

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 dataset

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)..

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 DIA data from Orbitrap Exploris 480, pFind/pQuant analysis results and MSCohort report results.

Demo DDA folds contains Skyline and SpectroDive analysis results and MSCohort report results.

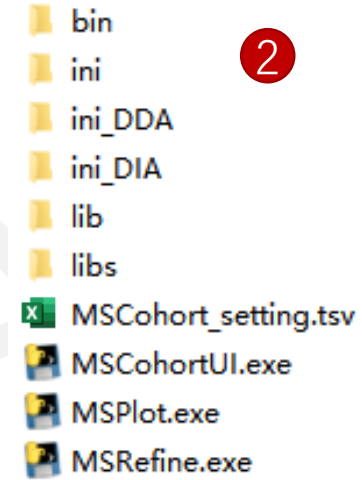
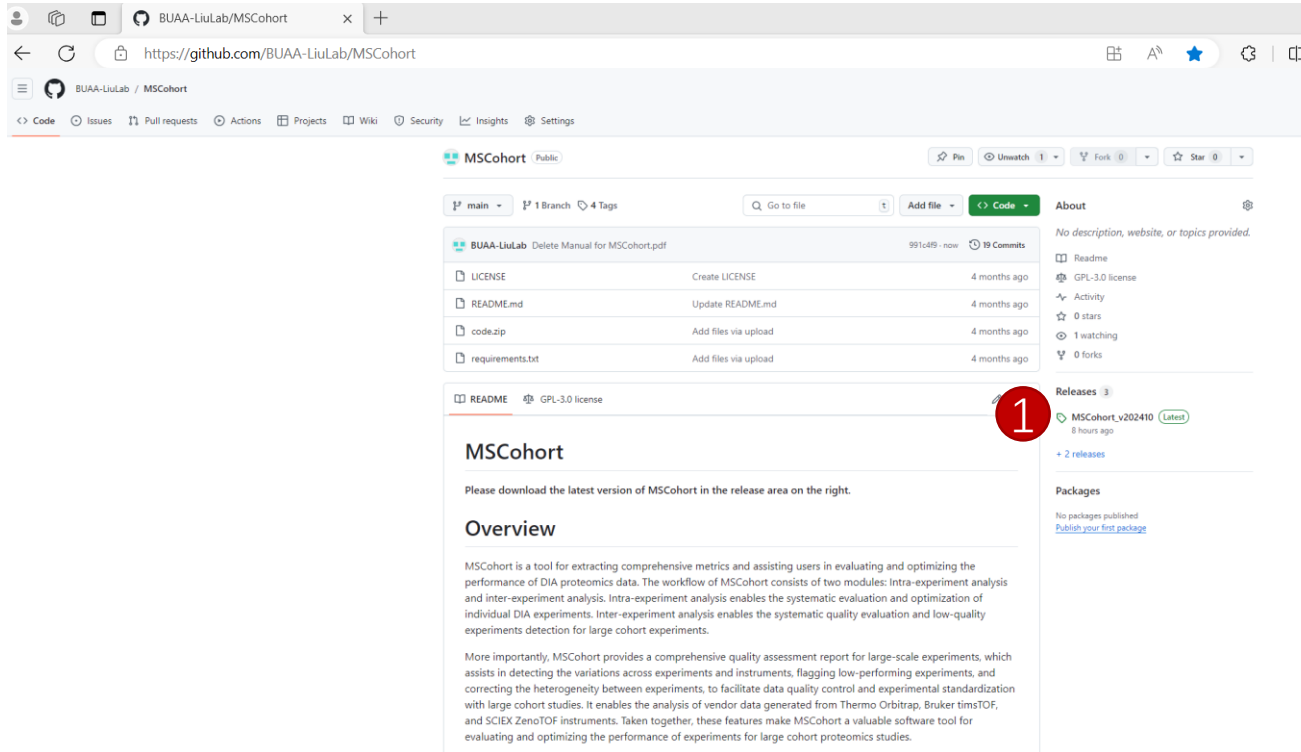
CONTENTS

- [MSCohort download](#)
- [Spectronaut custom report for MSCohort](#)
- [MSCohort manual for DIA intra-experiment analysis](#)
- [MSCohort manual for DIA inter-experiment analysis](#)
- [Notes for modifying the scoring criteria](#)
- [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.

Spectronaut customized report for MSCohort



1. Spectronaut customized report

The columns of information that MSCohort needs to read

The screenshot shows the Spectronaut software interface. At the top, there are icons for Analysis, Post Analysis, Report (highlighted with a red circle 1), Databases, Settings, and About. Below these, a file named 'hela7/min' is open. On the left, the 'Standard Report' and 'PTM Site Report' sections are visible. The 'Normal Report' section is expanded, showing various report formats. A red circle 2 highlights the 'Normal Report' section. The 'Columns' section on the right is expanded, showing a list of columns. A red circle 3 highlights the 'Columns' section. The 'Save As' dialog at the bottom is open, showing the file name 'MSCohort' and the 'OK' button. A red circle 4 highlights the 'Save As' dialog.

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 **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**”.

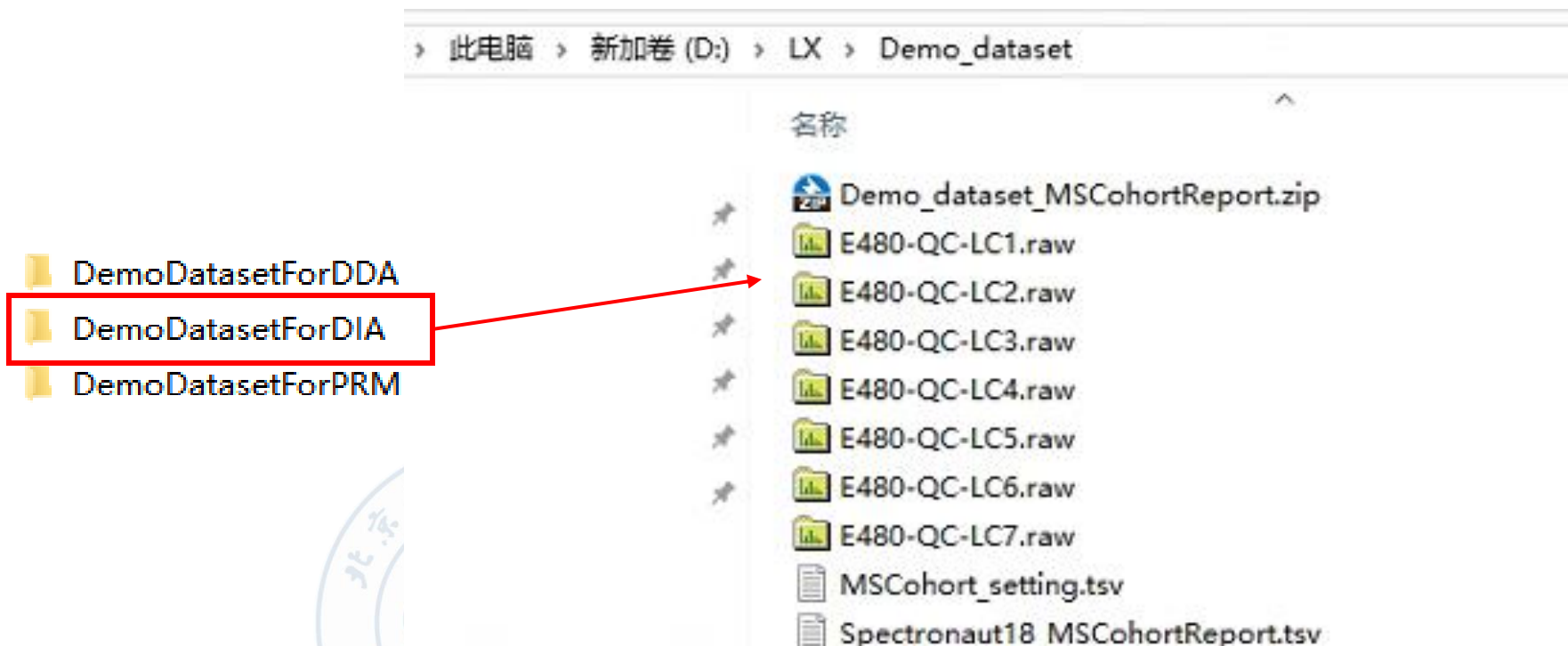
[illegible]

- 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.

MSCohort manual for DIA intra- experiment analysis



Demo dataset



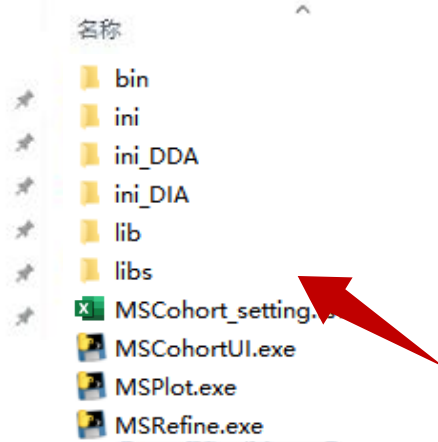
Taking the E480-QC-LC1.raw file as an example to demonstrate the workflow of intra-experiment analysis.

