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)

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

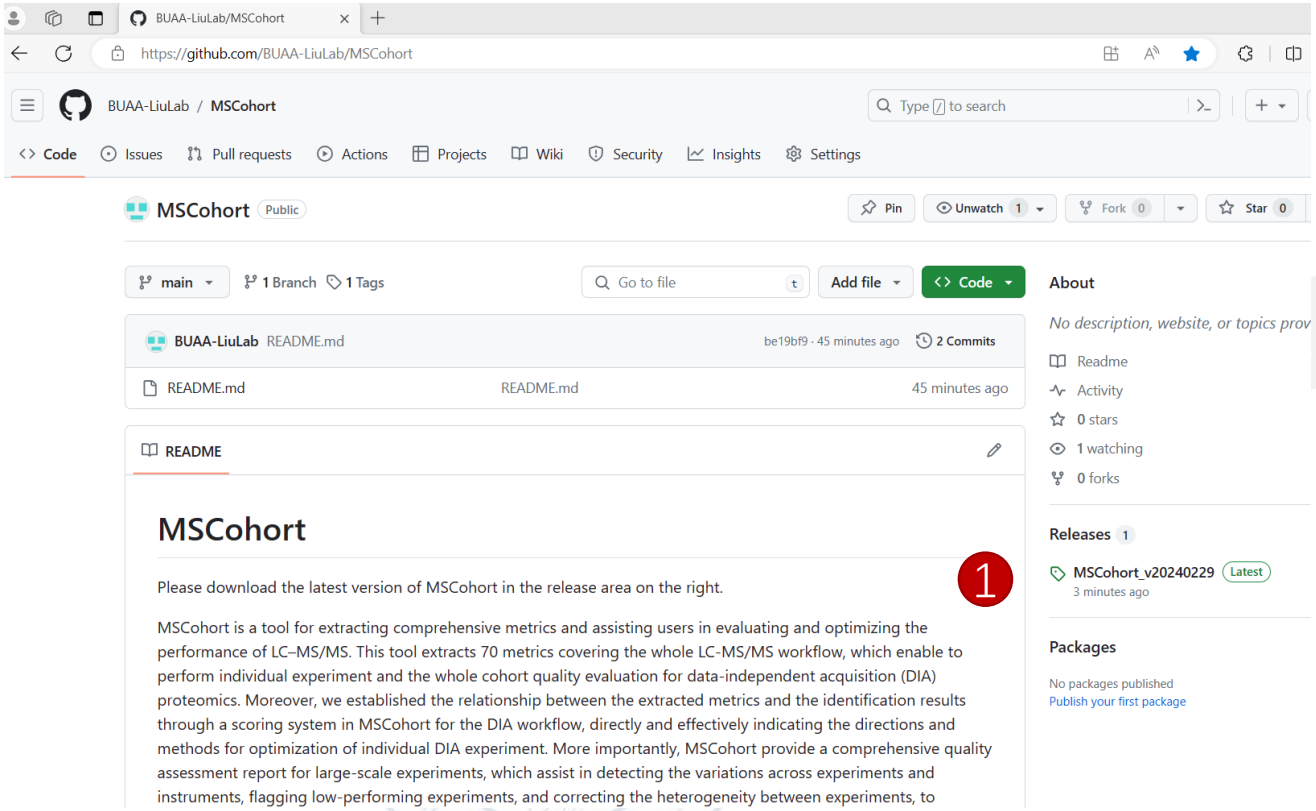
- Demo DIA folds contains raw files of 7 human urine QC DIA data from Orbitrap Exploris 480, Spectronaut analysis results and MSCohort report results.
- Demo DDA folds contains raw files of 3 human urine QC DDA data from Orbitrap Exploris 480, pFind/pQuant analysis results and MSCohort report results.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

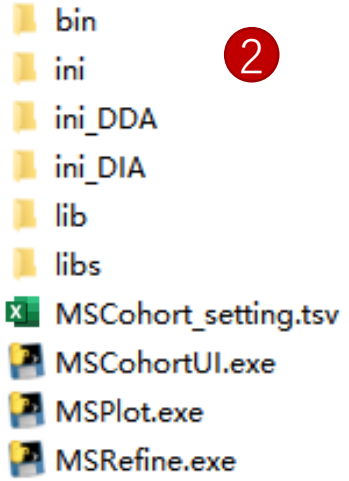
CONTENTS

- 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

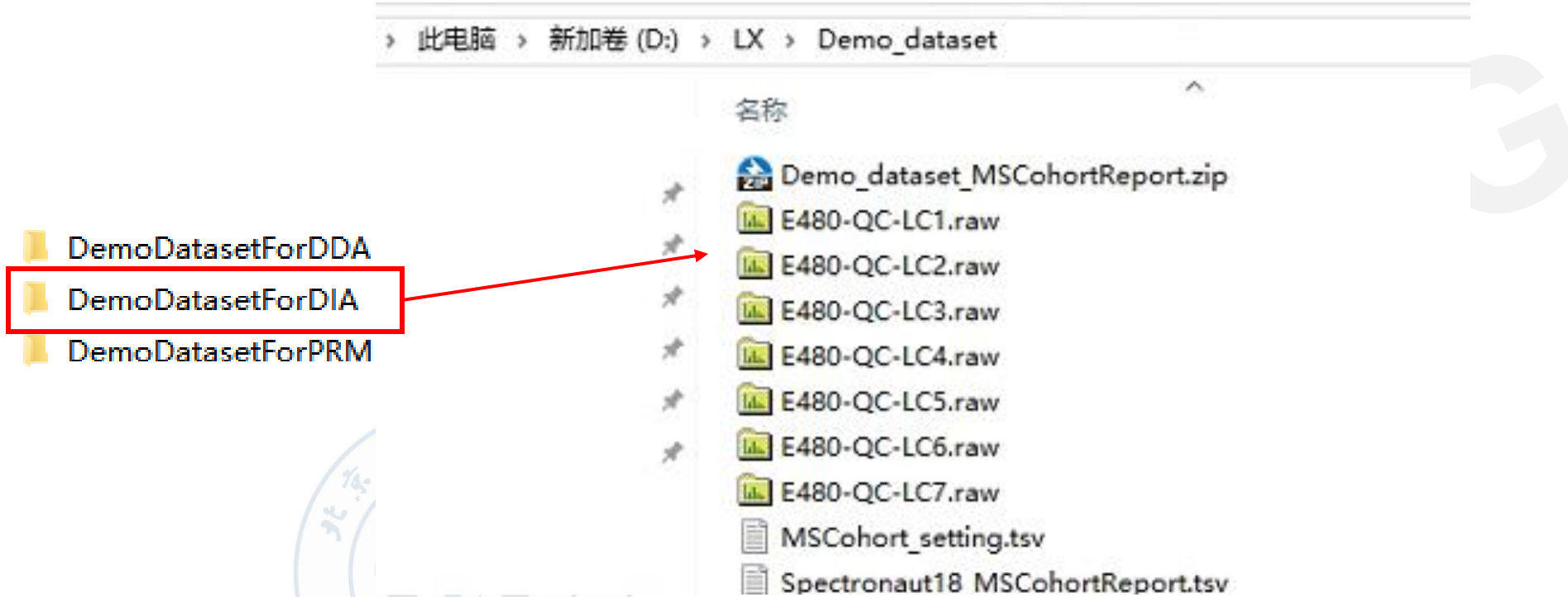
https://github.com/BUAA-LiuLab/MSCohort

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

② The unzipped MSCohort file.

MSCohort manual for DIA intra-experiment analysis

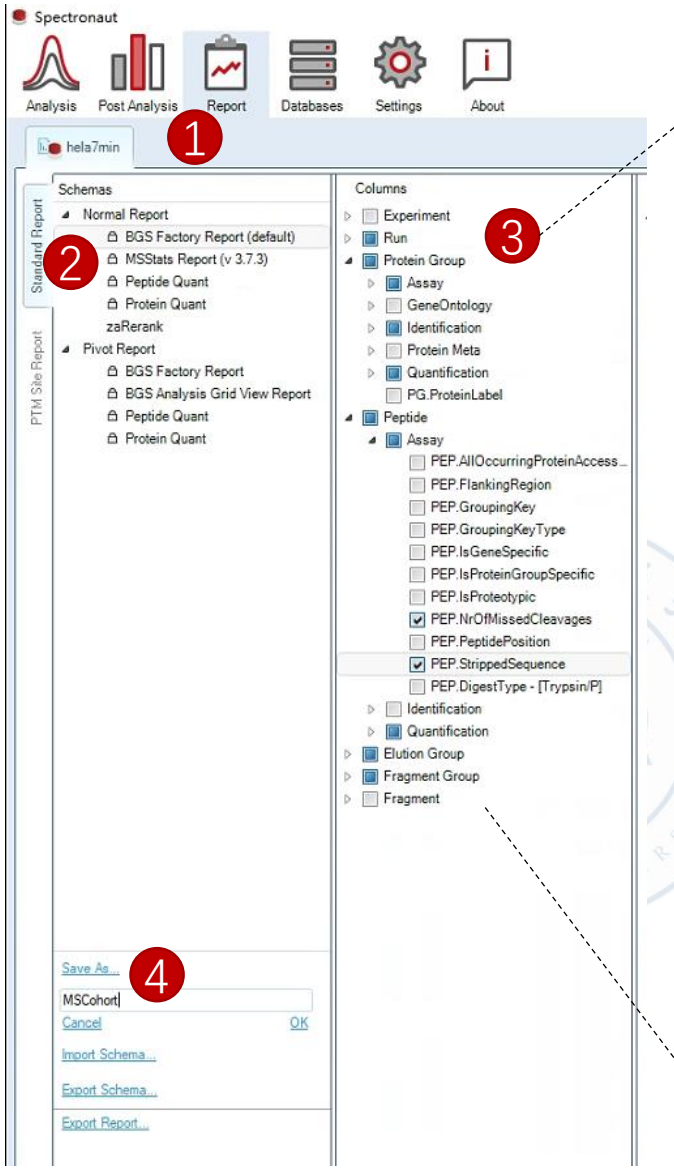
Demo dataset



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

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)

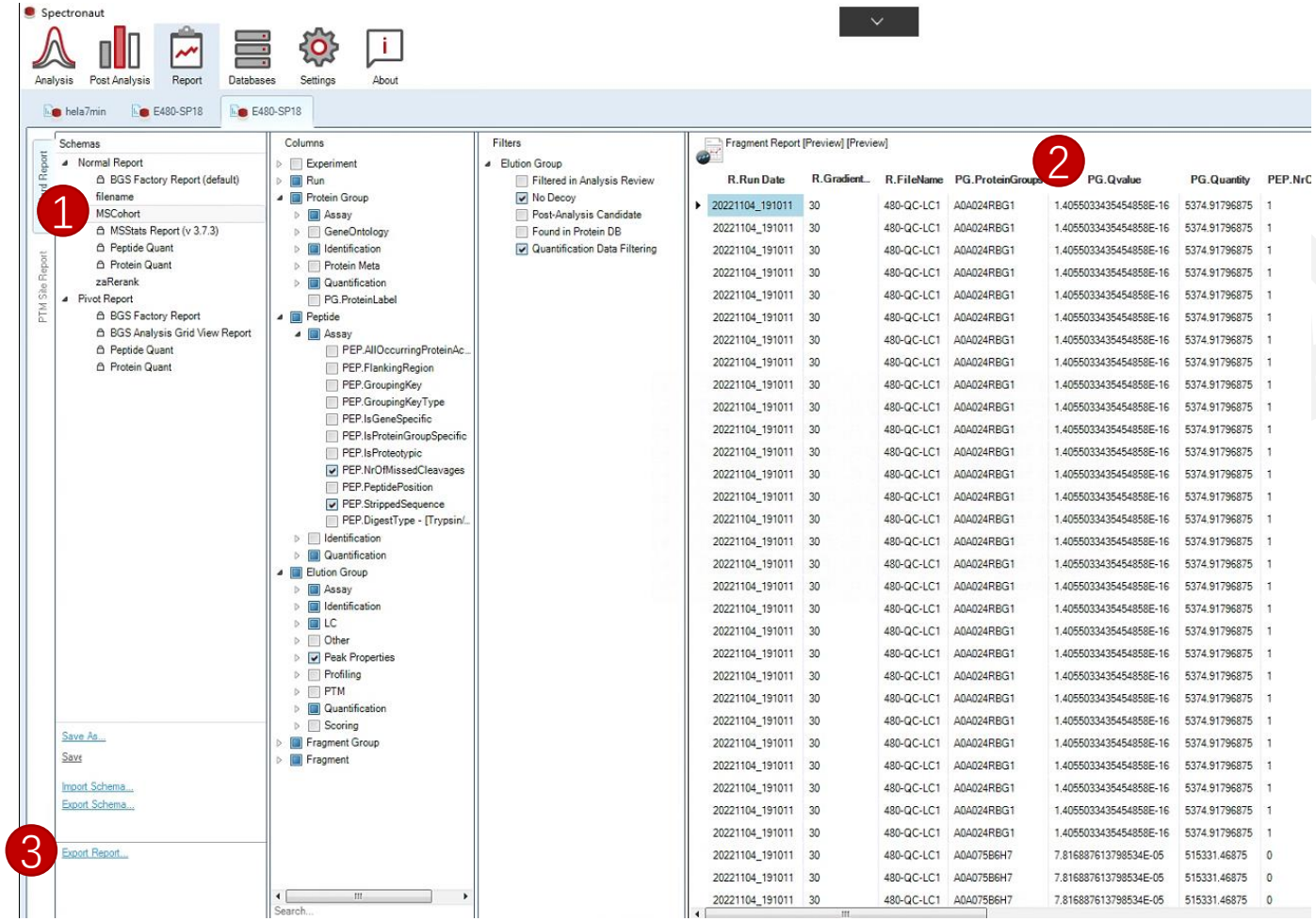
① Choose Report Perspective in Spectronaut;

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".

Export MSCohort report from Spectronaut

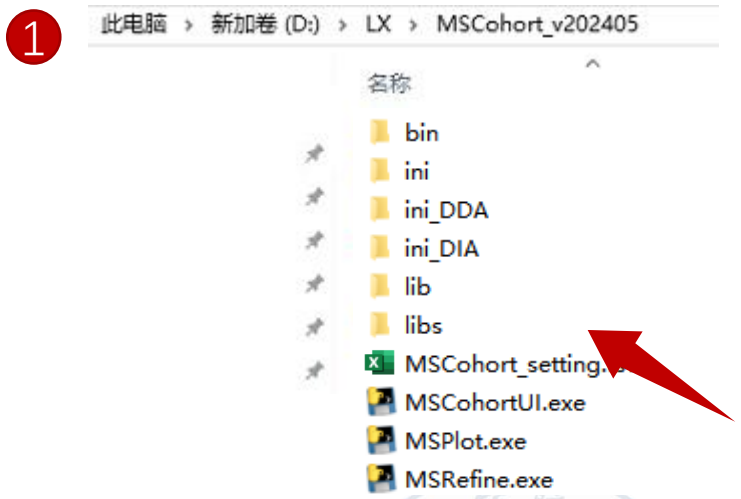


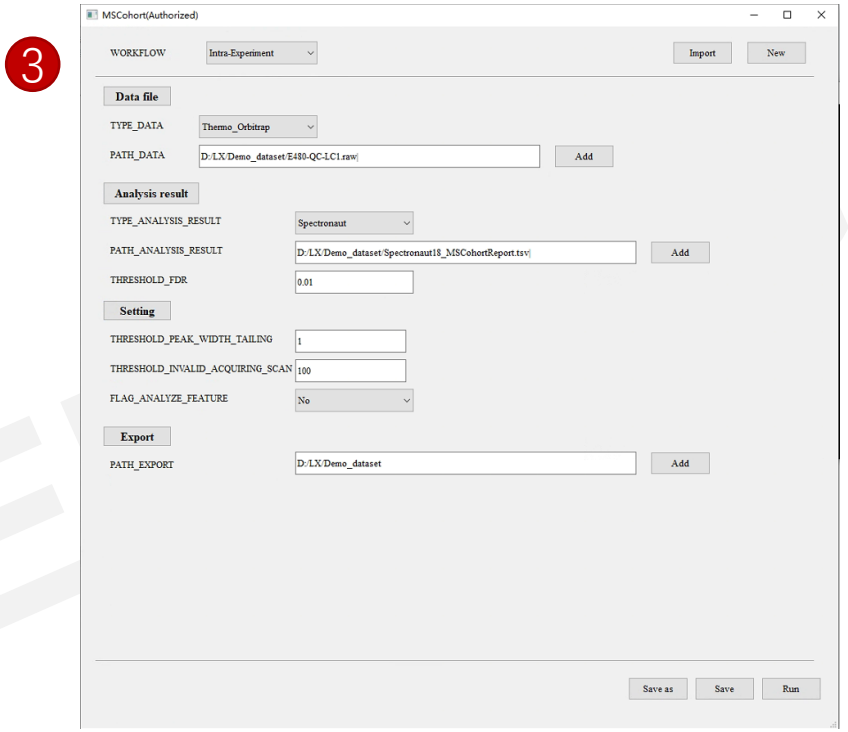
① Choose the MSCohort report schema;

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.

