

# MSCohort Manual



Version. 202410

### ### 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 (1G) 2G or higher is recommended

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

**Other:** At present, Spectronaut identification and quantification results can be used for quality analysis by MSCohort.



### ### Demo datasets

The demo dataset has been deposited to the ProteomeXchange Consortium (<https://proteomecentral.proteomexchange.org>) via the iProX partner repository with the dataset identifier PXD057133 (in ProteomeXchange) and IPX0010061000 (in iProX).

You can use the following link: <https://www.iprox.cn/page/project.html?id=IPX0010061000> to download the Demo datasets.

The demo datasets available for MSCohort analysis. You can download this dataset for testing and using MSCohort.

This dataset contains Demo DIA folds / Demo DDA folds / Demo PRM folds

- Demo DIA folds contains raw files of 7 human urine QC DIA data from Orbitrap Exploris 480, Spectronaut analysis results and MSCohort report results.
- Demo DDA folds contains raw files of 3 human urine QC DDA data from Orbitrap Exploris 480, pFind/pQuant analysis results and MSCohort report results.
- Demo PRM folds contains Skyline and SpectroDive analysis results and MSCohort report results.

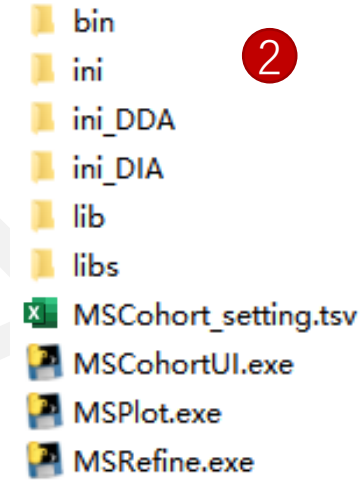
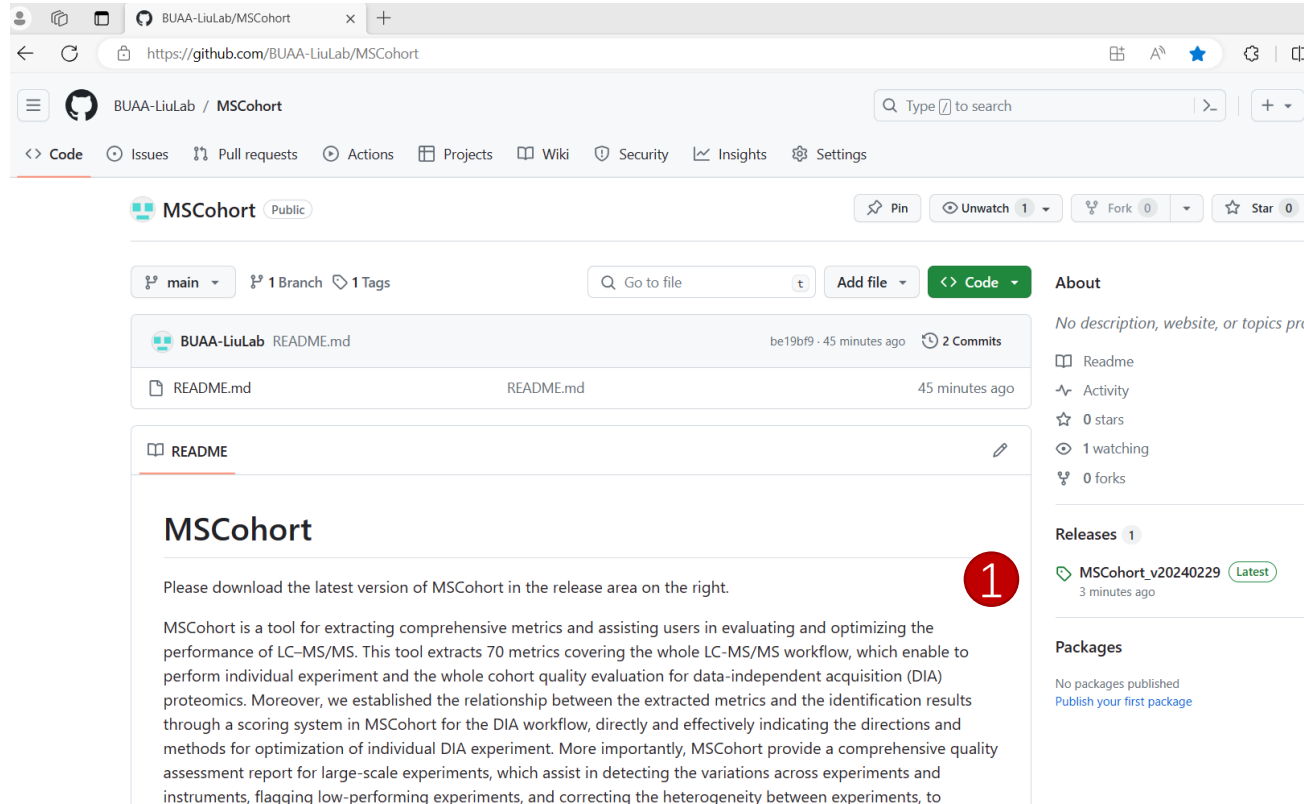
# CONTENTS

- [MSCohort download](#)
- [MSCohort manual for DIA intra-experiment analysis](#)
- [MSCohort manual for DIA inter-experiment analysis](#)
- [MSCohort manual for DDA intra-experiment analysis](#)
- [MSCohort manual for DDA inter-experiment analysis](#)
- [MSCohort manual for PRM inter-experiment analysis](#)

# **MSCohort download**



# 1. MSCohort download



## ① Login

<https://github.com/BUAA-LiuLab/MS Cohort>

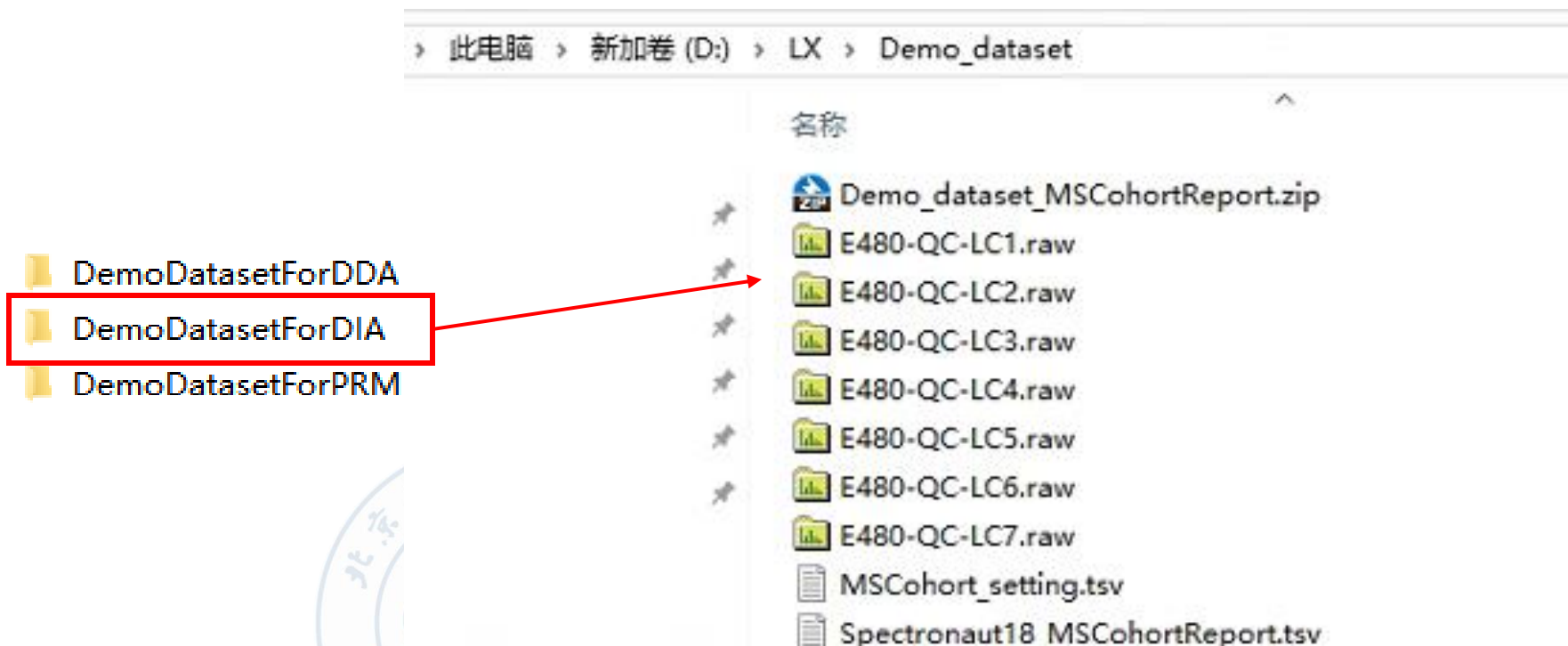
and download the latest version at the right of the page.

## ② The unzipped MSCohort file.

# **MSCohort manual for DIA intra- experiment analysis**



# Demo dataset

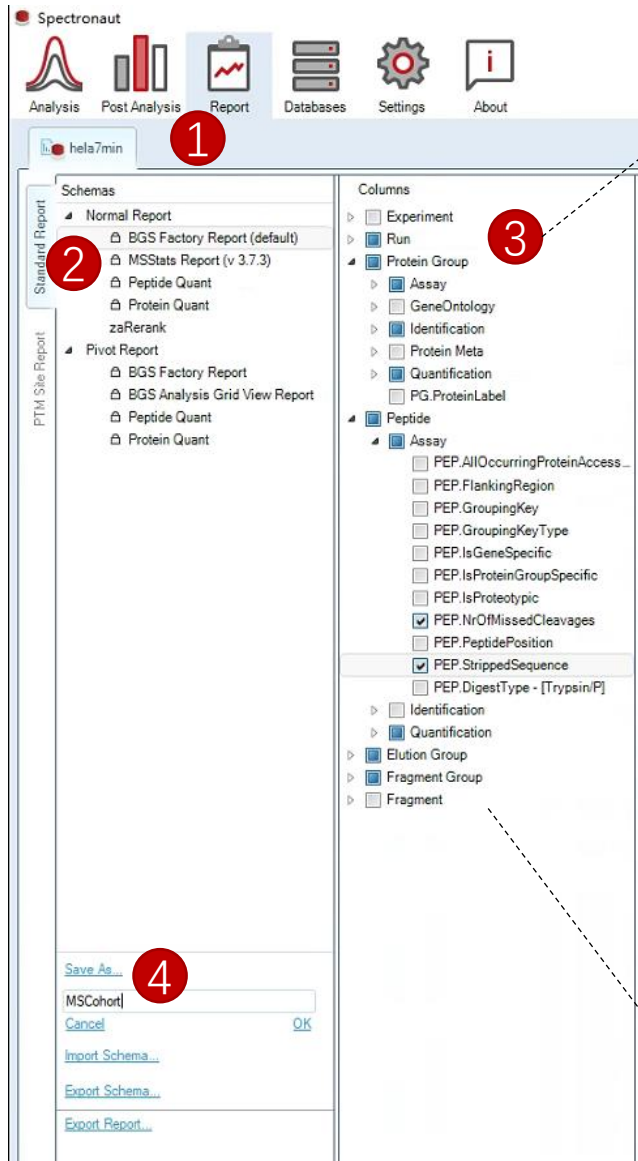


Taking the E480-QC-LC1.raw file as an example to demonstrate the workflow of intra-experiment analysis.

Intra-experiment analysis enables the systematic evaluation and optimization of individual (single-run) DIA experiments.



# Spectronaut customized report



The columns of information that MSCohort needs to read

<b>Run</b>	
	R.Run Date
	R.Gradient Length [min]
	R.FileName
<b>Protein Group</b>	
	PG.ProteinGroups
	PG.Qvalue
	PG.Quantity
<b>Peptide</b>	
	PEP.NrOfMissedCleavages
	PEP.StrippedSequence
	PEP.Quantity
<b>Elution Group</b>	
	EG.IsDecoy
	EG.ModifiedPeptide
	EG.PrecursorId
	EG.Qvalue
	EG.ApexRT
	EG.DatapointsPerPeak
	EG.DatapointsPerPeak (MS1)
	EG.DeltaRT
	EG.EndRT
	EG.FWHM
	EG.PeakWidth
	EG.StartRT
	EG.SignalToNoise
	EG.TotalQuantity (Settings)
<b>Fragment Group</b>	
	FG.Charge
	FG.PrecMz
	FG.PrecWindow
	FG.PrecWindowNumber
	FG.CalibratedMassAccuracy (PPM)
	FG.RawMassAccuracy (PPM)
<b>Fragment</b>	
	F.RawMassAccuracy (PPM)
	F.CalibratedMassAccuracy (PPM)

- 1 Choose **Report** Perspective in Spectronaut;
- 2 Choose a **Normal Report** format schema as a base to build MSCohort report;
- 3 Choose the **columns** of information that MSCohort needs to read;
- 4 Save the selected columns as a new schema, and name the new report schema as “**MSCohort**”.

# Export MSCohort report from Spectronaut

The screenshot shows the Spectronaut software interface. The top navigation bar includes Analysis, Post Analysis, Report, Databases, Settings, and About. The main window is divided into several panels:

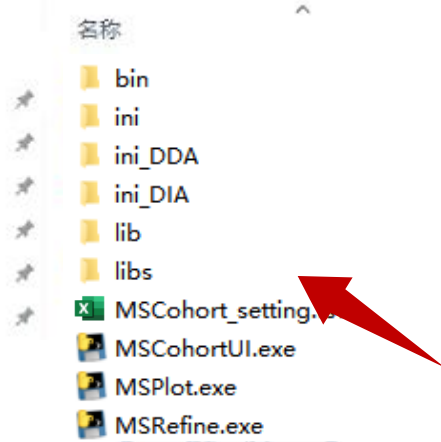
- Schemas Panel (Left):** Shows a tree view of report schemas. The **MSCohort** schema is selected under the **Normal Report** category. A red circle with the number 1 highlights this selection.
- Columns Panel (Middle-Left):** Shows a list of columns available for the report, such as Experiment, Run, Protein Group, Assay, Identification, and Elution Group. A red circle with the number 2 highlights the **Elution Group** section.
- Filters Panel (Middle-Right):** Shows filters for the report, including **Elution Group**, **Filtered in Analysis Review**, **No Decoy**, **Post-Analysis Candidate**, **Found in Protein DB**, and **Quantification Data Filtering**.
- Fragment Report [Preview] [Preview] (Right):** A table showing the preview of the MSCohort report. The table has columns: R.Run Date, R.Gradient, R.FileName, PG.ProteinGroup, PG.Qvalue, PG.Quantity, and PEP.NrC. A red circle with the number 2 highlights the **PG.ProteinGroup** column header.
- Export Report Button (Bottom Left):** A button labeled **Export Report...** is highlighted with a red circle and the number 3.

- ① Choose the MSCohort schema;
- ② Report preview: A preview of how MSCohort report will look like;
- ③ Export the matrix by clicking on "**Export Report...**" in the bottom left corner.

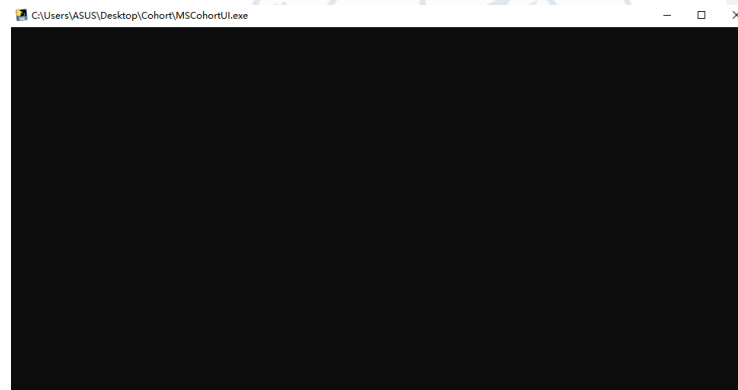
# 1. Analyzing with MSCohort

1

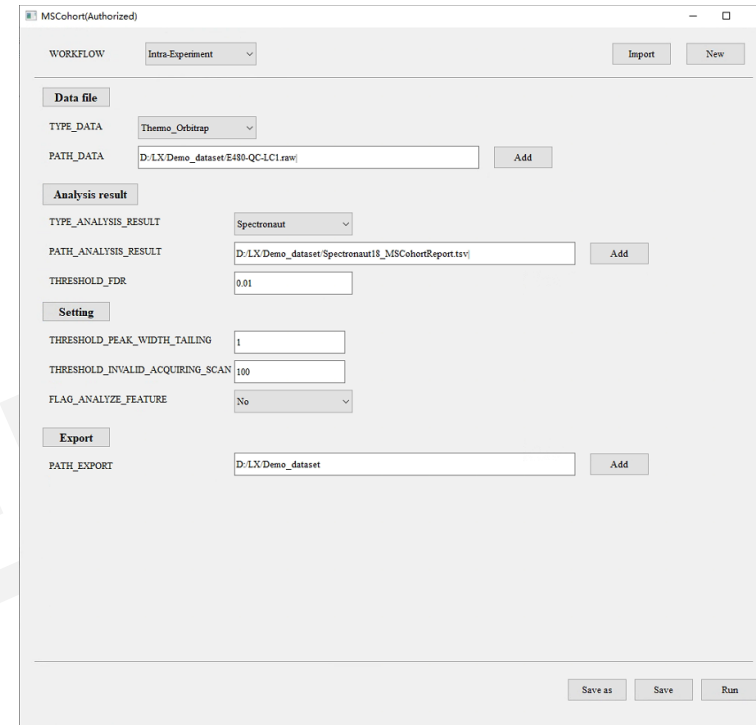
此电脑 > 新加卷 (D:) > LX > MSCohort\_v202405



2



3



- ① Double-click [MSCohortUI.exe](#);
- ② Wait for MSCohort program to load;
- ③ The MSCohortUI.exe settings screen is displayed.