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

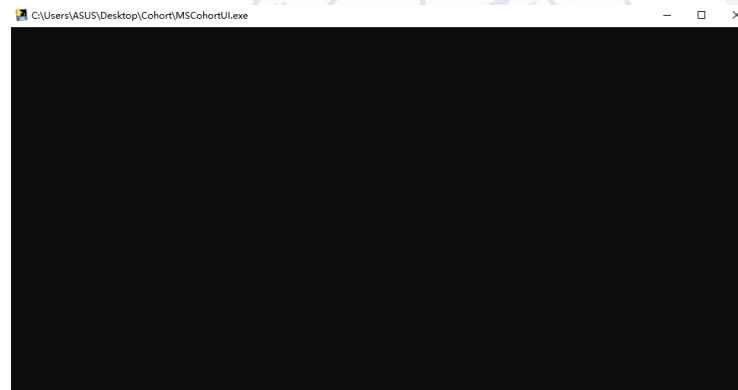
1. Analyzing with MSCohort

1

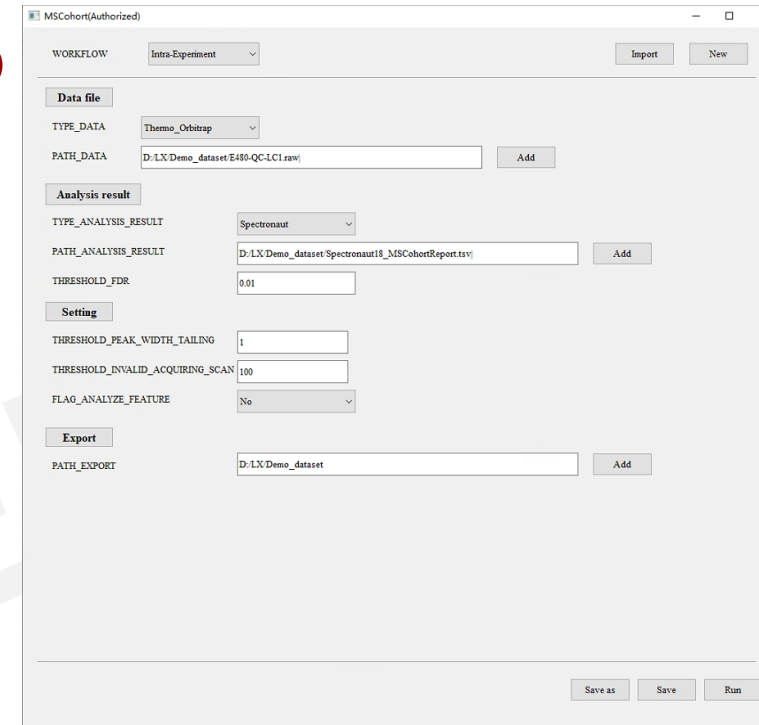
此电脑 > 新加卷 (D:) > LX > MSCohort_v202405



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 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-LC1.raw (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:

- Analysis result**: TYPE_ANALYSIS_RESULT: Spectronaut (dropdown menu).
- Setting**: THRESHOLD_FDR: 0.01 (text field).
- Setting**: THRESHOLD_PEAK_WIDTH_TAILING: 1 (text field).
- Setting**: THRESHOLD_INVALID_ACQUIRING_SCAN: 100 (text field).
- Setting**: FLAG_ANALYZE_FEATURE: No (dropdown menu).

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' set to 'D:/LX/Demo_dataset/E480-QC-LC1.raw'. The 'Analysis result' section has 'TYPE_ANALYSIS_RESULT' set to 'Spectronaut', 'PATH_ANALYSIS_RESULT' set to 'D:/LX/Demo_dataset/Spectronaut1\$ MSCohortReport.tsv', and 'THRESHOLD_FDR' set to '0.01'. The 'Setting' section has 'THRESHOLD_PEAK_WIDTH_TAILING' set to '1', 'THRESHOLD_INVALID_ACQUIRING_SCAN' set to '100', and 'FLAG_ANALYZE_FEATURE' set to 'Yes'. At the bottom, there's an 'Export' section with 'PATH_EXPORT' set to 'D:/LX/Demo_dataset'. At the very bottom, there are three buttons: 'Save as', 'Save', and 'Run'. Red numbered circles 1 through 5 are placed over the following elements: 1. 'THRESHOLD_PEAK_WIDTH_TAILING' input field; 2. 'THRESHOLD_INVALID_ACQUIRING_SCAN' input field; 3. 'FLAG_ANALYZE_FEATURE' dropdown menu; 4. 'Save as' button; 5. 'Run' button.

- 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 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 menu)
- Data file:**
 - TYPE_DATA: Thermo_Orbitrap (dropdown menu)
 - PATH_DATA: D:\X\Demo_dataset\E480-QC-LC1.raw (text input)
- Analysis result:**
 - TYPE_ANALYSIS_RESULT: Spectronaut (dropdown menu)
 - PATH_ANALYSIS_RESULT: D:\X\Demo_dataset\Spectronaut18_MSCohortReport.tsv (text input)
 - THRESHOLD_FDR: 0.01 (text input)
- Setting:**
 - THRESHOLD_PEAK_WIDTH_TAILING: 1 (text input)
 - THRESHOLD_INVALID_ACQUIRING_SCAN: 100 (text input)
 - FLAG_ANALYZE_FEATURE: Yes (dropdown menu)
- Export:**
 - PATH_EXPORT: D:\X\Demo_dataset (text input)

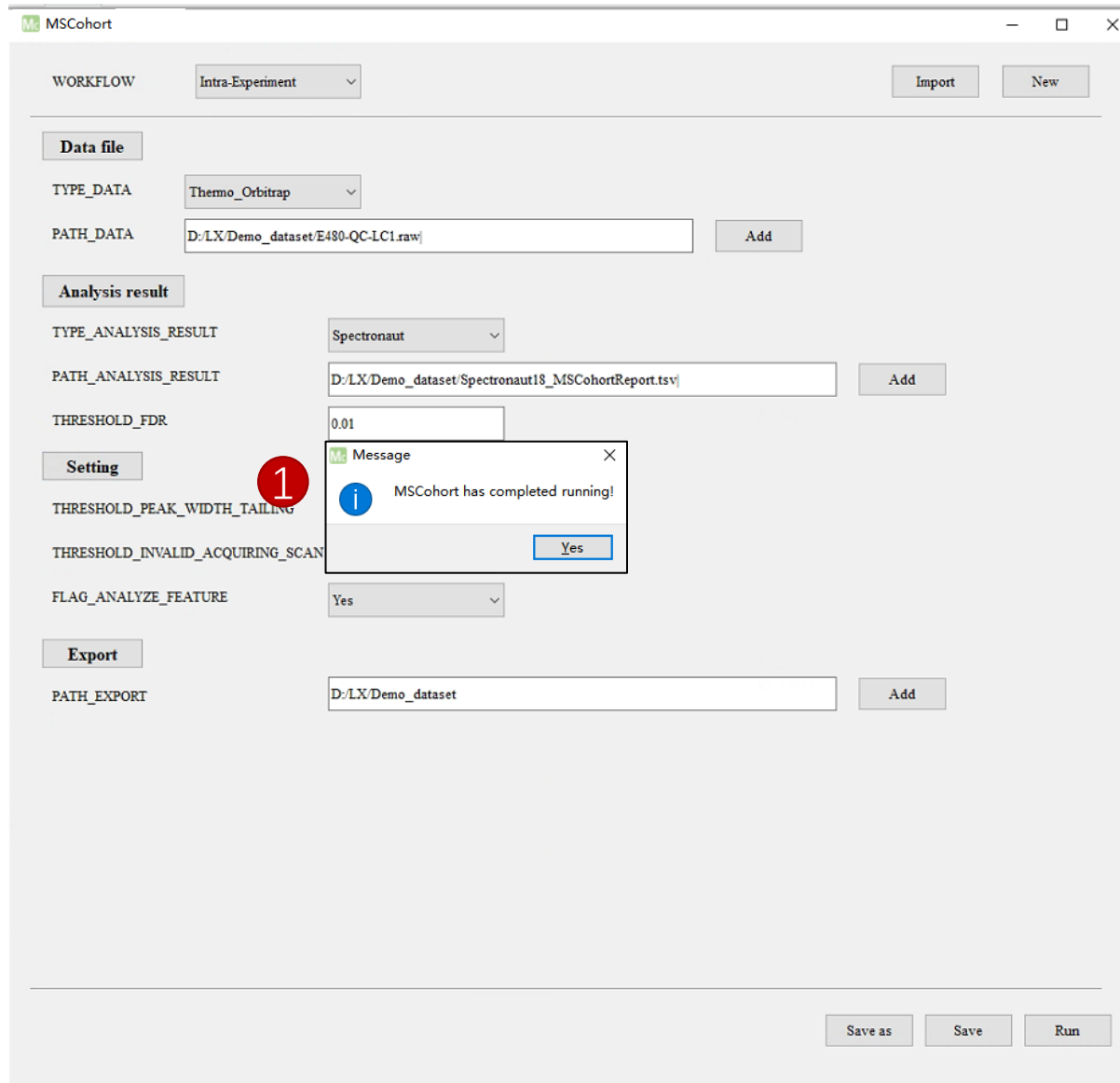
Buttons: Import, New, Add, Save as, Save, Run.

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.
Do not save the same PATH_DATA under a same folder, the results will be **overwritten.**
- ② 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 corresponding to Cohort and view it.

2. MSCohort Results

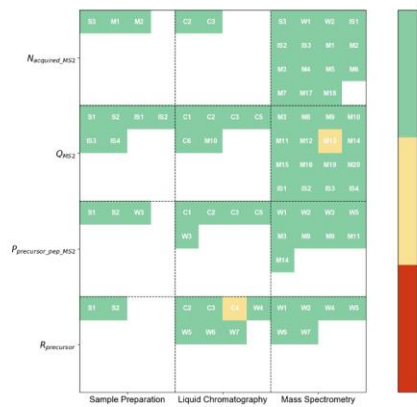
1

> 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 ;

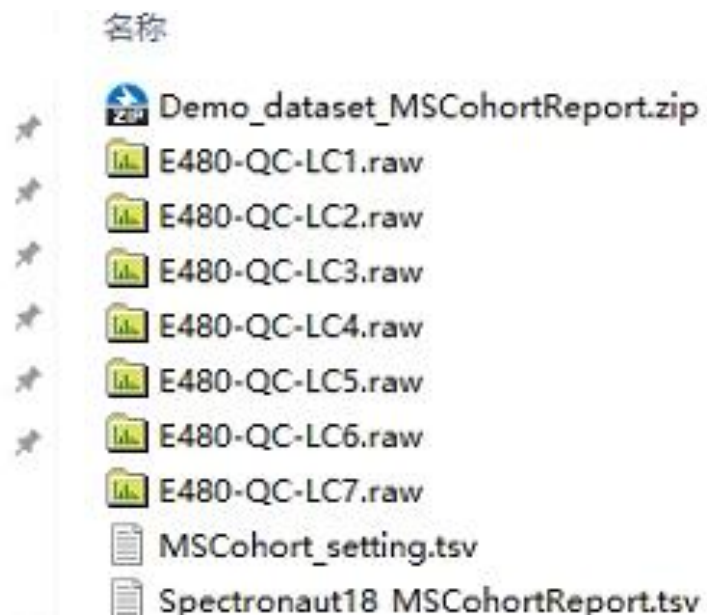
② Double clicking [Analysis_Report.html](#), the report will be preformed in the browser.

