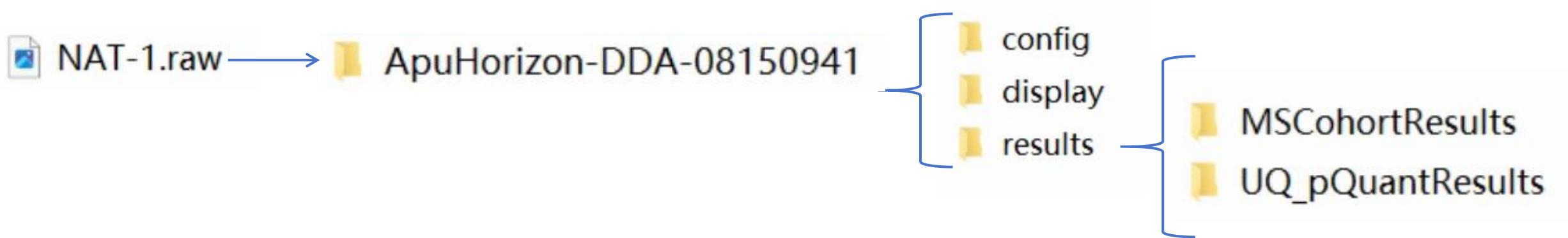
Go to the Run page to watch the program's progress.

- **STATE:** Shows the current status of the program's loop.
- **OUTPUT:** Here you can see live logs:
 1. Logs from ApuHorizon-DDA.
 2. More detailed logs from the tools it calls (like pQuant).



STEP-8 View analysis results

Navigate to "Monitored Path" and view the data analysis results in the folder named **ApuHorizon-DDA+[Timestamp]**.



ATTENTIONS

1. Always click **Save** or **Save As** before leaving a page. If you don't, your changes will be lost.
2. ApuHorizon automatically locates the analysis results from ApuPioneer. **Do not manually rename the identification result folder, as this will prevent the analysis from running correctly.**

INTRODUCTION

ApuHorizon: The protein quantification system of Apus. It is deployed on a dedicated server with substantial memory to handle the high-throughput quantification of a large cohort comprising hundreds to thousands of samples.

- **Input:** Receives identification results from ApuPioneer.
- **Process:** Performs LFQ using a proven engine (defaults to **DIA-NN** for DIA mode).
- **Output:** Generates a result file detailing analyte abundances per sample.
- **Integration:** Output is ready for immediate use in ApuInsight (for bioinformatics).

INTRODUCTION

Version: ApuHorizon-DIA_v202507

Release Date: 2025.07.15

Computer configuration

CPU: Intel or AMD processor, 32 cores or higher is recommended

RAM: 256GB or higher is recommended

ROM: 1TB or higher high-speed NVMe SSD is recommended

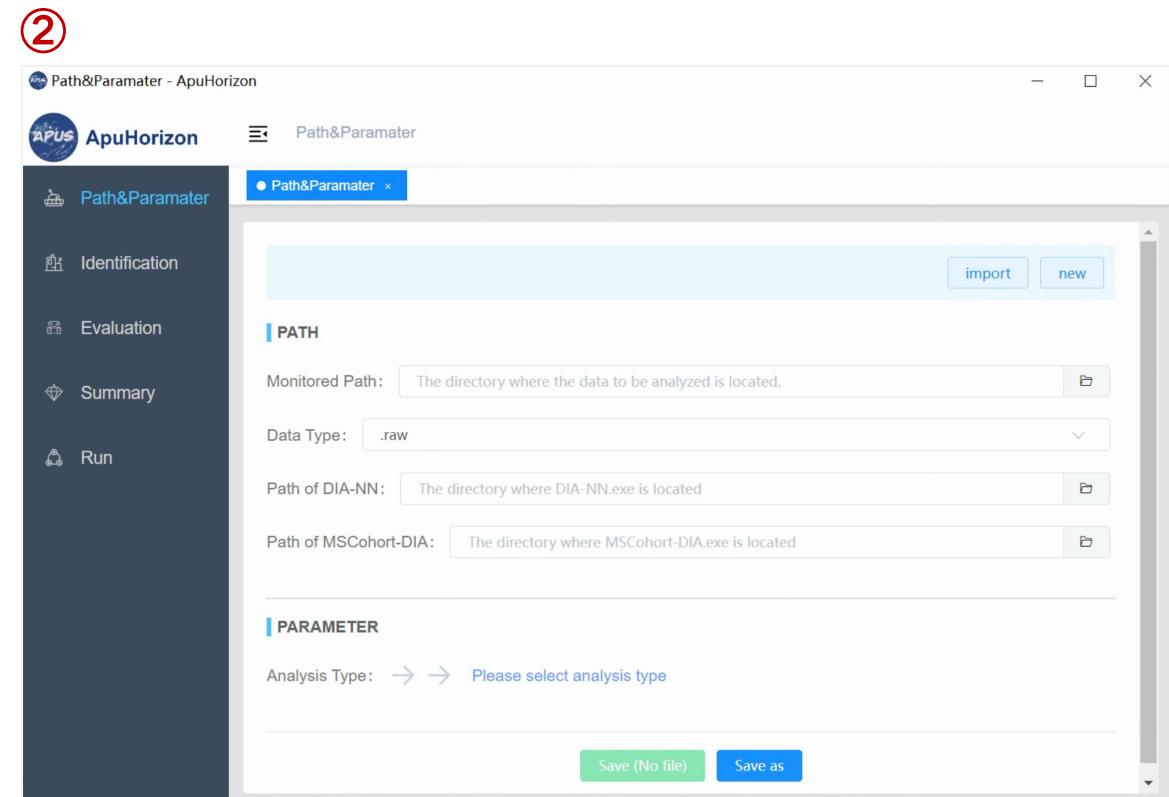
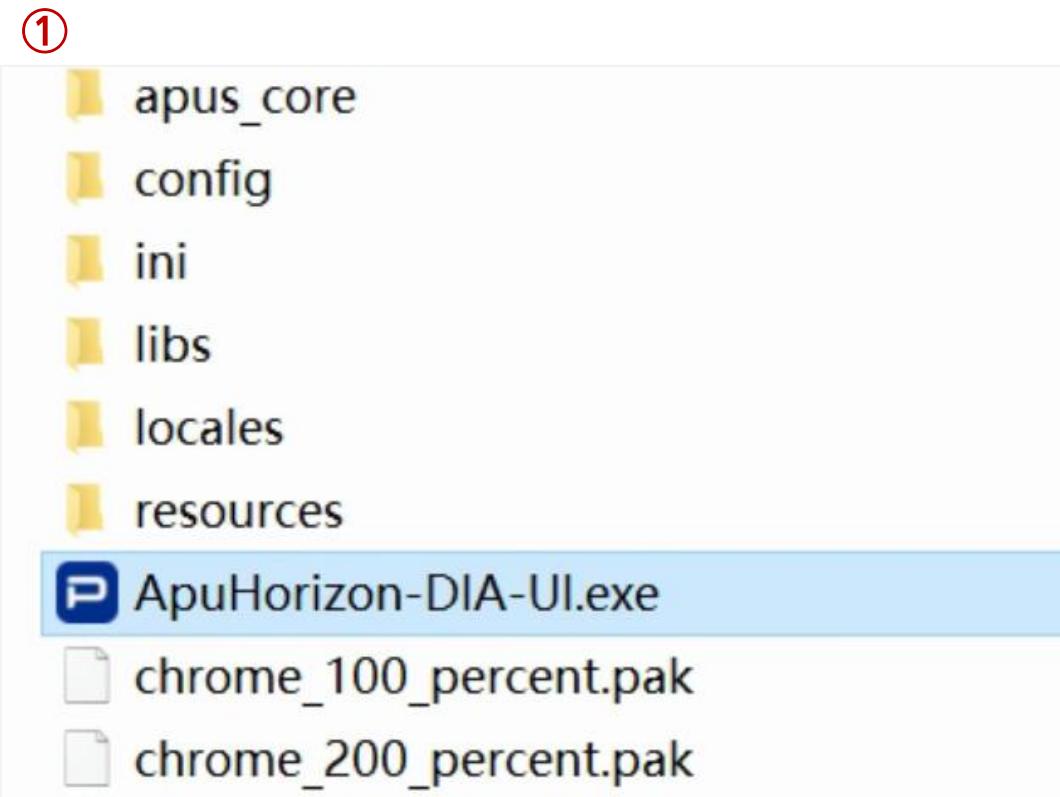
OS: Windows 10 (x64) or Windows 11 (x64)