# 1. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following fields and annotations:

- 1** WORKFLOW: Intra-Experiment(DIA) (dropdown menu is open showing options: Intra-Experiment(DIA), Inter-Experiment(DIA), Intra-Experiment(DDA), Inter-Experiment(DDA), Intra-Experiment(PRM), Inter-Experiment(PRM))
- 2** TYPE\_DATA: Thermo\_Orbitrap (dropdown menu is open showing options: Thermo\_Orbitrap, Bruker\_timsTOF, SCIEX\_ZenoTOF)
- 3** PATH\_DATA: D:/LX/Demo\_dataset/E480-QC (text field)
- 4** PATH\_ANALYSIS\_RESULT: D:/LX/Demo\_dataset/Spectronaut18\_MSCohortReport.tsv (text field)
- 5** PATH\_EXPORT: D:/LX/Demo\_dataset (text field)

Other visible fields include:

- THRESHOLD\_FDR: 0.01
- THRESHOLD\_PEAK\_WIDTH\_TAILING: 1
- THRESHOLD\_INVALID\_ACQUIRING\_SCAN: 100
- FLAG\_ANALYZE\_FEATURE: No

Buttons: Import, New, Add, Save as, Save, Run.

- ① Select **WORKFLOW** as **Intra-experiment**;
- ② Select **TYPE\_DATA** according to the data type ;
- ③ Click **Add** to select the raw file into the **PATH\_DATA**;
- ④ Click **Add** to select the Spectronaut customized report for MSCohort into the **PATH\_ANALYSIS\_RESULT**;
- ⑤ Click **Add** to set the **PATH\_EXPORT** for saving the results.

## Note:

Space (“ ”) cannot exist in the file directory (including PATH\_DATA, PATH\_ANALYSIS\_RESULT, and PATH\_EXPORT ), which will affect the normal running of the program.

# 1. Analyzing with MSCohort

The screenshot shows the MSCohort software interface. At the top, there's a 'WORKFLOW' dropdown set to 'Intra-Experiment(DIA)' and buttons for 'Import' and 'New'. Below this are three main sections: 'Data file', 'Analysis result', and 'Setting'. The 'Data file' section has 'TYPE\_DATA' set to 'Thermo\_Orbitrap' and 'PATH\_DATA' with a file path and an 'Add' button. The 'Analysis result' section has 'TYPE\_ANALYSIS\_RESULT' set to 'Spectronaut', 'PATH\_ANALYSIS\_RESULT' with a file path and an 'Add' button, and 'THRESHOLD\_FDR' set to '0.01'. The 'Setting' section has three items: 'THRESHOLD\_PEAK\_WIDTH\_TAILING' set to '1' (annotated with a red circle 1), 'THRESHOLD\_INVALID\_ACQUIRING\_SCAN' set to '100' (annotated with a red circle 2), and 'FLAG\_ANALYZE\_FEATURE' set to 'Yes' (annotated with a red circle 3). Below the settings is an 'Export' section with 'PATH\_EXPORT' and an 'Add' button. At the bottom right, there are three buttons: 'Save as' (annotated with a red circle 4), 'Save', and 'Run' (annotated with a red circle 5).

- ① **THRESHOLD\_PEAK\_WIDTH\_TAIL** (default setting 1, user adjustable parameter according to experimental condition), set as 1 represents that precursors with peak width more than 1 minutes are used to calculate the proportion of precursors with long eluting width;
- ② **THRESHOLD\_INVALID\_ACQUIRING** (default setting 100, user adjustable parameter according to experimental condition), set as 100 represents that the retention time length from the start to the number of identified MS2 scans per minute no more than 100 as chromatographic invalid acquiring time;
- ③ **FLAG\_ANALYZE\_FEATURE** (default setting **No**). This analysis may take a long time for timsTOF and zenoTOF data, it is recommended to set as **No** for for timsTOF and zenoTOF data;
- ④ Click **Save as** button to save the config file;
- ⑤ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

# 1. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following settings:

- WORKFLOW:** Intra-Experiment(DIA) (dropdown)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap (dropdown)
  - PATH\_DATA: D:\X\Demo\_dataset\E480-QC-LC1.raw (text field)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: Spectronaut (dropdown)
  - PATH\_ANALYSIS\_RESULT: D:\X\Demo\_dataset\Spectronaut18\_MSCohortReport.tsv (text field)
  - THRESHOLD\_FDR: 0.01 (text field)
- Setting:**
  - THRESHOLD\_PEAK\_WIDTH\_TAILING: 1 (text field)
  - THRESHOLD\_INVALID\_ACQUIRING\_SCAN: 100 (text field)
  - FLAG\_ANALYZE\_FEATURE: Yes (dropdown)
- Export:**
  - PATH\_EXPORT: D:\X\Demo\_dataset (text field)

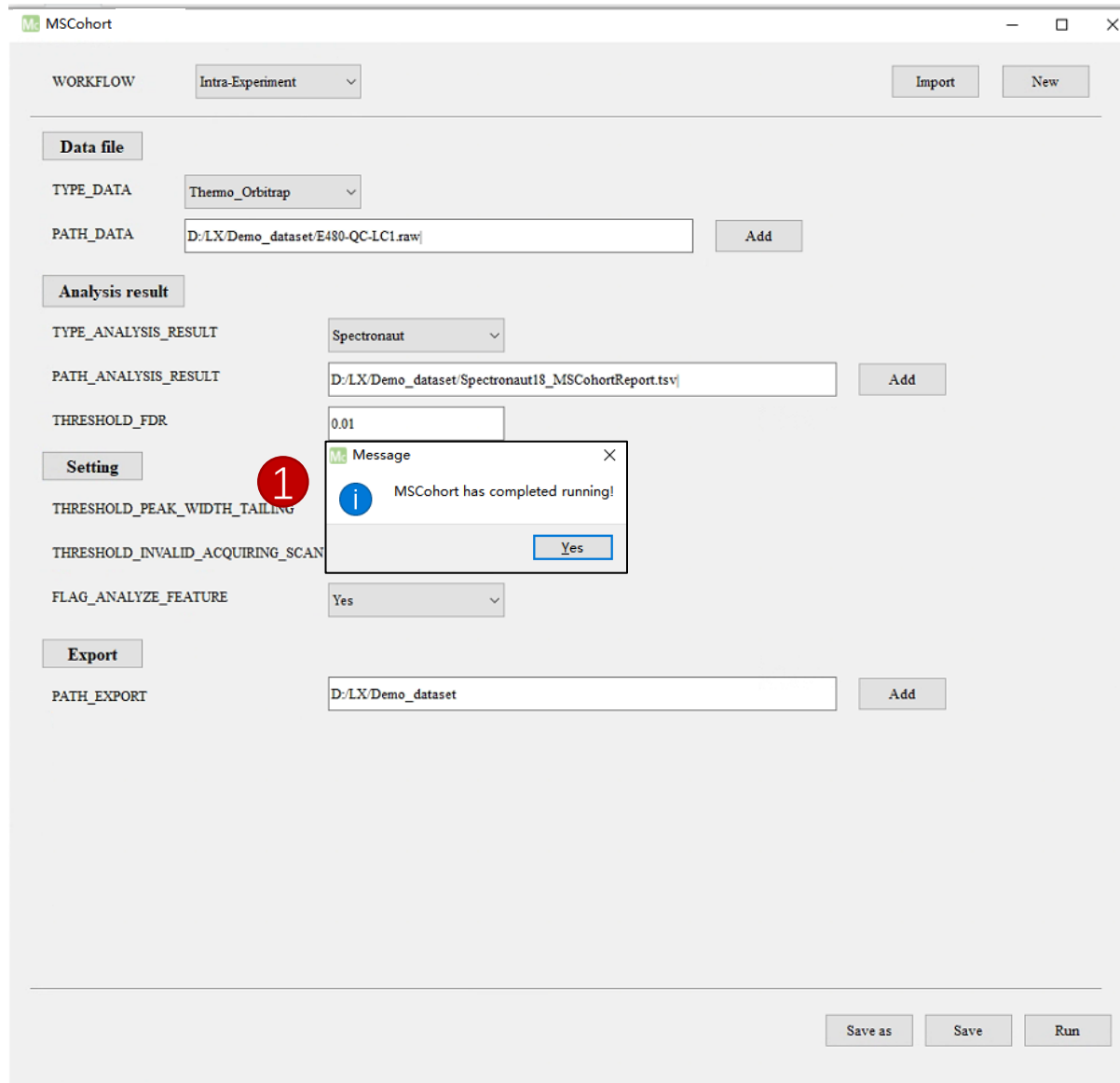
Buttons: Import, New, Add, Save as, Save, Run.

Annotations: A red circle with the number '1' is next to the PATH\_EXPORT field. A red circle with the number '2' is next to the New button.

## Note:

- ① **PATH\_EXPORT** : The output MSCohort analysis results will be stored in a new folder under the PATH\_EXPORT directory, named **MSCohort\_“filename” folder**.
- ② Do not save the same **PATH\_DATA** under a same folder, the results will be **overwritten**.
- ③ Choose **New** for a new experiments
- ③ Check the filename in **PATH\_DATA** must in the Spectronaut report in **PATH\_ANALYSIS\_RESULT**.

## 2. Analyzing with MSCohort



- ① When the program popup window shows “MSCohort has completed running!” Indicates that the current process is complete. Please open the result file under the PATH\_EXPORT directory and view it.



## 2. MSCohort Results

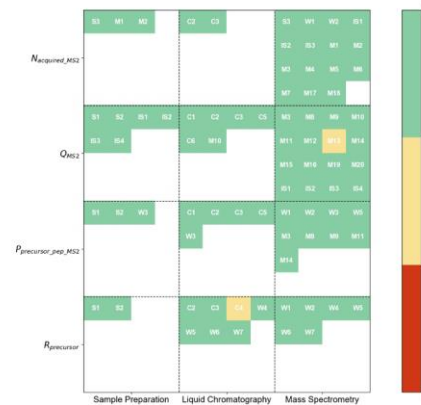
1

> LX > Demo\_dataset > MSCohort\_E480-QC-LC1

名称	修改日期
picture	2024/
Analysis_Report.html	2024/
INFO_Summary.txt	2024/
INFO_Chromatography.txt	2024/
INFO_Cycle_MS1.txt	2024/
INFO_Cycle_MS2.txt	2024/
INFO_Feature.txt	2024/
INFO_ID.txt	2024/
INFO_Mass_Deviation.txt	2024/
INFO_MS1.txt	2024/
INFO_MS1_PEAKS.txt	2024/
INFO_MS2.txt	2024/
INFO_MS2_PEAKS.txt	2024/
INFO_peptides.txt	2024/
INFO_Scans.txt	2024/

2

1.3 Metric-Score Analysis



- ① The MSCohort results ;
- ② Double clicking [Analysis\\_Report.html](#), the report will be preformed in the browser.

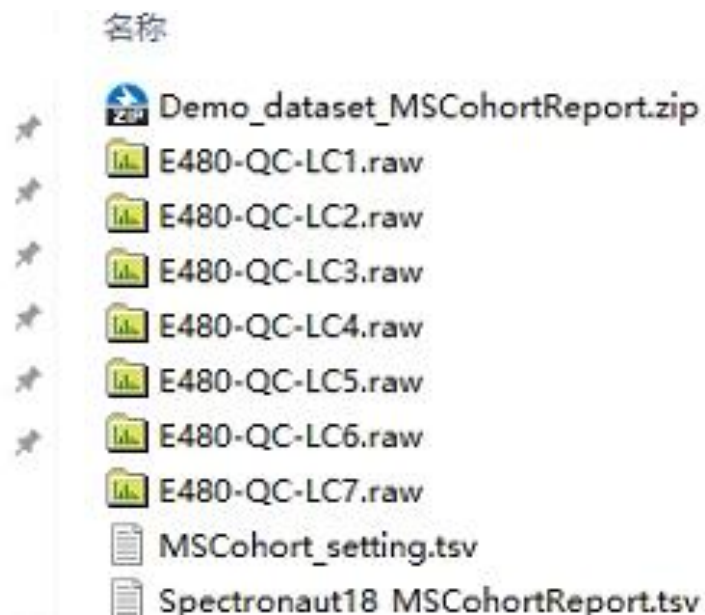


# **MSCohort manual for DIA inter- experiment analysis**



# Demo dataset

- DemoDatasetForDDA
- DemoDatasetForDIA**
- DemoDatasetForPRM



Taking the 7 E480-QC raw files as an example to demonstrate the workflow of inter-experiment analysis.

Inter-experiment analysis enables the systematic quality evaluation and low-quality experiments detection for multiple DIA experiments.

# 1. Preparation for MSCohort\_setting.tsv file

名称

bin

ini

lib

MSCohort\_cfg.txt

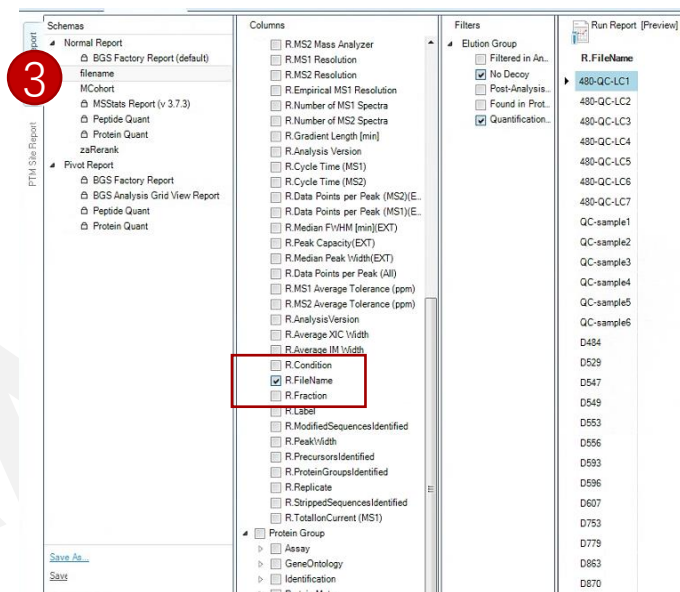
MSCohort\_setting.tsv

MSCohortUI.exe

2

Group Name	Raw Name	Experiment	Threshold
E480-QC	E480-QC-LC1	E480-QC1	10
E480-QC	E480-QC-LC2	E480-QC2	10
E480-QC	E480-QC-LC3	E480-QC3	10
E480-QC	E480-QC-LC4	E480-QC4	10
E480-QC	E480-QC-LC5	E480-QC5	10
E480-QC	E480-QC-LC6	E480-QC6	10
E480-QC	E480-QC-LC7	E480-QC7	10

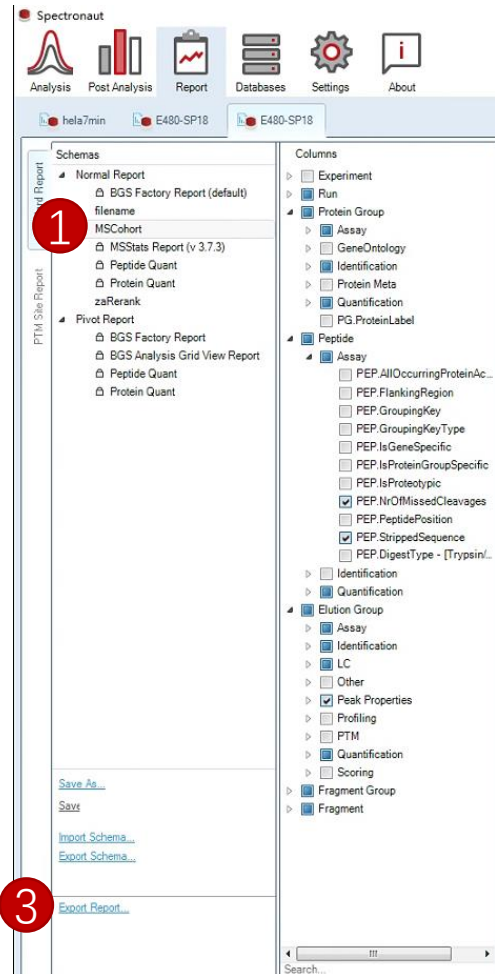
3



- ① Open [MSCohort\\_setting.tsv](#) with Excel;
- ② Fill the columns as the example file showed:
  - The first column is the Group Name.
  - The second column is the Raw Name, which is the same as [R.FileName](#) reported from Spectronaut.
  - The third column is the Raw Name for short, which is convenient for display in large MSCohort report. This column could also be the same as Raw Name column.
  - The fourth column is the intensity threshold, protein/peptide intensity that less than this threshold would be replaced as NaN, and would not be used for subsequent Pearson correlation analysis.
- ③ For large-scale cohorts, users can get the Raw Name list from Spectronaut by exporting only [R.FileName](#) column.