MSCohort manual for DIA inter- experiment analysis



Demo dataset

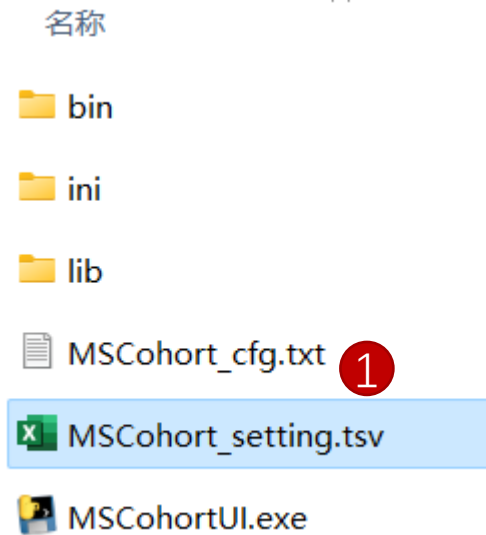
- DemoDatasetForDDA
- DemoDatasetForDIA**
- DemoDatasetForPRM



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

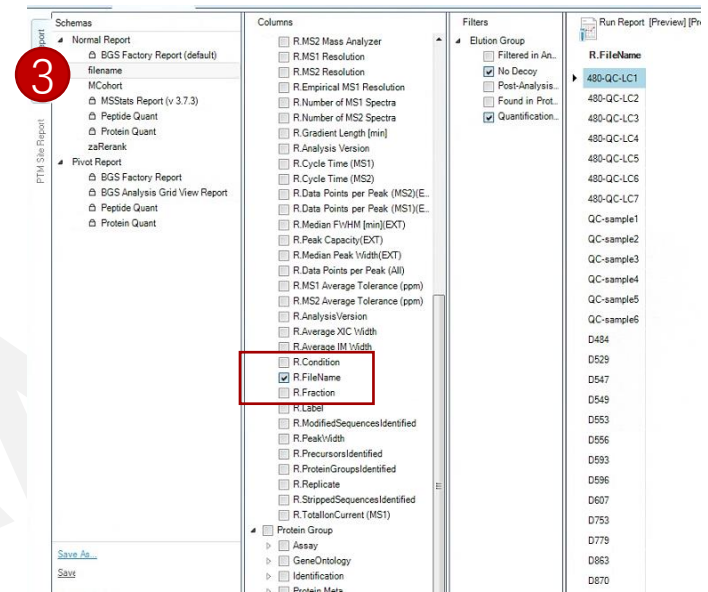
Inter-experiment analysis enables the systematic quality evaluation and low-quality experiments detection for multiple DIA experiments.

1. Preparation for MSCohort_setting.tsv file



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



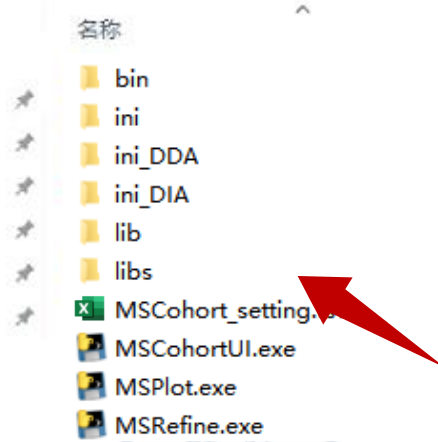
- ① Open [MSCohort_setting.tsv](#) with Excel;
- ② Fill the columns as the example file showed:
 - The first column is the Group Name.
 - The second column is the Raw Name, which is the same as [R.FileName](#) reported from Spectronaut.
 - The third column is the Raw Name for short, which is convenient 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

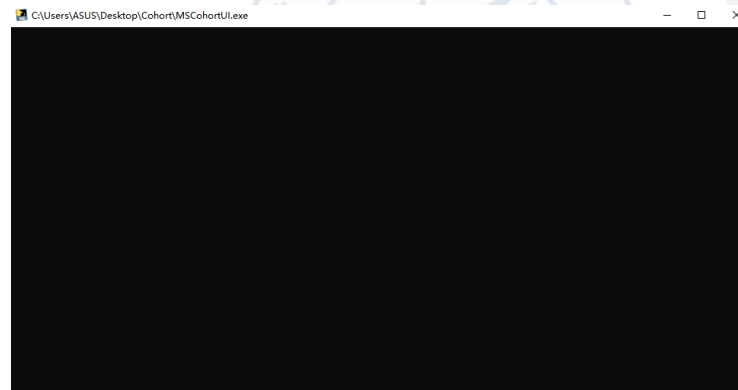
1. Analyzing with MSCohort

1

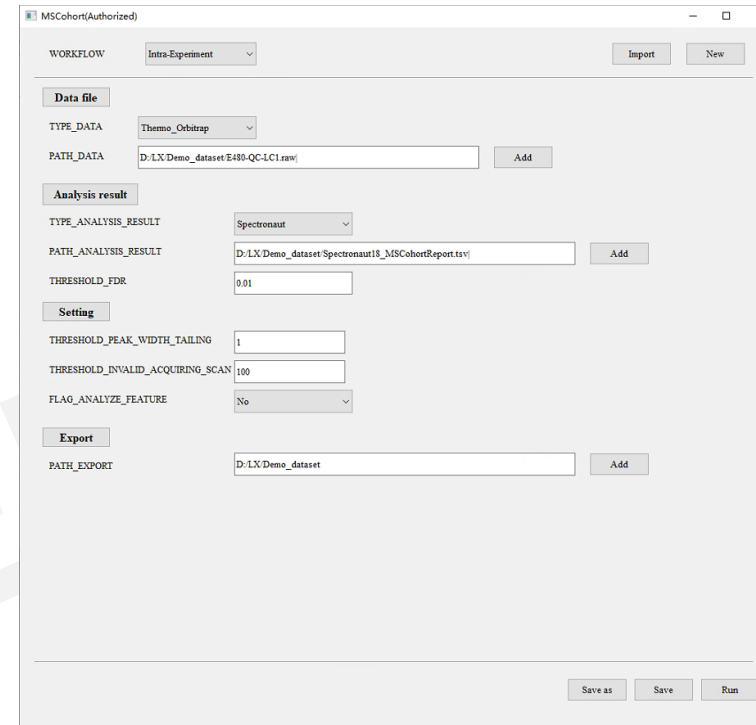
此电脑 > 新加卷 (D:) > LX > MSCohort_v202405



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.
- Analysis result: A dropdown menu showing "Spectronaut".
- THRESHOLD_FDR: A text input field containing "0.01".
- Setting: A section containing several settings:
 - TYPE_NORMALIZATION: A dropdown menu showing "DirectLFQ".
 - FLAG_OUTLIERS: A dropdown menu showing "2-SD".
 - FLAG_SHOW_ORDER: A dropdown menu showing "group series".
 - THRESHOLD_PEAK_WIDTH_TAILING: A text input field containing "1".
 - THRESHOLD_INVALID_ACQUIRING_SCAN: A text input field containing "100".
 - FLAG_ANALYZE_FEATURE: A dropdown menu showing "No".
- Export: A section containing:
 - PATH_EXPORT: A text input field containing "D:/LX/Demo_dataset" with an "Add" button.
- Buttons: "Import", "New", "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 placed over the interface: circle 1 is over the 'PATH_EXPORT' field, circle 2 is over the 'Save as' button, and circle 3 is over the 'Run' button.

MS Cohort

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

1 PATH_EXPORT: D:/LX/Demo_dataset [Add]

2 [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.

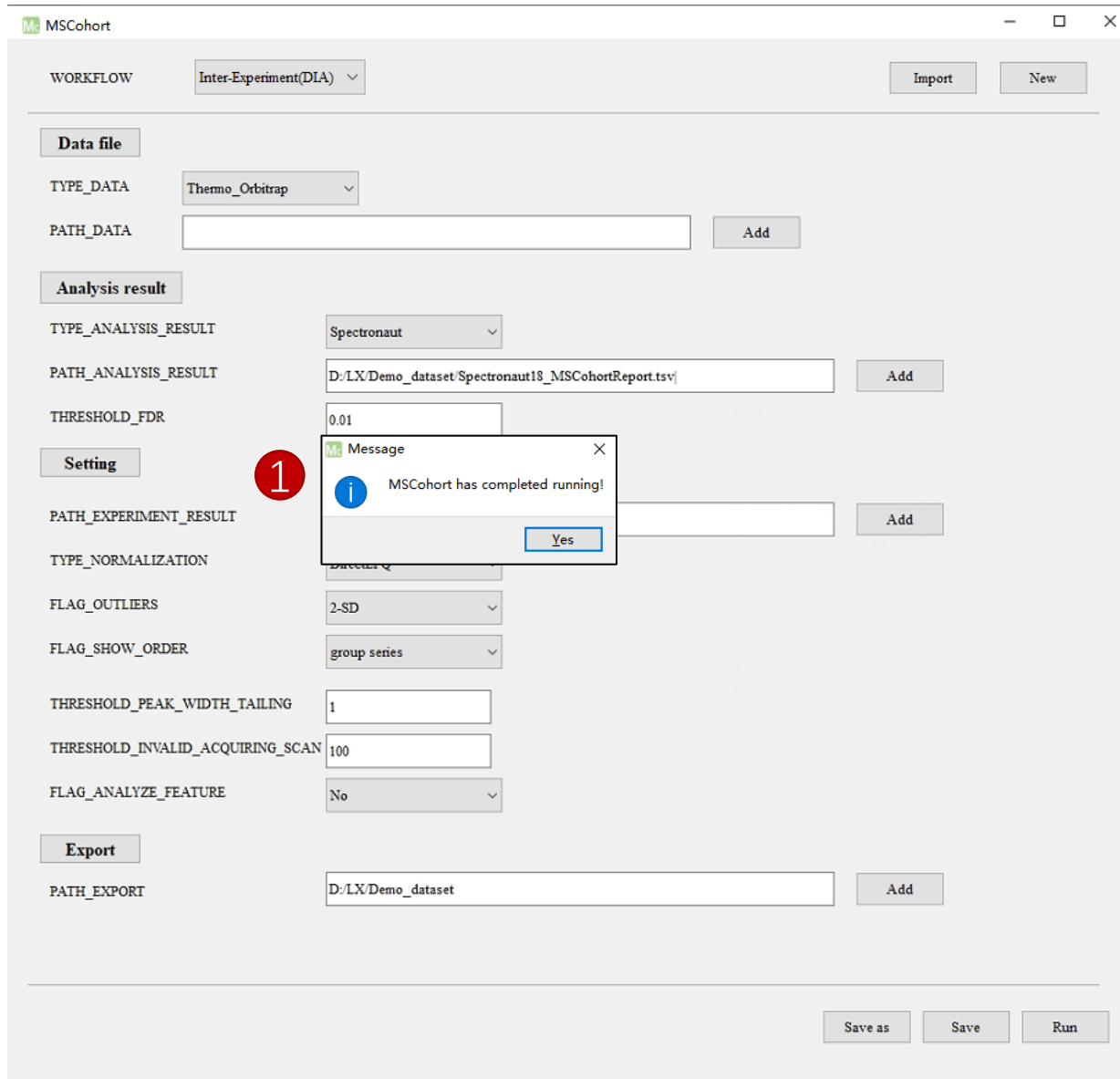
Do not save the different experiments results under a same folder, the results will be **overwritten.**