STEP-1 Preparation

1. **Install software:** Deploy ApuHorizon-DIA, DIA-NN, and MSCohort-DIA on the quantitative analysis server.
 - Download ApuHorizon-DIA and MSCohort-DIA from <https://github.com/BUAA-LiuLab/Apus>.
 - Download DIA-NN from <https://github.com/vdemichev/DiaNN>.
2. **Mount storage:** Use the **SMB** protocol to mount your data storage system to the local filesystem of the compute nodes, so it can be accessed like a local disk.

STEP-2

1. Navigate to the “/ApuHorizon-DIA_v202507/” directory.
2. Execute **ApuHorizon-DIA-UI.exe** to launch the application. This will open the main software interface.



STEP-3 Path&Paramater

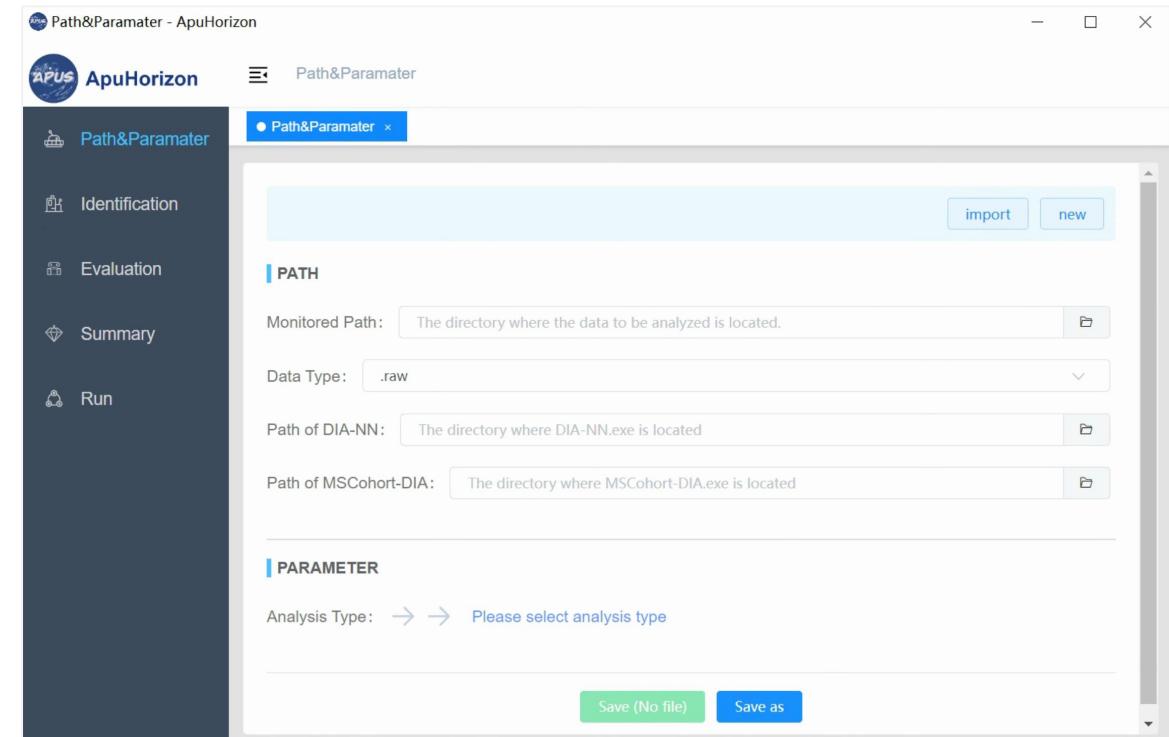
1. Fill in **data information**:

1. Monitored path: The folder path where your mass spectrometry data is stored.
2. Data type: The format of your data files (e.g., .raw, .d, .wiff).

2. Fill in **software paths**:

Enter the installation directory paths for DIA-NN and MSCohort-DIA. This should be the folder containing the main executable file (e.g., MSCohort-DIA.exe).