1. Analyzing with MSCohort



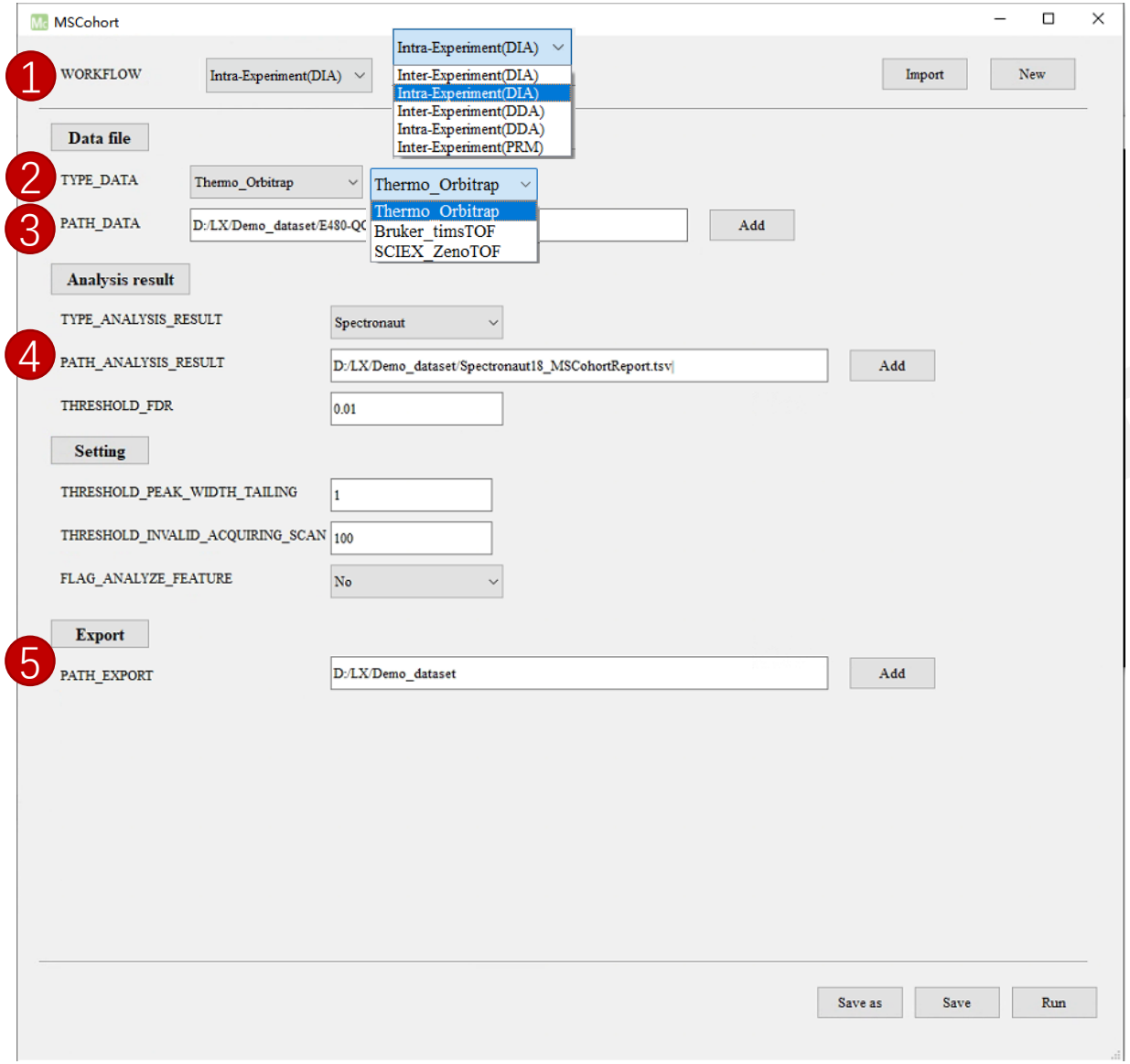


① Double-click MSCohortUI.exe;

② Wait for MSCohort program to load;

③ The MSCohortUI.exe settings screen is displayed.

1. Analyzing with MSCohort



① Select **WORKFLOW** as **Intra-experiment(DIA)**;

② Select **TYPE_DATA** according to the data type ;

③ Click Add to select the raw file into the **PATH_DATA**;

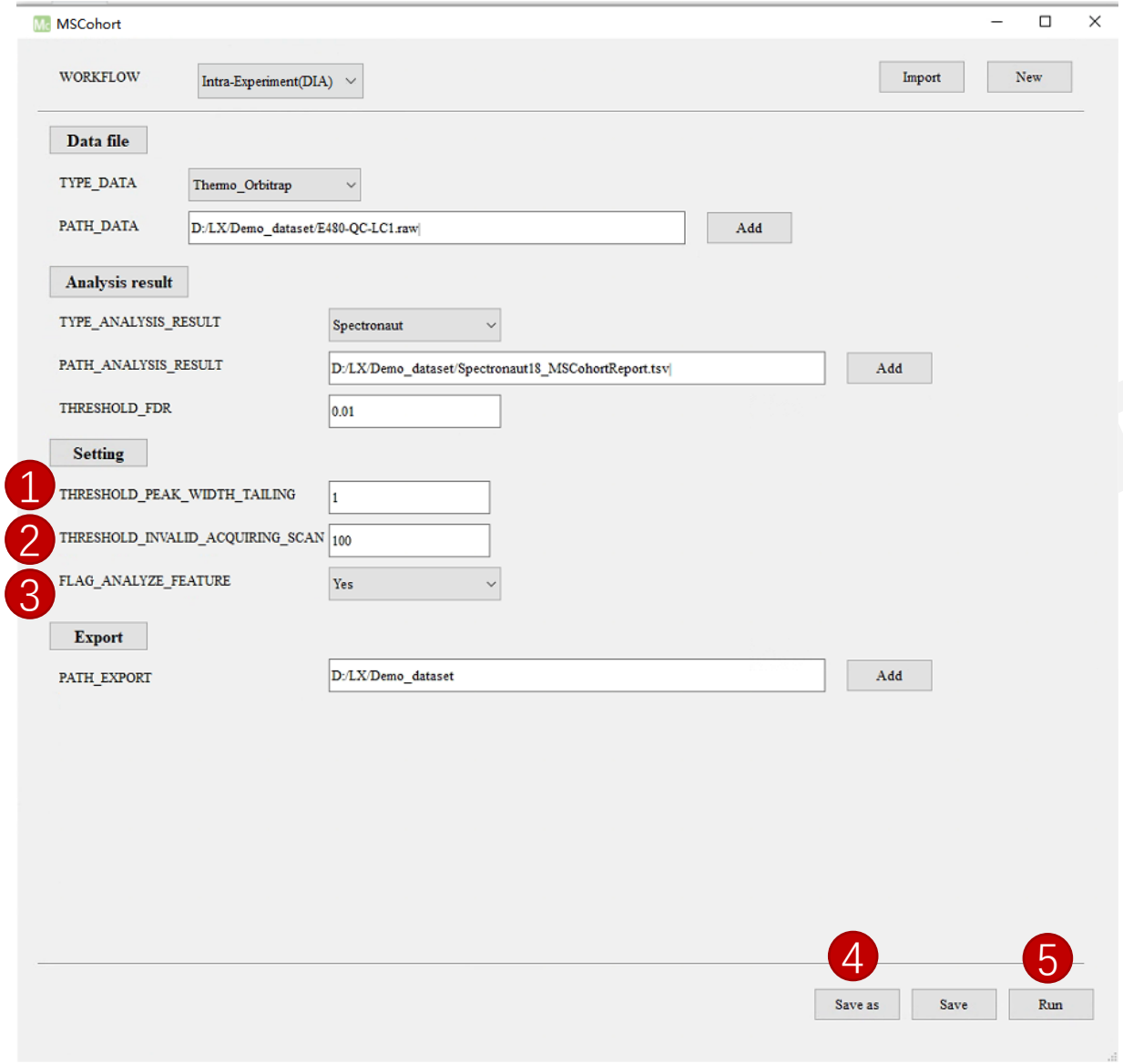
④ Click Add to select the Spectronaut customized report for MSCohort into the **PATH_ANALYSIS_RESULT**;

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



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;

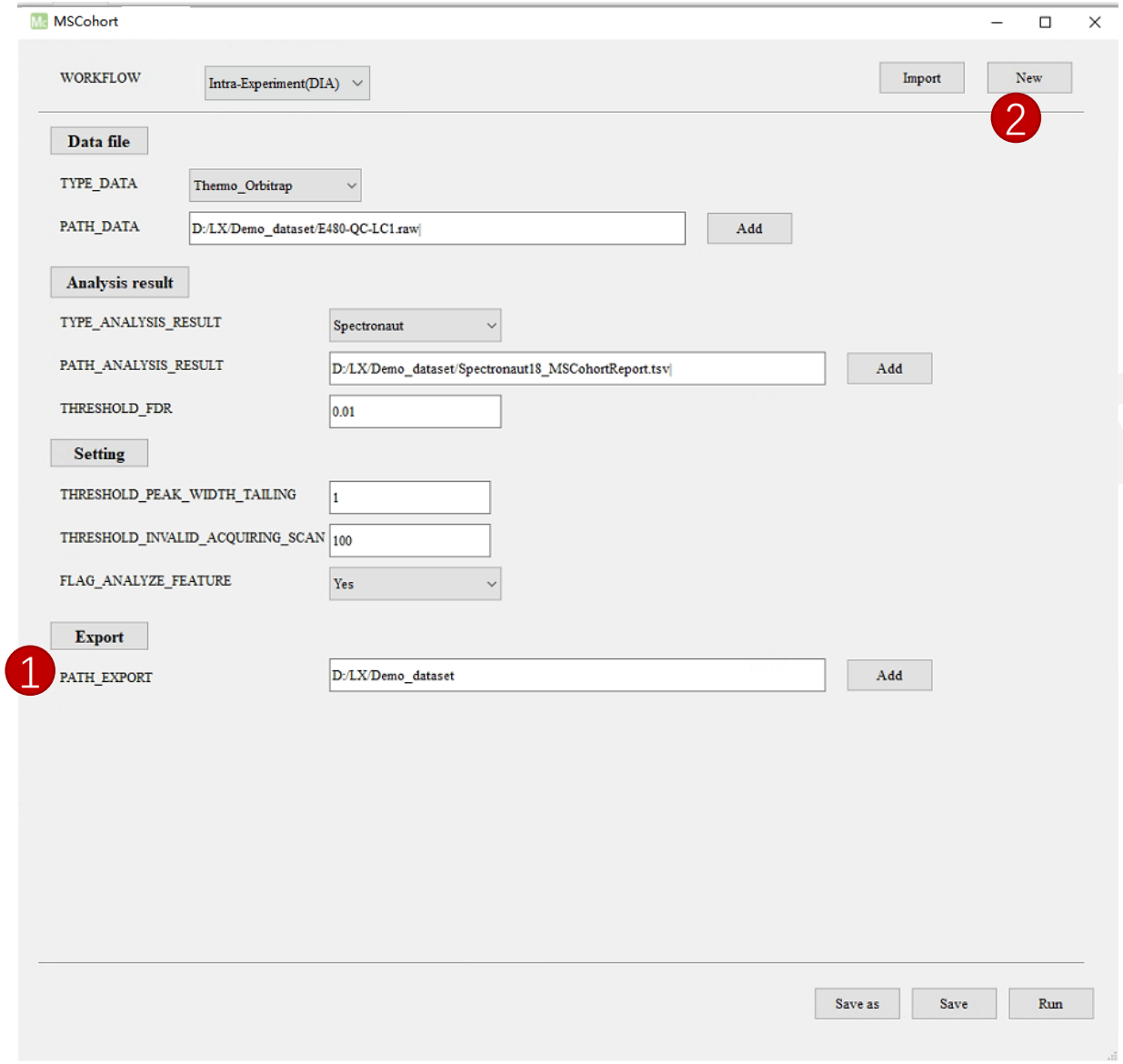
② **THRESHOLD_ INVALID_ACQUIRING** (default setting 100, user adjustable parameter according to experimental condition), set as 100 represents that the retention time length from the start to the number of identified MS2 scans per minute no more than 100 as chromatographic invalid acquiring time;

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



Note：

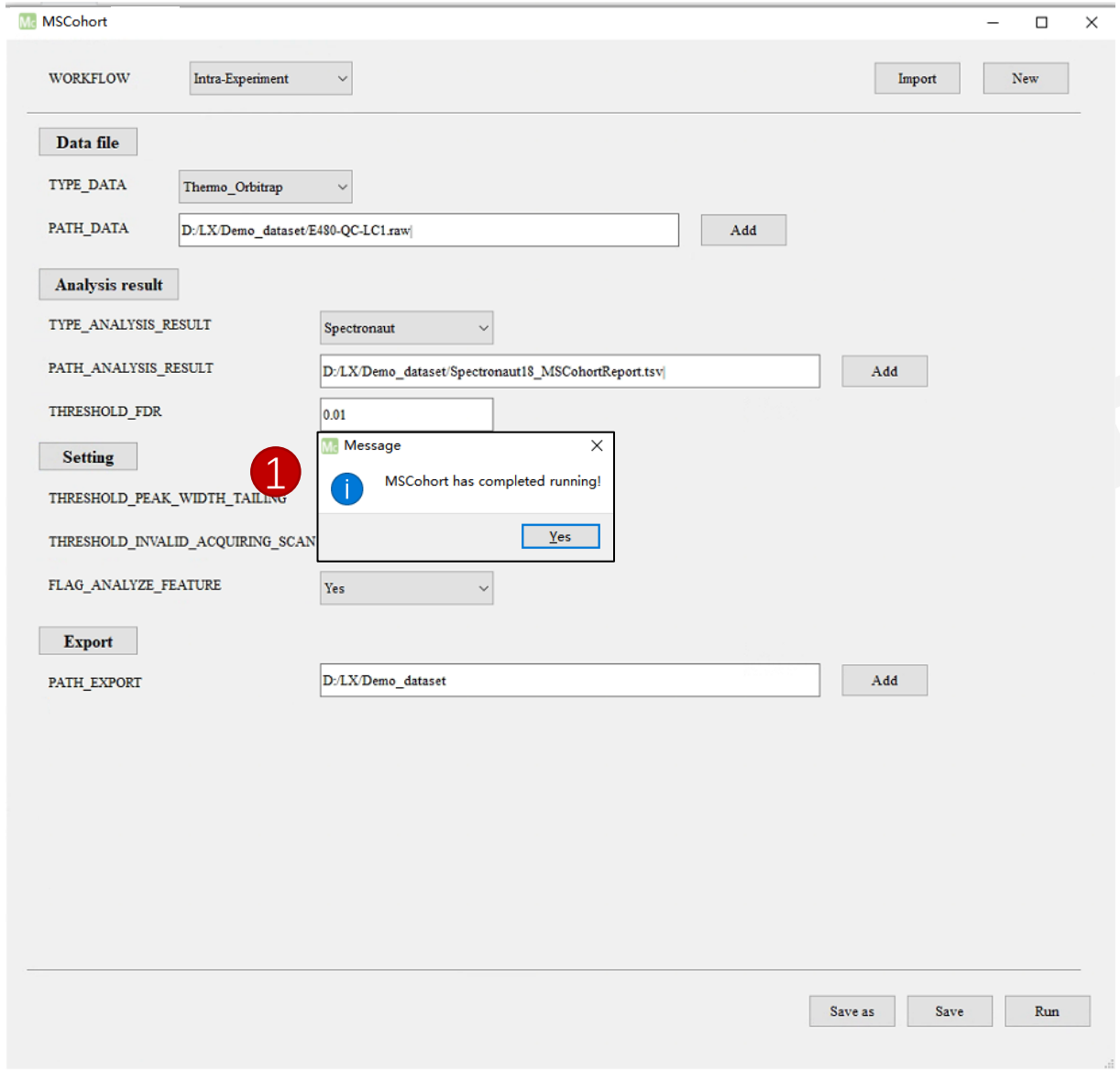
① **PATH_EXPORT：** The output MSCohort analysis results will be stored in a new folder under the PATH_EXPORT directory, named MSCohort _"filename" folder.

Do not save the same PATH_DATA under a same folder, the results will be overwritten.

② Choose New for a new experiments

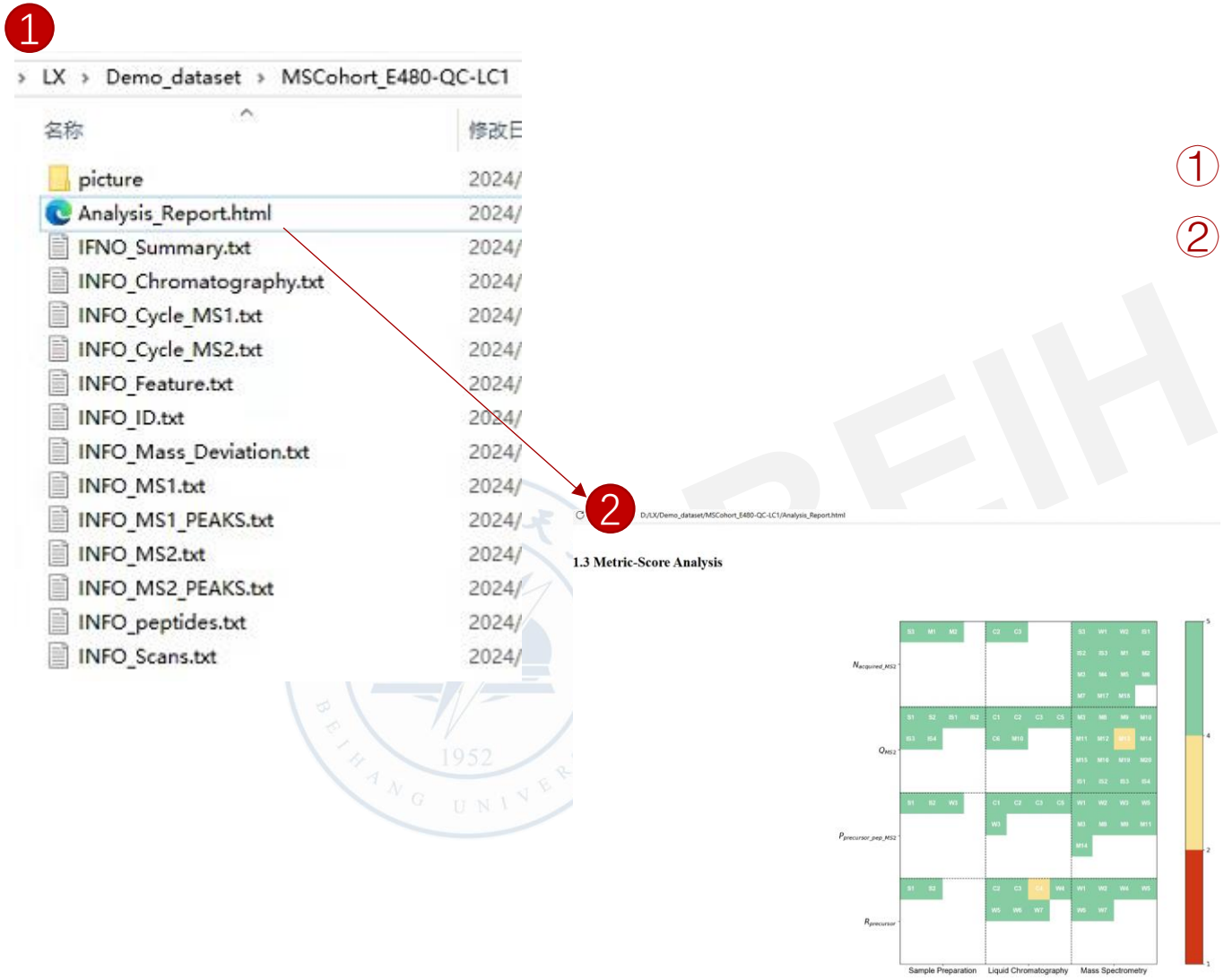
③ Check the filename in PATH_DATA must in the Spectronaut report in PATH_ANALYSIS_RESULT.

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 under the PATH_EXPORT directory and view it.