- 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

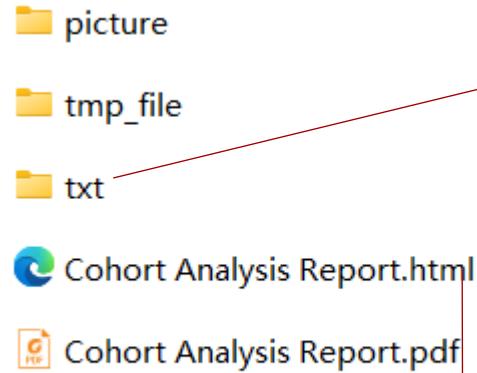


The screenshot displays the MSCohort software interface. A red circle with the number '1' is positioned next to a 'Message' popup window that reads 'MSCohort has completed running!' with a 'Yes' button. The interface includes several sections: 'Workflow' (Inter-Experiment(DIA)), 'Data file' (Thermo_Orbitrap), 'Analysis result' (Spectronaut), 'Setting' (2-SD, group series), and 'Export' (D:/LX/Demo_dataset). The 'Path' fields are populated with 'D:/LX/Demo_dataset/Spectronaut18_MSCohortReport.tsv' and 'D:/LX/Demo_dataset'. The 'Threshold' fields are set to 0.01, 1, and 100. The 'Flag' fields are set to 'No' and 'group series'. The 'Run' button is visible at the bottom right.

- 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.

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

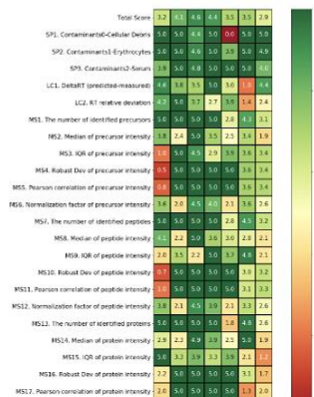
- ① The MSCohort results ;
- ② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.
- ③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

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.

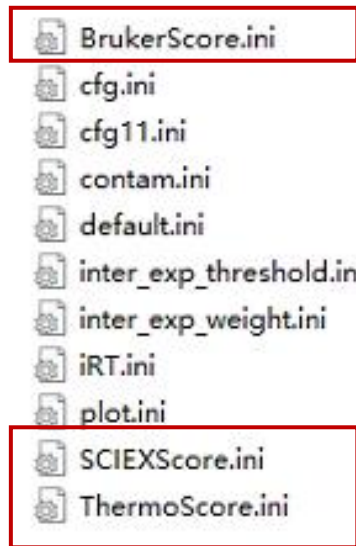


Notes for modifying the scoring criteria



1. Modifying the scoring criteria for intra-experiment analysis

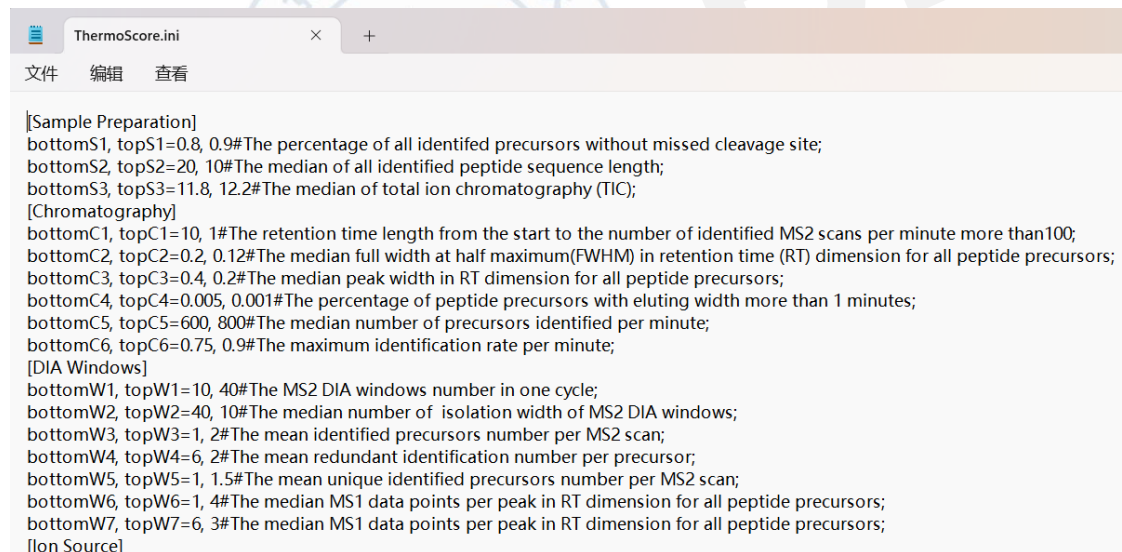
1



① Open the **ini** folder, there will be three parameter files related to intra-experiment scoring: ThermoScore.ini, BrukerScore.ini, SCIEXScore.ini;

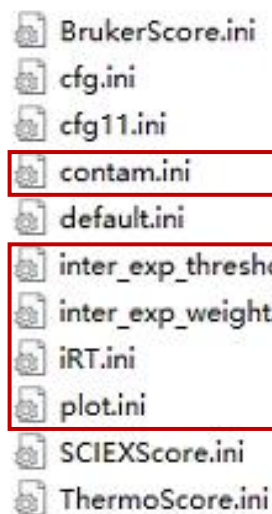
② Click the file to modify the second-level scoring threshold. If you do not modify it, it will be the default value. Closing the file and running the software, it will score according to the standard you set.

2

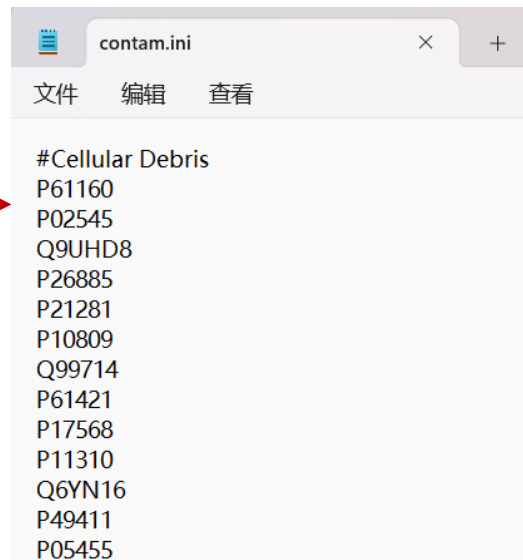


1. Modifying the scoring criteria for inter-experiment analysis

1



2

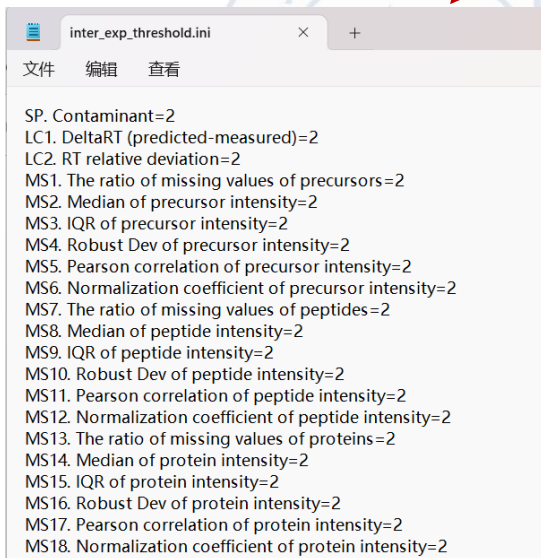


① Open the **ini** folder, there will be five parameter files related to inter-experiment scoring: **contam.ini**, **inter_exp_threshold.ini**, **inter_exp_weight.ini**, **iRT.ini**, **plot.ini**;