3. Fill in **analysis type**: "quantification only" or "quantification + quality control".



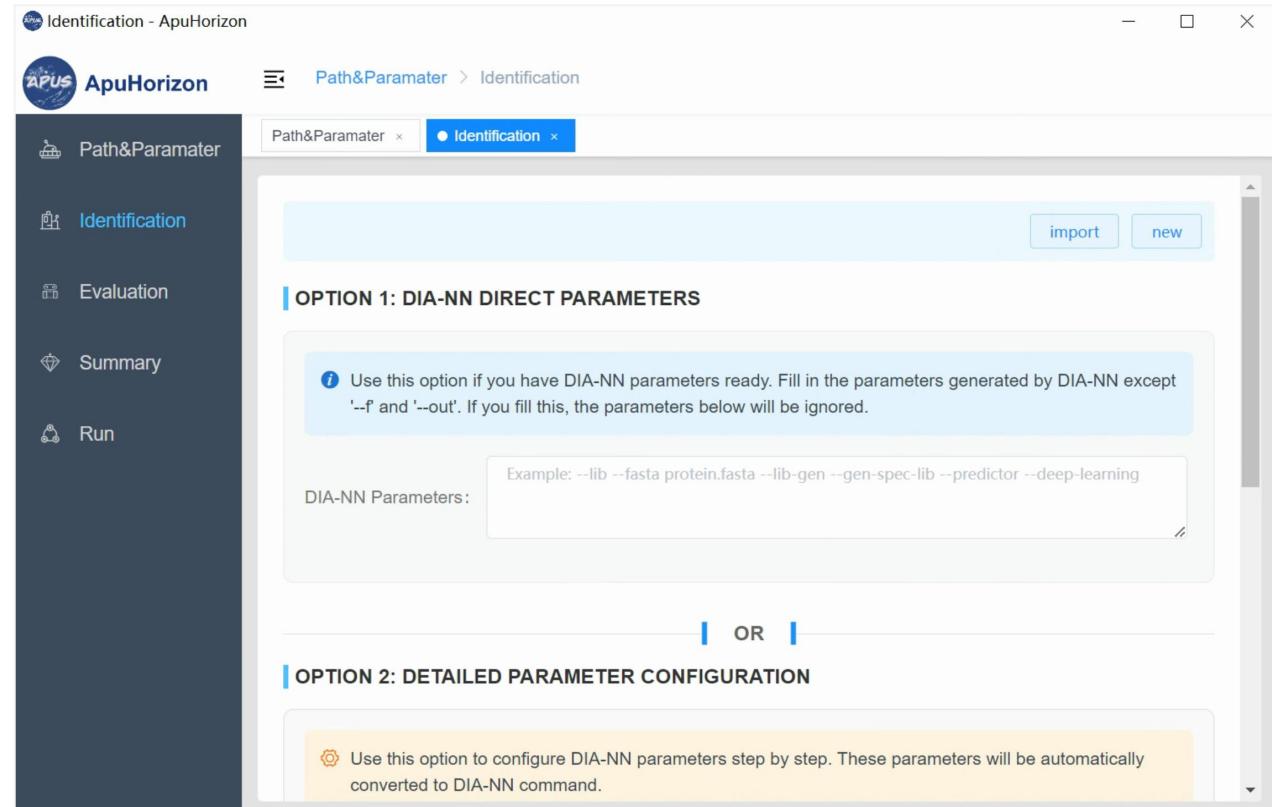
STEP-4 Quantification

Configure DIA-NN Parameters.

1. Many parameters are preset with optimized default values. Users only need to specify some essential parameters to run the analysis.
2. Notably, for DIA mode, besides setting common DIA-NN parameters via the GUI, ApuHorizon allows users to input a full command-line argument string.

This string should be generated from the native DIA-NN interface, excluding the `--f` and `--out` arguments.

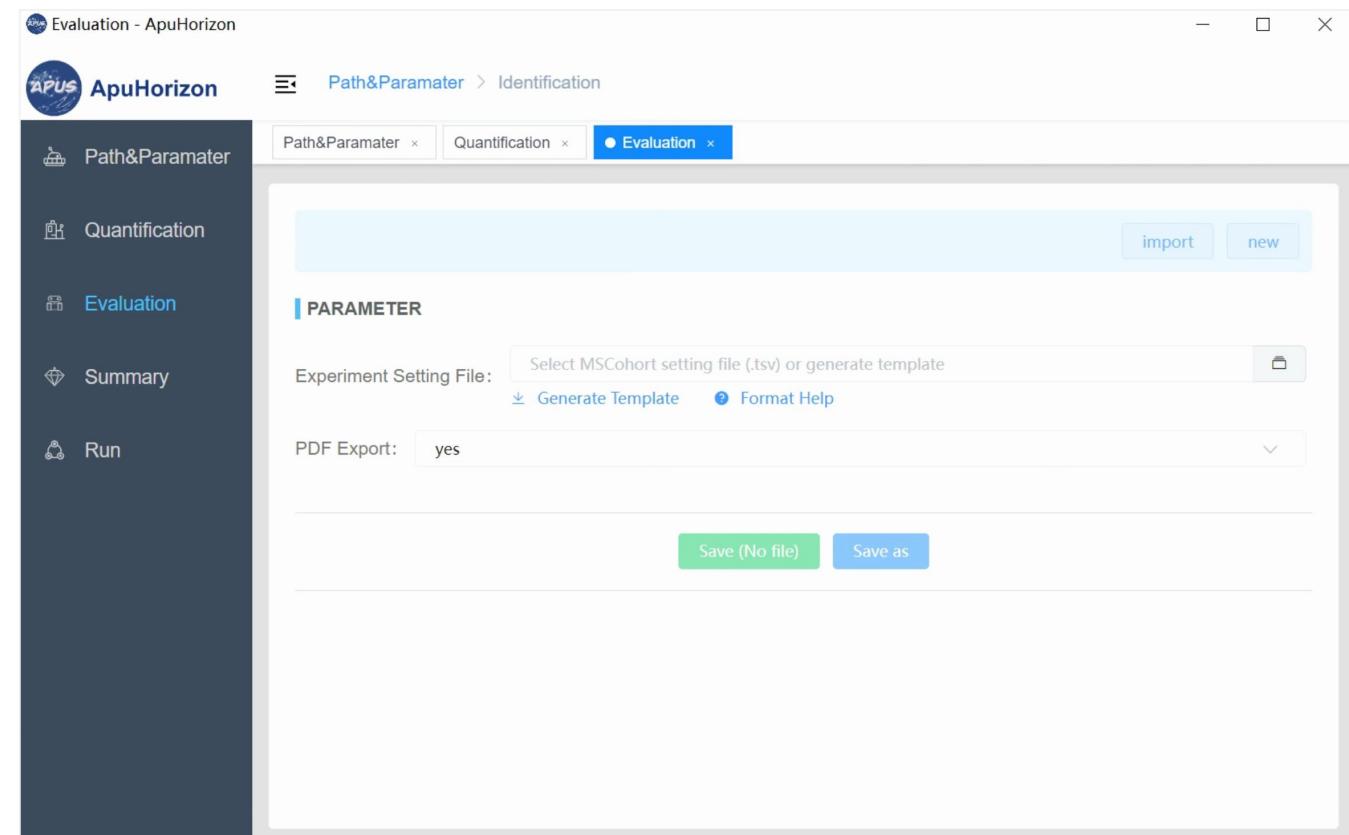
This feature provides fine-grained control over the quantification process.