2. MSCohort Results

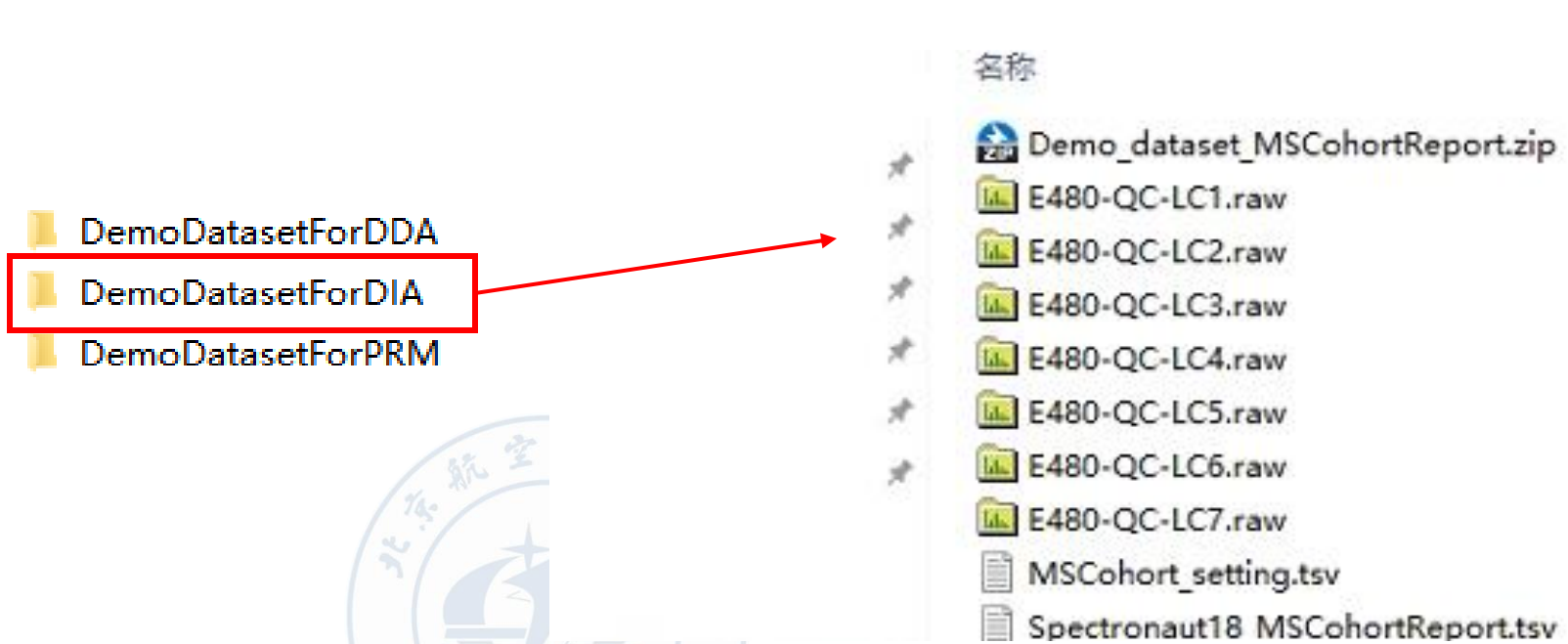


① The MSCohort results ;

② Double clicking **Analysis_Report.html**, the report will be preformed in the browser.

MSCohort manual for DIA inter-experiment analysis

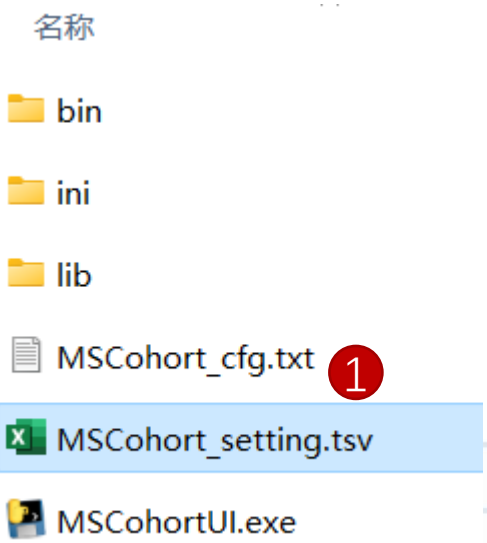
Demo dataset



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

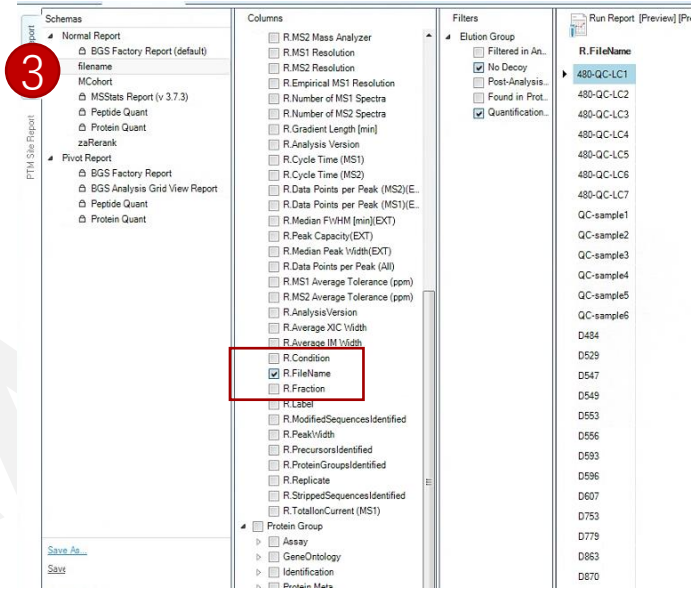


1

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;

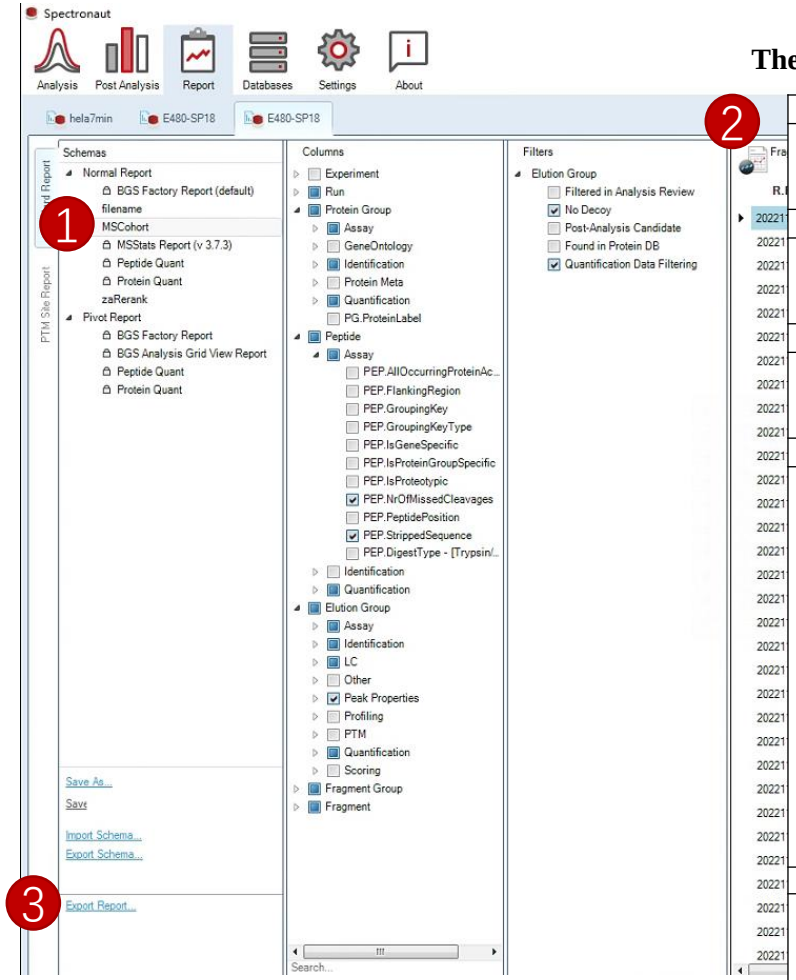
② Fill the columns as the example file showed:

- 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

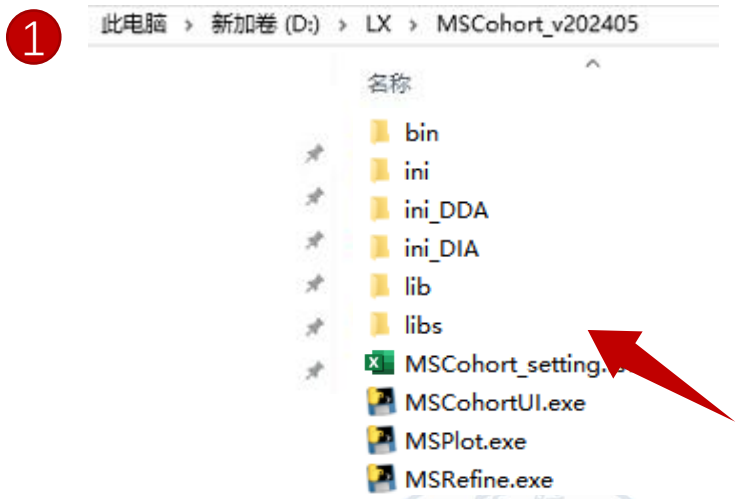
Run	
	R.Run Date
	R.Gradient Length [min]
	R.FileName
Protein Group	
	PG.ProteinGroups
	PG.Qvalue
	PG.Quantity
Peptide	
	PEP.NrOfMissedCleavages
	PEP.StrippedSequence
	PEP.Quantity
Elution Group	
	EG.IsDecoy
	EG.ModifiedPeptide
	EG.PrecursorId
	EG.Qvalue
	EG.ApexRT
	EG.DatapointsPerPeak
	EG.DatapointsPerPeak (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)

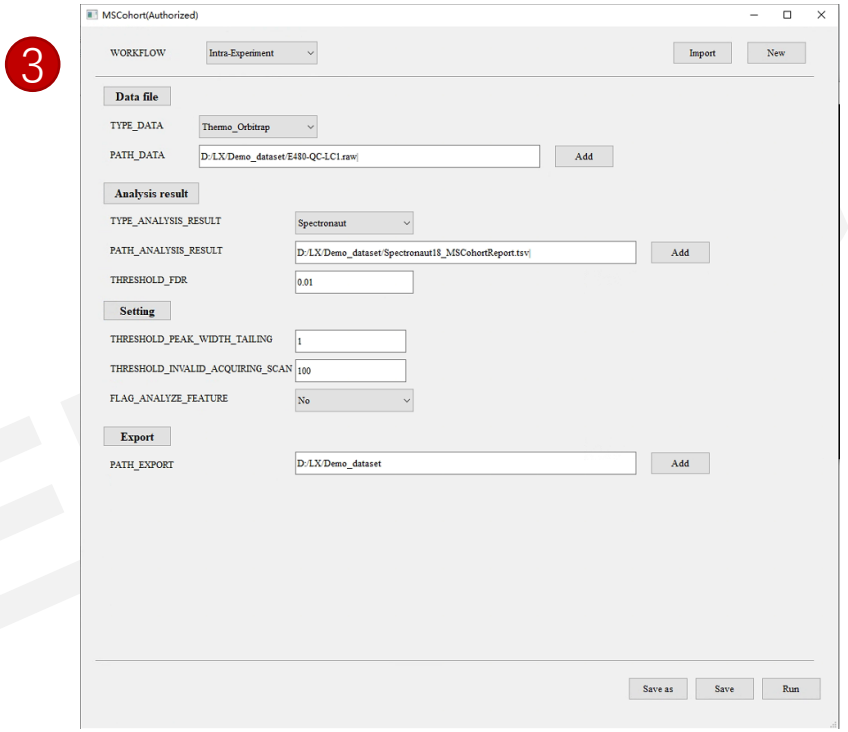
① Choose the MSCohort report schema;

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



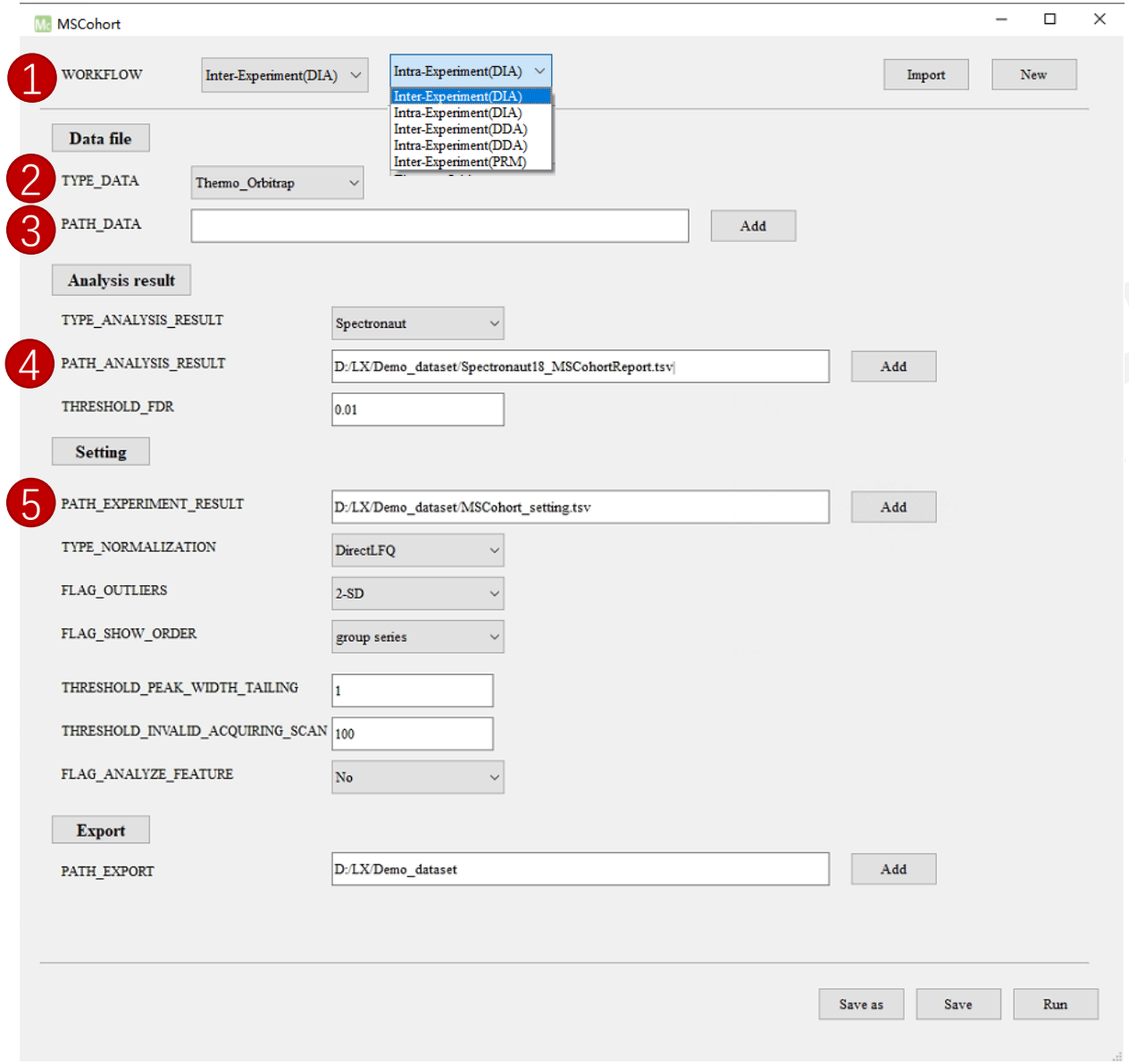


① Double-click MSCohortUI.exe;

② Wait for MSCohort program to load;

③ The MSCohortUI.exe settings screen is displayed.

2. Analyzing with MSCohort



① Set **WORKFLOW** as **Inter-experiment(DIA)**;

② Select **TYPE_DATA** according to the data type ;

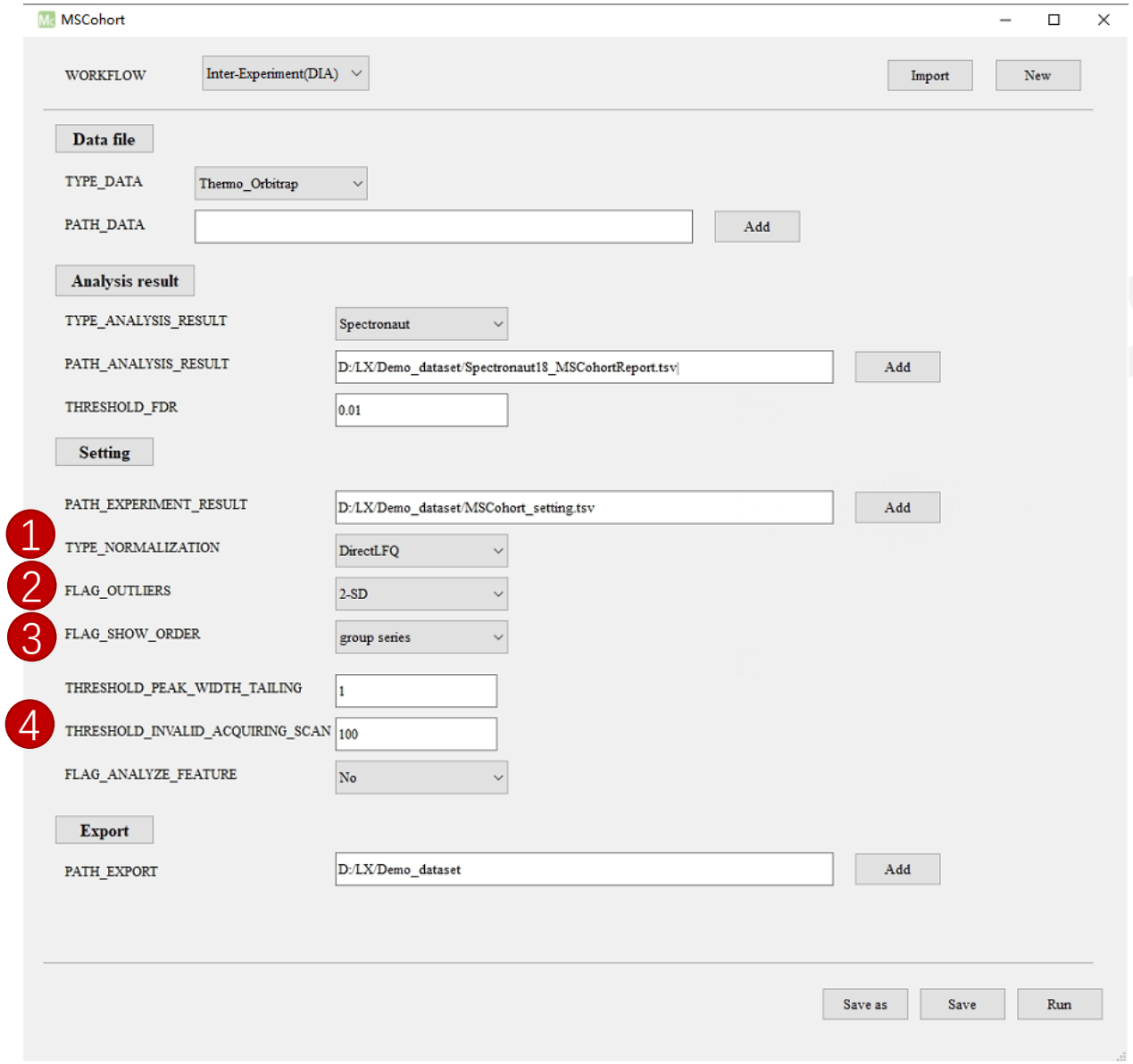
③ The **PATH_DATA** could be empty. MSCohort support the inter-experimental analysis mainly based on Spectronaut result, eliminating the need to submit raw files, which will obtain the result in a relatively **short time**. It is recommended for large cohort analysis.