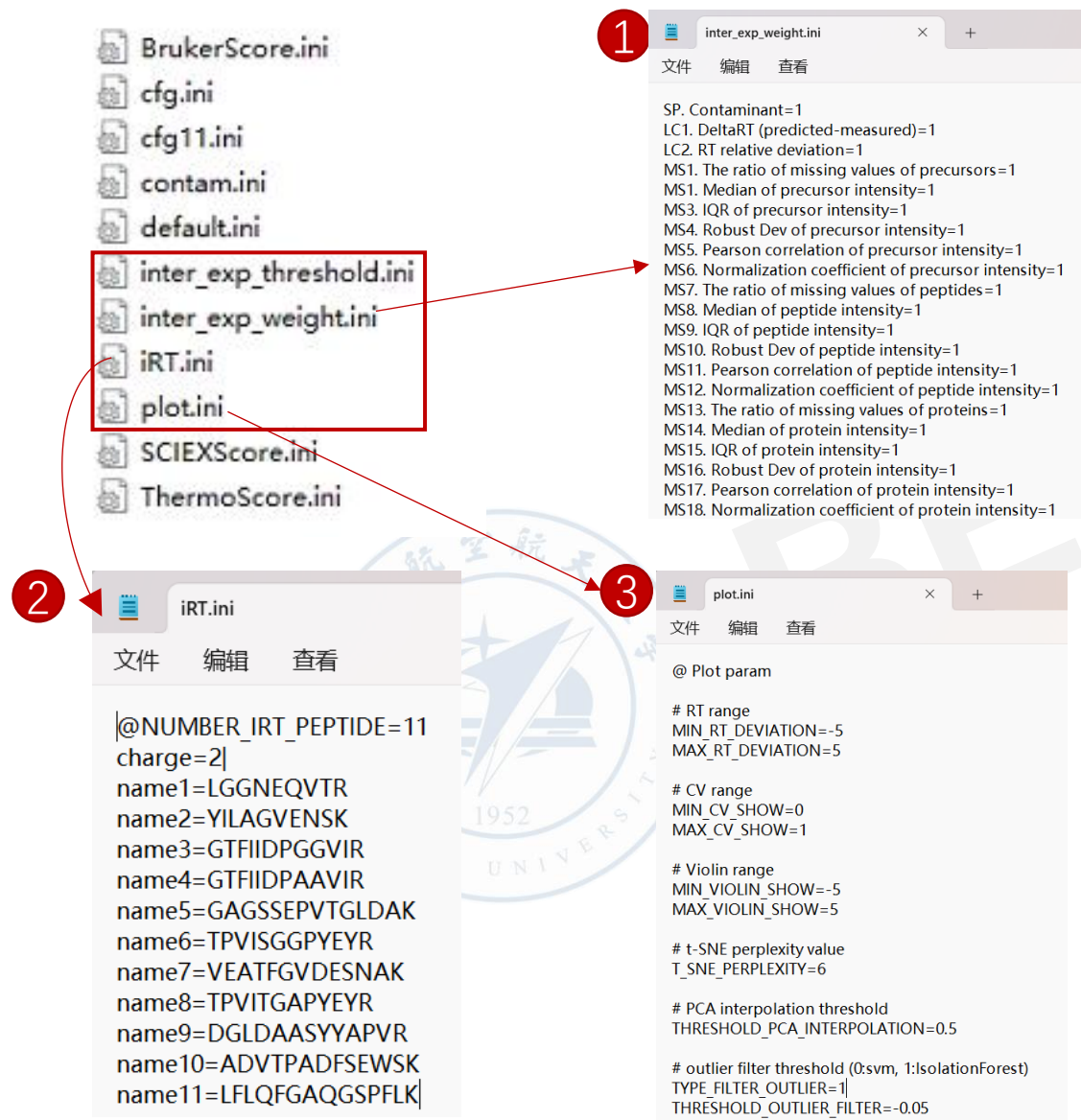
② Open the **contam.ini**, users can modify the list of contaminating proteins, # begins with contaminants category, enter the list of contaminating proteins under this category below.

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. Closing and saving the file, it will score according to the standard you set.

3



1. Modifying the scoring criteria for inter-experiment analysis

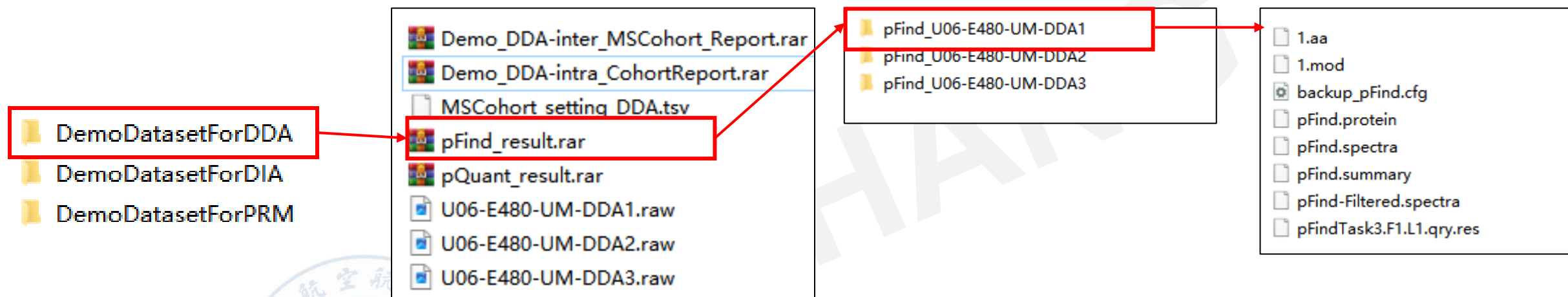


- ① Open the `inter_exp_weight.ini`, users can modify the scoring weight for each metric. We initially defined the weight as 1 for each metric. If you do modify it, closing and saving the file, it will score according to the standard you set.
- ② Open the `iRT.ini`, users can modify the list of iRT peptide 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



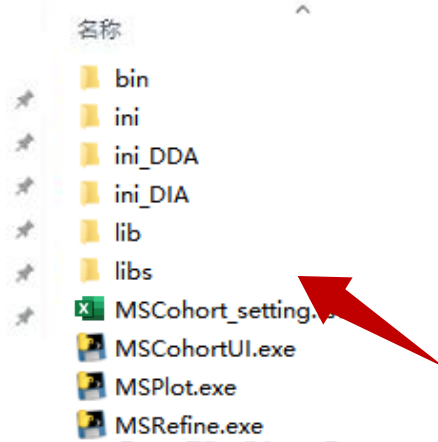
Taking the U06-E480-UM-DDA1.raw file as an example to demonstrate the workflow of DDA intra-experiment analysis.

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

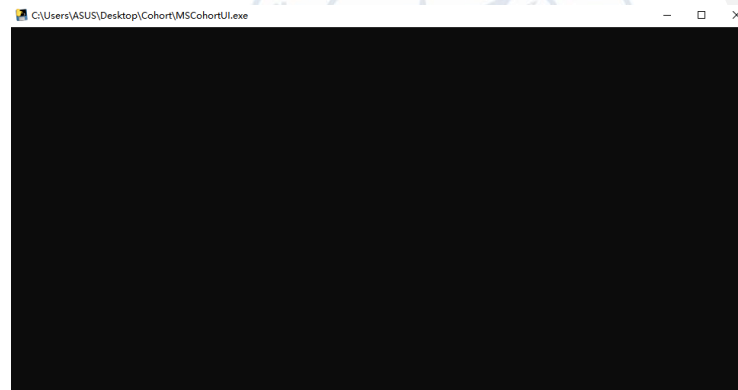
1. Analyzing with MSCohort

1

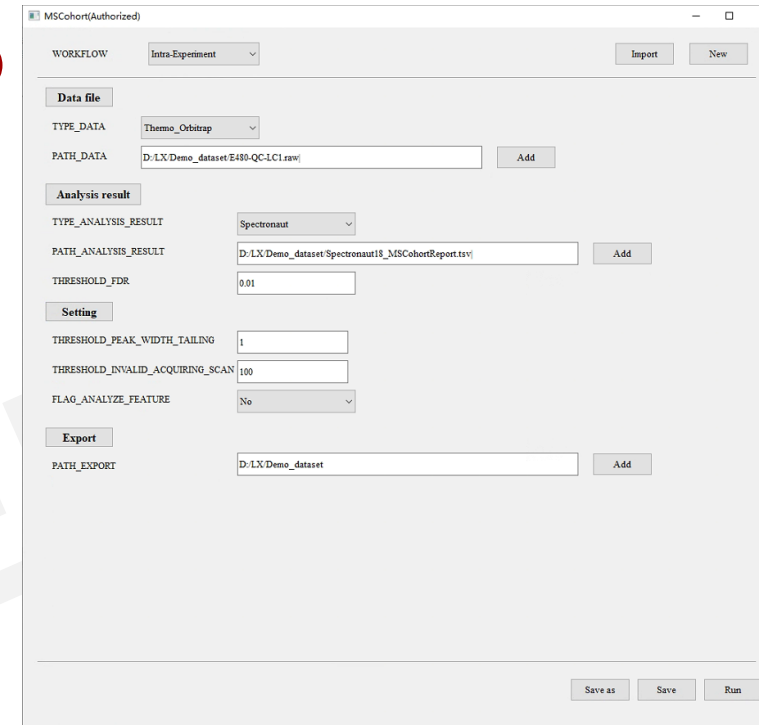
此电脑 > 新加卷 (D:) > LX > MSCohort_v202405



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: Intra-Experiment(DDA) (dropdown menu is open showing options: Intra-Experiment(DDA), Inter-Experiment(DIA), Intra-Experiment(DIA), Inter-Experiment(DDA), Intra-Experiment(DDA), Inter-Experiment(PRM))
- 2** TYPE_DATA: Thermo_Orbitrap (dropdown menu is open showing options: Thermo_Orbitrap, Bruker_timsTOF, SCIEX_ZenoTOF)
- 3** PATH_DATA: sers/ASUS/Desktop/paper_test/test/DemoDatasetForDDA/U06-E480-UM-DDA1.raw (text field, Add button)
- 4** PATH_ANALYSIS_RESULT: DA_test_data_new/pFind_result/pFind_U06-E480-UM-DDA1/pFind-Filtered.spectra (text field, Add button)
- 5** PATH_EXPORT: C:/Users/ASUS/Desktop/paper_test/test/DemoDatasetForDDA (text field, Add button)

Other visible fields include:

- THRESHOLD_FDR: 0.01
- THRESHOLD_PEAK_WIDTH_TAILING: 1
- THRESHOLD_INVALID_ACQUIRING_SCAN: 100
- FLAG_ANALYZE_FEATURE: Yes
- FLAG_ANALYZE_PIF: Yes
- FLAG_ANALYZE_TAG: Yes

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 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