**All subsequent analyses are based only on the information in the Raw Name column provided by MSCohort\_setting.tsv**

# 1. Preparation of MSCohort report from Spectronaut



The columns of information that MSCohort needs to read

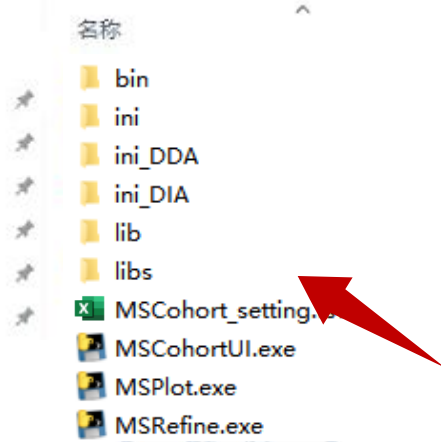
Run	R.Run Date R.Gradient Length [min] R.FileName
Protein Group	PG.ProteinGroups PG.Qvalue PG.Quantity
Peptide	PEP.NrOfMissedCleavages PEP.StrippedSequence PEP.Quantity
Elution Group	EG.IsDecoy EG.ModifiedPeptide EG.PrecursorId EG.Qvalue EG.ApexRT EG.DatapointsPerPeak EG.DatapointsPerPeak (MS1) EG.DeltaRT EG.EndRT EG.FWHM EG.PeakWidth EG.StartRT EG.SignalToNoise EG.TotalQuantity (Settings)
Fragment Group	FG.Charge FG.PrecMz FG.PrecWindow FG.PrecWindowNumber FG.CalibratedMassAccuracy (PPM) FG.RawMassAccuracy (PPM)
Fragment	F.RawMassAccuracy (PPM) F.CalibratedMassAccuracy (PPM)

- ① Choose the MSCohort report schema;
- ② Report preview: A preview of how MSCohort report will look like;
- ③ Export the matrix by clicking on "[Export Report...](#)" in the bottom left corner.

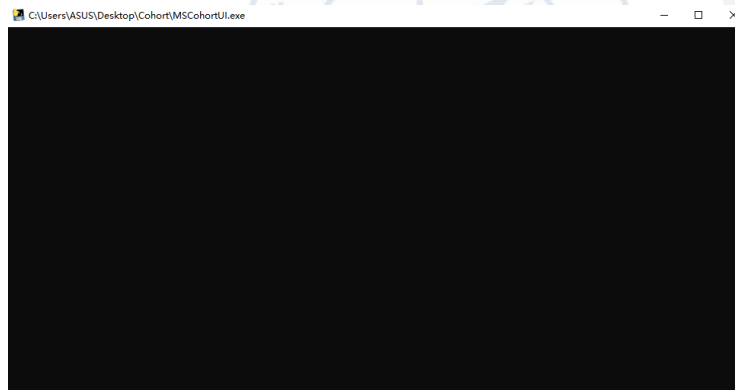
## 2. Analyzing with MSCohort

1

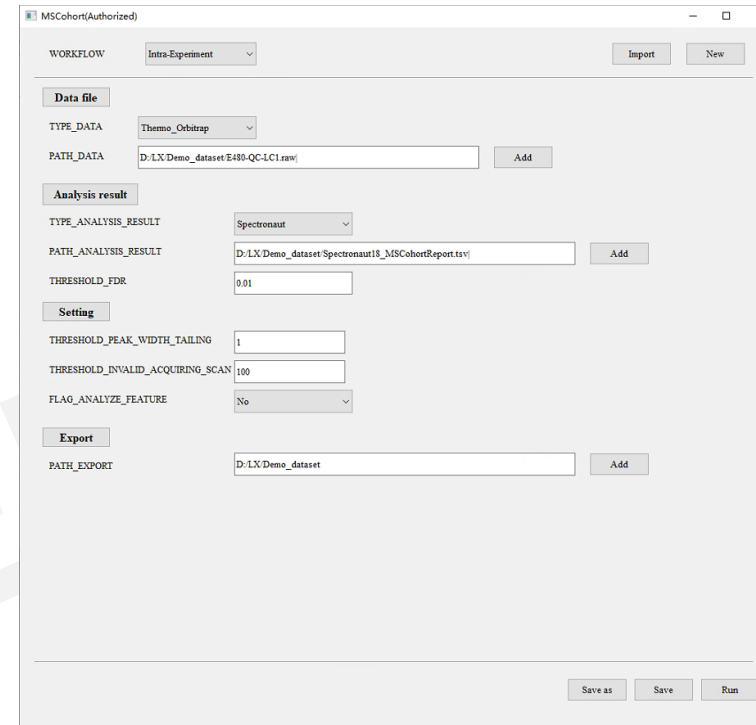
此电脑 > 新加卷 (D:) > LX > MSCohort\_v202405



2



3



- ① Double-click [MSCohortUI.exe](#);
- ② Wait for MSCohort program to load;
- ③ The MSCohortUI.exe settings screen is displayed.

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following components and annotations:

- 1** WORKFLOW: A dropdown menu showing "Inter-Experiment(DIA)" and a sub-menu with options: "Intra-Experiment(DIA)", "Inter-Experiment(DIA)", "Intra-Experiment(DDA)", "Inter-Experiment(DDA)", and "Inter-Experiment(PRM)".
- 2** TYPE\_DATA: A dropdown menu showing "Thermo\_Orbitrap".
- 3** PATH\_DATA: An empty text input field with an "Add" button.
- 4** PATH\_ANALYSIS\_RESULT: A text input field containing "D:/LX/Demo\_dataset/Spectronaut18\_MSCohortReport.tsv" with an "Add" button.
- 5** PATH\_EXPERIMENT\_RESULT: A text input field containing "D:/LX/Demo\_dataset/MSCohort\_setting.tsv" with an "Add" button.
- Other settings include: TYPE\_ANALYSIS\_RESULT (Spectronaut), THRESHOLD\_FDR (0.01), TYPE\_NORMALIZATION (DirectLFQ), FLAG\_OUTLIERS (2-SD), FLAG\_SHOW\_ORDER (group series), THRESHOLD\_PEAK\_WIDTH\_TAILING (1), THRESHOLD\_INVALID\_ACQUIRING\_SCAN (100), FLAG\_ANALYZE\_FEATURE (No), and PATH\_EXPORT (D:/LX/Demo\_dataset).
- Buttons at the bottom: "Save as", "Save", and "Run".

- ① Set **WORKFLOW** as **Inter-experiment**;
- ② Select **TYPE\_DATA** according to the data type ;
- ③ The **PATH\_DATA** could be empty. MSCohort support the inter-experimental analysis mainly based on Spectronaut result, eliminating the need to submit raw files, which will obtain the result in a relatively **short time**. It is recommended for large cohort analysis.  
  
In addition, users could choose to add the raw data. MSCohort would provide comprehensive analysis reports not only for inter-experiment, but also for intra-experiment analysis. This may **take a long time**.
- ④ Click **Add** to select the Spectronaut customized report for MSCohort into the **PATH\_ANALYSIS\_RESULT**;
- ⑤ Click **Add** to select the **MSCohort\_setting.tsv** into the **PATH\_EXPERIMENT\_RESULT**;

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration options:

- WORKFLOW:** Inter-Experiment(DIA) (dropdown)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap (dropdown)
  - PATH\_DATA: (empty text field) [Add]
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: Spectronaut (dropdown)
  - PATH\_ANALYSIS\_RESULT: D:/LX/Demo\_dataset/Spectronaut18\_MSCohortReport.tsv [Add]
  - THRESHOLD\_FDR: 0.01 (text field)
- Setting:**
  - PATH\_EXPERIMENT\_RESULT: D:/LX/Demo\_dataset/MSCohort\_setting.tsv [Add]
  - TYPE\_NORMALIZATION: DirectLFQ (dropdown)
  - FLAG\_OUTLIERS: 2-SD (dropdown)
  - FLAG\_SHOW\_ORDER: group series (dropdown)
  - THRESHOLD\_PEAK\_WIDTH\_TAILING: 1 (text field)
  - THRESHOLD\_INVALID\_ACQUIRING\_SCAN: 100 (text field)
  - FLAG\_ANALYZE\_FEATURE: No (dropdown)
- Export:**
  - PATH\_EXPORT: D:/LX/Demo\_dataset [Add]

At the bottom, there are buttons for "Save as", "Save", and "Run".

- ① Choose normalization strategies in **TYPE\_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);
- ② Choose **FLAG\_OUTLIERS threshold** according to the experiment condition (default is 2\*SD, users could adjust the threshold (Notes for modifying the scoring criteria));
- ③ Choose **FLAG\_SHOW\_ORDER**. **group series** represents the experiment order showed in MSCohort report is the same as the **MSCohort\_setting.tsv**; **time series** represents the experiment order showed in MSCohort report is sorted by run date.
- ④ Set **THRESHOLD\_PEAK\_WIDTH\_TAIL**, **THRESHOLD INVALID\_ACQUIRING**, **FLAG\_ANALYZE\_FEATURE** as intra-experiment analysis?;



## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface. At the top, there is a 'WORKFLOW' dropdown set to 'Inter-Experiment(DIA)' and buttons for 'Import' and 'New'. Below this are three main sections: 'Data file', 'Analysis result', and 'Setting'. The 'Data file' section has 'TYPE\_DATA' set to 'Thermo\_Orbitrap' and an empty 'PATH\_DATA' field with an 'Add' button. The 'Analysis result' section has 'TYPE\_ANALYSIS\_RESULT' set to 'Spectronaut', 'PATH\_ANALYSIS\_RESULT' set to 'D:/LX/Demo\_dataset/Spectronaut18\_MSCohortReport.tsv', and 'THRESHOLD\_FDR' set to '0.01'. The 'Setting' section has 'PATH\_EXPERIMENT\_RESULT' set to 'D:/LX/Demo\_dataset/MSCohort\_setting.tsv', 'TYPE\_NORMALIZATION' set to 'DirectLFQ', 'FLAG\_OUTLIERS' set to '2-SD', 'FLAG\_SHOW\_ORDER' set to 'group series', 'THRESHOLD\_PEAK\_WIDTH\_TAILING' set to '1', 'THRESHOLD\_INVALID\_ACQUIRING\_SCAN' set to '100', and 'FLAG\_ANALYZE\_FEATURE' set to 'No'. At the bottom, there is an 'Export' section with 'PATH\_EXPORT' set to 'D:/LX/Demo\_dataset'. Three red circles with numbers 1, 2, and 3 are overlaid on the interface. Circle 1 is next to the 'PATH\_EXPORT' field. Circle 2 is above the 'Save as' button. Circle 3 is above the 'Run' button.

WORKFLOW: Inter-Experiment(DIA) [Import] [New]

**Data file**

TYPE\_DATA: Thermo\_Orbitrap

PATH\_DATA: [Add]

**Analysis result**

TYPE\_ANALYSIS\_RESULT: Spectronaut

PATH\_ANALYSIS\_RESULT: D:/LX/Demo\_dataset/Spectronaut18\_MSCohortReport.tsv [Add]

THRESHOLD\_FDR: 0.01

**Setting**

PATH\_EXPERIMENT\_RESULT: D:/LX/Demo\_dataset/MSCohort\_setting.tsv [Add]

TYPE\_NORMALIZATION: DirectLFQ

FLAG\_OUTLIERS: 2-SD

FLAG\_SHOW\_ORDER: group series

THRESHOLD\_PEAK\_WIDTH\_TAILING: 1

THRESHOLD\_INVALID\_ACQUIRING\_SCAN: 100

FLAG\_ANALYZE\_FEATURE: No

**Export**

PATH\_EXPORT: D:/LX/Demo\_dataset [Add]

[1] [2] [3]

[Save as] [Save] [Run]

- ① Set the **PATH\_EXPORT** for saving the results. The output MSCohort analysis results will be stored in a new folder under the PATH\_EXPORT directory, named MSCohort\_“year+month+day” folder.

**Do not save the different experiments results under a same folder, the results will be **overwritten**.**

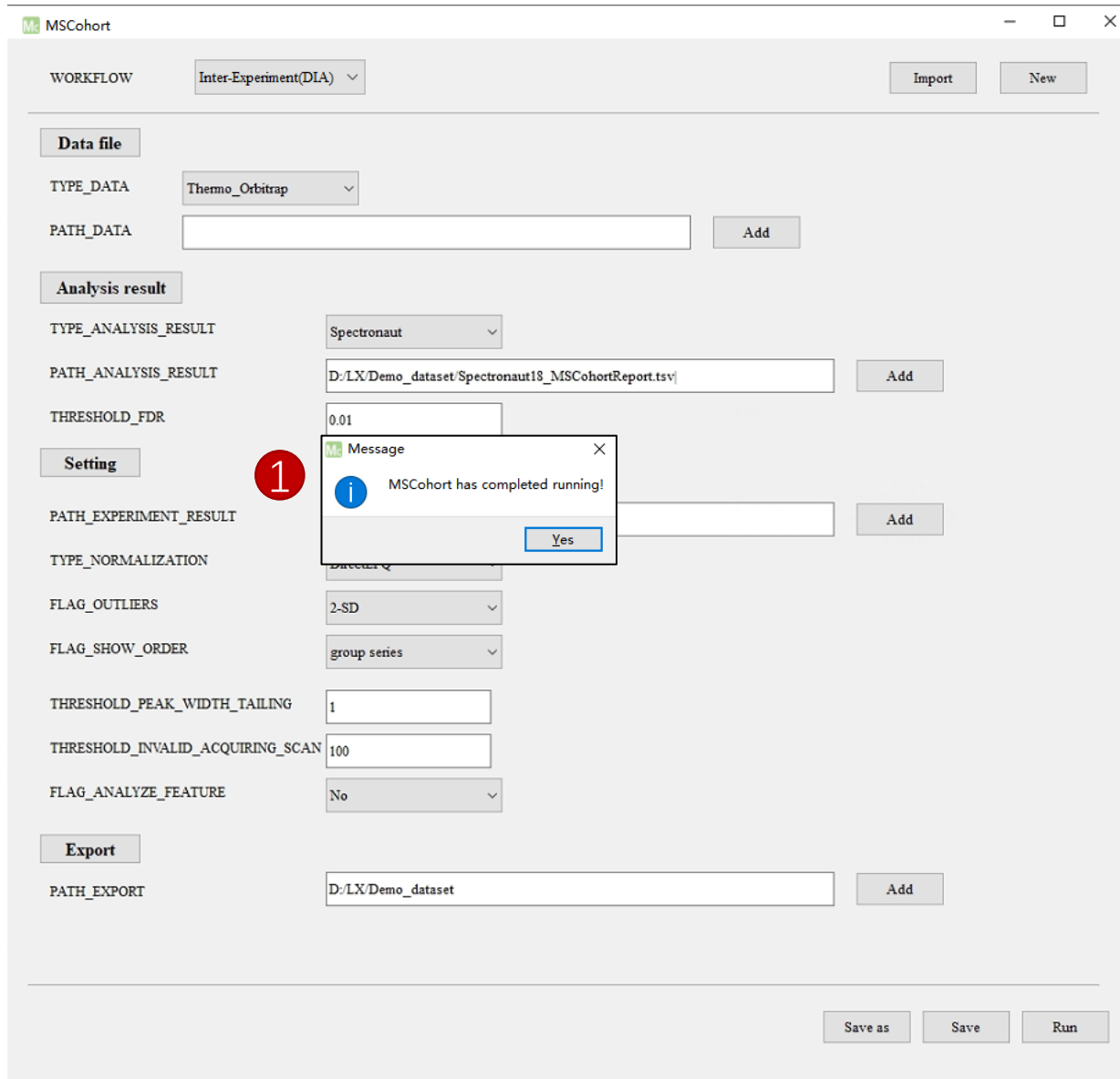
- ② Click **Save as** button to save the config file;
- ③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

### Note:

Space (“ ”) cannot exist in the file directory (including PATH\_DATA, PATH\_ANALYSIS\_RESULT, PATH\_EXPERIMENT\_RESULT, and PATH\_EXPORT ), which will affect the normal running of the program.