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



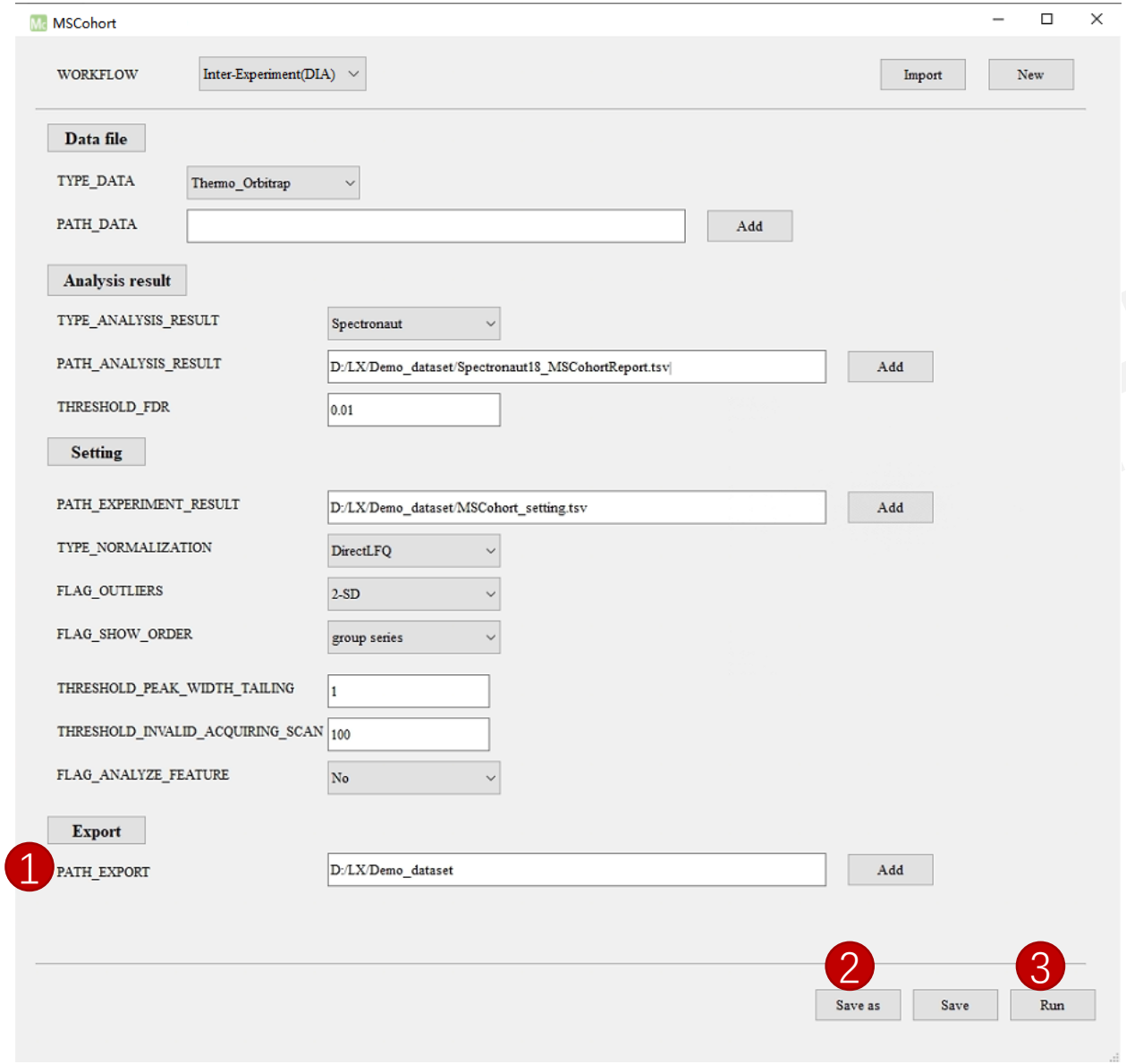
① Choose normalization strategies in **TYPE_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);

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 **THERSHOLD_PEAK_WIDTH_TAIL, THRESHOLD INVALID_ACQUIRING, FLAG_ANALYZE_FEATURE** as intra-experiment analysis;

2. Analyzing with MSCohort



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.

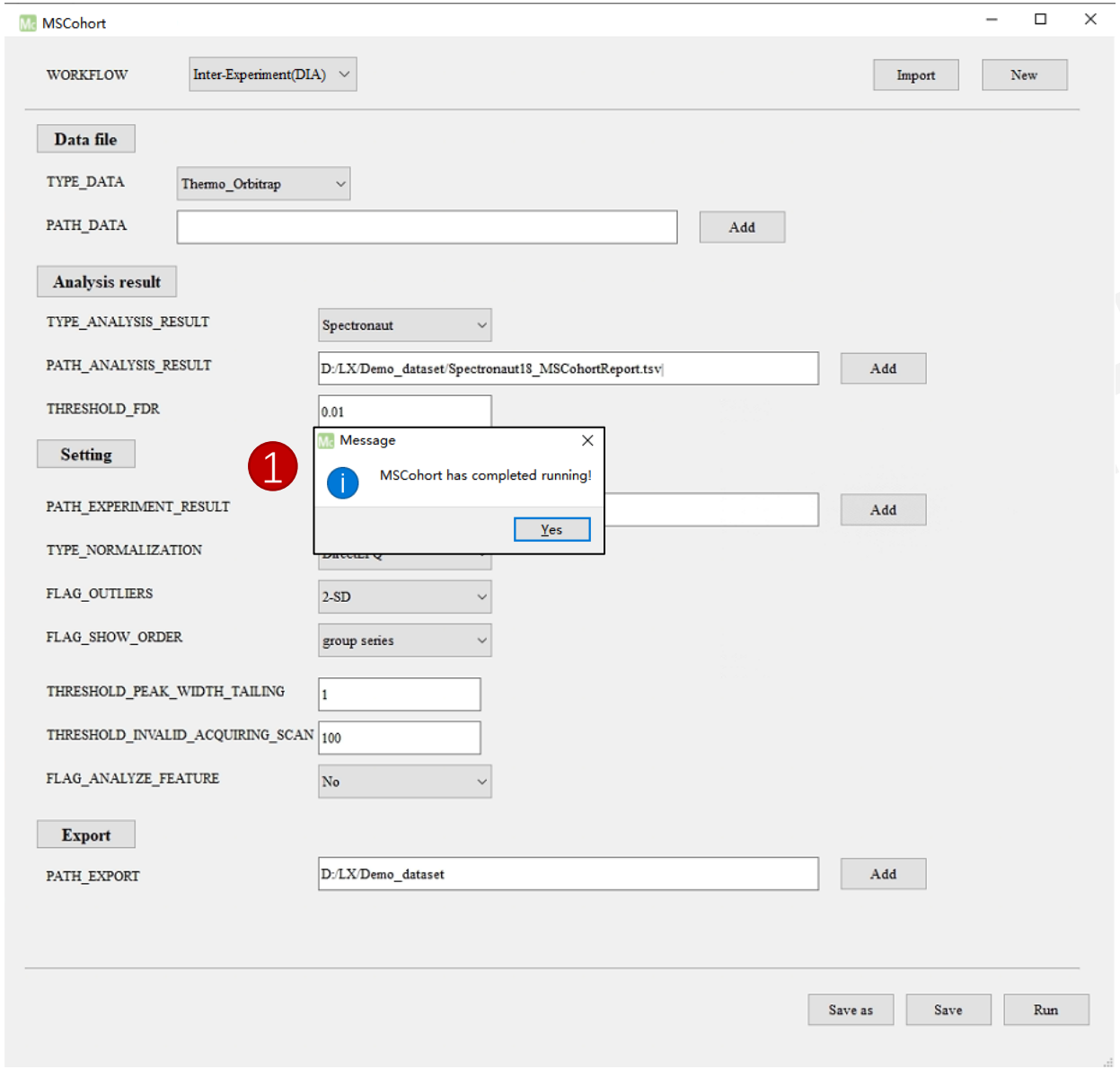
② Click **Save as** button to save the config file;

③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

Note：

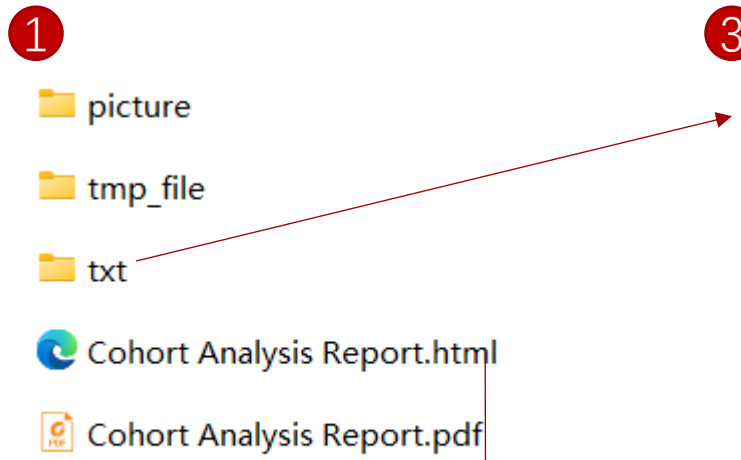
Space (" ") cannot exist in the file directory (including PATH_DATA, PATH_ANALYSIS_RESULT, PATH_EXPERIMENT_RESULT, and PATH_EXPORT), which will affect the normal running of the program.

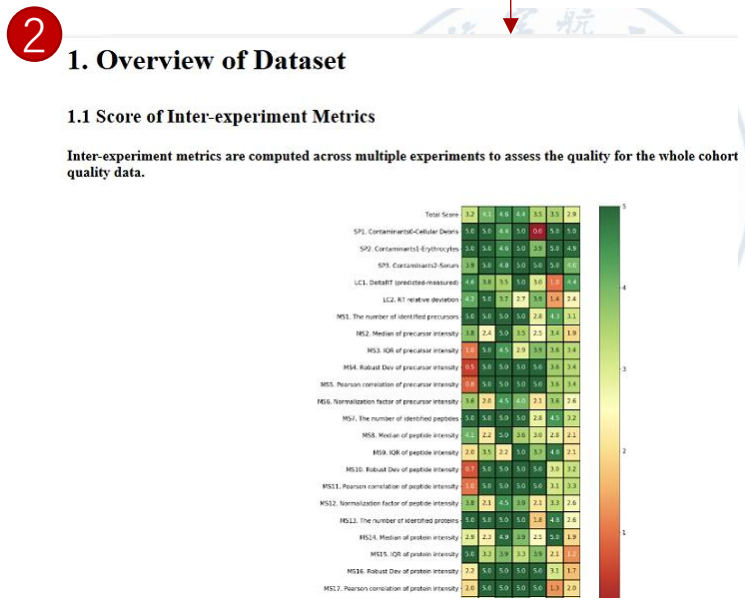
2. Analyzing with MSCohort

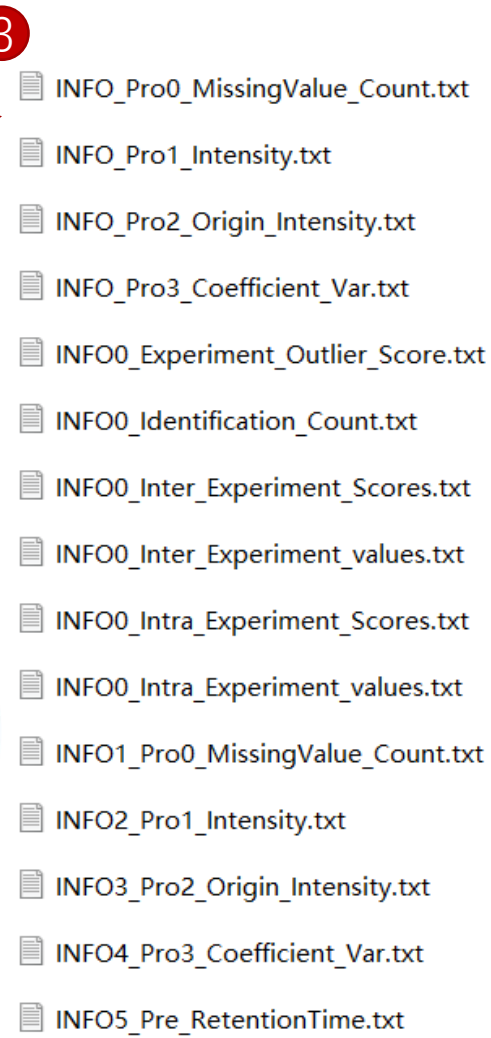


① When the program popup window shows " MSCohort has completed running! "Indicates that the current process is complete. Please open the result file under the PATH_EXPORT directory and view it.

3. MSCohort Inter-experiment analysis results







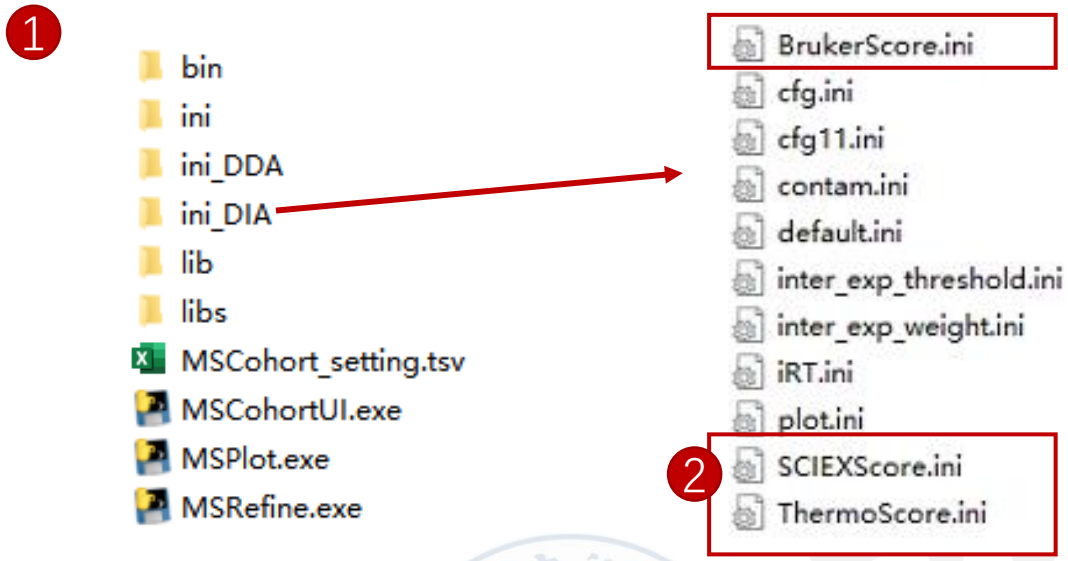
① The MSCohort results ;

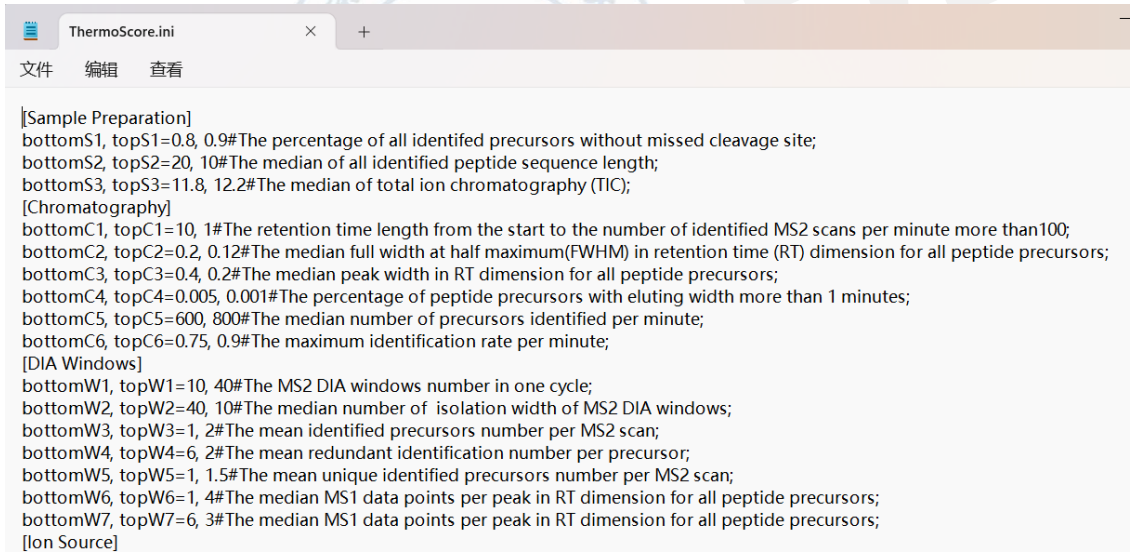
② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.