STEP-5 Evaluation

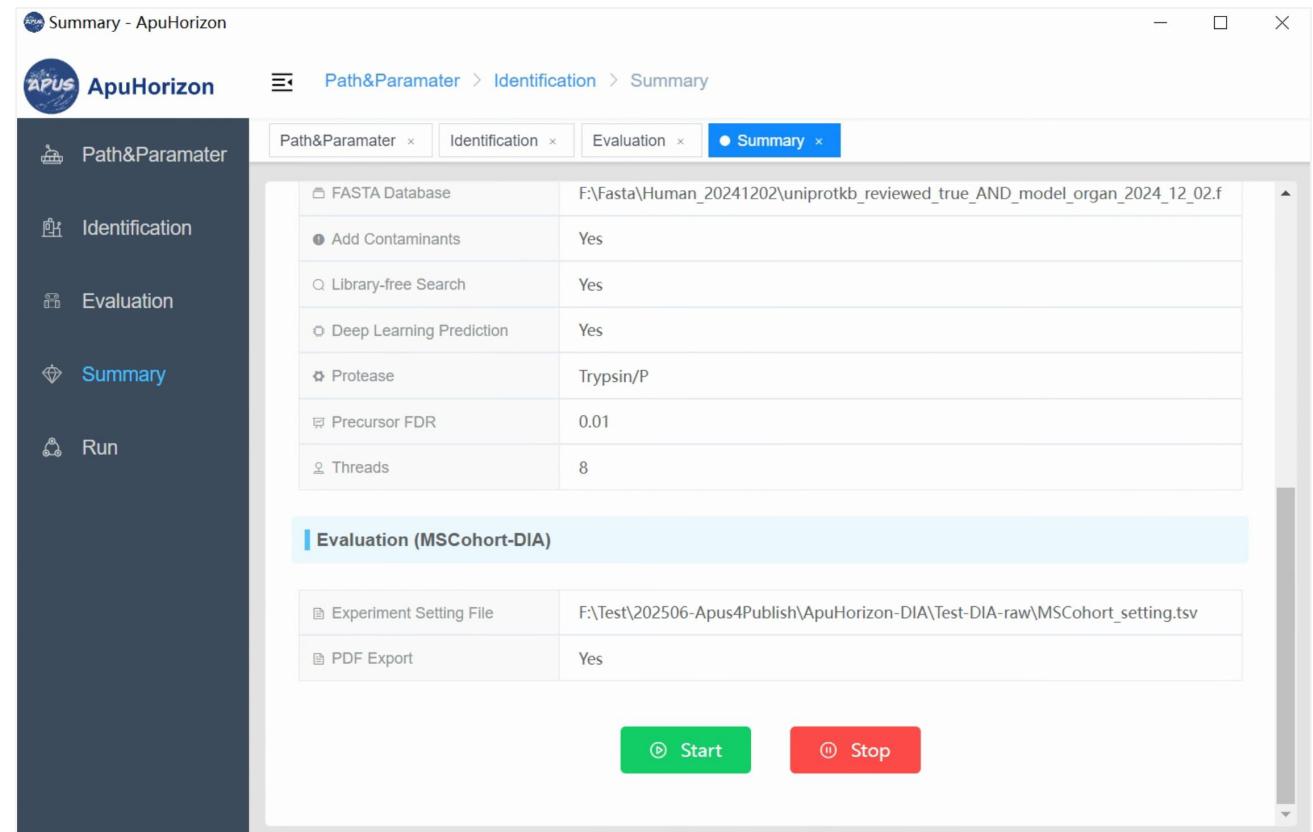
Fill in the following **MSCohort-DIA parameters** for quality control.

More advanced **MSCohort-DIA settings are already set to default**. You can view or edit them in the file: “/ApuHorizon-DIA_v202507/config/MSCohort-DIA_cfg.txt”.



STEP-6 Summary

- Review configuration:** Navigate to the Summary page to review all previously configured parameters.
- Start process:** After verifying that all settings are correct, click the **Start** button to launch the application.



STEP-7 Run

Go to the Run page to watch the program's progress.

- **STATE:** Shows the current status of the program's loop.
- **OUTPUT:** Here you can see live logs:
 1. Logs from ApuHorizon-DIA.
 2. More detailed logs from the tools it calls (like DIA-NN).

The screenshot shows the ApuHorizon software interface. The main window title is "Run - ApuHorizon". The left sidebar has tabs for "Path&Parameter", "Identification", "Evaluation", "Summary", and "Run". The "Run" tab is selected. The main area has two sections: "STATE" and "OUTPUT". The "STATE" section shows three buttons: "PARSING", "ANALYZING ...", and "WAITING". The "OUTPUT" section shows a log file path: "E:\czk\Work\Project\202309-Apus\Software\ApuHorizon-DIA_v202507\ApuHorizon-DIA-UI.log". Below this, there is a scrollable text area displaying log entries from line 45 to 58. The log entries include various system messages and file paths related to the DIA-NN analysis.

```
45 2025-08-15 15:09:59 "PATH_MSCOHORT_DDA_EXE": ""
46 2025-08-15 15:09:59 "PATH_MSCOHORT_DDA_CONFIG": ""
47 2025-08-15 15:09:59 "PATH_MSCOHORT_DIA_EXE": "E:\LabSoftware\MSCohort-DIANN_v202504"
48 2025-08-15 15:09:59 "PATH_MSCOHORT_DIA_CONFIG": "E:\czk\Work\Project\202309-Apus\Software\ApuHorizon-DIA_v202507\ApuHorizon-DIA-UI.log"
49 2025-08-15 15:09:59 "PATH_QUEUE_LIST": "Apus_Queue.txt"
50 2025-08-15 15:09:59 "PATH_SUCCESSFUL_LIST": "Apus_Successful.txt"
51 2025-08-15 15:09:59 "PATH_FAILED_LIST": "Apus_Failed.txt"
52 2025-08-15 15:09:59 "TYPE_FLOW": "1"
53 2025-08-15 15:09:59 "TYPE_ANALYSIS": "2"
54 2025-08-15 15:09:59 "TIME_WAIT": "60"
55 2025-08-15 15:09:59 [Apus] [stat] Reading parameters...
56 2025-08-15 15:09:59 [Apus] [stat] Parsing folder...
57 2025-08-15 15:10:00 [Apus] [info] The total number of files need to be analyzed: 3
58 2025-08-15 15:10:01 [Apus] [stat] Data detected, automatic analysis begins!
```