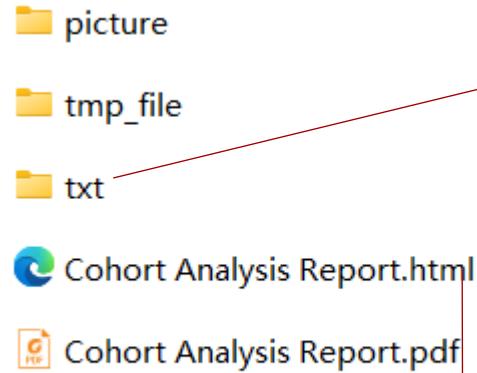
## 2. Analyzing with MSCohort



- ① When the program popup window shows “MSCohort has completed running!” Indicates that the current process is complete. Please open the result file under the PATH\_EXPORT directory and view it.

# 3. MSCohort Results

1



3

- INFO\_Pro0\_MissingValue\_Count.txt
- INFO\_Pro1\_Intensity.txt
- INFO\_Pro2\_Origin\_Intensity.txt
- INFO\_Pro3\_Coefficient\_Var.txt
- INFO0\_Experiment\_Outlier\_Score.txt
- INFO0\_Identification\_Count.txt
- INFO0\_Inter\_Experiment\_Scores.txt
- INFO0\_Inter\_Experiment\_values.txt
- INFO0\_Intra\_Experiment\_Scores.txt
- INFO0\_Intra\_Experiment\_values.txt
- INFO1\_Pro0\_MissingValue\_Count.txt
- INFO2\_Pro1\_Intensity.txt
- INFO3\_Pro2\_Origin\_Intensity.txt
- INFO4\_Pro3\_Coefficient\_Var.txt
- INFO5\_Pre\_RetentionTime.txt

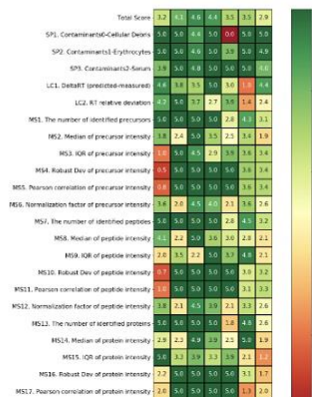
- ① The MSCohort results ;
- ② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.
- ③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

2

## 1. Overview of Dataset

### 1.1 Score of Inter-experiment Metrics

Inter-experiment metrics are computed across multiple experiments to assess the quality for the whole cohort quality data.

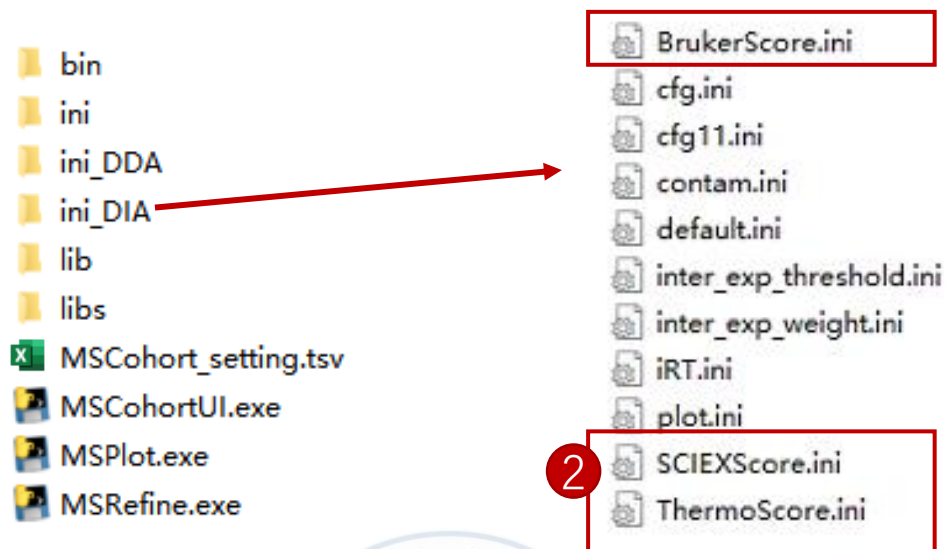


# **Notes for modifying the scoring criteria**



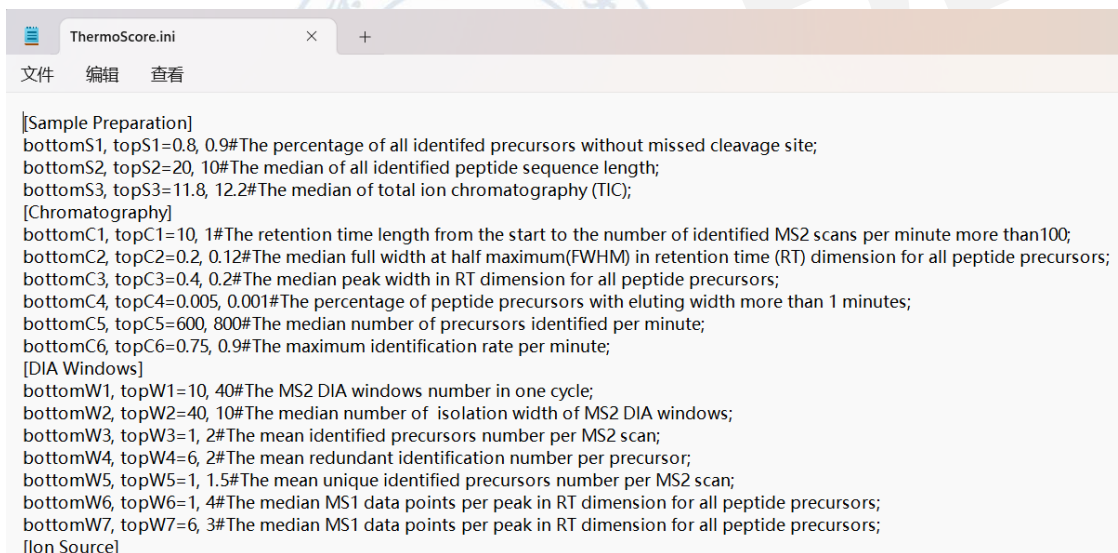
# 1. Modifying the scoring standards for intra-experiment analysis

①



- ① Open the **ini\_DIA** folder, there will be three parameter files related to intra-experiment scoring: **ThermoScore.ini**, **BrukerScore.ini**, **SCIEXScore.ini**;
- ② Click the file to modify the metrics scoring standards. If you do not modify it, it will be the default value.
- ③ Users can adjust the scoring standards for each metric according the actual situation. Where “bottom” represents the scoring standards for 1 point, and “top” represents the scoring standards for 5 points.

Then, save and close the file, and run the software, it will score according to the standard you set.



## 2. Modifying the scoring standards for inter-experiment analysis

1

BrokerScore.ini  
cfg.ini  
cfg11.ini  
contam.ini  
default.ini  
inter\_exp\_threshold.ini  
inter\_exp\_weight.ini  
iRT.ini  
plot.ini  
SCIEXScore.ini  
ThermoScore.ini

2

contam.ini

文件 编辑 查看

```
#Cellular Debris  
P61160  
P02545  
Q9UHD8  
P26885  
P21281  
P10809  
Q99714  
P61421  
P17568  
P11310  
Q6YN16  
P49411  
P05455
```

1

Open the [ini\\_DIA](#) folder, there will be five parameter files related to inter-experiment scoring: [contam.ini](#), [inter\\_exp\\_threshold.ini](#), [inter\\_exp\\_weight.ini](#), [iRT.ini](#), [plot.ini](#);

2

Open the [contam.ini](#), users can modify the list of contaminating proteins, # begins with contaminants category, enter the list of contaminating proteins under this category below.

3

Open the [inter\\_exp\\_threshold.ini](#), users can modify the scoring threshold. For each metric, we initially defined potentially outlier experiments as those with a value more than two standard deviations ( $2 \times SD$ ) from the median. If you do not modify it, it will be the default value as 2. You can change the value and save the file, it will score according to the standard you set.

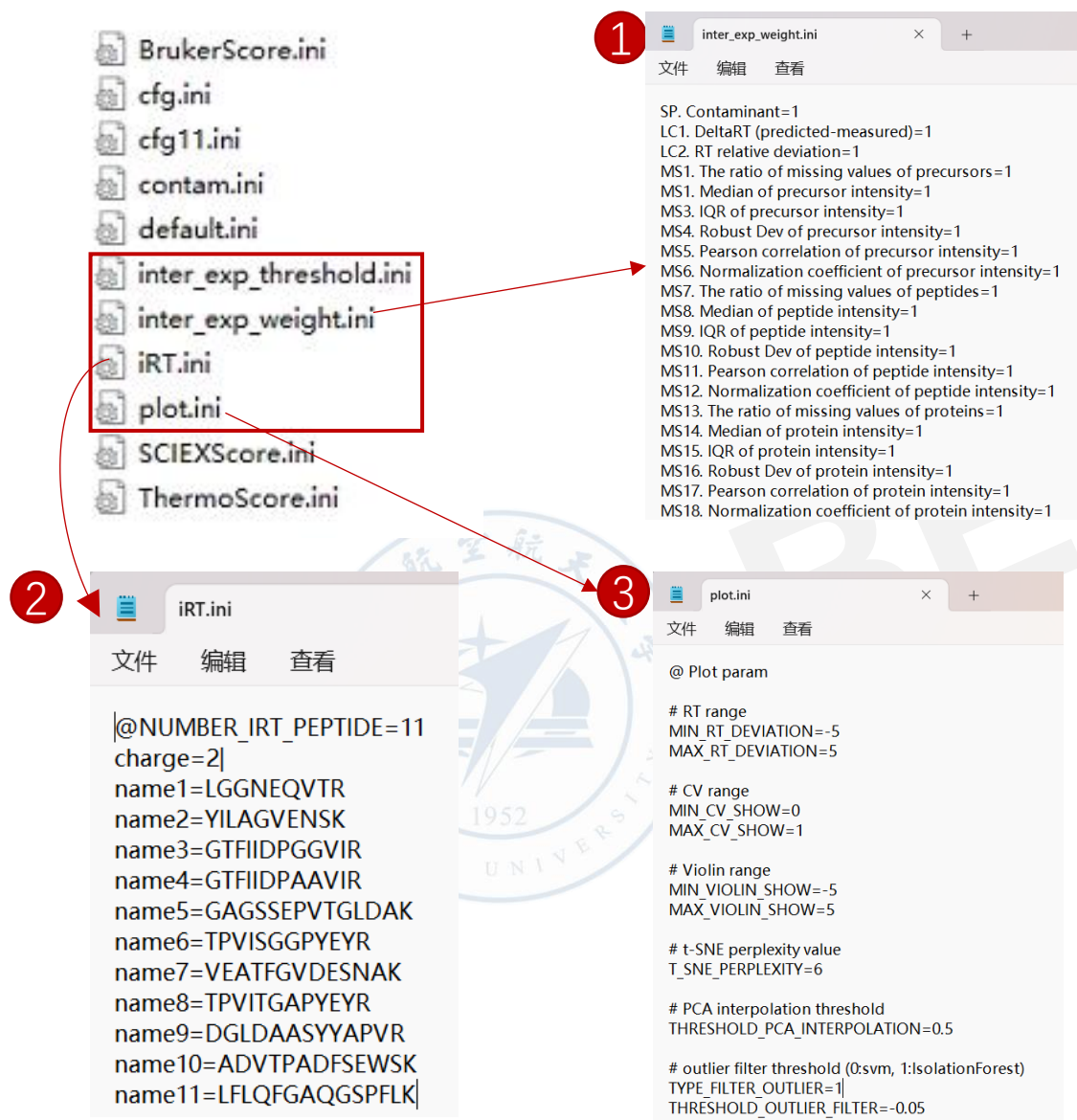
3

inter\_exp\_threshold.ini

文件 编辑 查看

```
SP. Contaminant=2  
LC1. DeltaRT (predicted-measured)=2  
LC2. RT relative deviation=2  
MS1. The ratio of missing values of precursors=2  
MS2. Median of precursor intensity=2  
MS3. IQR of precursor intensity=2  
MS4. Robust Dev of precursor intensity=2  
MS5. Pearson correlation of precursor intensity=2  
MS6. Normalization coefficient of precursor intensity=2  
MS7. The ratio of missing values of peptides=2  
MS8. Median of peptide intensity=2  
MS9. IQR of peptide intensity=2  
MS10. Robust Dev of peptide intensity=2  
MS11. Pearson correlation of peptide intensity=2  
MS12. Normalization coefficient of peptide intensity=2  
MS13. The ratio of missing values of proteins=2  
MS14. Median of protein intensity=2  
MS15. IQR of protein intensity=2  
MS16. Robust Dev of protein intensity=2  
MS17. Pearson correlation of protein intensity=2  
MS18. Normalization coefficient of protein intensity=2
```

### 3. Modifying the scoring criteria for inter-experiment analysis



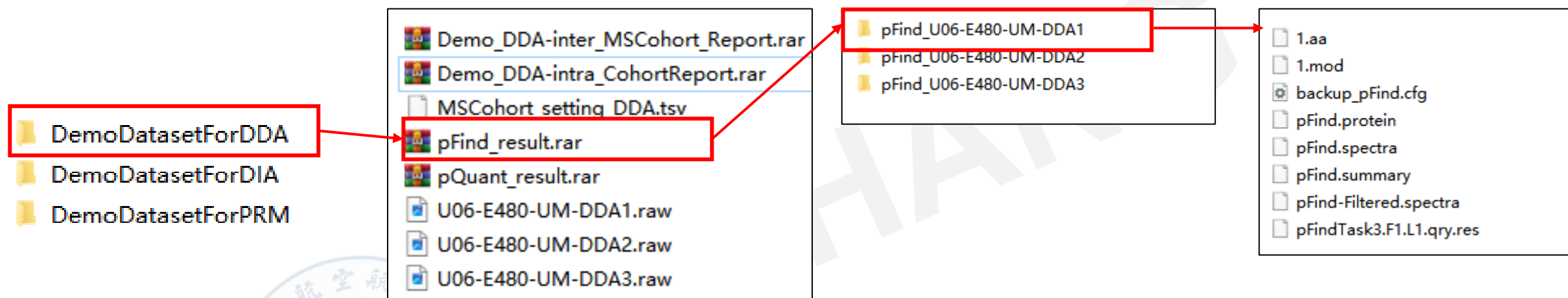
- ① Open the [inter\\_exp\\_weight.ini](#), users can modify the scoring weight for each metric. We initially defined the weight as 1 for each metric. If you do modify it, closing and saving the file, it will score according to the standard you set.
- ② Open the [iRT.ini](#), users can modify the list of iRT peptide sequence. The default iRT peptide sequence is the 11 non-naturally occurring synthetic peptides from the iRT kit (Biognosys). Users can modify the sequence according to experimental conditions.
- ③ Open the [plot.ini](#), users can modify the parameters related to plot. RT range is y axis range showed in RT deviation analysis plot; CV range is the y axis range showed in CV plot; Violin range is the y axis range showed in Intensity ratio distribution plot; TYPE\_FILTER\_OUTLIERS is the approach to detect outliers (0: one-class SVM; 1: isolation forest); THRESHOLD\_OUTLIER\_FILTER is threshold to flag as outliers (In the data we tested, the SVM model threshold was set to -1; the isolation forest was set to -0.05).

# **MSCohort manual for DDA intra- experiment analysis**





# Demo dataset



Taking the U06-E480-UM-DDA1.raw file as an example to demonstrate the workflow of DDA intra-experiment analysis.