③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

Notes for modifying the scoring criteria

1. Modifying the scoring standards for intra-experiment analysis





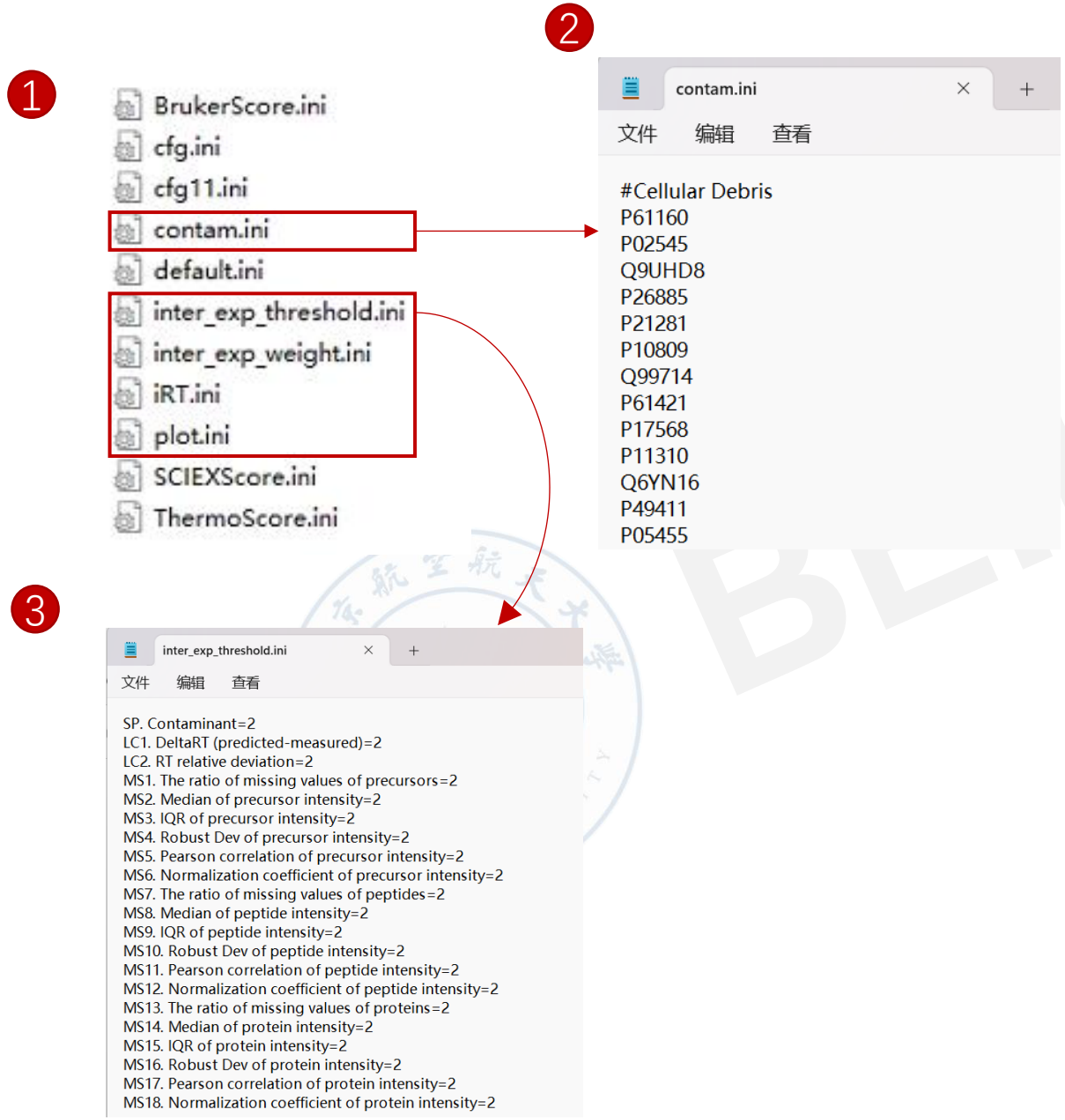
① Open the **ini_DIA** folder, there will be three parameter files related to intra-experiment scoring: ThermoScore.ini, BrukerScore.ini, SCIEXScore.ini;

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

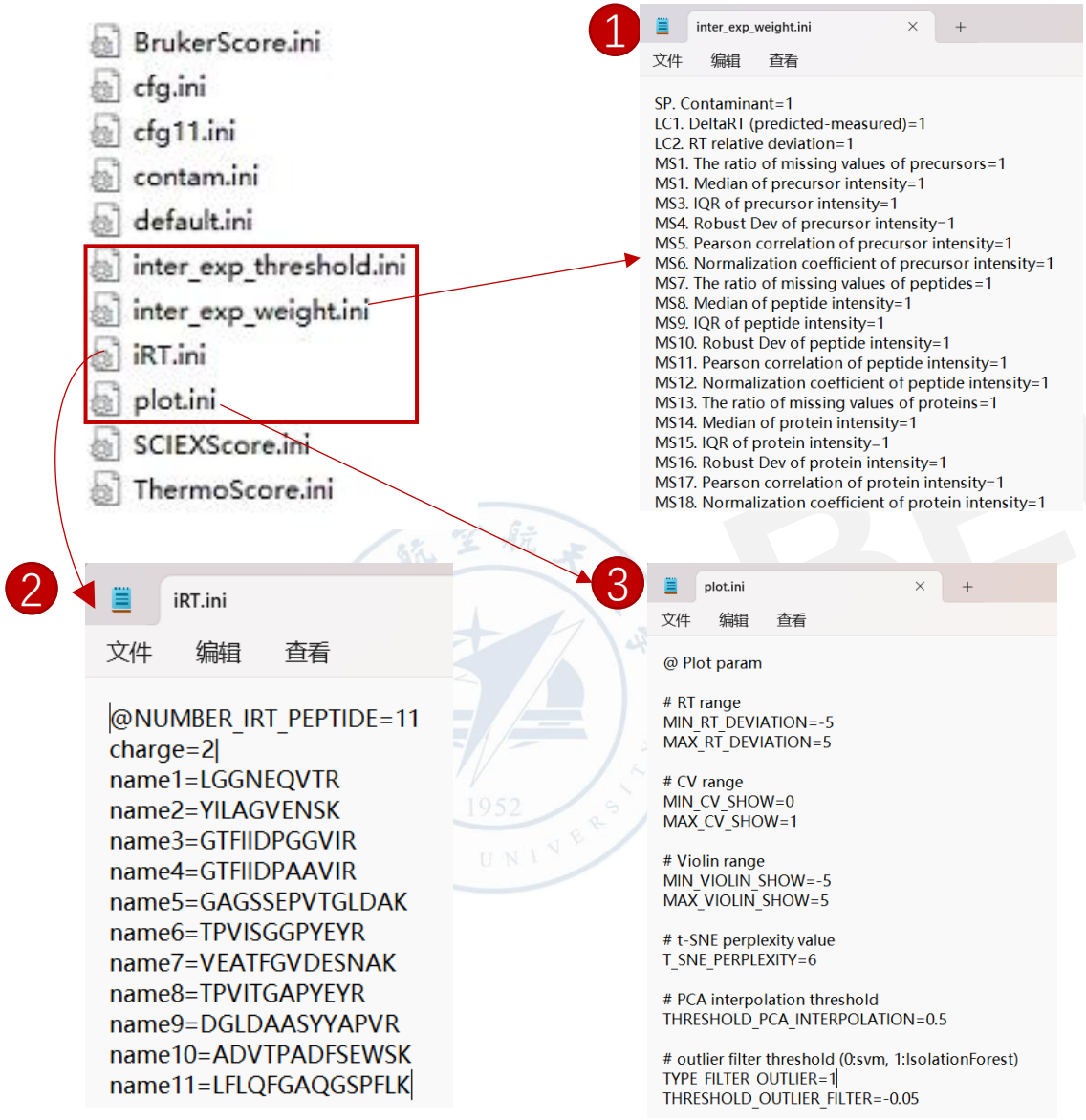


① Open the **ini_DIA** folder, there will be five parameter files related to inter-experiment scoring: contam.ini, inter_exp_threshold.ini, inter_exp_weight.ini, iRT.ini, plot.ini;

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*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. Modifying the scoring criteria for inter-experiment analysis



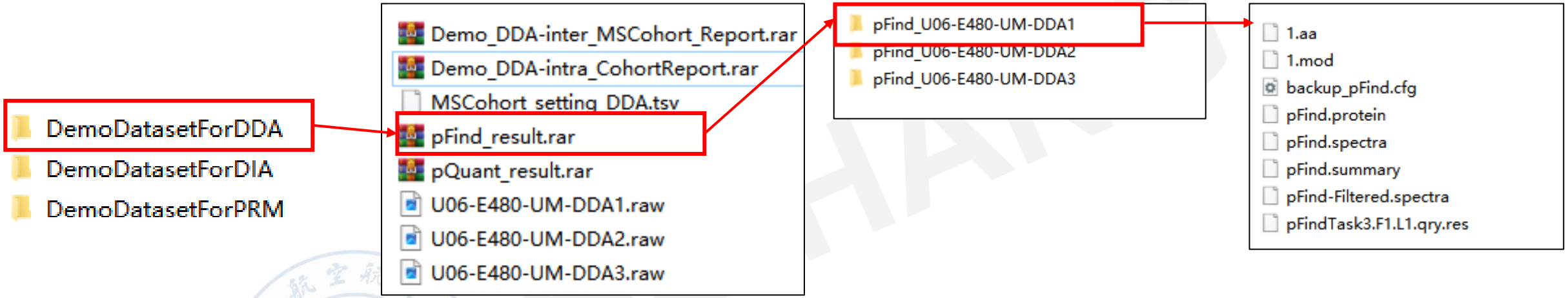
① Open the inter_exp_weight.ini, users can modify the scoring weight for each metric. We initially defined the weight as 1 for each metric. If you do modify it, closing and saving the file, it will score according to the standard you set.

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

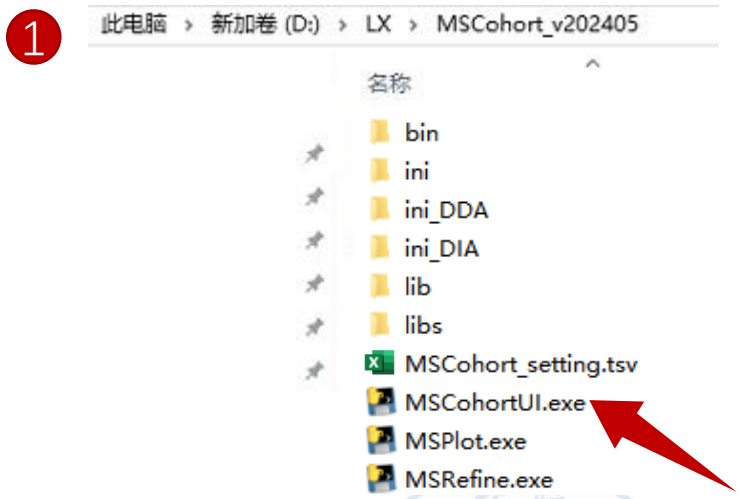
Demo dataset

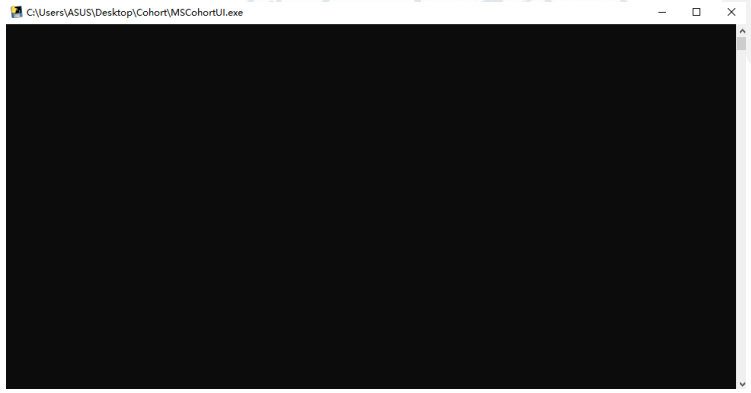


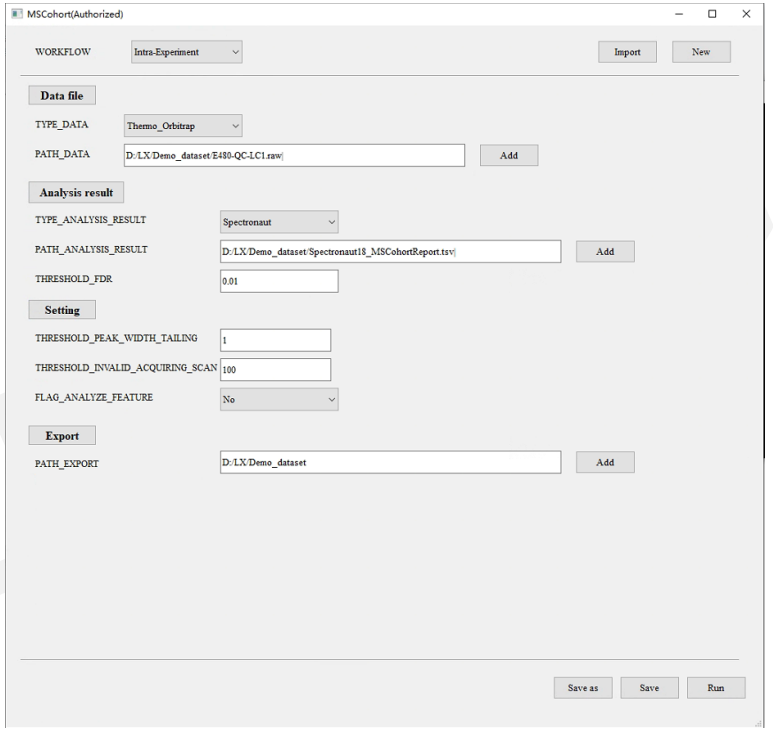
Taking the U06-E480-UM-DDA1.raw file as an example to demonstrate the workflow of DDA intra-experiment analysis.

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

1. Analyzing with MSCohort





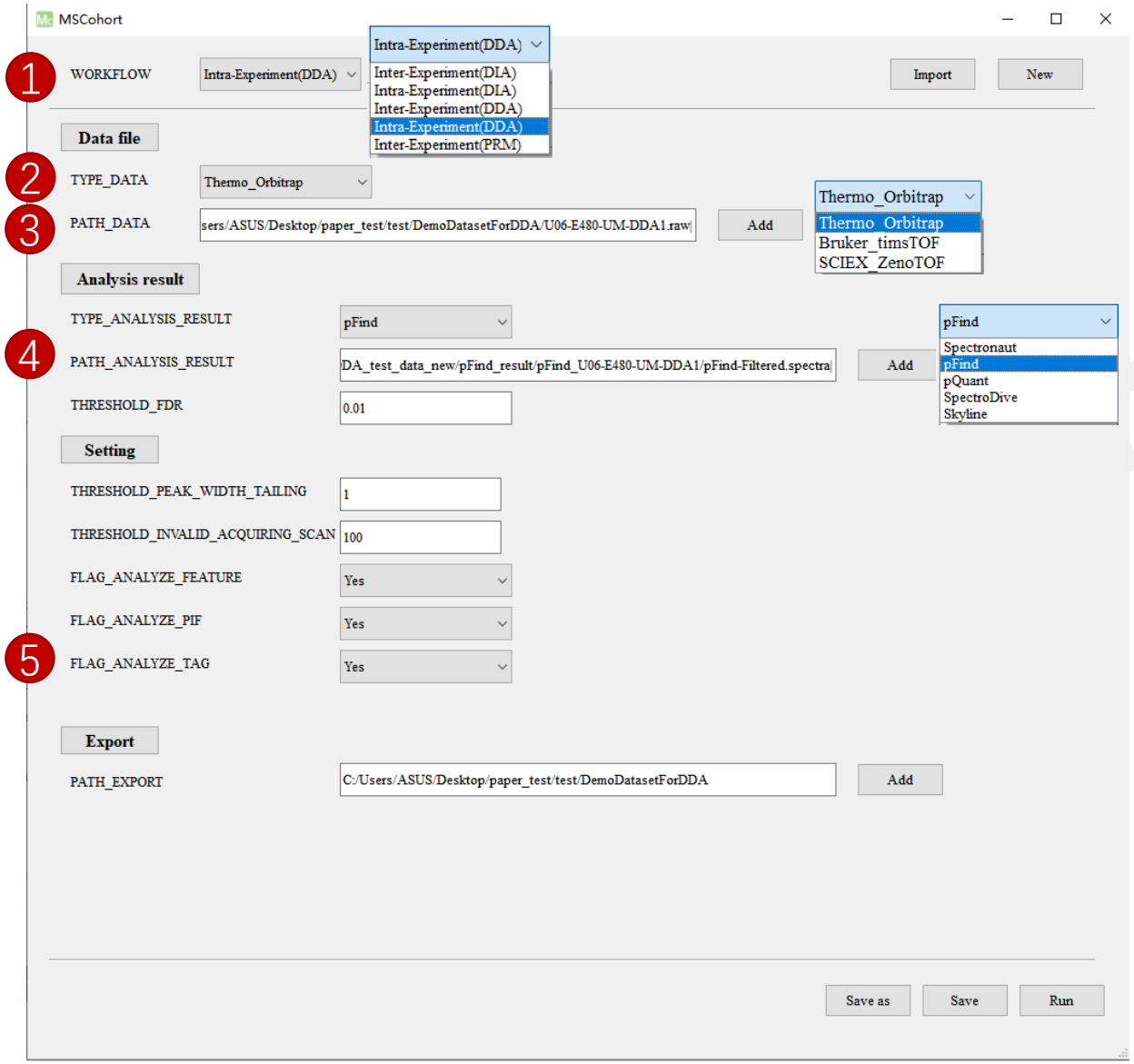


① Double-click MSCohortUI.exe;

② Wait for MSCohort program to load;

③ The MSCohortUI.exe settings screen is displayed.