The screenshot shows the ApuHorizon software interface. The main window title is "Run - ApuHorizon". The left sidebar has tabs for "Path&Parameter", "Identification", "Evaluation", "Summary", and "Run". The "Run" tab is selected. The main area has a section titled "Real-time Software Output" which is "Running". There are buttons for "Clear", "Scroll to Bottom", and "Auto Scroll". The log area shows a series of log entries from 15:10:02 to 15:10:22. The log entries are primarily INFO level messages from the DIA-NN process, including loading FASTA files, processing, assembling elution groups, and predicting spectra and IMs. Some WARNING messages are also present, such as one about mass accuracy optimization.

```
15:10:02 [INFO] DIA-NN will optimise the mass accuracy automatically using the first run in the experiment. This is useful primarily for quick initial analyses, when it is not yet known which mass accuracy setting works best for a particular acquisition scheme.
WARNING: it is strongly recommended to first generate an in silico-predicted library in a separate pipeline step and then use it to process the raw data, now without activating FASTA digest
15:10:02 [INFO] 3 files will be processed
[0:00] Loading FASTA
F:\Fasta\Human_20241202\uniprotkb_reviewed_true_AND_model_organ_2024_12_02.fasta
15:10:04 [INFO] [0:02] Loading FASTA camprotR_240512_cRAP_20190401_full_tags.fasta
15:10:05 [INFO] [0:02] Processing FASTA
15:10:10 [INFO] [0:07] Assembling elution groups
15:10:14 [INFO] [0:11] 4302395 precursors generated
15:10:14 [INFO] [0:11] Gene names missing for some isoforms
15:10:14 [INFO] [0:11] Library contains 20522 proteins, and 20284 genes
15:10:16 [INFO] [0:13] Encoding peptides for spectra and RTs prediction
15:10:22 [INFO] [0:20] Predicting spectra and IMs
```

STEP-8 View analysis results

Navigate to "Monitored Path" and view the data analysis results in the folder named **ApuHorizon-DIA+[Timestamp]**.



ATTENTIONS

1. Always click **Save** or **Save As** before leaving a page. If you don't, your changes will be lost.
2. ApuHorizon automatically locates the analysis results from ApuPioneer. **Do not manually rename the identification result folder, as this will prevent the analysis from running correctly.**

ApuMonitor

INTRODUCTION

Version: ApuMonitor_v202507

Release Date: 2025.07.15

OS: Windows 10 (x64) or Windows 11 (x64)

The **ApuMonitor** system is deployed on the identification nodes, the quantification server, and a dedicated monitoring server to perform both intra- and inter-experiment quality control (QC).

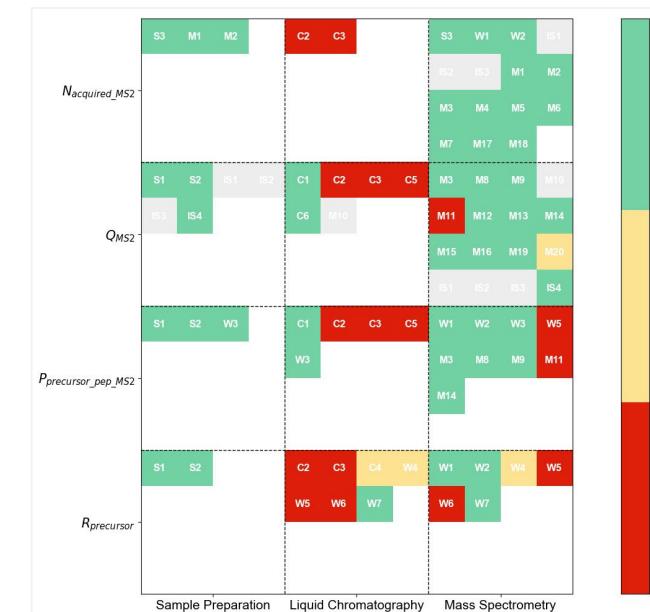
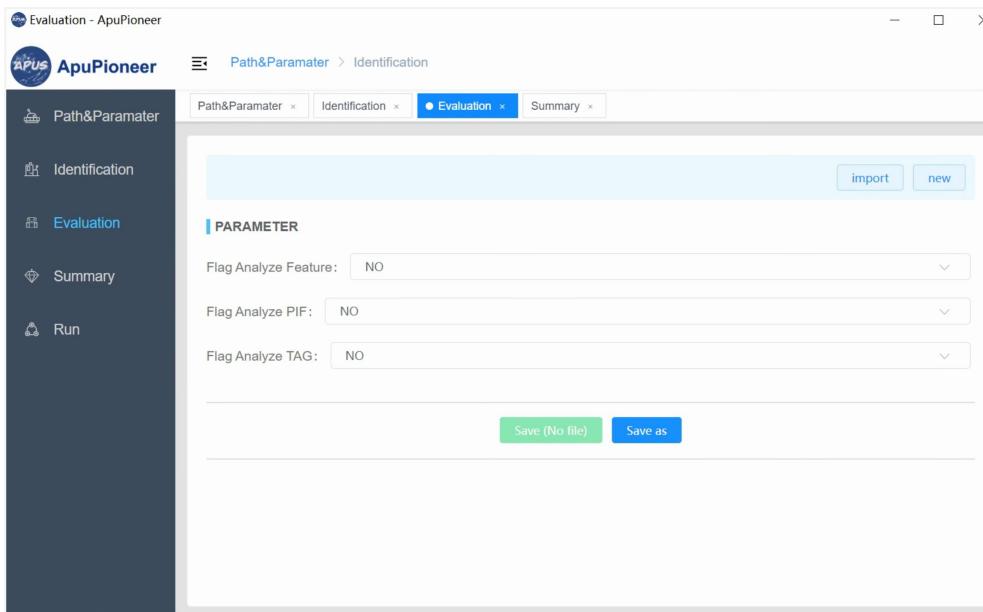
- Intra-experiment QC is automatically triggered for each run after database searching is completed. This process tracks key metrics related to data acquisition and identification and generates a single-run QC report.
- Inter-experiment QC is performed following the quantification of the entire cohort, producing a cohort-level QC report to identify low-quality samples.
- Additionally, Apus provides a visualization dashboard that allows researchers to monitor data quality in real time from PC/Pad/Phone, thereby ensuring the transparency and traceability of the entire analytical workflow.