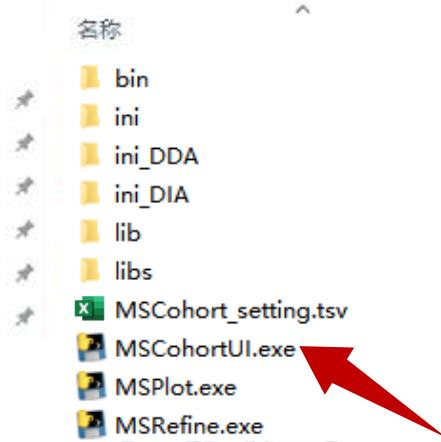
Intra-experiment analysis enables the systematic evaluation and optimization of individual DDA experiments.



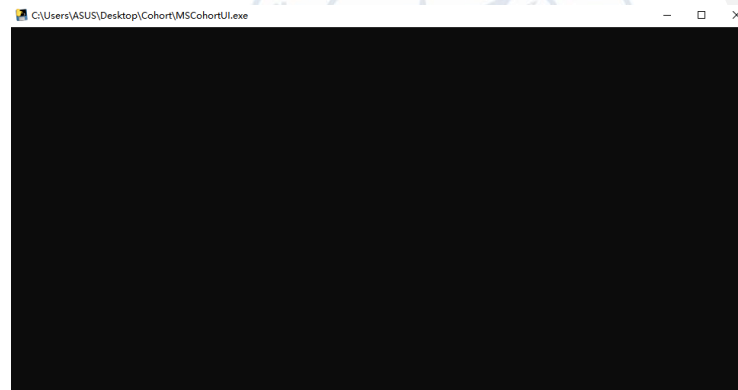
# 1. Analyzing with MSCohort

1

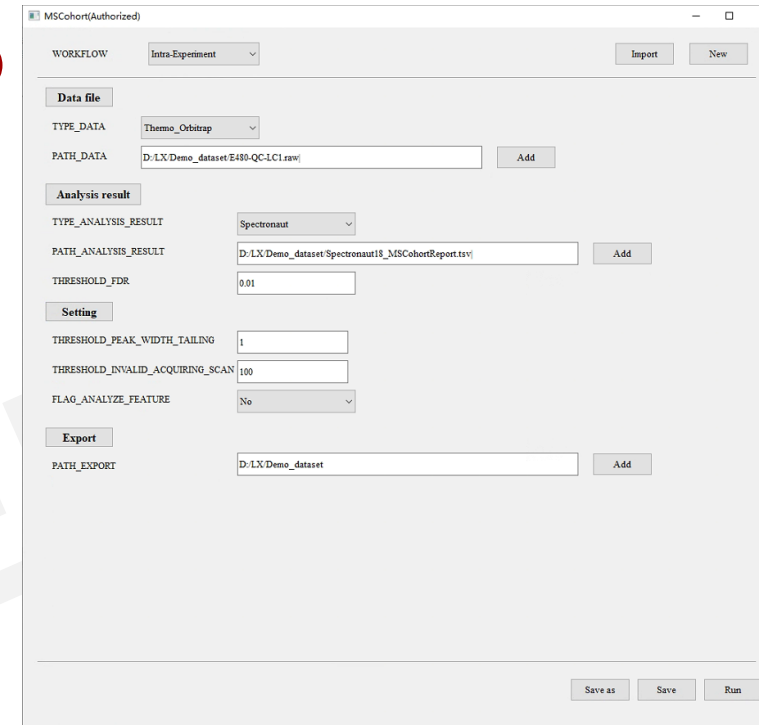
此电脑 > 新加卷 (D:) > LX > MSCohort\_v202405



2



3



- ① Double-click [MSCohortUI.exe](#);
- ② Wait for MSCohort program to load;
- ③ The MSCohortUI.exe settings screen is displayed.

# 1. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following components and annotations:

- 1** **WORKFLOW**: A dropdown menu showing options: Intra-Experiment(DDA), Inter-Experiment(DIA), Intra-Experiment(DIA), Inter-Experiment(DDA), Intra-Experiment(DDA) (highlighted), and Inter-Experiment(PRM).
- 2** **Data file**: A button to import data.
- 3** **TYPE\_DATA**: A dropdown menu showing options: Thermo\_Orbitrap (highlighted), Bruker\_timsTOF, and SCIEX\_ZenoTOF.
- 4** **PATH\_DATA**: A text field containing the path "sers/ASUS/Desktop/paper\_test/test/DemoDatasetForDDA/U06-E480-UM-DDA1.raw" and an **Add** button.
- 5** **Analysis result**: A section containing:
  - TYPE\_ANALYSIS\_RESULT**: A dropdown menu showing options: pFind (highlighted), Spectronaut, pQuant, SpectroDive, and Skyline.
  - PATH\_ANALYSIS\_RESULT**: A text field containing the path "DA\_test\_data\_new/pFind\_result/pFind\_U06-E480-UM-DDA1/pFind-Filtered.spectra" and an **Add** button.
  - THRESHOLD\_FDR**: A text field containing the value "0.01".
- Setting**: A section containing:
  - THRESHOLD\_PEAK\_WIDTH\_TAILING**: A text field containing the value "1".
  - THRESHOLD\_INVALID\_ACQUIRING\_SCAN**: A text field containing the value "100".
  - FLAG\_ANALYZE\_FEATURE**: A dropdown menu showing "Yes".
  - FLAG\_ANALYZE\_PIF**: A dropdown menu showing "Yes".
  - FLAG\_ANALYZE\_TAG**: A dropdown menu showing "Yes".
- Export**: A section containing:
  - PATH\_EXPORT**: A text field containing the path "C:/Users/ASUS/Desktop/paper\_test/test/DemoDatasetForDDA" and an **Add** button.

At the bottom of the interface are buttons for **Save as**, **Save**, and **Run**.

- ① Select **WORKFLOW** as **Intra-experiment**;
- ② Select **TYPE\_DATA** according to the data type ;
- ③ Click **Add** to select the raw file into the **PATH\_DATA**;
- ④ Click **Add** to select the pFind report (pFind-Filtered.spectra) for MSCohort into the **PATH\_ANALYSIS\_RESULT**;
- ⑤ Click **Add** to set the **PATH\_EXPORT** for saving the results.

## Note:

Space (“ ”) cannot exist in the file directory (including **PATH\_DATA**, **PATH\_ANALYSIS\_RESULT**, and **PATH\_EXPORT** ), which will affect the normal running of the program.

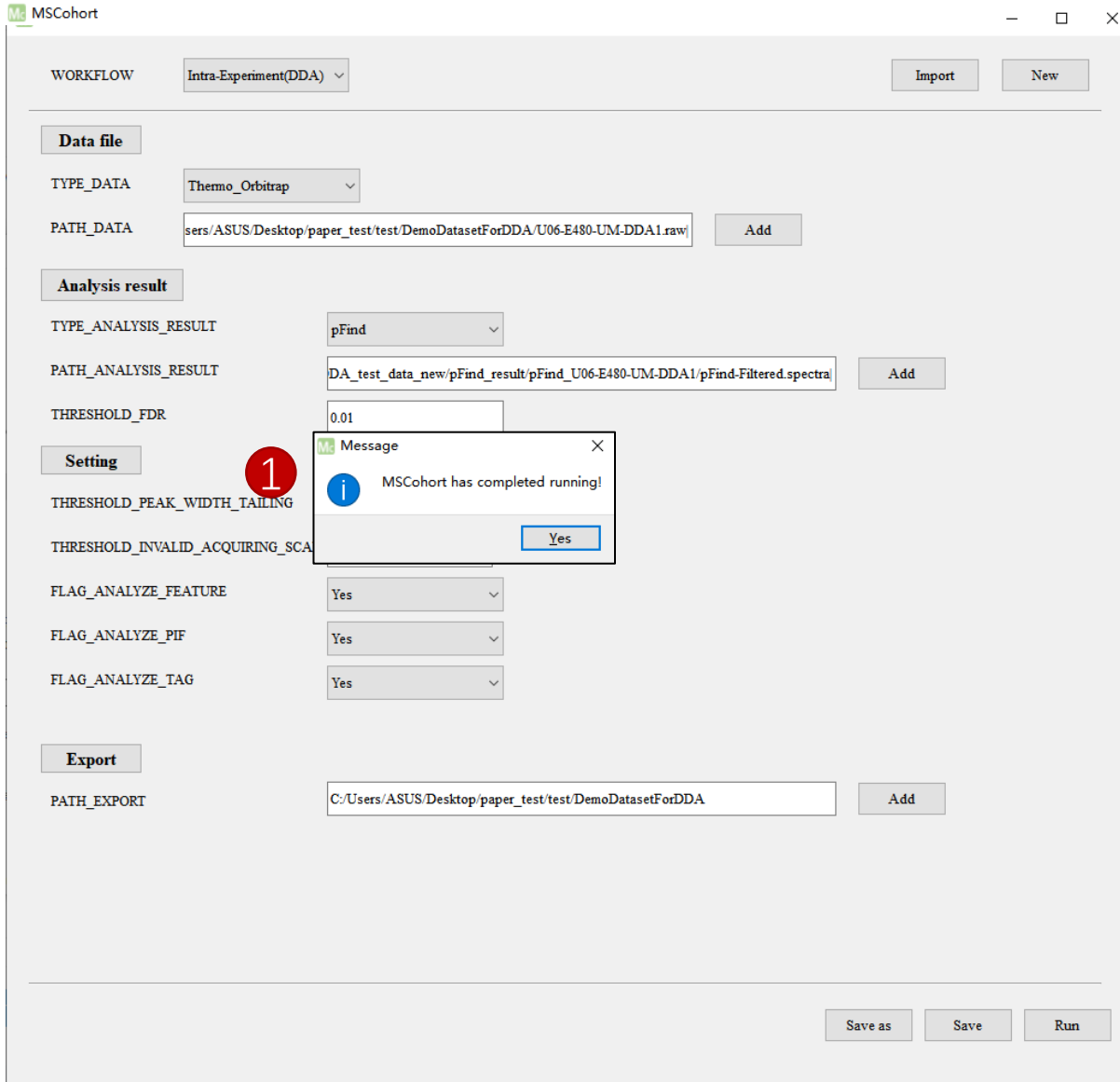
# 1. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following sections and annotations:

- WORKFLOW:** Intra-Experiment(DDA) (dropdown), Import, New
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap (dropdown)
  - PATH\_DATA: sers/ASUS/Desktop/paper\_test/test/DemoDatasetForDDA/U06-E480-UM-DDA1.raw (text input), Add
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: pFind (dropdown)
  - PATH\_ANALYSIS\_RESULT: DA\_test\_data\_new/pFind\_result/pFind\_U06-E480-UM-DDA1/pFind-Filtered.spectra (text input), Add
  - THRESHOLD\_FDR: 0.01 (text input)
- Setting:**
  - 1** THRESHOLD\_PEAK\_WIDTH\_TAILING: 1 (text input)
  - 2** THRESHOLD\_INVALID\_ACQUIRING\_SCAN: 100 (text input)
  - 3** FLAG\_ANALYZE\_FEATURE: Yes (dropdown)
  - FLAG\_ANALYZE\_PIF: Yes (dropdown)
  - FLAG\_ANALYZE\_TAG: Yes (dropdown)
- Export:**
  - PATH\_EXPORT: C:/Users/ASUS/Desktop/paper\_test/test/DemoDatasetForDDA (text input), Add
- Bottom:** Save as, Save, Run (annotated with **4** and **5** respectively)

- ① THRESHOLD\_PEAK\_WIDTH\_TAIL** (default setting 1, user adjustable parameter according to experimental condition), set as 1 represents that precursors with peak width more than 1 minutes are used to calculate the proportion of precursors with long eluting width;
- ② THRESHOLD\_INVALID\_ACQUIRING** (default setting 100, user adjustable parameter according to experimental condition), set as 100 represents that the retention time length from the start to the number of identified MS2 scans per minute no more than 100 as chromatographic invalid acquiring time;
- ③ FLAG\_ANALYZE\_FEATURE** (default setting No). This analysis may take a long time for timsTOF and zenoTOF data, it is recommended to set as No for for timsTOF and zenoTOF data;
- ④** Click **Save as** button to save the config file;
- ⑤** Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

## 2. Analyzing with MSCohort



The screenshot displays the MSCohort software interface. A red circle with the number '1' is positioned over a 'Message' popup window. The popup window contains the text 'MSCohort has completed running!' and a 'Yes' button. The background interface includes sections for 'Workflow' (Intra-Experiment(DDA)), 'Data file' (Thermo\_Orbitrap, PATH\_DATA: sers/ASUS/Desktop/paper\_test/test/DemoDatasetForDDA/U06-E480-UM-DDA1.raw), 'Analysis result' (pFind, PATH\_ANALYSIS\_RESULT: DA\_test\_data\_new/pFind\_result/pFind\_U06-E480-UM-DDA1/pFind-Filtered.spectra, THRESHOLD\_FDR: 0.01), 'Setting' (THRESHOLD\_PEAK\_WIDTH\_TAILING, THRESHOLD\_INVALID\_ACQUIRING\_SCA, FLAG\_ANALYZE\_FEATURE, FLAG\_ANALYZE\_PIF, FLAG\_ANALYZE\_TAG), and 'Export' (PATH\_EXPORT: C:/Users/ASUS/Desktop/paper\_test/test/DemoDatasetForDDA). Buttons for 'Import', 'New', 'Add', 'Save as', 'Save', and 'Run' are visible.

- ① When the program popup window shows “MSCohort has completed running!” Indicates that the current process is complete. Please open the result file corresponding to Cohort and view it.

## 2. MSCohort Results

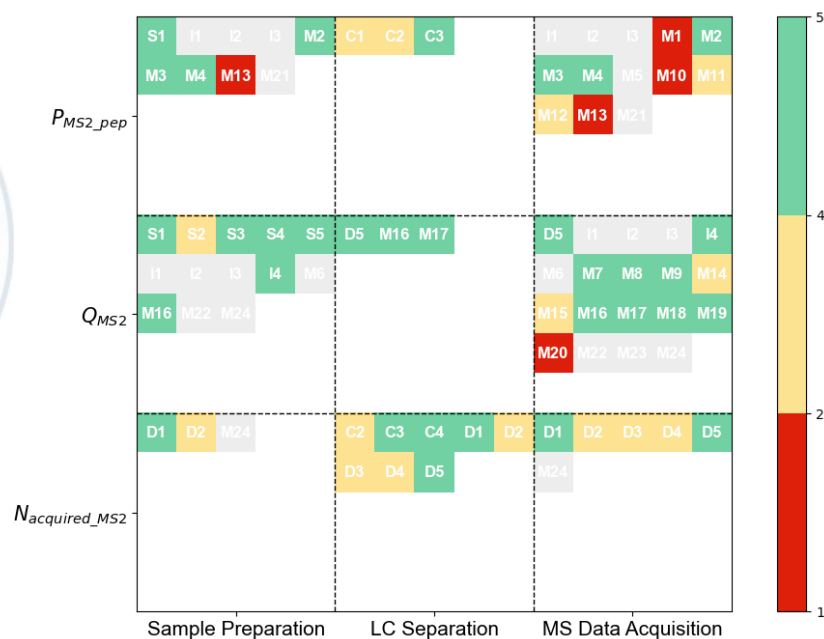
1

名称	修改
temp	202
Analysis_Report.html	202
INFO_Chromatography.txt	202
INFO_Cycle.txt	202
INFO_Feature.txt	202
INFO_ID.txt	202
INFO_Mass_Deviation.txt	202
INFO_MS1.txt	202
INFO_MS1_PEAKS.txt	202
INFO_MS2.txt	202
INFO_MS2_Dev.txt	202
INFO_MS2_PEAKS.txt	202
INFO_PIF.txt	202
INFO_Protein.txt	202
INFO_Summary.txt	202
INFO_Tag.txt	202
INFO_Tag_Sequence.txt	202

2

① The MSCohort results ;

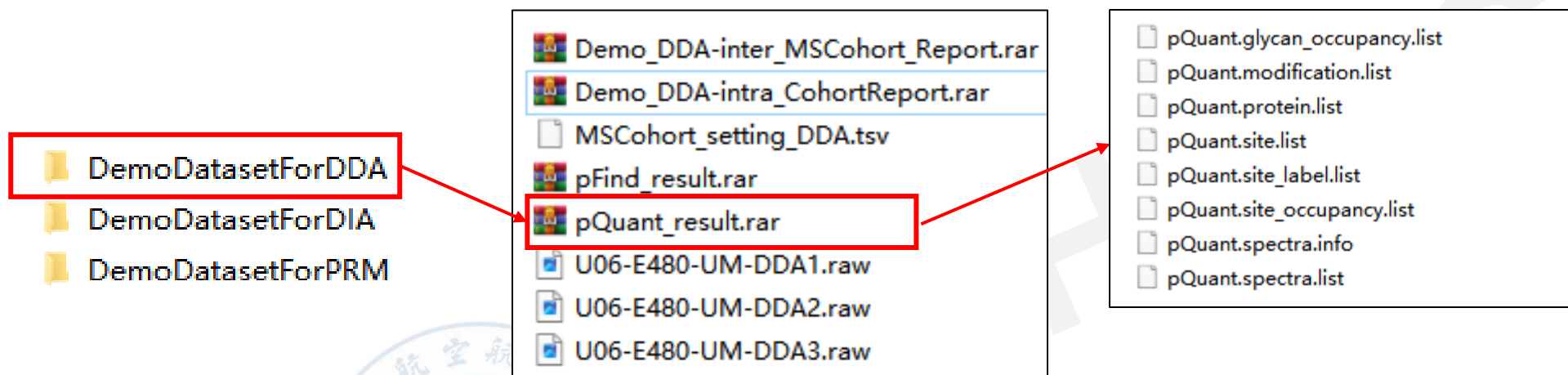
② Double clicking [Analysis\\_Report.html](#), the report will be preformed in the browser.