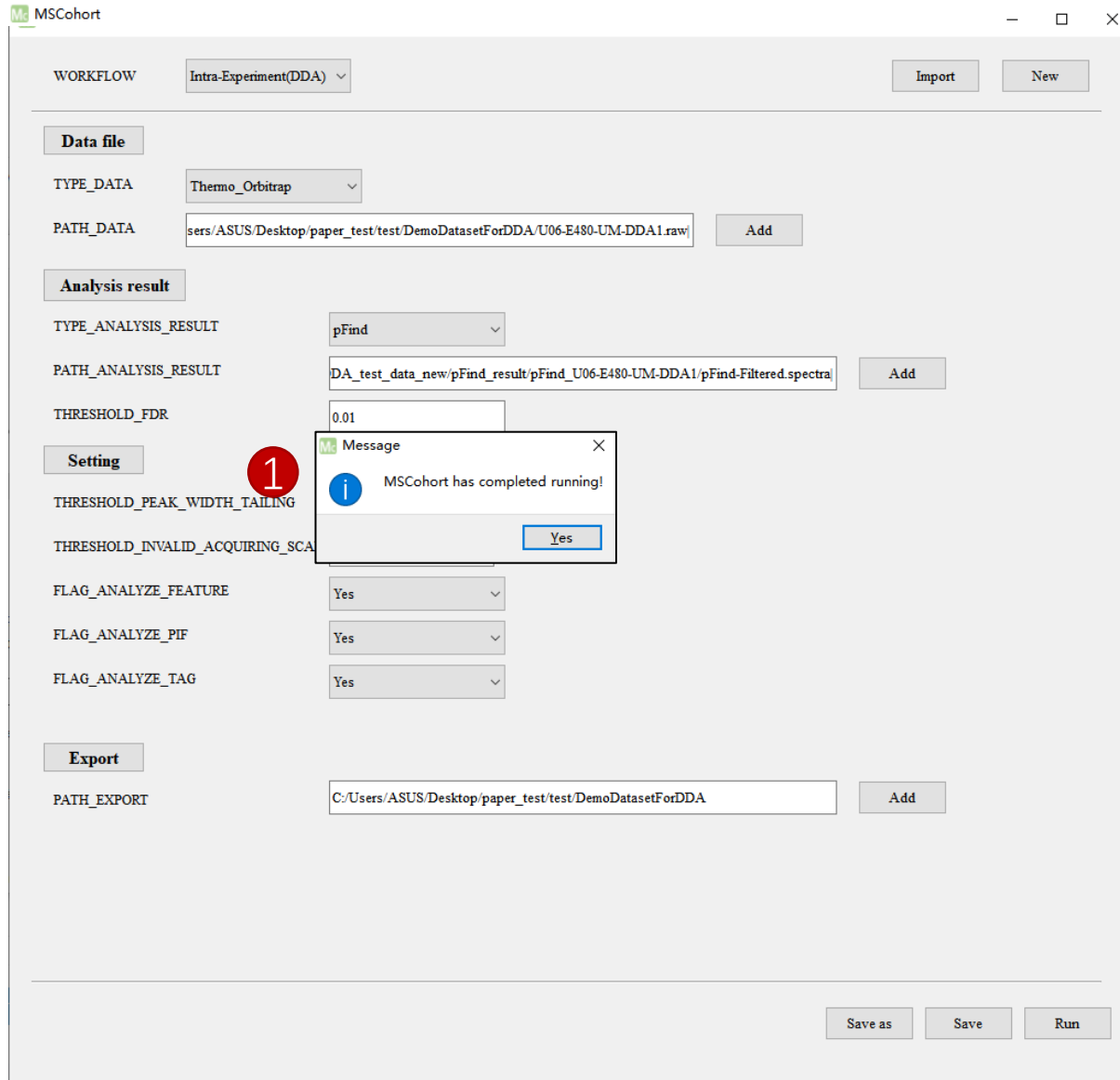
The screenshot shows the MSCohort software interface with the following sections and annotations:

- WORKFLOW:** Intra-Experiment(DDA) (dropdown)
- 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]
- Buttons:** Save as, Save, Run (at the bottom right)

Annotations 4 and 5 point to the 'Save as' and 'Run' buttons 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 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



- 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

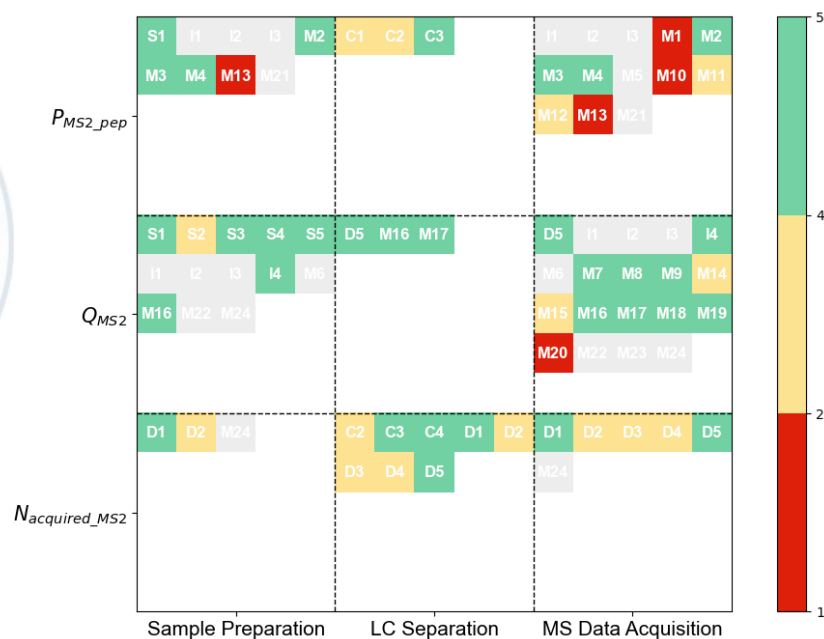
1

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 ;

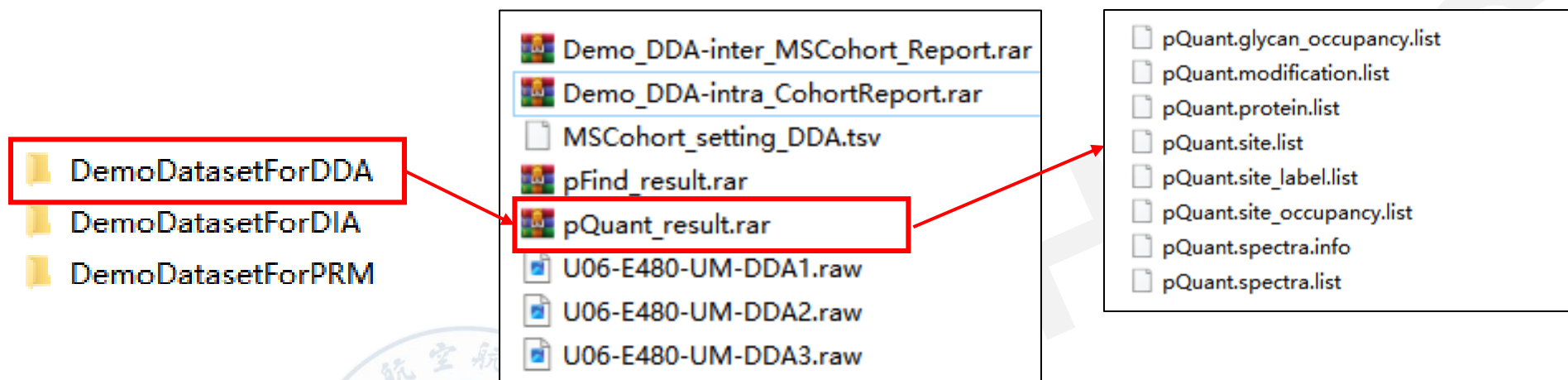
② Double clicking [Analysis_Report.html](#), the report will be preformed in the browser.



MSCohort manual for DDA inter- experiment analysis



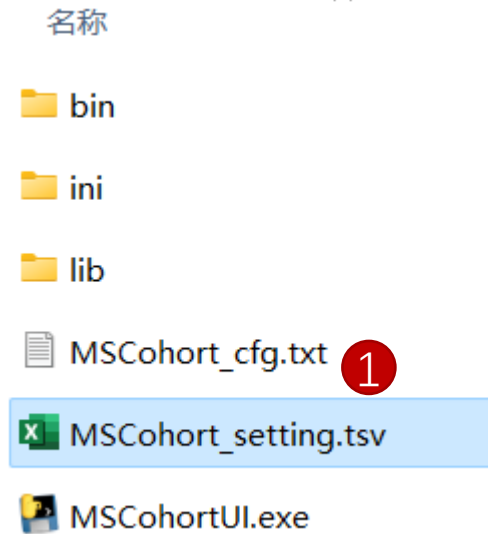
Demo dataset



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



2

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

	A	B	C	D	E	F	G
1	ProteinName	GroupType	GroupID	Number	Empty_Separate	JD_06232014_sample1_A	JD_06232014_sample1_B
2	sp P1909 Leading	sp P1909		415		1.18038E+11	1.38222E+11
3	sp P0725 Leading	sp P0725		333		65636351994	75009105469
4	sp P0714 Leading	sp P0714		365		1.15039E+11	1.29925E+11
5	sp Q0095 Leading	sp Q0095		266		56144848386	60354899605
6	sp P3389 Leading	sp P3389		237		35768011003	40616103258
7	sp P0054 Leading	sp P0054		675		1.79459E+12	2.05969E+12
8	sp P0092 Leading	sp P0092		662		2.7155E+12	3.10134E+12

① Open MSCohort_setting.tsv with Excel;

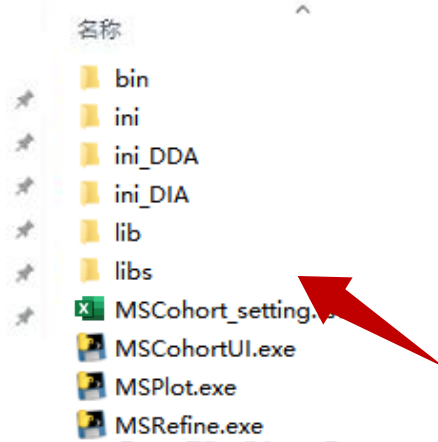
② Fill the columns as the example file (MSCohort_setting_DDA.tsv) showed or alternatively, directly use the MSCohort_setting_DDA.tsv file:

- 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.