MODE-1 Intra-experiment QC

When performing identification analysis with ApuPioneer, users can select the "database search + QC" mode and configure its parameters on the **Evaluation** interface.

The system will then automatically perform intra-experiment QC for each dataset upon the completion of database searching.

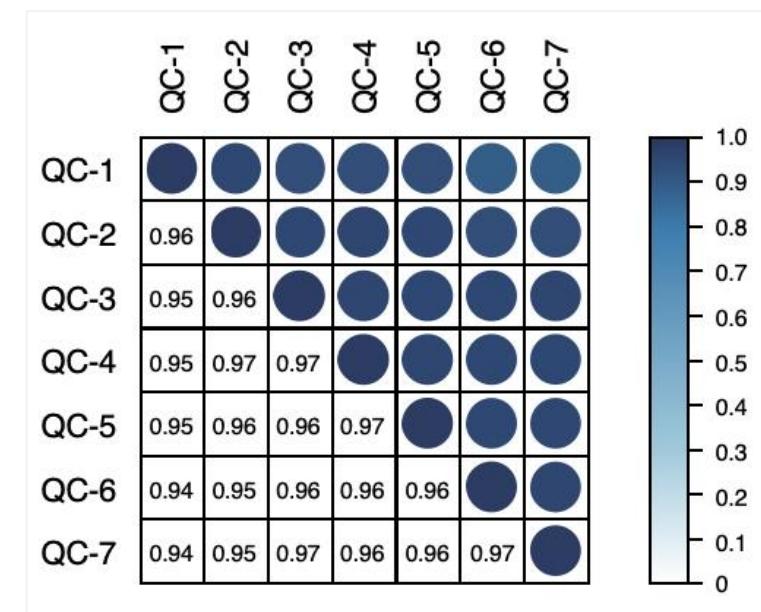
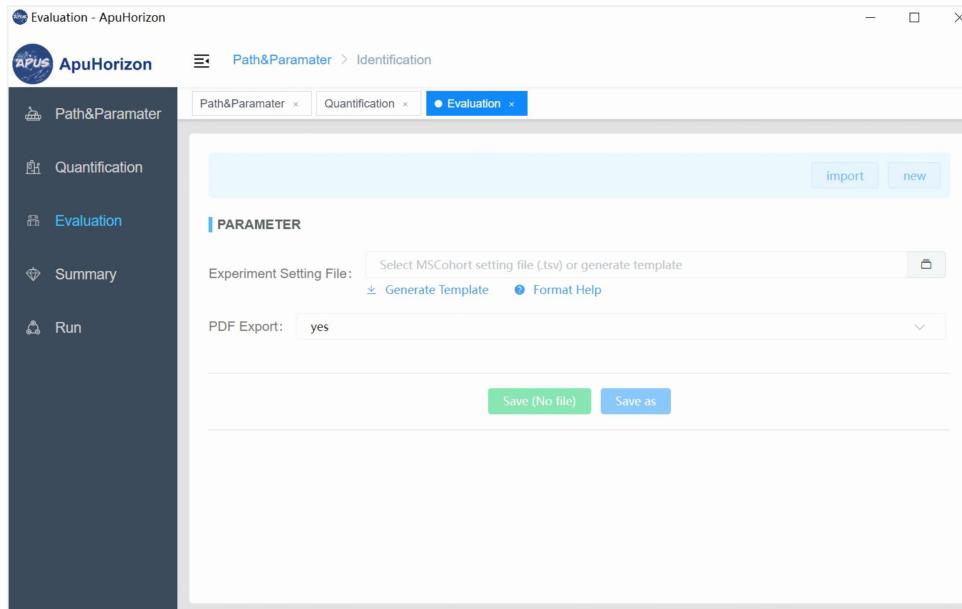
This process tracks QC metrics throughout data acquisition and identification, ultimately generating a QC report for each individual sample.



MODE-2 Inter-experiment QC

When performing quantification analysis with ApuHorizon, users can select the "quantification + QC" mode and configure parameters on the **Evaluation** interface.

After ApuHorizon completes quantification for the entire cohort, ApuMonitor automatically extracts 23 inter-experiment QC metrics, evaluates the quality of the entire cohort, and generates a cross-experiment QC report.



MODE-3 Visualization platform

ApuMonitor provides a dedicated visualization platform for intra-experiment QC. This platform extracts key QC metrics in real-time from a pool of 58 available metrics and transfers them to a local MySQL database.

By configuring a data connection in FineReport, the system reads these metrics from the MySQL database in real-time and renders them onto a FineReport dashboard. This architecture enables the visualization to be displayed on any device.

The following section outlines the procedure for using ApuMonitor to achieve real-time monitoring of data quality.