# **MSCohort manual for DDA inter- experiment analysis**



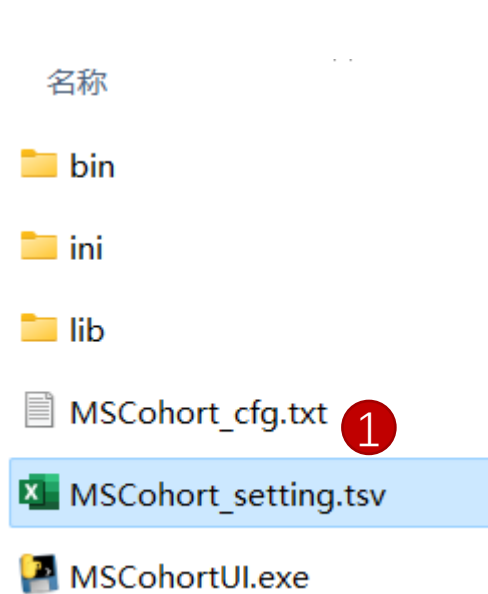
# Demo dataset



Taking the 3 E480 raw files as an example to demonstrate the workflow of DDA inter-experiment analysis.

Inter-experiment analysis enables the systematic quality evaluation and low-quality experiments detection for multiple DDA experiments.

# 1. Preparation for MSCohort\_setting.tsv file



Group Name	Raw Name	Experiment	Threshold
U06	U06-E480-UM-DDA1	U06-E480-UM-DDA1	10
U06	U06-E480-UM-DDA2	U06-E480-UM-DDA2	10
U06	U06-E480-UM-DDA3	U06-E480-UM-DDA3	10

① Open [MSCohort\\_setting.tsv](#) with Excel;

② Fill the columns as the example file (MSCohort\_setting\_DDA.tsv) showed or alternatively, directly use the MSCohort\_setting\_DDA.tsv file:

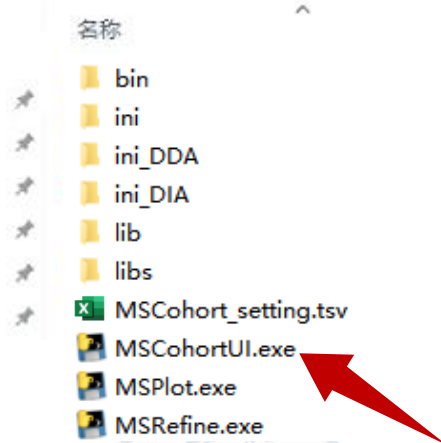
- The first column is the Group Name.
- The second column is the Raw Name, which is the same as [sample](#) reported from skyline.
- The third column is the Raw Name for short, which is convenient for display in large MSCohort report. This column could also be the same as Raw Name column.
- The fourth column is the intensity threshold, protein/peptide intensity that less than this threshold would be replaced as NaN, and would not be used for subsequent Pearson correlation analysis.



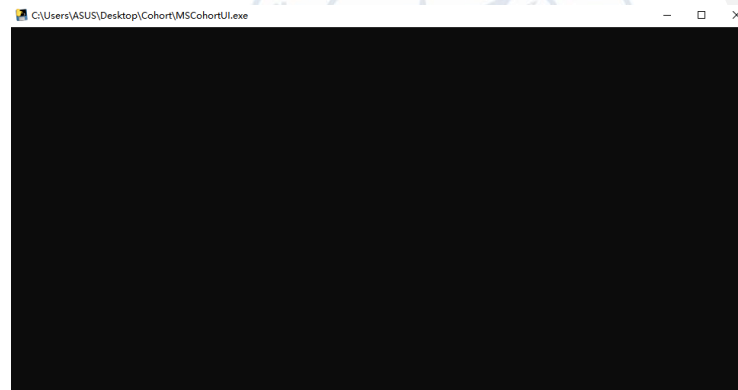
# 1. Analyzing with MSCohort

1

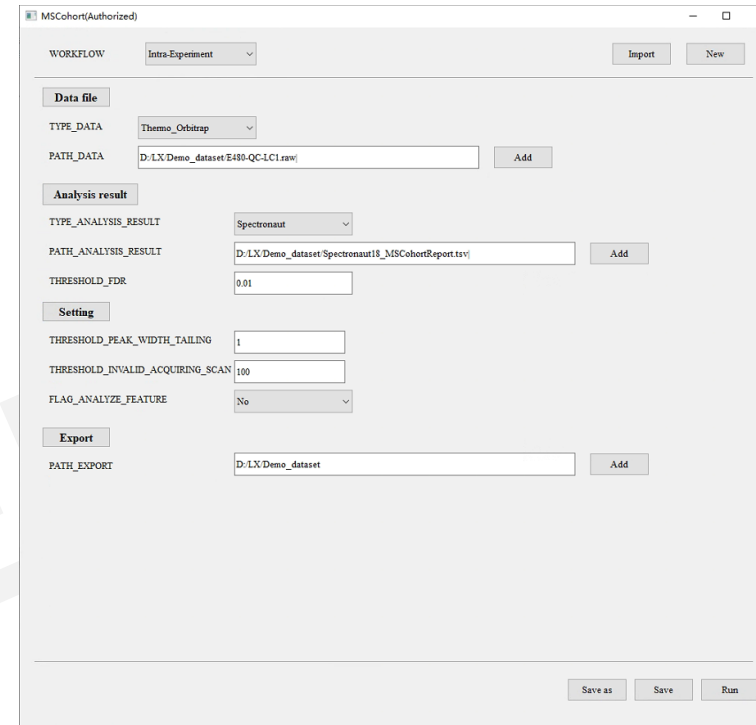
此电脑 > 新加卷 (D:) > LX > MSCohort\_v202405



2



3



- ① Double-click [MSCohortUI.exe](#);
- ② Wait for MSCohort program to load;
- ③ The MSCohortUI.exe settings screen is displayed.

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration:

- WORKFLOW:** Inter-Experiment(DDA) (Annotated 1)
- TYPE\_DATA:** Thermo\_Orbitrap (Annotated 2)
- PATH\_DATA:** (Empty field) (Annotated 3)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT:** pQuant
  - PATH\_ANALYSIS\_RESULT:** \_test/DDA\_test\_data\_new/pQuant\_result/UQ\_20241024\_013044/pQuant.protein.list (Annotated 4)
  - THRESHOLD\_FDR:** 0.01
- Setting:**
  - PATH\_EXPERIMENT\_RESULT:** op/paper\_test/DDA\_test/DDA\_test\_data\_new/COhort\_result/MSCohort\_setting.tsv (Annotated 5)
  - TYPE\_NORMALIZATION:** DirectLFQ
  - FLAG\_OUTLIERS:** 2-SD
  - FLAG\_SHOW\_ORDER:** group series
  - THRESHOLD\_PEAK\_WIDTH\_TAILING:** 1
  - THRESHOLD\_INVALID\_ACQUIRING\_SCAN:** 100
  - FLAG\_ANALYZE\_FEATURE:** No
- Export:**
  - PATH\_EXPORT:** C:/Users/ASUS/Desktop/paper\_test/DDA\_test/DDA\_test\_data\_new/COhort\_result

Buttons at the bottom: Save as, Save, Run.

- ① Set **WORKFLOW** as **Inter-experiment**;
- ② Select **TYPE\_DATA** according to the data type ;
- ③ The **PATH\_DATA** is empty.
- ④ Click **Add** to select the pQuant report file (pQuant.protein.list) for MSCohort into the **PATH\_ANALYSIS\_RESULT**;
- ⑤ Click **Add** to select the **MSCohort\_setting.tsv** into the **PATH\_EXPERIMENT\_RESULT**;

## 2. Analyzing with MSCohort

The screenshot displays the MSCohort software interface with the following configuration details:

- Workflow:** Inter-Experiment(DDA)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap
  - PATH\_DATA: (empty field with an 'Add' button)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: pQuant
  - PATH\_ANALYSIS\_RESULT: \_test/DDA\_test\_data\_new/pQuant\_result/UQ\_20241024\_013044/pQuant.protein.list (with an 'Add' button)
  - THRESHOLD\_FDR: 0.01
- Setting:**
  - PATH\_EXPERIMENT\_RESULT: op/paper\_test/DDA\_test/DDA\_test\_data\_new/COhort\_result/MSCOhort\_setting.tsv (with an 'Add' button)
  - TYPE\_NORMALIZATION: DirectLFQ
  - FLAG\_OUTLIERS: 2-SD
  - FLAG\_SHOW\_ORDER: group series
  - THRESHOLD\_PEAK\_WIDTH\_TAILING: 1
  - THRESHOLD\_INVALID\_ACQUIRING\_SCAN: 100
  - FLAG\_ANALYZE\_FEATURE: No
- Export:**
  - PATH\_EXPORT: C:/Users/ASUS/Desktop/paper\_test/DDA\_test/DDA\_test\_data\_new/COhort\_result (with an 'Add' button)

At the bottom of the interface are buttons for 'Save as', 'Save', and 'Run'.

- ① Choose normalization strategies in **TYPE\_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);
- ② Choose **FLAG\_OUTLIERS threshold** according to the experiment condition (default is 2\*SD, users could adjust the threshold (Notes for modifying the scoring criteria));
- ③ Choose **FLAG\_SHOW\_ORDER. group series** represents the experiment order showed in MSCohort report is the same as the **MSCOhort\_setting.tsv**; **time series** represents the experiment order showed in MSCohort report is sorted by run date.
- ④ Set **THRESHOLD\_PEAK\_WIDTH\_TAIL**, **THRESHOLD INVALID\_ACQUIRING**, **FLAG\_ANALYZE\_FEATURE** as intra-experiment analysis?;

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration options:

- WORKFLOW:** Inter-Experiment(DDA)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap
  - PATH\_DATA: (empty field)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: pQuant
  - PATH\_ANALYSIS\_RESULT: \_test/DDA\_test\_data\_new/pQuant\_result/UQ\_20241024\_013044/pQuant.protein.list
  - THRESHOLD\_FDR: 0.01
- Setting:**
  - PATH\_EXPERIMENT\_RESULT: op/paper\_test/DDA\_test/DDA\_test\_data\_new/COhort\_result/MSCohort\_setting.tsv
  - TYPE\_NORMALIZATION: DirectLFQ
  - FLAG\_OUTLIERS: 2-SD
  - FLAG\_SHOW\_ORDER: group series
  - THRESHOLD\_PEAK\_WIDTH\_TAILING: 1
  - THRESHOLD\_INVALID\_ACQUIRING\_SCAN: 100
  - FLAG\_ANALYZE\_FEATURE: No
- Export:**
  - PATH\_EXPORT: C:/Users/ASUS/Desktop/paper\_test/DDA\_test/DDA\_test\_data\_new/COhort\_result

At the bottom right, there are three buttons: "Save as" (labeled 2), "Save", and "Run" (labeled 3). A red circle labeled 1 is next to the PATH\_EXPORT field.

- ① Set the **PATH\_EXPORT** for saving the results. The output MSCohort analysis results will be stored in a new folder under the PATH\_EXPORT directory.

**Do not save the different experiments results under a same folder, the results will be **overwritten**.**

- ② Click **Save as** button to save the config file;
- ③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

### Note:

Space (“ ”) cannot exist in the file directory (including PATH\_DATA, PATH\_ANALYSIS\_RESULT, PATH\_EXPERIMENT\_RESULT, and PATH\_EXPORT ), which will affect the normal running of the program.

# 3. MSCohort Results

1

iRT\_Visual  
 model\_train  
 MSCohort  
 picture  
 tmp\_file  
 txt  
 Cohort Analysis Report.html  
 Cohort Analysis Report.pdf  
 INFO19\_iRT\_QC\_Summary.txt

3

INFO\_Pro0\_MissingValue\_Count.txt  
 INFO\_Pro1\_Intensity.txt  
 INFO\_Pro2\_Log2\_Intensity.txt  
 INFO\_Pro3\_Coefficient\_Var.txt  
 INFO0\_Identification\_Count.txt  
 INFO5\_Pre\_RetentionTime.txt  
 INFO6\_iRT\_XICDetail.txt  
 INFO7\_iRT\_LabelDetail.txt  
 INFO8\_Pro4\_Norm\_Intensity.txt  
 INFO9\_Pro5\_Log2\_Norm\_Intensity.txt  
 INFO10\_Pro6\_Norm\_Coefficient\_Var.txt  
 INFO11\_Pro7\_Norm\_PCA\_Coordinate.txt  
 INFO12\_Pro8\_Pearson\_Correlation.txt  
 INFO13\_Pro9\_Violin\_Statistic.txt  
 INFO14\_Pro10\_Norm\_Violin\_Statistic.txt  
 INFO15\_Pro11\_Cluster\_Hotmap.csv  
 INFO16\_Pro12\_Norm\_Cluster\_Hotmap.csv  
 INFO17\_Pro13\_Summary\_CoefficientVar.txt  
 INFO18\_Pro14\_Summary\_Norm\_CoefficientVar.txt  
 INFO19\_iRTSummary.txt  
 INFO20\_Pro\_Norm\_Pearson\_Correlation.txt  
 INFO21\_Empty.txt  
 INFO22\_Pro\_Log10iBAQ.txt  
 INFO23\_Peptide0\_MissingValue\_Count.txt  
 INFO24\_Peptide1\_Intensity.txt  
 INFO25\_Peptide2\_Log2\_Intensity.txt  
 INFO26\_Peptide3\_Norm\_Intensity.txt  
 INFO27\_Peptide4\_Norm\_Log2\_Intensity.txt

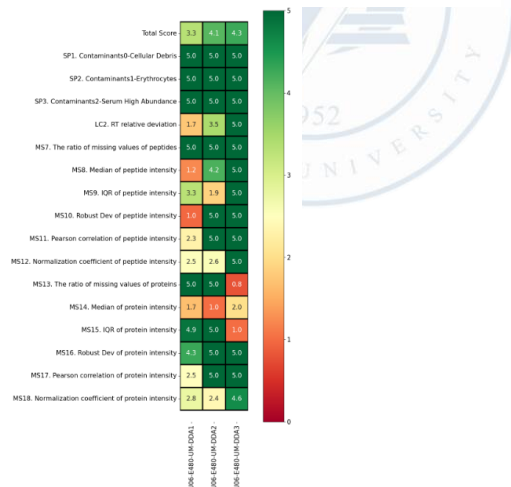
- ① The MSCohort results ;
- ② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.
- ③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

2

## 1. Overview of Dataset

### 1.1 Score of Inter-experiment Metrics

Inter-experiment metrics are computed across multiple experiments to assess the quality for the whole cohort quality data.