1. Analyzing with MSCohort



① Select **WORKFLOW** as **Intra-experiment(DDA)**;

② Select **TYPE_DATA** according to the data type ;

③ Click Add to select the raw file into the **PATH_DATA**;

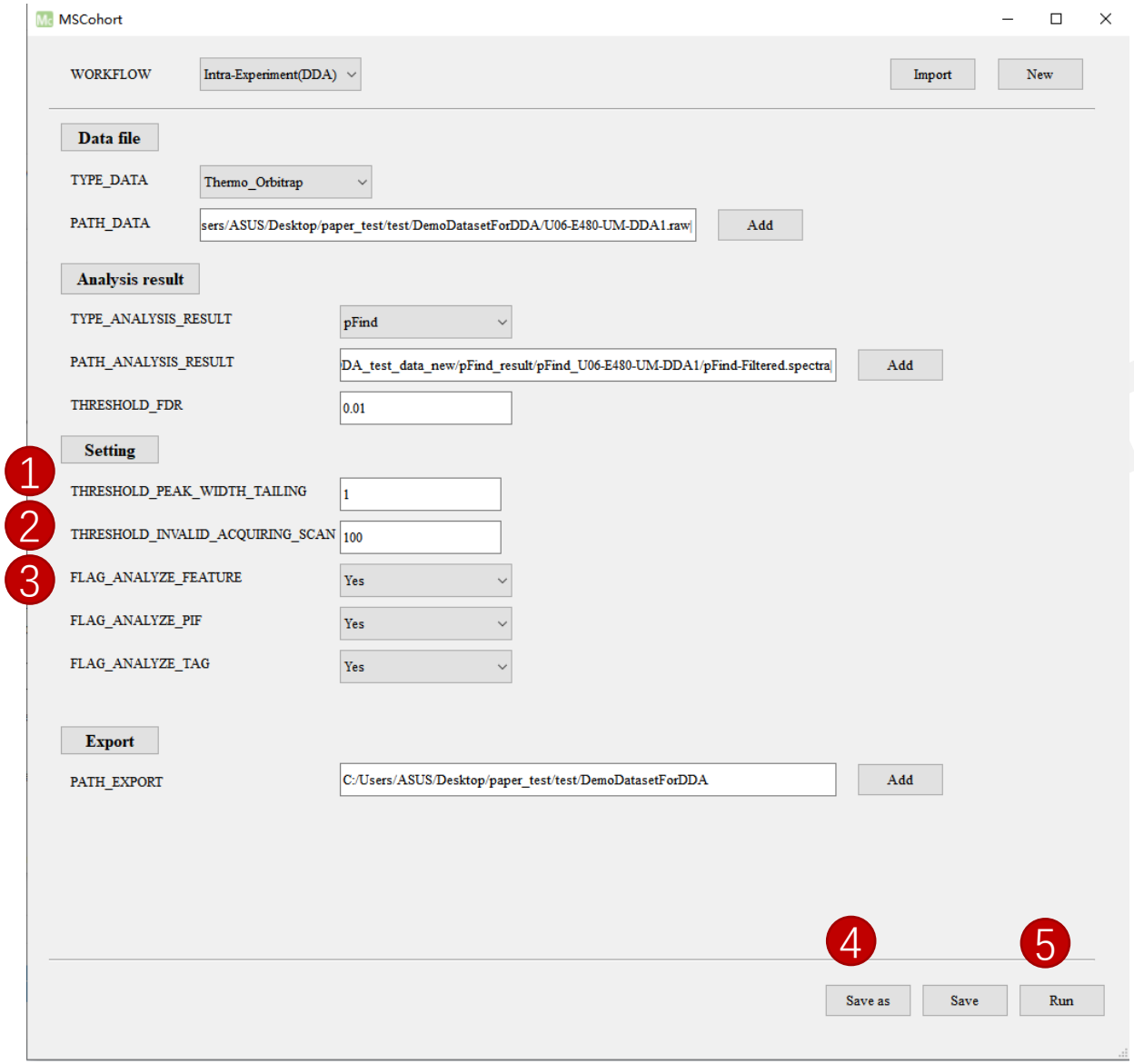
④ Click Add to select the pFind report (pFind-Filtered.spectra) for MSCohort into the **PATH_ANALYSIS_RESULT**;

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



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;

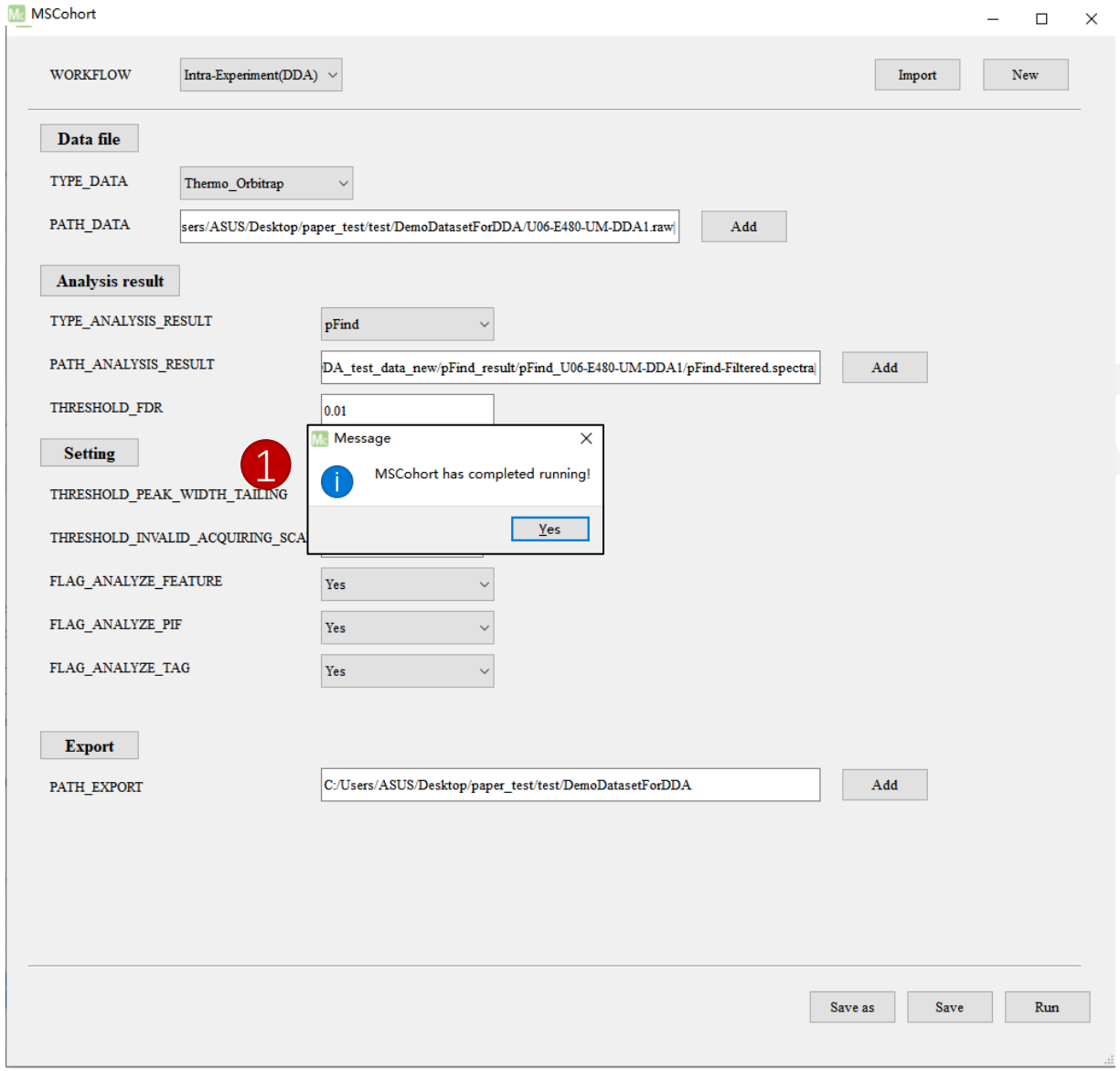
② **THRESHOLD_ INVALID_ACQUIRING** (default setting 100, user adjustable parameter according to experimental condition), set as 100 represents that the retention time length from the start to the number of identified MS2 scans per minute no more than 100 as chromatographic invalid acquiring time;

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;

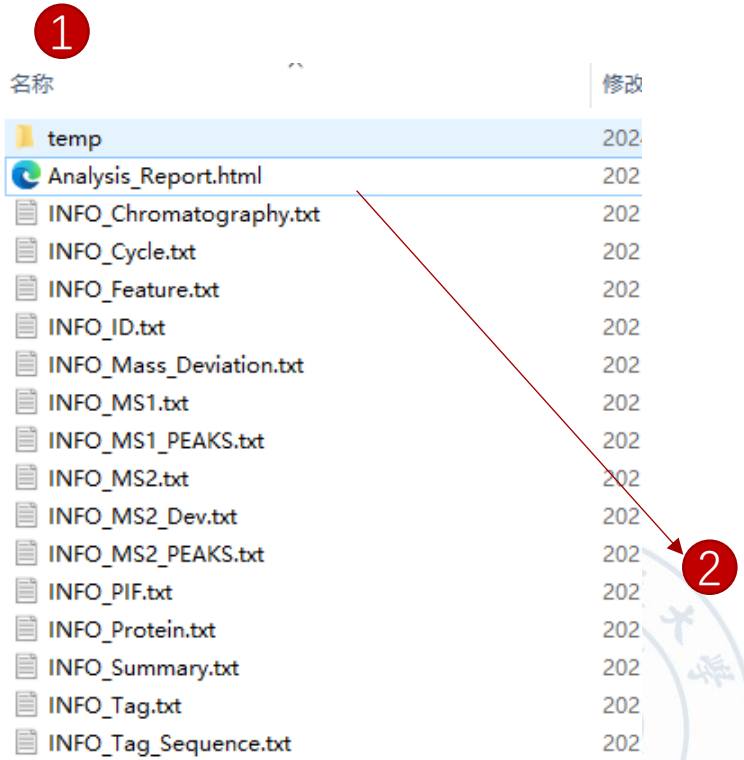
⑤ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

2. Analyzing with MSCohort



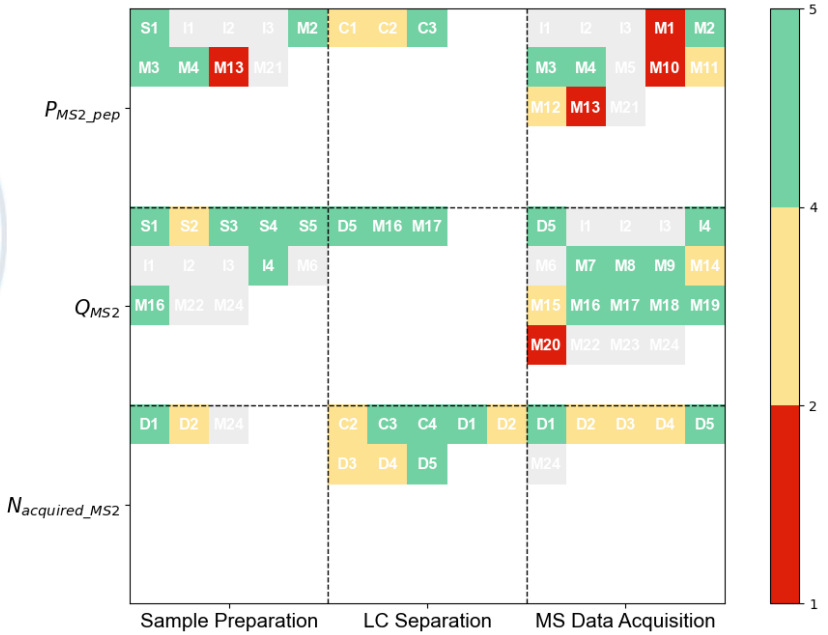
① When the program popup window shows “ MSCohort has completed running! ”Indicates that the current process is complete. Please open the result file corresponding to Cohort and view it.

2. MSCohort Results



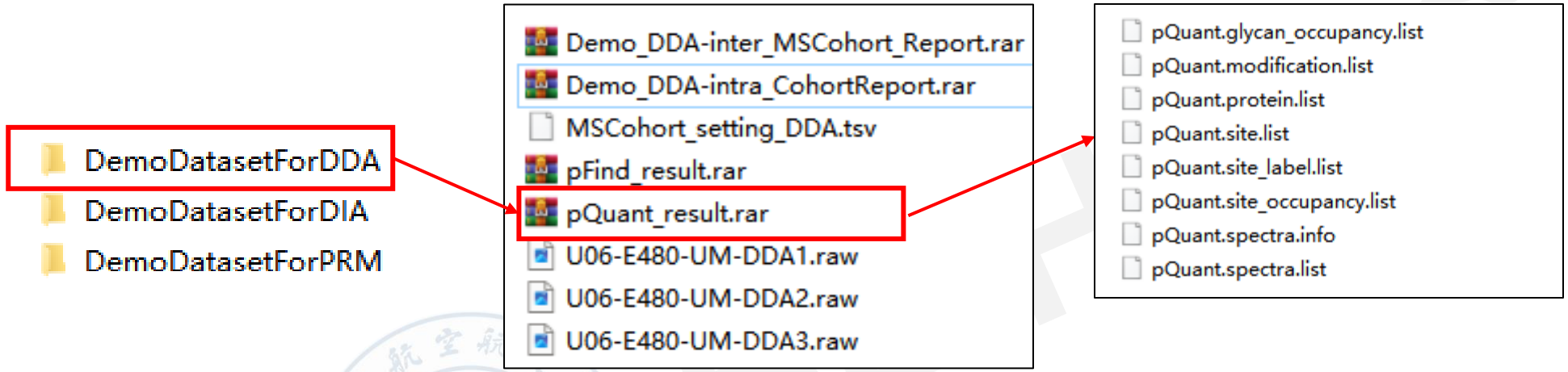
① The MSCohort results ;

② Double clicking **Analysis_Report.html**, the report will be preformed in the browser.



MSCohort manual for DDA inter-experiment analysis

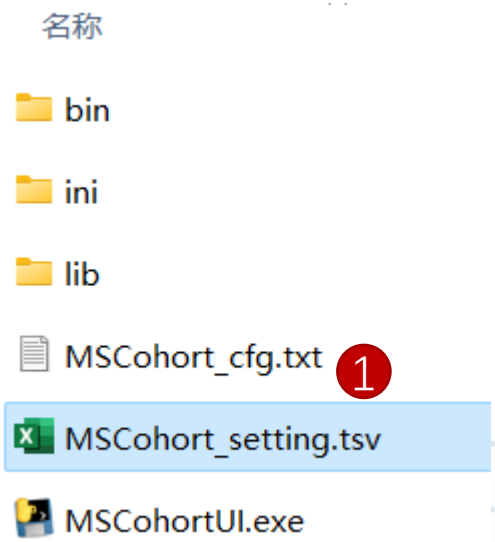
Demo dataset



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

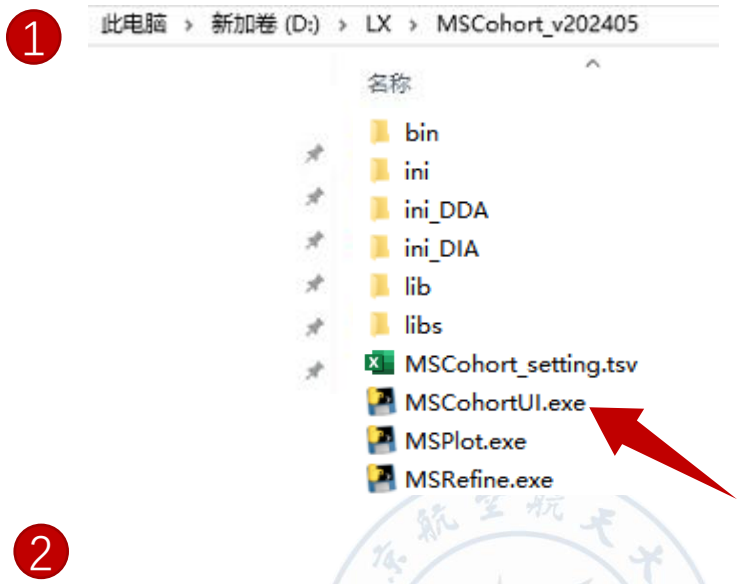
Group Name	Raw Name	Experiment	Threshold
U06	U06-E480-UM-DDA1	U06-E480-UM-DDA1	10
U06	U06-E480-UM-DDA2	U06-E480-UM-DDA2	10
U06	U06-E480-UM-DDA3	U06-E480-UM-DDA3	10

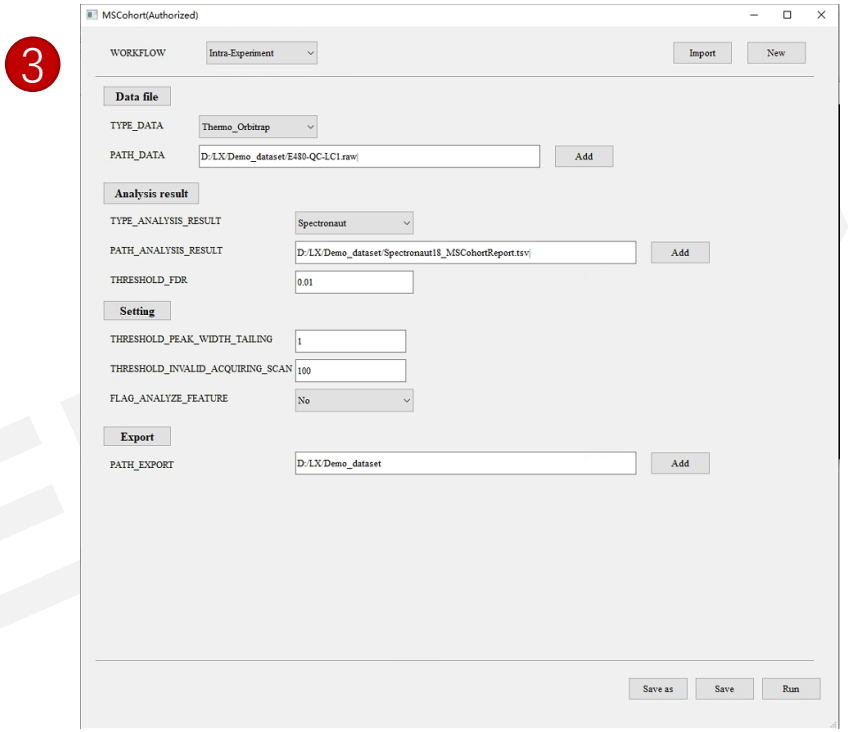
① Open MSCohort_setting.tsv with Excel;

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



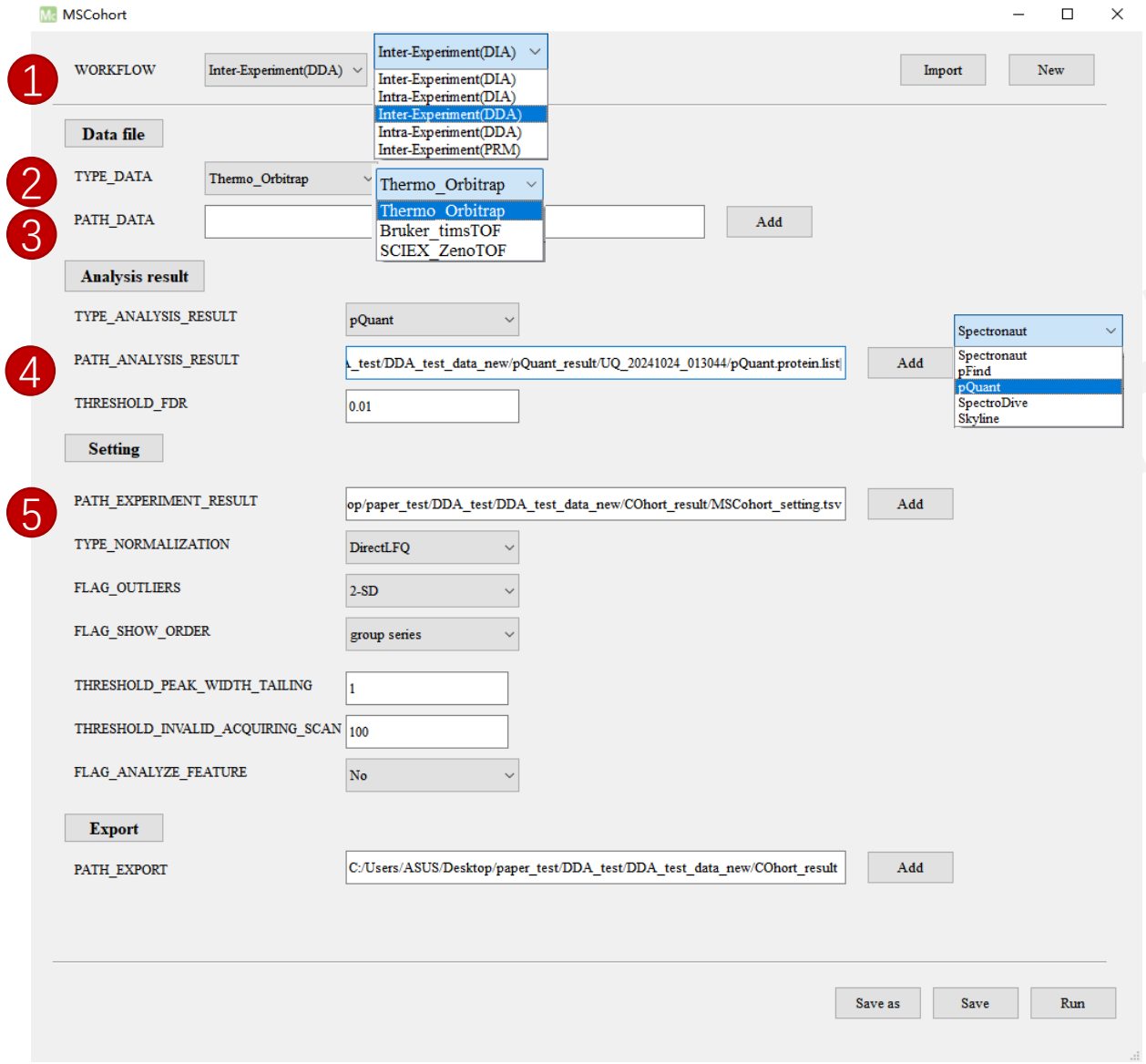


① Double-click MSCohortUI.exe;

② Wait for MSCohort program to load;

③ The MSCohortUI.exe settings screen is displayed.