MODE-3 Visualization platform

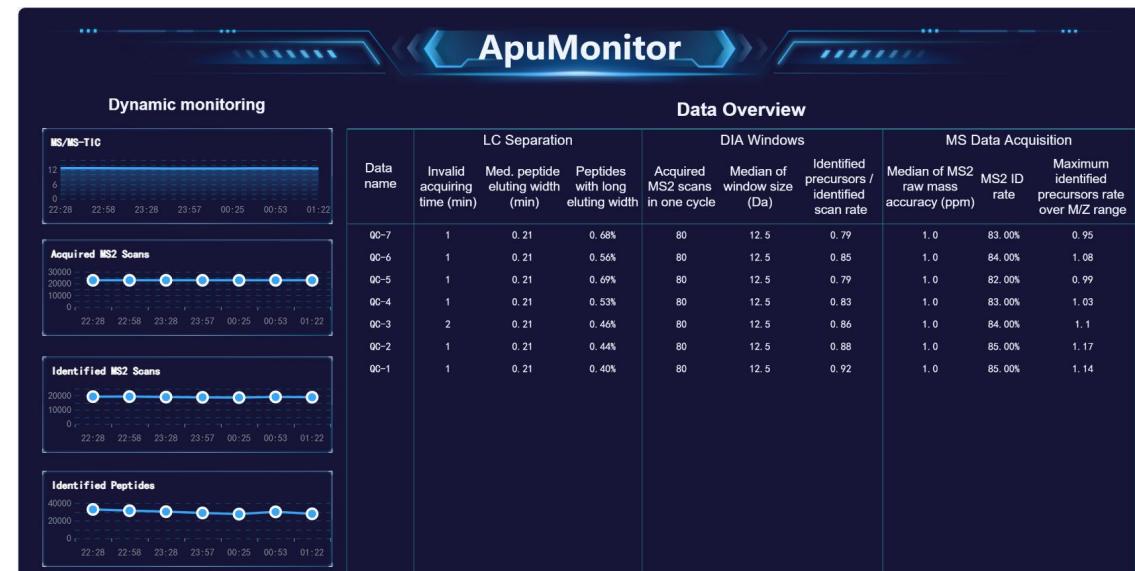
1. Download [ApuMonitor_v202507](#).
2. Navigate into the downloaded folder and select either the DDA or DIA subdirectory, depending on your experimental design. This guide will use ApuMonitor-DDA as an example.
3. Install [MySQL](#) on a computer that is on the same network as the data storage system.
4. Install a [graphical database management tool](#), such as [Navicat](#).
5. Create a new [connection](#) (you can assign it any name).
6. Within that connection, [create a new database named apumonitor](#). (Character set: utf8mb3; Collation: utf8mb3_general_ci)
7. [Import the apumonitor.sql file](#) located in the ApuMonitor-DDA folder into the newly created apumonitor database. (Note: Ensure you are importing the .sql file directly into the apumonitor database, not into the connection level.)
8. At this point, you can inspect the apumonitor database in Navicat. It should contain two tables: [data_overview](#) and [figure](#).

MODE-3 Visualization platform

9. Mount the data storage system to your local file system using the SMB protocol. This will make the remote storage accessible as a local disk on your machine.
10. Open the [ApuMonitor_cfg.txt](#) file. You need to [customize the following five parameters](#):
 - ① PATH_MONITORED: The directory where the mass spectrometry data is stored.
 - ② DB_HOST: The hostname of the database server (e.g., localhost).
 - ③ DB_DATABASE: The name of the database (e.g., apumonitor).
 - ④ DB_USERNAME: The username for database access.
 - ⑤ DB_PASSWORD: The password for the specified user.
11. Execute the software from the command line with the following command: [ApuMonitor.exe](#) [ApuMonitor_cfg.txt](#)
12. Once running, ApuMonitor will establish communication with the local MySQL database. It will periodically check for new results, read metrics from the MSRefineResult folder (located within the PATH_MONITORED directory), and upload the data into the data_overview and figure tables.

MODE-3 Visualization platform

13. Download and install the **FineReport** reporting tool (version 11.0 is recommended). After installation, create a new project and import the ApuMonitor-DDA.frm report template.
14. Within FineReport, establish a data connection to the MySQL database you configured previously.
15. Open the report. It will be rendered as a web report in your default browser.
16. If ApuMonitor.exe is running, local data will be continuously uploaded to the MySQL database. Consequently, the web report will automatically refresh every 10 minutes to display the latest data.



ApulInsight

INTRODUCTION

Version: ApulInsight_v202507

Release Date: 2025.07.15

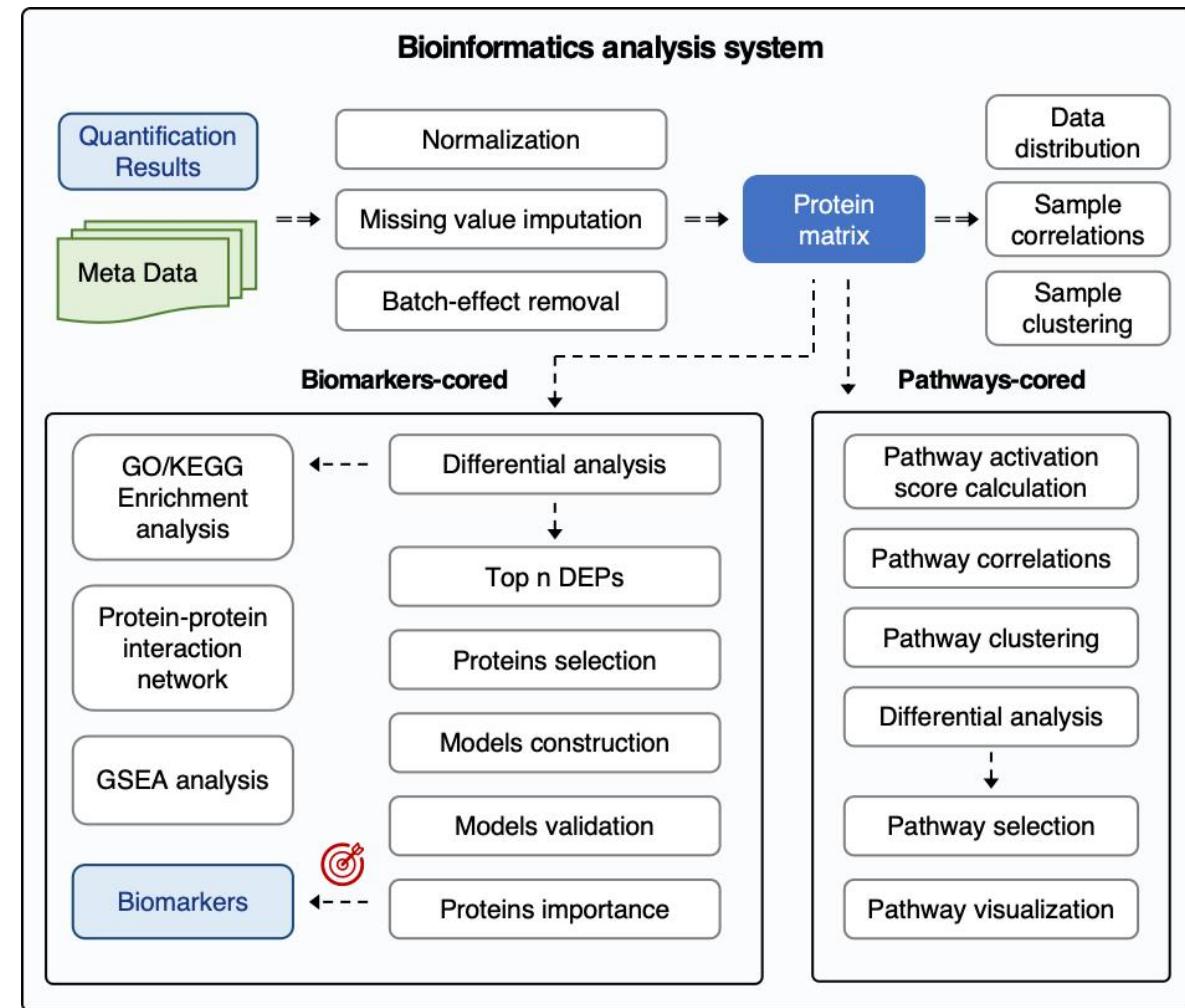
OS: Windows 10 (x64) or Windows 11 (x64)