# **MSCohort manual for PRM inter- experiment analysis**



# Demo dataset (Skyline)

skyline\_result.xlsx

	A	B	C	D	E	F	G	H	I	J	K	L
1	peptide	protein	sample	parent	charge	daught	charge	by	rt	area		
2	AADDWEPFASGK	sp P0276 AH1030		697.815	2	735.367	1	y7	30.52	49840	1338	4
3	AADDWEPFASGK	sp P0276 AH1030		697.815	2	606.325	1	y6	30.52	177060	5203	1
4	AADDWEPFASGK	sp P0276 AH1030		697.815	2	362.203	1	y4	30.52	53853	0	3
5	AADDWEPFASGK	sp P0276 AH1030		697.815	2	291.166	1	y3	30.52	57368	1457	2
6	AADDWEPFASGK	sp P0276 AH1107		697.815	2	735.367	1	y7	30.26	16332	0	4
7	AADDWEPFASGK	sp P0276 AH1107		697.815	2	606.325	1	y6	30.26	69586	1399	1
8	AADDWEPFASGK	sp P0276 AH1107		697.815	2	362.203	1	y4	30.26	17011	0	3
9	AADDWEPFASGK	sp P0276 AH1107		697.815	2	291.166	1	y3	30.36	17852	372	2
10	AADDWEPFASGK	sp P0276 AH1108		697.815	2	735.367	1	y7	30.39	27186	0	4
11	AADDWEPFASGK	sp P0276 AH1108		697.815	2	606.325	1	y6	30.49	123939	0	1
12	AADDWEPFASGK	sp P0276 AH1108		697.815	2	362.203	1	y4	30.49	36596	0	3
13	AADDWEPFASGK	sp P0276 AH1108		697.815	2	291.166	1	y3	30.49	39138	0	2
14	AADDWEPFASGK	sp P0276 AH1110		697.815	2	735.367	1	y7	30.42	54703	25	4
15	AADDWEPFASGK	sp P0276 AH1110		697.815	2	606.325	1	y6	30.42	252829	8339	1
16	AADDWEPFASGK	sp P0276 AH1110		697.815	2	362.203	1	y4	30.31	77988	151	3
17	AADDWEPFASGK	sp P0276 AH1110		697.815	2	291.166	1	y3	30.31	80837	1492	2
18	AADDWEPFASGK	sp P0276 AH1152		697.815	2	735.367	1	y7	30.36	49964	813	4
19	AADDWEPFASGK	sp P0276 AH1152		697.815	2	606.325	1	y6	30.36	218723	4769	1
20	AADDWEPFASGK	sp P0276 AH1152		697.815	2	362.203	1	y4	30.26	59048	0	2
21	AADDWEPFASGK	sp P0276 AH1152		697.815	2	291.166	1	y3	30.36	56004	0	3
22	AADDWEPFASGK	sp P0276 AH1155		697.815	2	735.367	1	y7	30.26	36350	1583	4
23	AADDWEPFASGK	sp P0276 AH1155		697.815	2	606.325	1	y6	30.36	162031	3167	1
24	AADDWEPFASGK	sp P0276 AH1155		697.815	2	362.203	1	y4	30.36	42934	1794	3
25	AADDWEPFASGK	sp P0276 AH1155		697.815	2	291.166	1	y3	30.36	45335	0	2
26	AADDWEPFASGK	sp P0276 AH1168		697.815	2	735.367	1	y7	30.56	71960	0	4
27	AADDWEPFASGK	sp P0276 AH1168		697.815	2	606.325	1	y6	30.56	301837	2283	1
28	AADDWEPFASGK	sp P0276 AH1168		697.815	2	362.203	1	y4	30.56	85678	0	3
29	AADDWEPFASGK	sp P0276 AH1168		697.815	2	291.166	1	y3	30.56	93287	0	2
30	AADDWEPFASGK	sp P0276 AH1180		697.815	2	735.367	1	y7	30.28	177166	0	4
31	AADDWEPFASGK	sp P0276 AH1180		697.815	2	606.325	1	y6	30.28	733783	4147	1
32	AADDWEPFASGK	sp P0276 AH1180		697.815	2	362.203	1	y4	30.28	200999	0	3
33	AADDWEPFASGK	sp P0276 AH1180		697.815	2	291.166	1	y3	30.28	200999	0	2

名称	类型	大小
MSCohort_setting_Skyline.tsv	TSV 文件	2 KB
skyline_result.xlsx	Microsoft Excel 工作表	5,347 KB

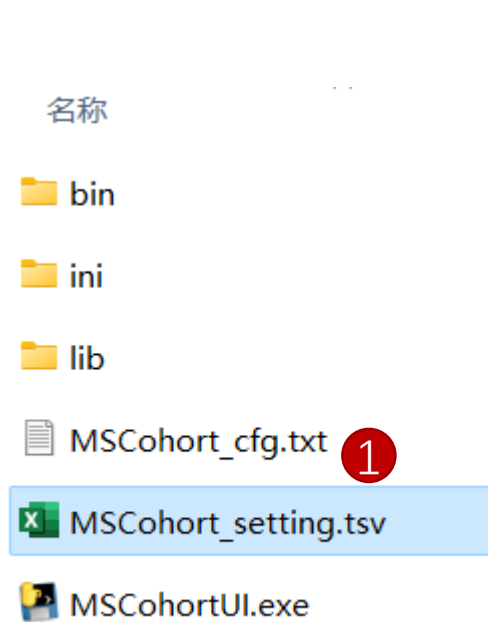
MSCohort\_setting\_Skyline.tsv

	A	B	C	D
1	Group Name	Raw Name	Experiment	Threshold
2	Group1	AH-QC5	AH-QC5	0
3	Group1	AH-QC7	AH-QC7	0
4	Group1	AH-QC10	AH-QC10	0
5	Group1	AH-QC11	AH-QC11	0
6	Group1	AH-QC12	AH-QC12	0
7	Group1	AH-QC14	AH-QC14	0
8	Group1	AH-QC15	AH-QC15	0
9	Group1	AH-QC16	AH-QC16	0
10	Group1	AH-QC17	AH-QC17	0
11	Group1	AH-QC18	AH-QC18	0
12	Group1	AH-QC19	AH-QC19	0
13	Group1	AH-QC20	AH-QC20	0
14	Group1	AH-QC21	AH-QC21	0
15	Group1	AH-QC22	AH-QC22	0
16	Group1	AH-QC23	AH-QC23	0
17	Group1	AH-QC24	AH-QC24	0
18	Group1	AH-QC25	AH-QC25	0
19	Group1	AH-QC26	AH-QC26	0
20	Group1	AH-QC27	AH-QC27	0
21	Group1	AH-QC28	AH-QC28	0
22	Group1	AH-QC29	AH-QC29	0
23	Group1	AH-QC30	AH-QC30	0

Taking the skyline result files as an example to demonstrate the workflow of PRM inter-experiment analysis.

Inter-experiment analysis enables the systematic quality evaluation and low-quality experiments detection for multiple PRM experiments.

# 1. Preparation for MSCohort\_setting.tsv file



Group Name	Raw Name	Experiment	Threshold
Group1	AH-QC5	AH-QC5	0
Group1	AH-QC6	AH-QC6	0
Group1	AH-QC7	AH-QC7	0
Group1	AH-QC8	AH-QC8	0
Group1	AH-QC10	AH-QC10	0
Group1	AH-QC11	AH-QC11	0
Group1	AH-QC12	AH-QC12	0
Group1	AH-QC14	AH-QC14	0
Group1	AH-QC15	AH-QC15	0
Group1	AH-QC16	AH-QC16	0
Group1	AH-QC17	AH-QC17	0
Group1	AH-QC18	AH-QC18	0
Group1	AH-QC19	AH-QC19	0

- ① Open [MSCohort\\_setting.tsv](#) with Excel;
- ② Fill the columns as the example file (MSCohort\_setting\_Skyline.tsv) showed or alternatively, directly use the MSCohort\_setting\_Skyline.tsv file:
  - The first column is the Group Name.
  - The second column is the Raw Name, which is the same as [sample](#) reported from skyline.
  - The third column is the Raw Name for short, which is convenient for display in large MSCohort report. This column could also be the same as Raw Name column.
  - The fourth column is the intensity threshold, protein/peptide intensity that less than this threshold would be replaced as NaN, and would not be used for subsequent Pearson correlation analysis.

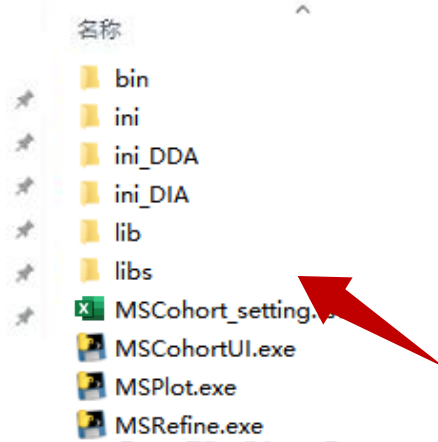
**All subsequent analyses are based only on the information in the Raw Name column provided by MSCohort\_setting.tsv**



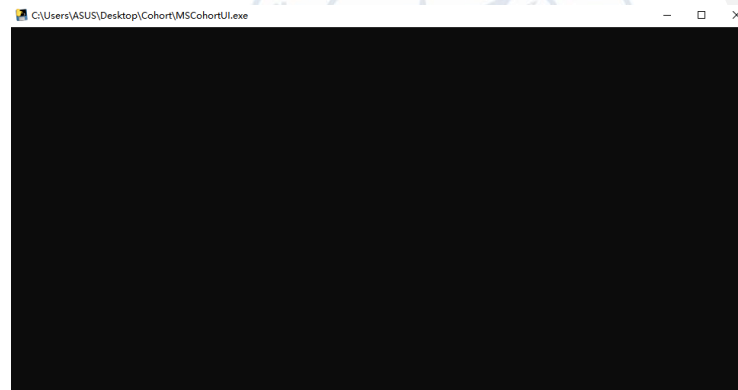
# 1. Analyzing with MSCohort

1

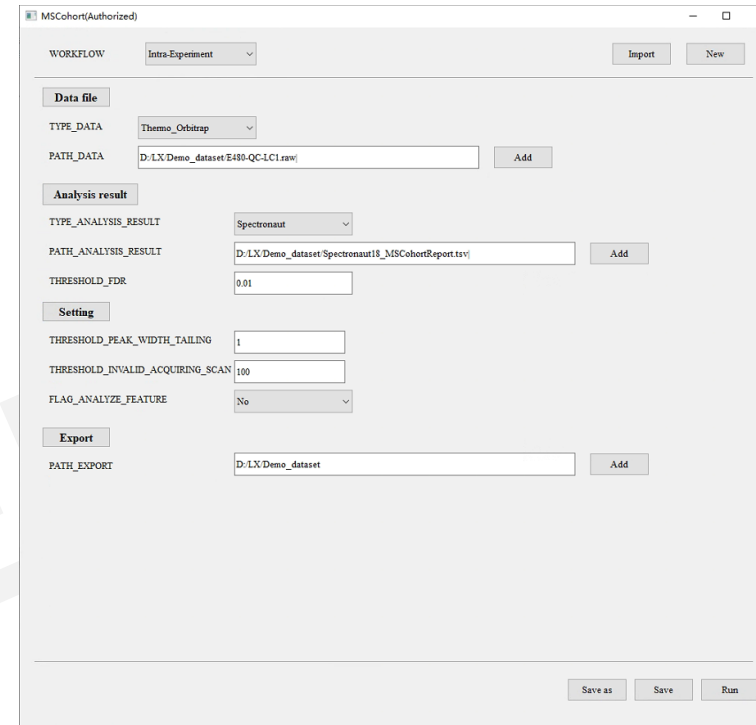
此电脑 > 新加卷 (D:) > LX > MSCohort\_v202405



2



3



- ① Double-click [MSCohortUI.exe](#);
- ② Wait for MSCohort program to load;
- ③ The MSCohortUI.exe settings screen is displayed.

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration steps highlighted by red circles and numbers:

- WORKFLOW** is set to **Inter-Experiment(PRM)**. A dropdown menu is open showing options: Inter-Experiment(DIA), Intra-Experiment(DIA), Inter-Experiment(DDA), Intra-Experiment(DDA), and Inter-Experiment(PRM) (selected).
- TYPE\_DATA** is set to **Thermo\_Orbitrap**. A dropdown menu is open showing options: Thermo\_Orbitrap (selected), Bruker\_timsTOF, and SCIEX\_ZenoTOF.
- PATH\_DATA** is empty.
- Analysis result** section: **TYPE\_ANALYSIS\_RESULT** is set to **Skyline**. A dropdown menu is open showing options: Spectronaut, pFind, pQuant, SpectroDive, and Skyline (selected). **PATH\_ANALYSIS\_RESULT** is set to **E:/RAW/PRM/skyline\_result.xlsx**. **THRESHOLD\_FDR** is set to **0.01**.
- Setting** section: **PATH\_EXPERIMENT\_RESULT** is set to **E:/RAW/PRM/MSCohort\_setting.tsv**. **TYPE\_NORMALIZATION** is set to **DirectLFQ**. **FLAG\_OUTLIERS** is set to **2-SD**. **FLAG\_SHOW\_ORDER** is set to **group series**.

At the bottom, the **Export** section shows **PATH\_EXPORT** set to **E:/RAW/PRM**. Buttons for **Import**, **New**, **Add**, **Save as**, **Save**, and **Run** are visible.

- ① Set **WORKFLOW** as **Inter-experiment**;
- ② Select **TYPE\_DATA** according to the data type ;
- ③ The **PATH\_DATA** is empty.
- ④ Click **Add** to select the Skyline\_result for MSCohort into the **PATH\_ANALYSIS\_RESULT**;
- ⑤ Click **Add** to select the **MSCohort\_setting.tsv** into the **PATH\_EXPERIMENT\_RESULT**;

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration:

- WORKFLOW:** Inter-Experiment(PRM)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap
  - PATH\_DATA: (empty field)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: Skyline
  - PATH\_ANALYSIS\_RESULT: E:/RAW/PRM/skyline\_result.xlsx
  - THRESHOLD\_FDR: 0.01
- Setting:**
  - PATH\_EXPERIMENT\_RESULT: E:/RAW/PRM/MSCohort\_setting.tsv
  - TYPE\_NORMALIZATION: DirectLFQ
  - FLAG\_OUTLIERS: 2-SD
  - FLAG\_SHOW\_ORDER: group series
- Export:**
  - PATH\_EXPORT: E:/RAW/PRM

Buttons at the bottom: Save as, Save, Run.

- ① Choose normalization strategies in **TYPE\_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);
- ② Choose **FLAG\_OUTLIERS threshold** according to the experiment condition (default is 2\*SD, users could adjust the threshold (Notes for modifying the scoring criteria));
- ③ Choose **FLAG\_SHOW\_ORDER**. **group series** represents the experiment order showed in MSCohort report is the same as the **MSCohort\_setting.tsv**; **time series** represents the experiment order showed in MSCohort report is sorted by run date.

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration options:

- WORKFLOW:** Inter-Experiment(PRM)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap
  - PATH\_DATA: (empty field with an Add button)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: Skyline
  - PATH\_ANALYSIS\_RESULT: E:/RAW/PRM/skyline\_result.xlsx (with an Add button)
  - THRESHOLD\_FDR: 0.01
- Setting:**
  - PATH\_EXPERIMENT\_RESULT: E:/RAW/PRM/MSCohort\_setting.tsv (with an Add button)
  - TYPE\_NORMALIZATION: DirectLFQ
  - FLAG\_OUTLIERS: 2-SD
  - FLAG\_SHOW\_ORDER: group series
- Export:**
  - PATH\_EXPORT: E:/RAW/PRM (with an Add button)

At the bottom right, there are three buttons: "Status" (highlighted with a red circle 2), "Save", and "Run" (highlighted with a red circle 3).

- ① Set the **PATH\_EXPORT** for saving the results. The output MSCohort analysis results will be stored in a new folder under the PATH\_EXPORT directory, named MSCohort \_“year+month+day” folder.

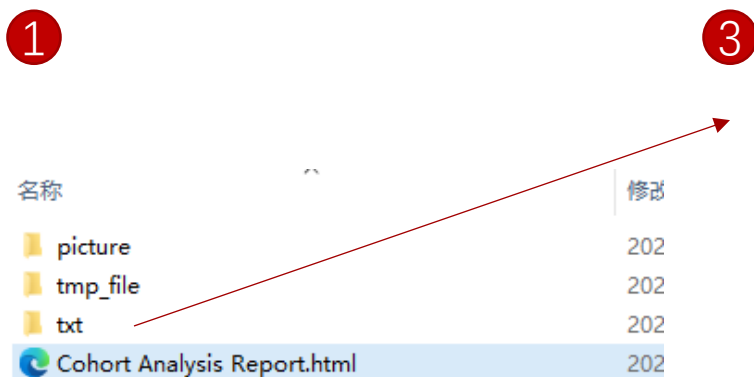
**Do not save the different experiments results under a same folder, the results will be **overwritten**.**

- ② Click **Save as** button to save the config file;
- ③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

### Note:

**Space (“ ”) cannot exist in the file directory** (including PATH\_DATA, PATH\_ANALYSIS\_RESULT, PATH\_EXPERIMENT\_RESULT, and PATH\_EXPORT ), which will affect the normal running of the program.