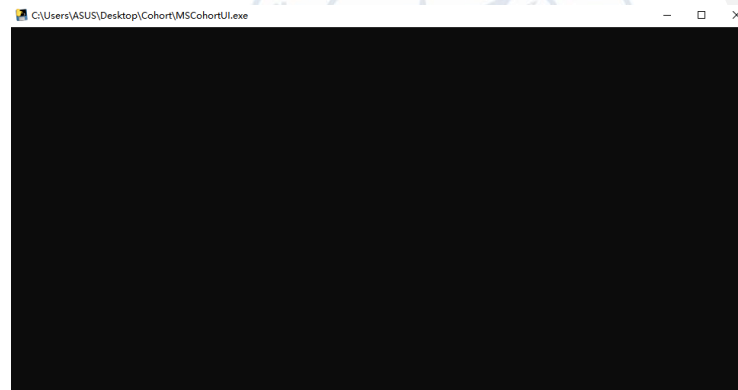
1. Analyzing with MSCohort

1

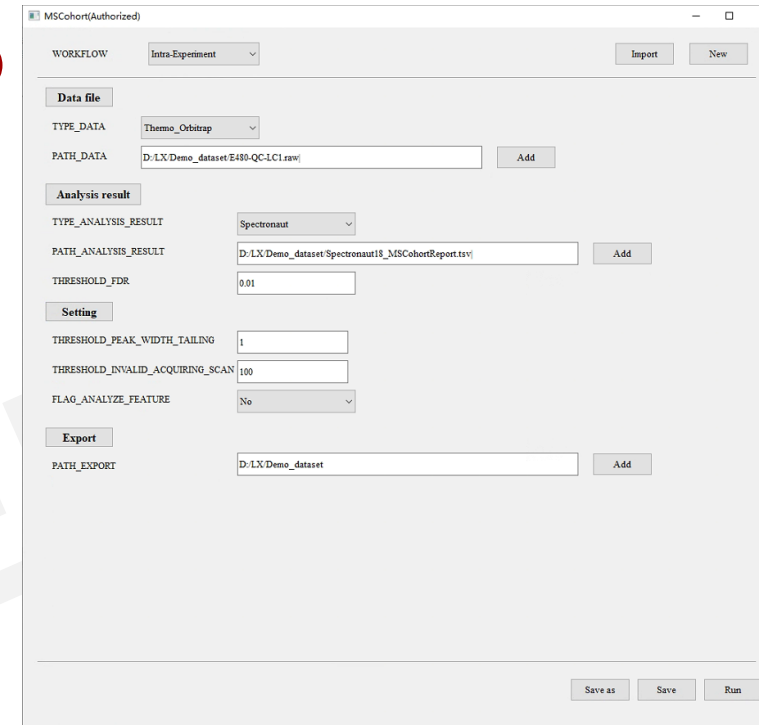
此电脑 > 新加卷 (D:) > LX > MSCohort_v202405



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**;
- ④ Click **Add** to select the **MSCohort_setting.tsv** into the **PATH_EXPERIMENT_RESULT**;

2. Analyzing with MSCohort

Workflow: Inter-Experiment(DDA) [Import] [New]

Data file

TYPE_DATA: Thermo_Orbitrap

PATH_DATA: [Add]

Analysis result

TYPE_ANALYSIS_RESULT: pQuant

PATH_ANALYSIS_RESULT: _test/DDA_test_data_new/pQuant_result/UQ_20241024_013044/pQuant.protein.list [Add]

THRESHOLD_FDR: 0.01

Setting

PATH_EXPERIMENT_RESULT: op/paper_test/DDA_test/DDA_test_data_new/COhort_result/MSCOhort_setting.tsv [Add]

1 TYPE_NORMALIZATION: DirectLFQ

2 FLAG_OUTLIERS: 2-SD

3 FLAG_SHOW_ORDER: group series

4 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 [Add]

[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.
- 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

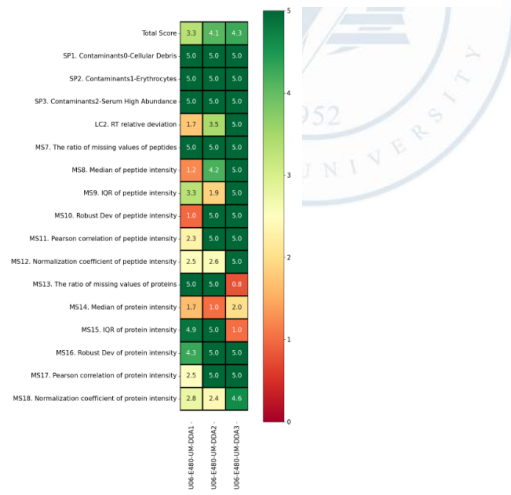
- ① The MSCohort results ;
- ② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.
- ③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

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)

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

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

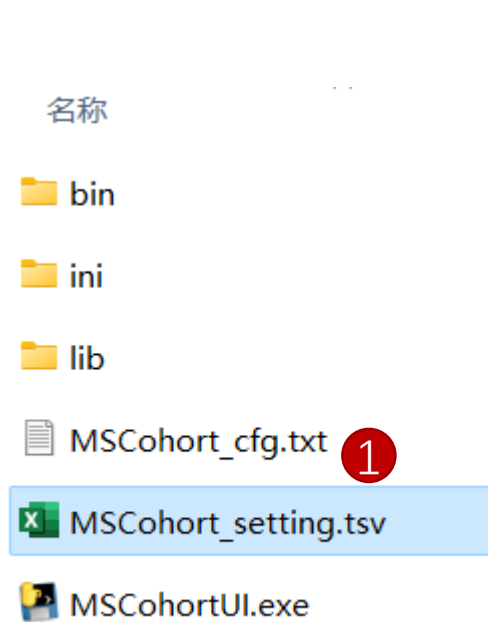
MSCohort_setting_Skyline.tsv

	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

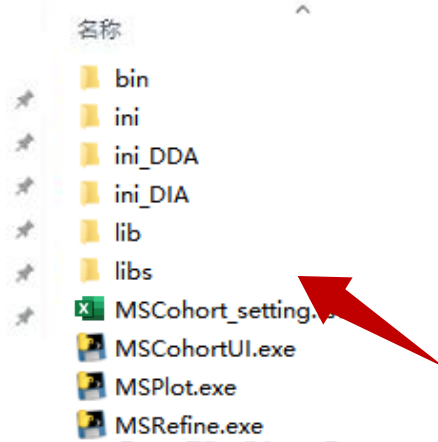
- ① Open [MSCohort_setting.tsv](#) with Excel;
- ② Fill the columns as the example file (MSCohort_setting_Skyline.tsv) showed or alternatively, directly use the MSCohort_setting_Skyline.tsv file:
 - The first column is the Group Name.
 - The second column is the Raw Name, which is the same as [sample](#) 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

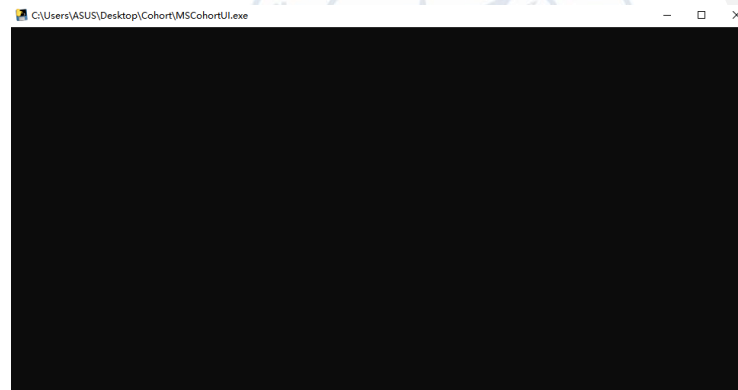
1. Analyzing with MSCohort

1

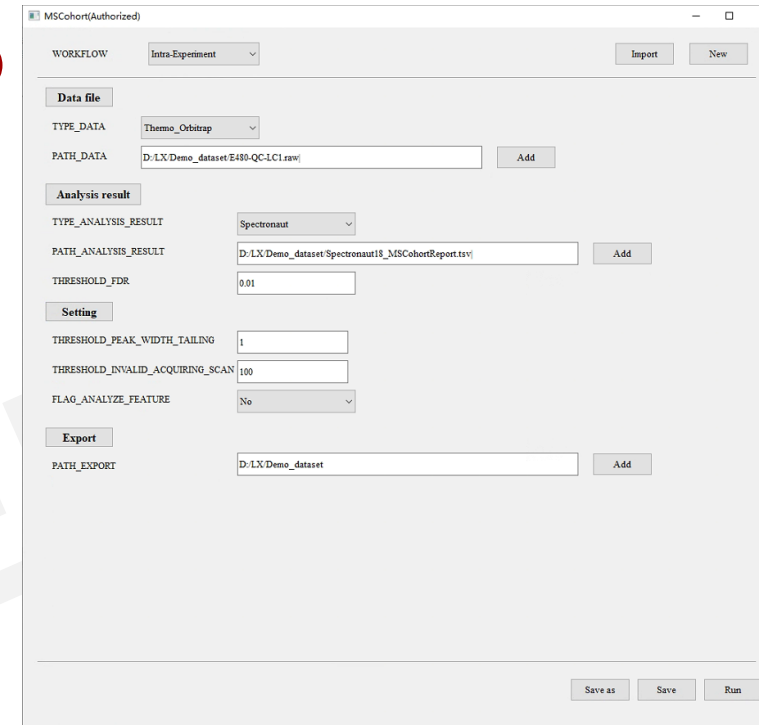
此电脑 > 新加卷 (D:) > LX > MSCohort_v202405



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:

- 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:** Skyline
 - PATH_ANALYSIS_RESULT:** E:/RAW/PRM/skyline_result.xlsx (Annotated 4)
 - THRESHOLD_FDR:** 0.01
- Setting:**
 - PATH_EXPERIMENT_RESULT:** E:/RAW/PRM/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.