ApulInsight is an automated bioinformatics workflow for large-cohort proteomics analysis. Users only need to provide a metadata file based on the protein quantification file. Subsequently, the automated analysis process can be initiated.

We developed two workflows to meet the needs of large-cohort proteomics analysis.

(1) Biomarker-cored workflow

(2) Pathway-cored workflow



STEP-1 Install

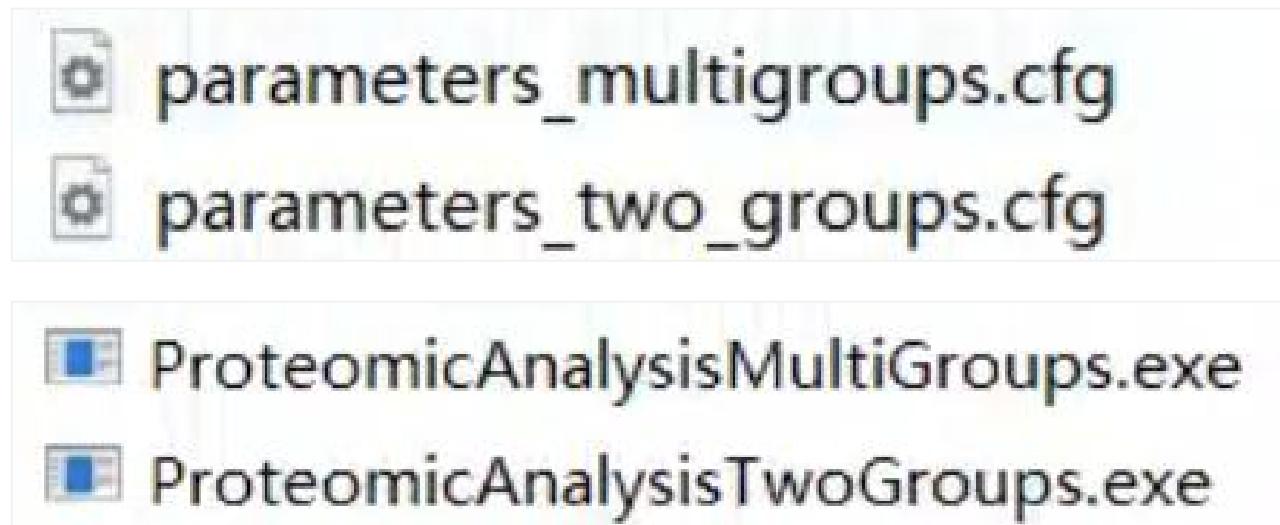
1. Download the installation package of Apulnsight from [*https://github.com/BUAA-LiuLab/Apus*](https://github.com/BUAA-LiuLab/Apus).
2. Enter the “Apulnsight_v202507” directory
3. Click on “**install.exe**” to install, and after completion, a folder named "**proteomicAnalysis**" will be generated, containing all contents such as scripts, dependency packages, configuration files, etc.

STEP-2 Fill in the configuration file

Configuration files and their corresponding exe programs are categorized into two types:

1. Two groups analysis: The configuration file is "parameters_two_groups.cfg", and the program is "ProteomicAnalysisMultiGroups.exe"
2. Multiple group analysis: The configuration file is "parameters_multigroups.cfg", and the program is "ProteomicAnalysisTwoGroups.exe"

According to the specific analysis task, fill in the configuration file (see the file "ApuInsight-Parameter Description.docx" for detailed parameter explanations)



STEP-3 Fill in meta information

1. The filename must be exactly "metadata.xlsx".
2. The file must contain four columns: **Sample**, **Condition**, **Batch**, and **PairedID**, corresponding to sample names, group labels, batch information, and pairing information, respectively.
 - The Condition value should be an abbreviation derived from the Sample name (e.g., for Sample "CTRL1" / "MSA1", the Condition would be "CTRL" / "MSA").
3. If batch information is available, fill it in the Batch column; otherwise, leave it blank.
4. For paired differential analysis, provide IDs in the PairedID column. The format should be the Condition value followed by a number (e.g., "CTRL_1", "MSA_1"). Paired samples must share the same numeric suffix. If this analysis is not needed, leave the column blank.
5. Avoid using special characters such as hyphens (-) in the file to prevent potential errors.
6. The image on the right is a simple example.

	A	B	C	D
1	Sample	Condition	Batch	PairedID
2	NAT_1	NAT		
3	NAT_2	NAT		
4	NAT_3	NAT		
5	NAT_4	NAT		
6	NAT_5	NAT		
7	T_1	T		
8	T_2	T		
9	T_3	T		
10	T_4	T		
11	T_5	T		

STEP-4 Run

After filling in the parameter configuration file and metadata.xlsx, run the program exe in cmd, followed by the corresponding cfg file name.

Example: **ProteomicAnalysisTwoGroups.exe parameters_two_groups.comg**

After waiting for the completion of the run, view the analysis results in the custom result output directory.