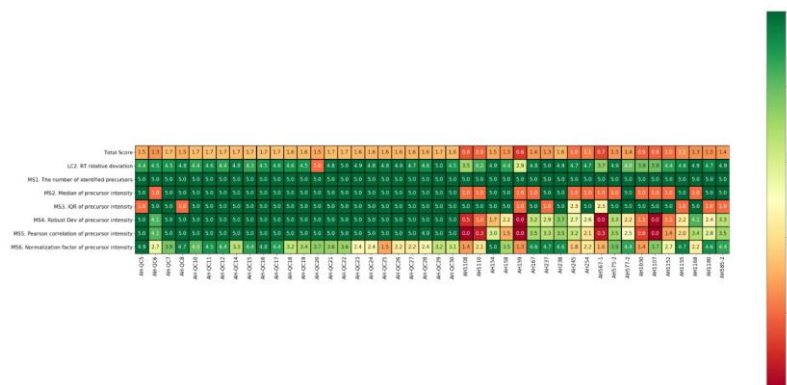
# 3. MSCohort Results



## 2 1. Overview of Dataset

### 1.1 Score of Inter-experiment Metrics

Inter-experiment metrics are computed across multiple experiments to assess the quality for the whole cohort quality data.



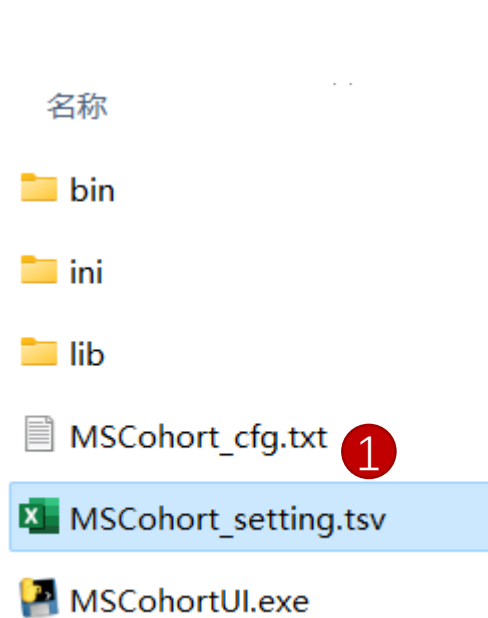
- INFO0\_Experiment\_Outlier\_Score.txt
- INFO0\_Identification\_Count.txt
- INFO0\_Inter\_Experiment\_Scores.txt
- INFO0\_Inter\_Experiment\_values.txt
- INFO0\_Intra\_Experiment\_Scores.txt
- INFO0\_Intra\_Experiment\_values.txt
- INFO1\_Pro0\_MissingValue\_Count.txt
- INFO2\_Pro1\_Intensity.txt
- INFO3\_Pro2-Origin\_Intensity.txt
- INFO4\_Pro3\_Coefficient\_Var.txt
- INFO5\_Pre\_RetentionTime.txt
- INFO8\_Pro4\_Norm\_Intensity.txt
- INFO9\_Pro5-Origin\_Norm\_Intensity.txt
- INFO10\_Pro6\_Norm\_Coefficient\_Var.txt
- INFO11\_Pro7\_Norm\_PCA\_Coordinate.txt
- INFO12\_Pro8\_Pearson\_Correlation.txt
- INFO13\_Pro9\_Violin\_Statistic.txt
- INFO14\_Pro10\_Norm\_Violin\_Statistic.txt
- INFO15\_Pro11\_Cluster\_Hotmap.csv
- INFO16\_Pro12\_Norm\_Cluster\_Hotmap.csv
- INFO17\_Pro13\_Summary\_CoefficientVar.txt
- INFO18\_Pro14\_Summary\_Norm\_CoefficientVar.txt
- INFO19\_iRTSummary.txt
- INFO20\_Pro\_Norm\_Pearson\_Correlation.txt
- INFO21\_Empty.txt
- INFO22\_Pro\_Log10iBAQ.txt
- INFO23\_Peptide0\_MissingValue\_Count.txt

- ① The MSCohort results ;
- ② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.
- ③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

## 大小



# 1. Preparation for MSCohort\_setting.tsv file



2

Group Name	Raw Name	Experiment	Threshold
Group1	LXO-mix1.raw	LXO-mix1.raw	0
Group1	LXO-mix2.raw	LXO-mix2.raw	0
Group1	LXO-mix3.raw	LXO-mix3.raw	0
Group1	LXO-mix4.raw	LXO-mix4.raw	0
Group1	LXO-mix5.raw	LXO-mix5.raw	0
Group1	LXO-mix6.raw	LXO-mix6.raw	0
Group1	LXO-mix7.raw	LXO-mix7.raw	0
Group1	LXO-mix8.raw	LXO-mix8.raw	0
Group1	LXO-mix9.raw	LXO-mix9.raw	0
Group1	LXO-mix10.raw	LXO-mix10.raw	0
Group1	LXO-mix11.raw	LXO-mix11.raw	0
Group1	LXO-mix12.raw	LXO-mix12.raw	0
Group1	LXO-mix13.raw	LXO-mix13.raw	0

① Open [MSCohort\\_setting.tsv](#) with Excel;

② Fill the columns as the example file (MSCohort\_setting\_SpectroDive.tsv) showed or alternatively, directly use the MSCohort\_setting\_SpectroDive.tsv file:

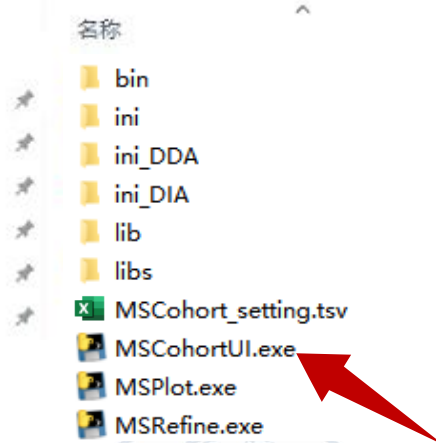
- The first column is the Group Name.
- The second column is the Raw Name, which is the same as [R.FileName](#) reported from SpectroDive.
- The third column is the Raw Name for short, which is convenient for display in large MSCohort report. This column could also be the same as Raw Name column.
- The fourth column is the intensity threshold, protein/peptide intensity that less than this threshold would be replaced as NaN, and would not be used for subsequent Pearson correlation analysis.

**All subsequent analyses are based only on the information in the Raw Name column provided by MSCohort\_setting.tsv**

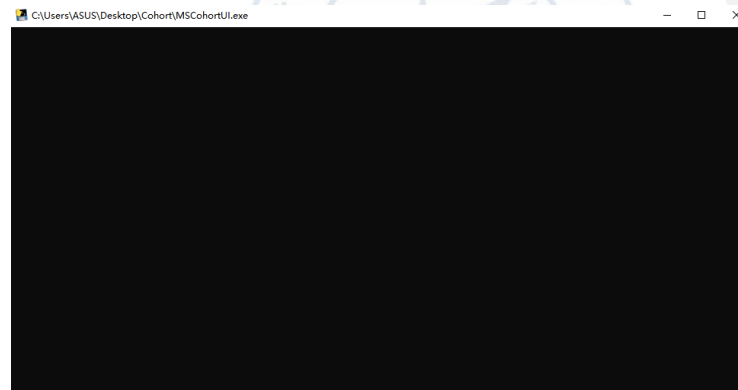
# 1. Analyzing with MSCohort

1

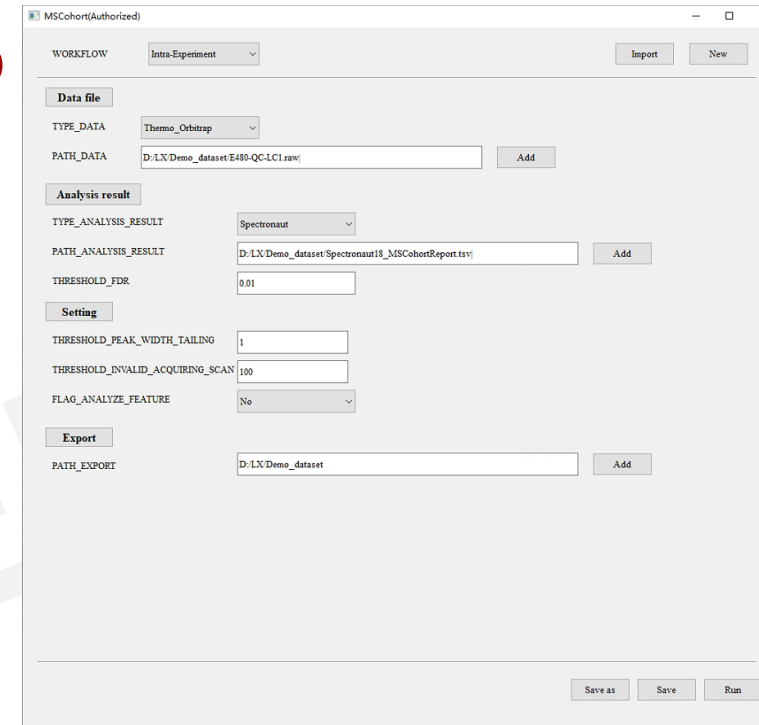
此电脑 > 新加卷 (D:) > LX > MSCohort\_v202405



2



3



- ① Double-click [MSCohortUI.exe](#);
- ② Wait for MSCohort program to load;
- ③ The MSCohortUI.exe settings screen is displayed.



## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration:

- WORKFLOW:** Inter-Experiment(PRM) (Annotated 1)
- Data file:** Thermo\_Orbitrap (Annotated 2)
- TYPE\_DATA:** Thermo\_Orbitrap (Annotated 3)
- PATH\_DATA:** (Empty field) (Annotated 3)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT:** SpectroDive
  - PATH\_ANALYSIS\_RESULT:** E:/RAW/PRM/spectroDive\_result/SpectroDive\_result (Annotated 4)
  - THRESHOLD\_FDR:** 0.01
- Setting:**
  - PATH\_EXPERIMENT\_RESULT:** E:/RAW/PRM/spectroDive\_result/MSCohort\_setting.tsv (Annotated 5)
  - TYPE\_NORMALIZATION:** DirectLFQ
  - FLAG\_OUTLIERS:** 2-SD
  - FLAG\_SHOW\_ORDER:** group series
- Export:**
  - PATH\_EXPORT:** E:/RAW/PRM

Buttons at the bottom: Save as, Save, Run.

- ① Set **WORKFLOW** as **Inter-experiment**;
- ② Select **TYPE\_DATA** according to the data type ;
- ③ The **PATH\_DATA** is empty.
- ④ Click **Add** to select the SpectroDive report for MSCohort into the **PATH\_ANALYSIS\_RESULT**;
- ⑤ Click **Add** to select the **MSCohort\_setting.tsv** into the **PATH\_EXPERIMENT\_RESULT**;

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration details:

- WORKFLOW:** Inter-Experiment(PRM)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap
  - PATH\_DATA: (empty field with an 'Add' button)
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: SpectroDive
  - PATH\_ANALYSIS\_RESULT: E:/RAW/PRM/spectroDive\_result/SpectroDive\_result.txt (with an 'Add' button)
  - THRESHOLD\_FDR: 0.01
- Setting:**
  - PATH\_EXPERIMENT\_RESULT: E:/RAW/PRM/spectroDive\_result/MSCohort\_setting.tsv (with an 'Add' button)
  - TYPE\_NORMALIZATION: DirectLFQ
  - FLAG\_OUTLIERS: 2-SD
  - FLAG\_SHOW\_ORDER: group series
- Export:**
  - PATH\_EXPORT: E:/RAW/PRM (with an 'Add' button)

At the bottom of the interface are buttons for 'Save as', 'Save', and 'Run'.

- ① Choose normalization strategies in **TYPE\_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);
- ② Choose **FLAG\_OUTLIERS threshold** according to the experiment condition (default is 2\*SD, users could adjust the threshold (Notes for modifying the scoring criteria));
- ③ Choose **FLAG\_SHOW\_ORDER**. **group series** represents the experiment order showed in MSCohort report is the same as the **MSCohort\_setting.tsv**; **time series** represents the experiment order showed in MSCohort report is sorted by run date.

## 2. Analyzing with MSCohort

The screenshot shows the MSCohort software interface with the following configuration options:

- WORKFLOW:** Inter-Experiment(PRM) (dropdown)
- Data file:**
  - TYPE\_DATA: Thermo\_Orbitrap (dropdown)
  - PATH\_DATA: (empty text field) [Add]
- Analysis result:**
  - TYPE\_ANALYSIS\_RESULT: SpectroDive (dropdown)
  - PATH\_ANALYSIS\_RESULT: E:/RAW/PRM/spectroDive\_result/SpectroDive\_result.txt [Add]
  - THRESHOLD\_FDR: 0.01 (text field)
- Setting:**
  - PATH\_EXPERIMENT\_RESULT: E:/RAW/PRM/spectroDive\_result/MSCohort\_setting.tsv [Add]
  - TYPE\_NORMALIZATION: DirectLFQ (dropdown)
  - FLAG\_OUTLIERS: 2-SD (dropdown)
  - FLAG\_SHOW\_ORDER: group series (dropdown)
- Export:**
  - PATH\_EXPORT: E:/RAW/PRM [Add]

At the bottom right, there are three buttons: "Save as" (labeled with red circle 2), "Save", and "Run" (labeled with red circle 3). A red circle 1 is placed next to the PATH\_EXPORT field.

- ① Set the **PATH\_EXPORT** for saving the results. The output MSCohort analysis results will be stored in a new folder under the PATH\_EXPORT directory, named MSCohort \_“year+month+day” folder.

**Do not save the different experiments results under a same folder, the results will be **overwritten**.**

- ② Click **Save as** button to save the config file;
- ③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

### Note:

Space (“ ”) cannot exist in the file directory (including PATH\_DATA, PATH\_ANALYSIS\_RESULT, PATH\_EXPERIMENT\_RESULT, and PATH\_EXPORT ), which will affect the normal running of the program.

# 3. MSCohort Results

1

名称	修改
picture	202
tmp_file	202
txt	202
Cohort Analysis Report.html	202

3

INFO0\_Experiment\_Outlier\_Score.txt  
 INFO0\_Identification\_Count.txt  
 INFO0\_Inter\_Experiment\_Scores.txt  
 INFO0\_Inter\_Experiment\_values.txt  
 INFO0\_Intra\_Experiment\_Scores.txt  
 INFO0\_Intra\_Experiment\_values.txt  
 INFO1\_Pro0\_MissingValue\_Count.txt  
 INFO2\_Pro1\_Intensity.txt  
 INFO3\_Pro2\_Origin\_Intensity.txt  
 INFO4\_Pro3\_Coefficient\_Var.txt  
 INFO5\_Pre\_RetentionTime.txt  
 INFO8\_Pro4\_Norm\_Intensity.txt  
 INFO9\_Pro5\_Origin\_Norm\_Intensity.txt  
 INFO10\_Pro6\_Norm\_Coefficient\_Var.txt  
 INFO11\_Pro7\_Norm\_PCA\_Coordinate.txt  
 INFO12\_Pro8\_Pearson\_Correlation.txt  
 INFO13\_Pro9\_Violin\_Statistic.txt  
 INFO14\_Pro10\_Norm\_Violin\_Statistic.txt  
 INFO15\_Pro11\_Cluster\_Hotmap.csv  
 INFO16\_Pro12\_Norm\_Cluster\_Hotmap.csv  
 INFO17\_Pro13\_Summary\_CoefficientVar.txt  
 INFO18\_Pro14\_Summary\_Norm\_CoefficientVar.txt  
 INFO19\_iRTSummary.txt  
 INFO20\_Pro\_Norm\_Pearson\_Correlation.txt  
 INFO21\_Empty.txt  
 INFO22\_Pro\_Log10iBAQ.txt  
 INFO23\_Peptide0\_MissingValue\_Count.txt

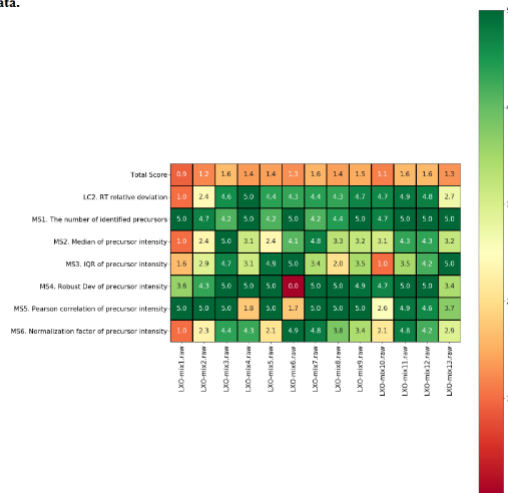
- ① The MSCohort results ;
- ② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.
- ③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

2

## 1. Overview of Dataset

### 1.1 Score of Inter-experiment Metrics

Inter-experiment metrics are computed across multiple experiments to assess the quality for the whole cohort quality data.



# Thanks!