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(PRM) (dropdown)
- Data file:**
 - TYPE_DATA: Thermo_Orbitrap (dropdown)
 - PATH_DATA: (text input) Add
- Analysis result:**
 - TYPE_ANALYSIS_RESULT: Skyline (dropdown)
 - PATH_ANALYSIS_RESULT: E:/RAW/PRM/skyline_result.xlsx Add
 - THRESHOLD_FDR: 0.01 (text input)
- Setting:**
 - PATH_EXPERIMENT_RESULT: E:/RAW/PRM/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, 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.



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 box) [Add]
- Analysis result:**
 - TYPE_ANALYSIS_RESULT: Skyline (dropdown)
 - PATH_ANALYSIS_RESULT: E:/RAW/PRM/skyline_result.xlsx [Add]
 - THRESHOLD_FDR: 0.01 (text box)
- Setting:**
 - PATH_EXPERIMENT_RESULT: E:/RAW/PRM/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: **1** (Run), **2** (Save), and **3** (Save).

- 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.

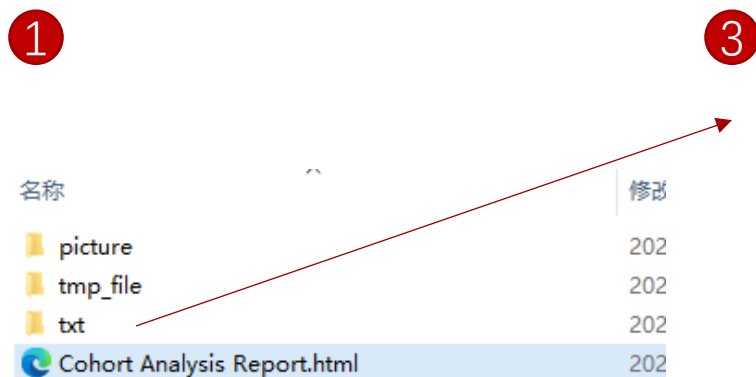
Do not save the different experiments results under a same folder, the results will be **overwritten.**

- 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.

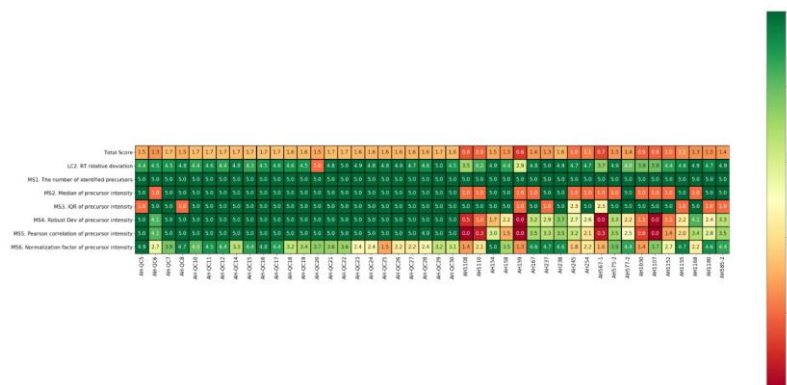
3. MSCohort Results



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.

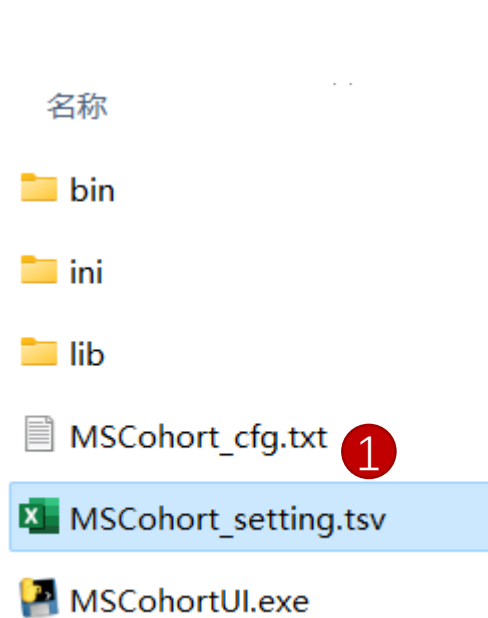


- INFO0_Experiment_Outlier_Score.txt
- INFO0_Identification_Count.txt
- INFO0_Inter_Experiment_Scores.txt
- INFO0_Inter_Experiment_values.txt
- INFO0_Intra_Experiment_Scores.txt
- INFO0_Intra_Experiment_values.txt
- INFO1_Pro0_MissingValue_Count.txt
- INFO2_Pro1_Intensity.txt
- INFO3_Pro2_Origin_Intensity.txt
- 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:

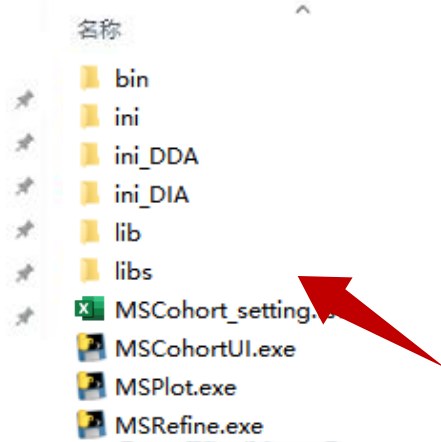
- The first column is the Group Name.
- The second column is the Raw Name, which is the same as [R.FileName](#) reported from SpectroDive.
- The third column is the Raw Name for short, which is convenient for display in large MSCohort report. This column could also be the same as Raw Name column.
- The 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

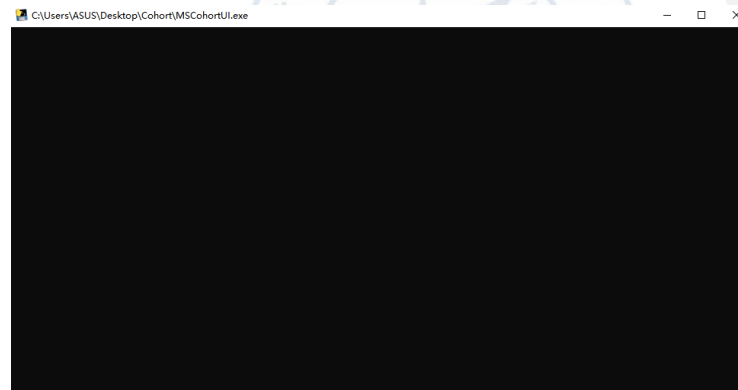
1. Analyzing with MSCohort

1

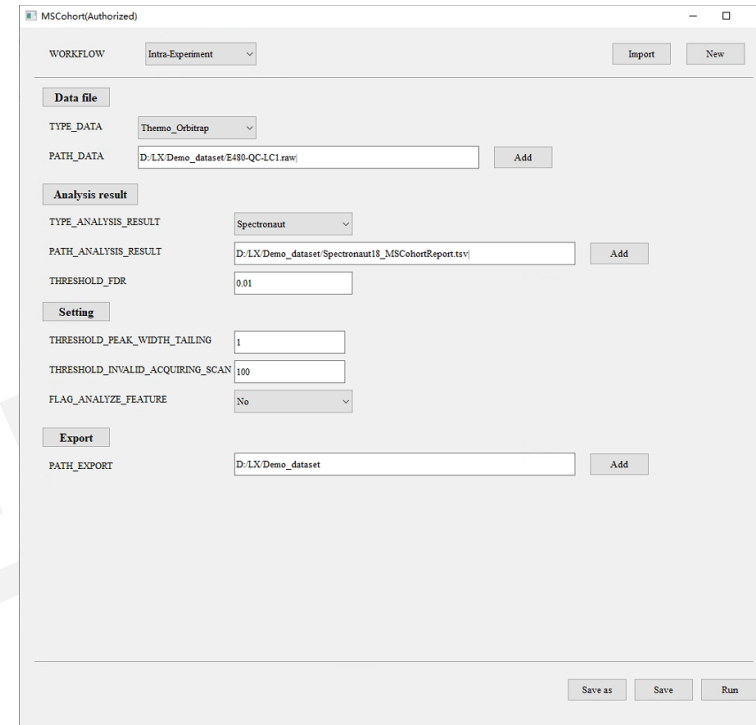
此电脑 > 新加卷 (D:) > LX > MSCohort_v202405



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:** (Empty field)
- TYPE_DATA:** Thermo_Orbitrap (Annotated 2)
- 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.

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 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)
- Data file:**
 - TYPE_DATA: Thermo_Orbitrap
 - PATH_DATA: (empty field)
- Analysis result:**
 - TYPE_ANALYSIS_RESULT: SpectroDive
 - PATH_ANALYSIS_RESULT: E:/RAW/PRM/spectroDive_result/SpectroDive_result.txt
 - THRESHOLD_FDR: 0.01
- Setting:**
 - PATH_EXPERIMENT_RESULT: E:/RAW/PRM/spectroDive_result/MSCohort_setting.tsv
 - TYPE_NORMALIZATION: DirectLFQ
 - FLAG_OUTLIERS: 2-SD
 - FLAG_SHOW_ORDER: group series
- Export:**
 - PATH_EXPORT: E:/RAW/PRM

Red circles highlight the following elements:

- 1:** The PATH_EXPORT field in the Export section.
- 2:** The "Save as" button at the bottom right.
- 3:** The "Run" button at the bottom right.

- 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

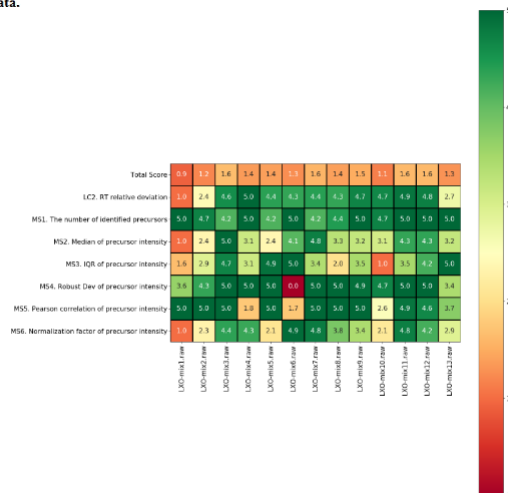
Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

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!