2. Analyzing with MSCohort



① Set **WORKFLOW** as **Inter-experiment(DDA)**;

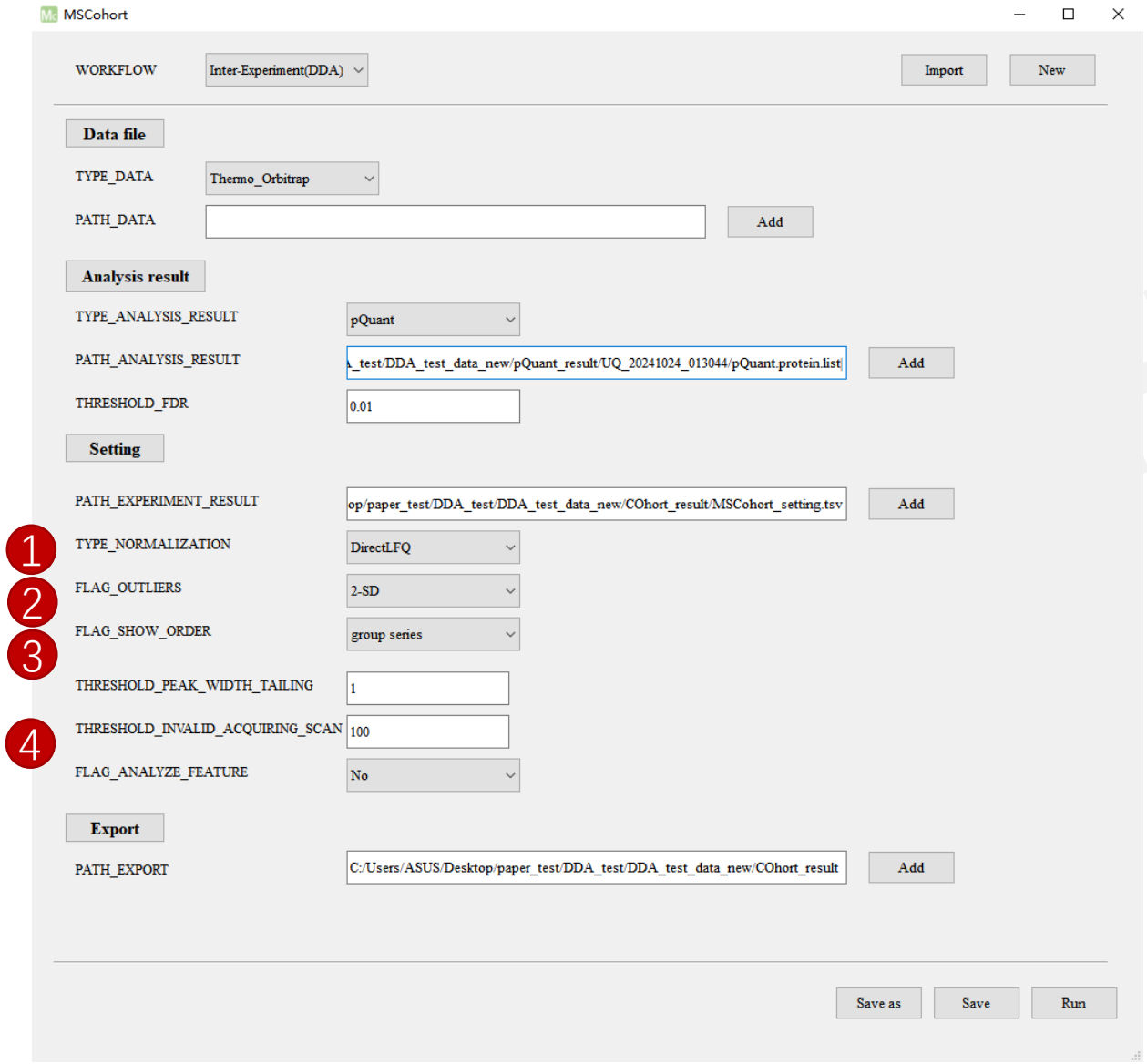
② Select **TYPE_DATA** according to the data type ;

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



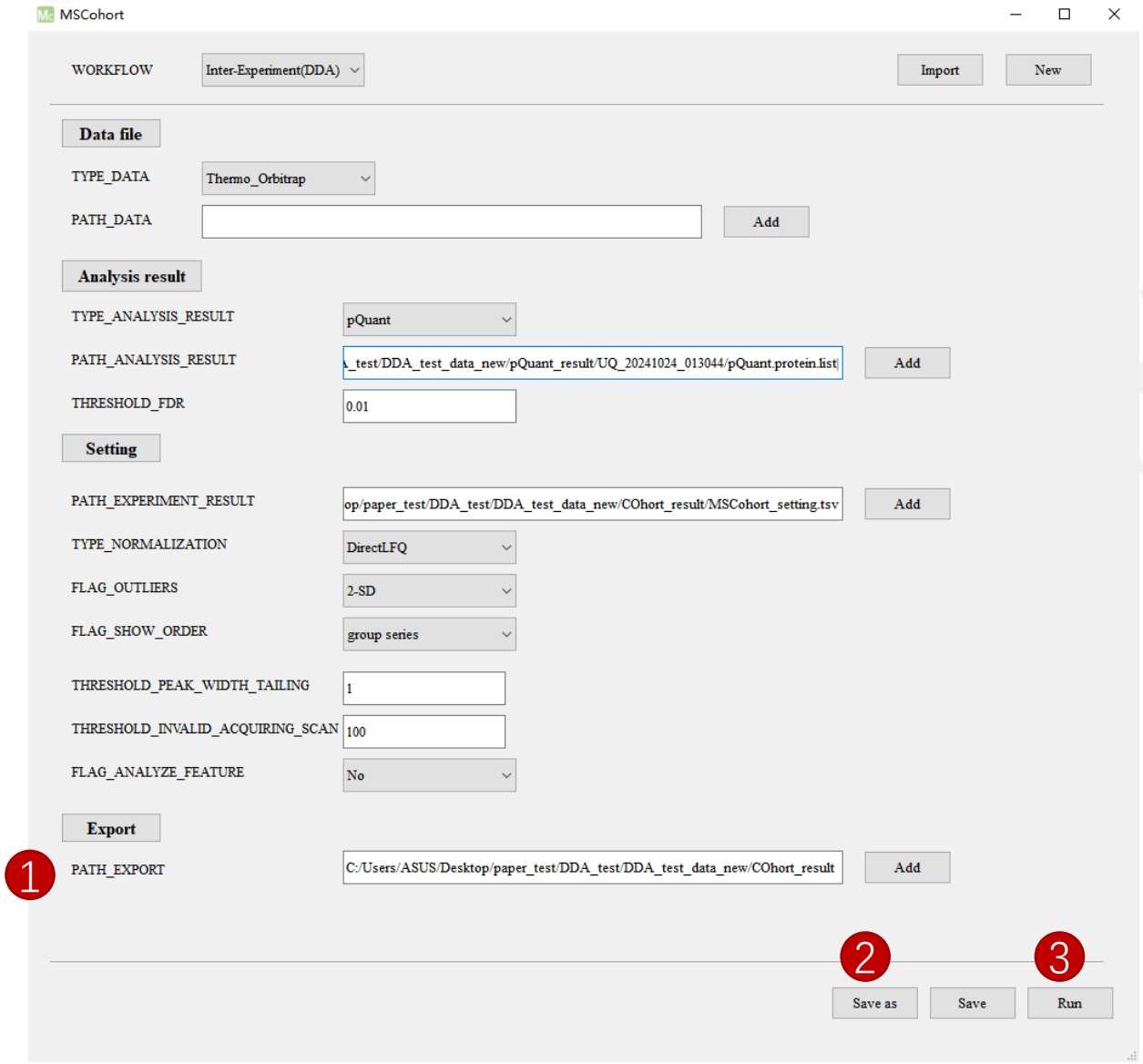
① Choose normalization strategies in **TYPE_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);

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 **THERSHOLD_PEAK_WIDTH_TAIL, THRESHOLD INVALID_ACQUIRING, FLAG_ANALYZE_FEATURE** as intra-experiment analysis;

2. Analyzing with MSCohort



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.

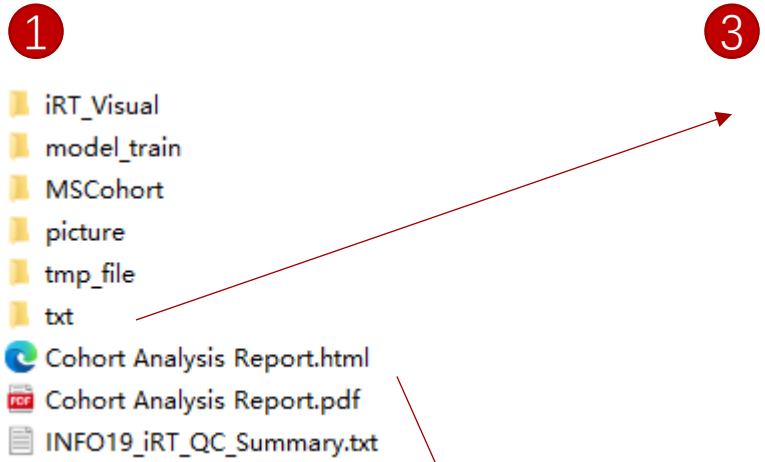
② Click **Save as** button to save the config file;

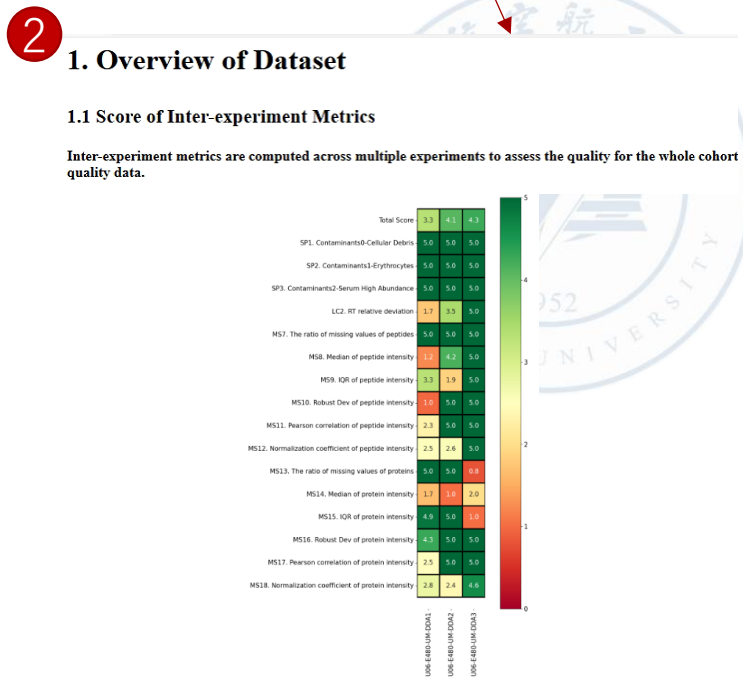
③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

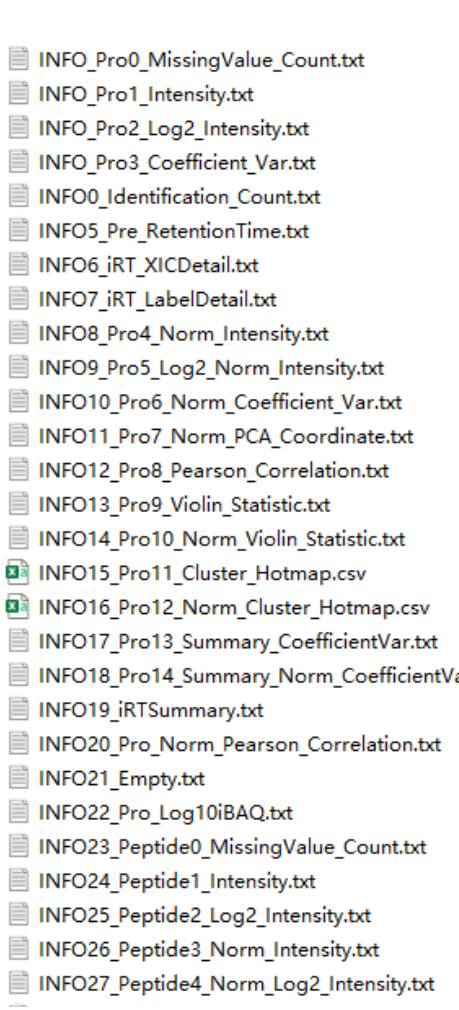
Note：

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 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.

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

Sheet1

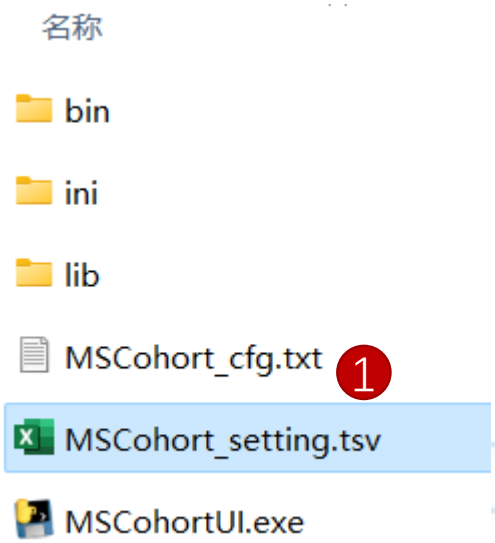
MSCohort_setting_Skyline.tsv

	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

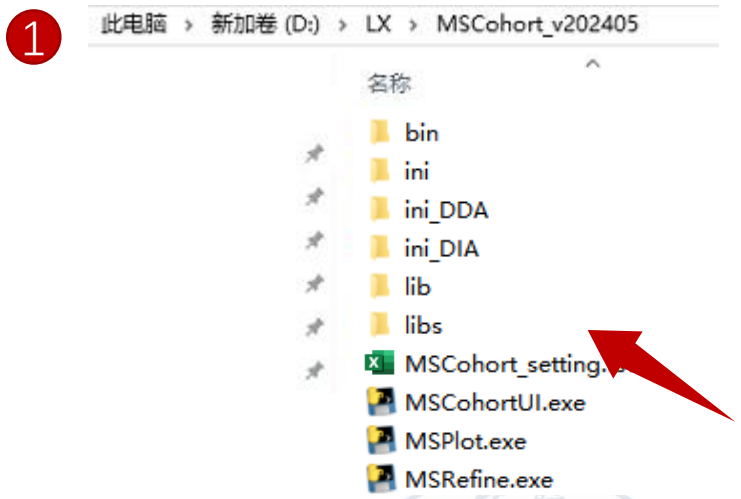
① Open MSCohort_setting.tsv with Excel;

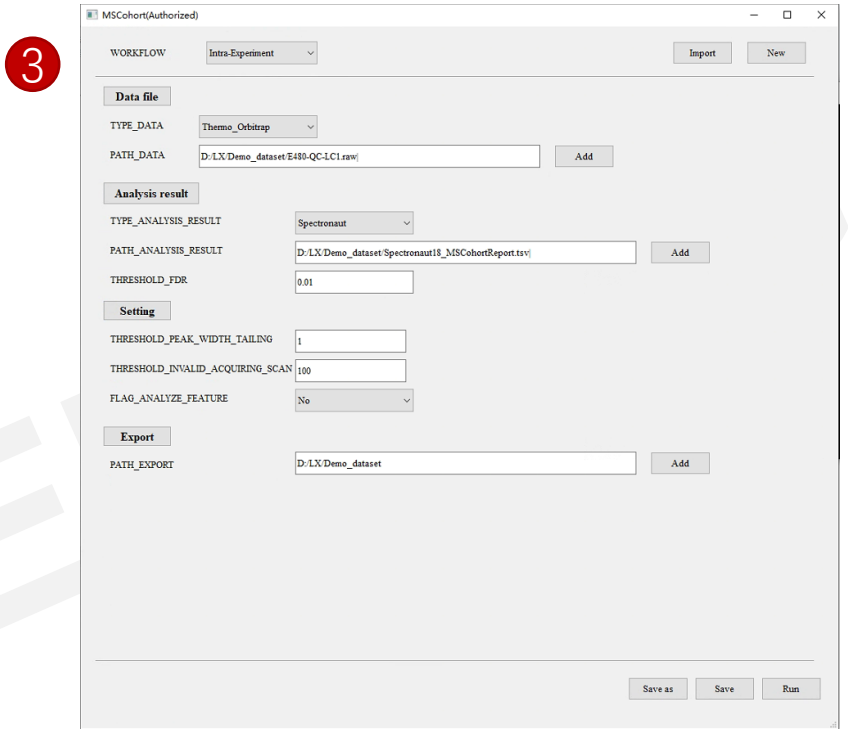
② Fill the columns as the example file (MSCohort_setting_Skyline.tsv) showed or alternatively, directly use the MSCohort_setting_Skyline.tsv file:

- 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



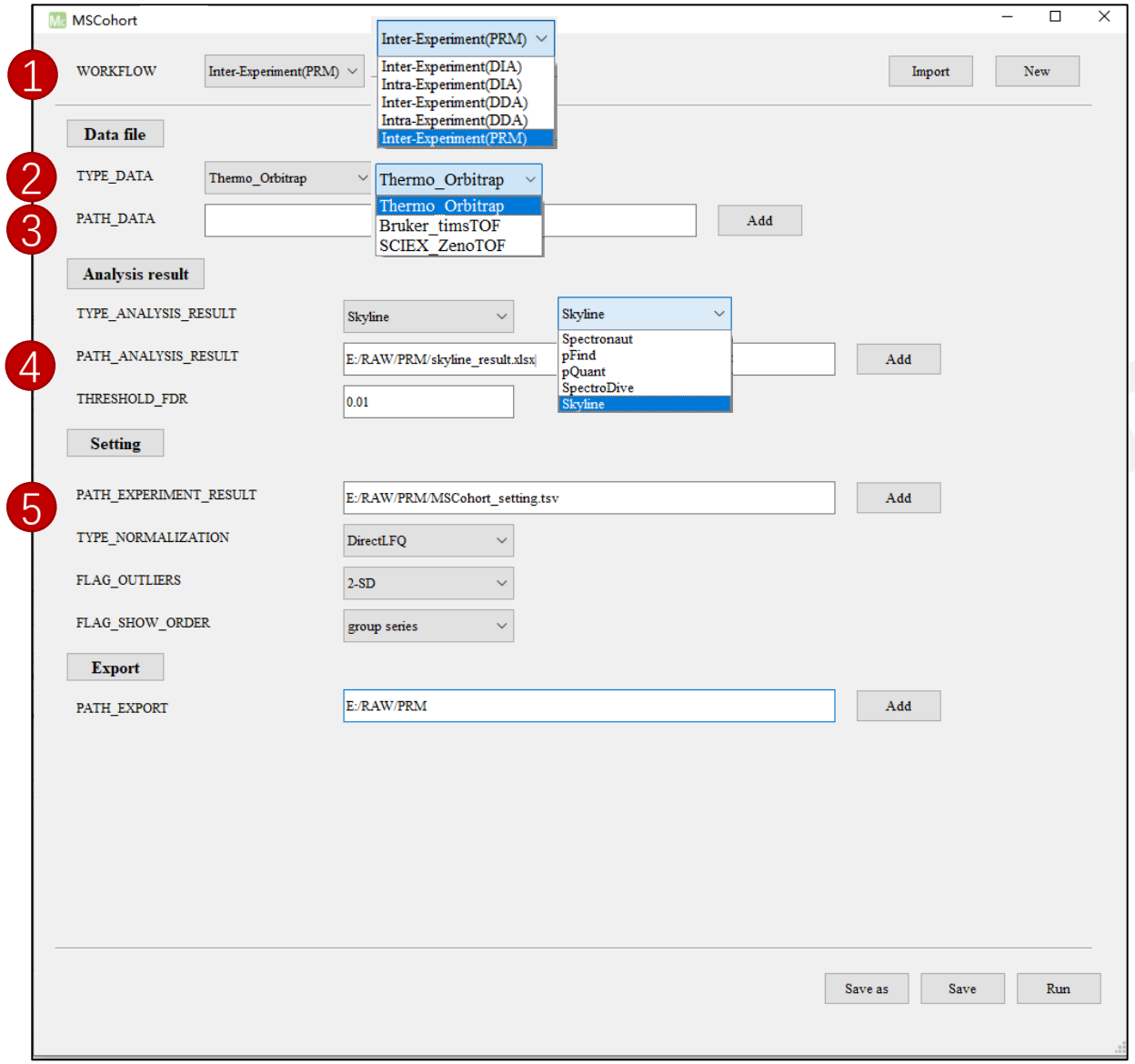


① Double-click MSCohortUI.exe;

② Wait for MSCohort program to load;

③ The MSCohortUI.exe settings screen is displayed.

2. Analyzing with MSCohort



① Set **WORKFLOW** as **Inter-experiment(PRM)**;

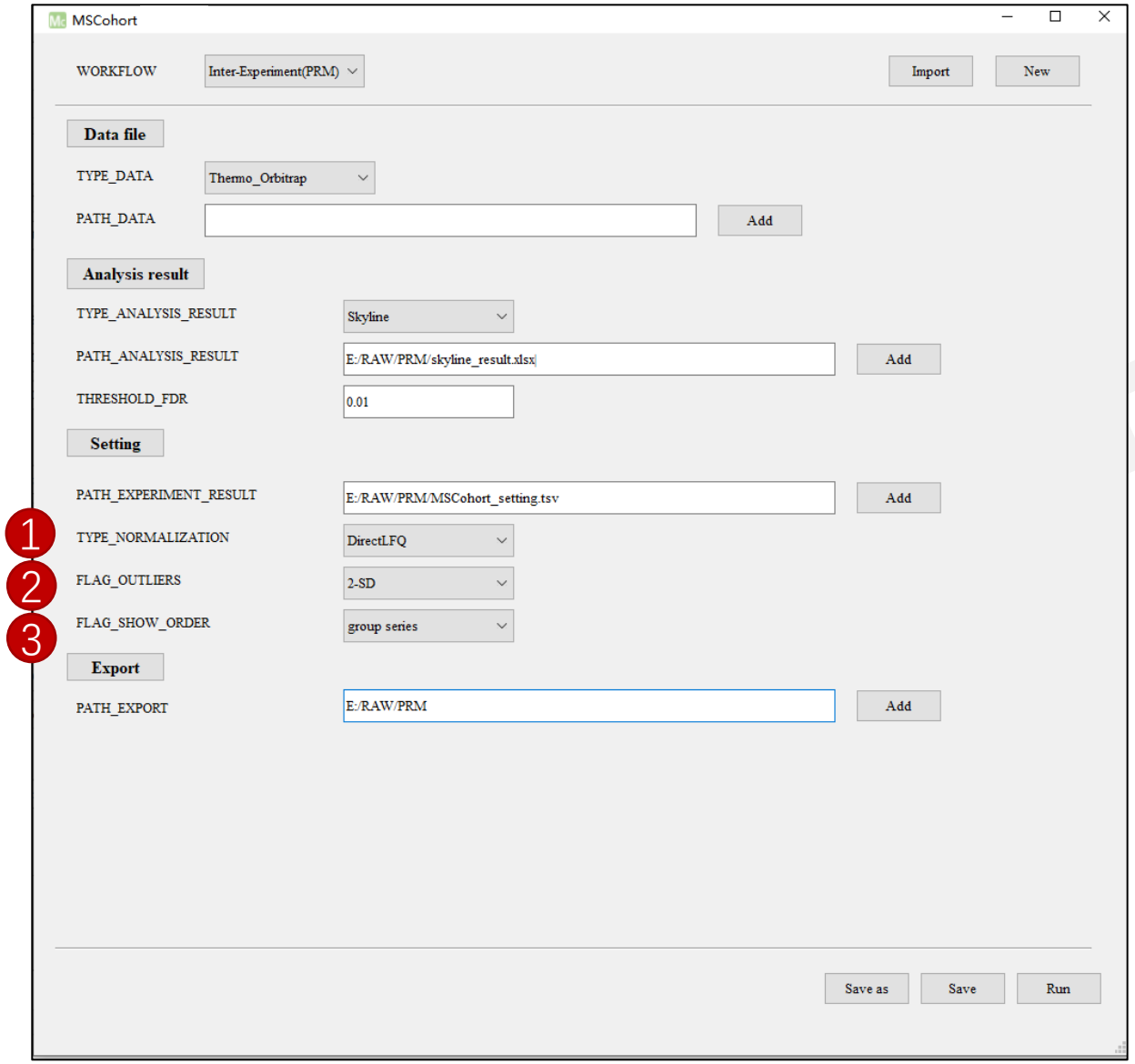
② Select **TYPE_DATA** according to the data type ;

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

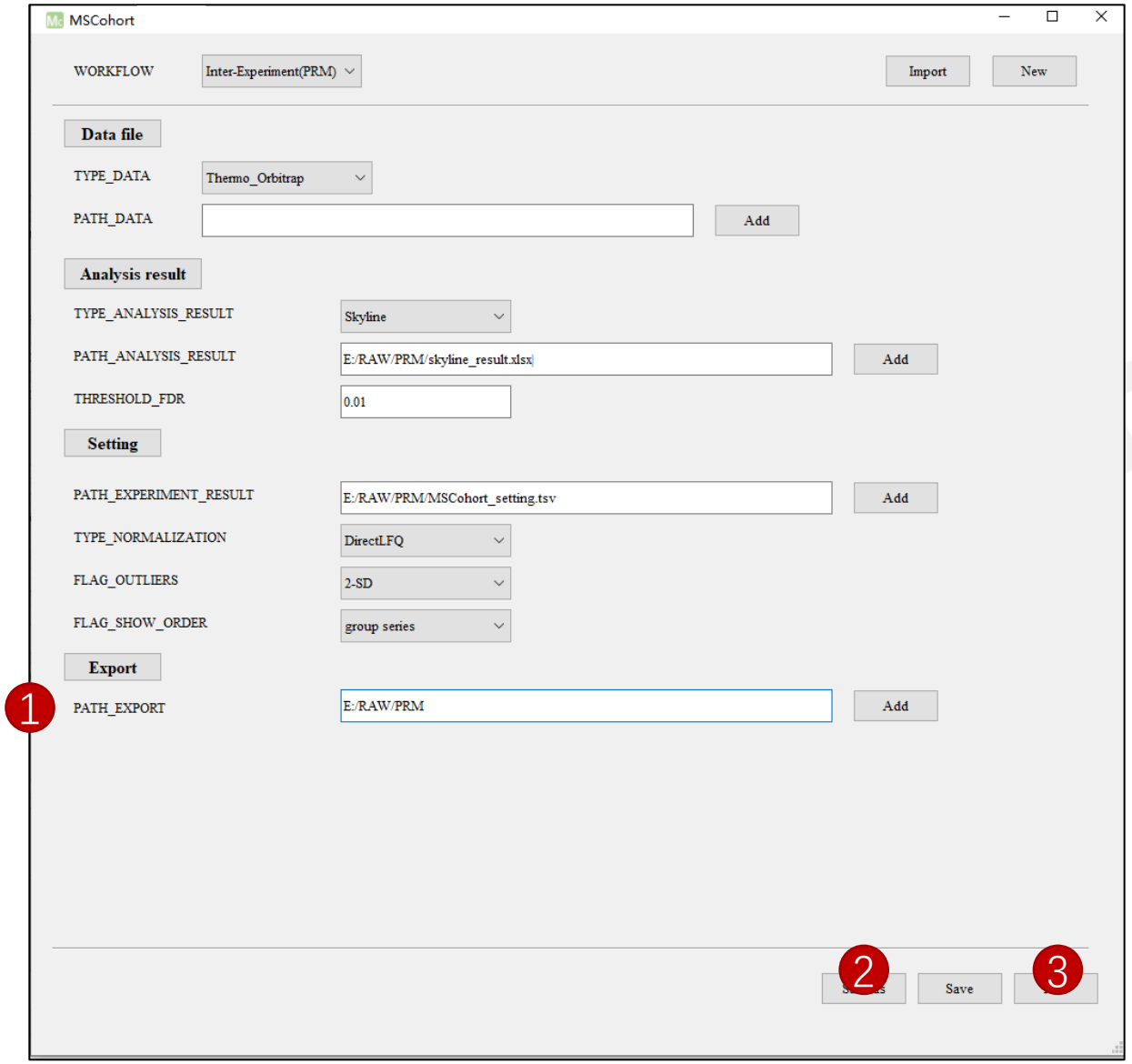


① Choose normalization strategies in **TYPE_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);

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



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.

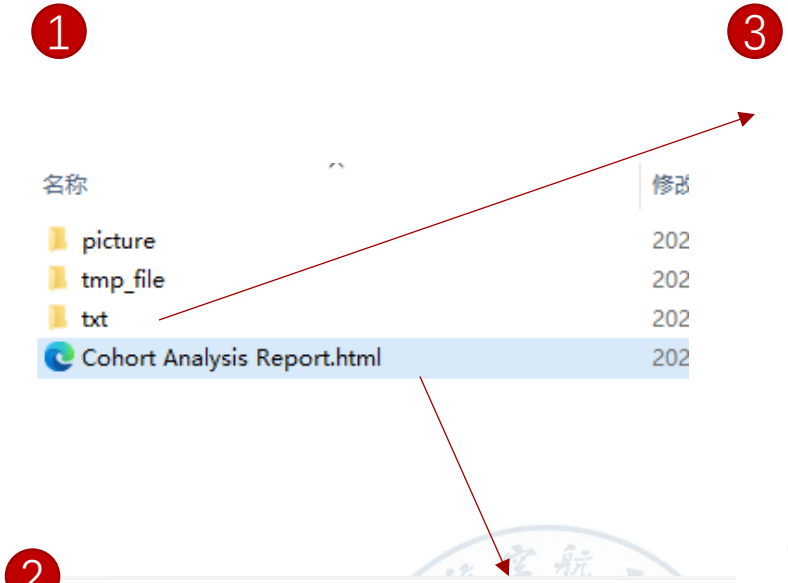
② Click **Save as** button to save the config file;

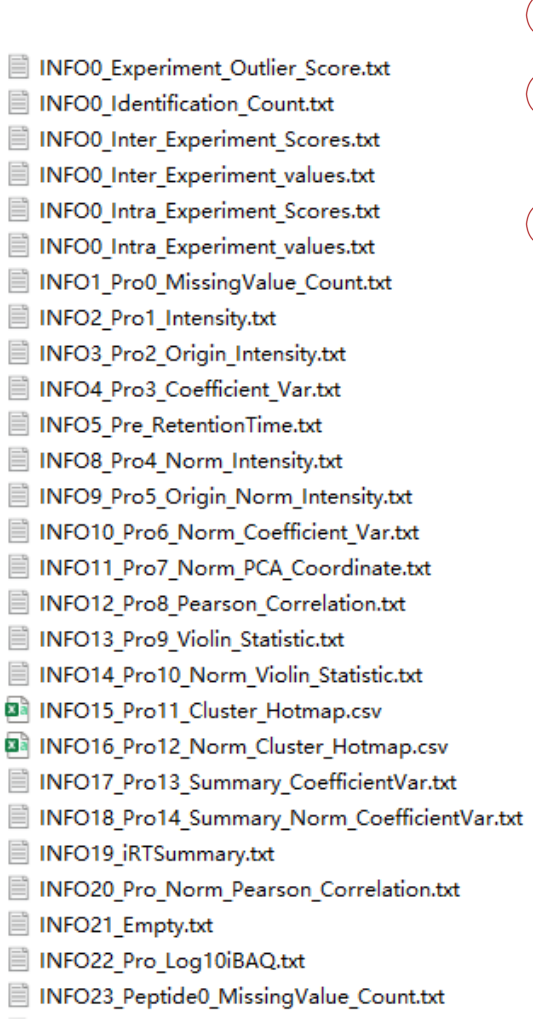
③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

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 The MSCohort results ;

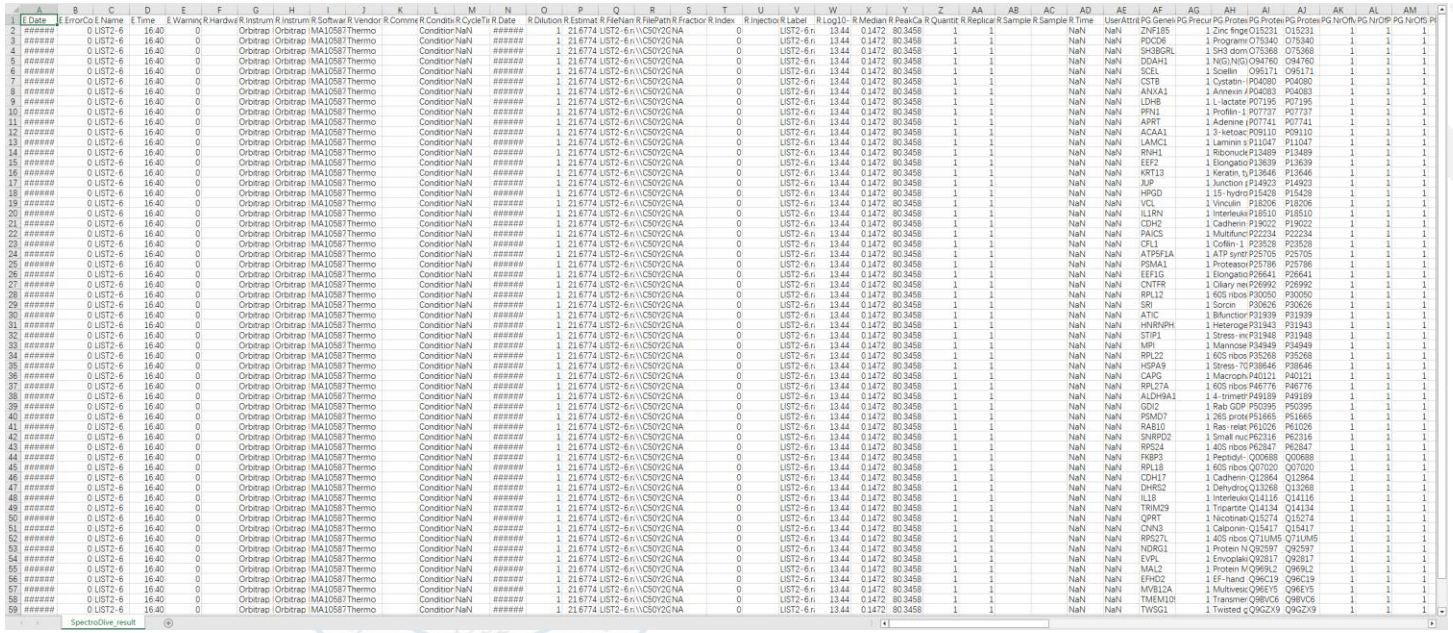
② Double clicking **Cohort Analysis Report.html**, the report will be showed in the browser.

③ Double clicking **txt** folder, the outputs are also exported to simple tab-delimited text files.

Demo dataset (SpectroDive)

名称	类型	大小
MSCohort_cfg_SpectroDive.tsv	TSV 文件	1 KB
SpectroDive_result.tsv	TSV 文件	31,426 KB

SpectroDive_result.tsv



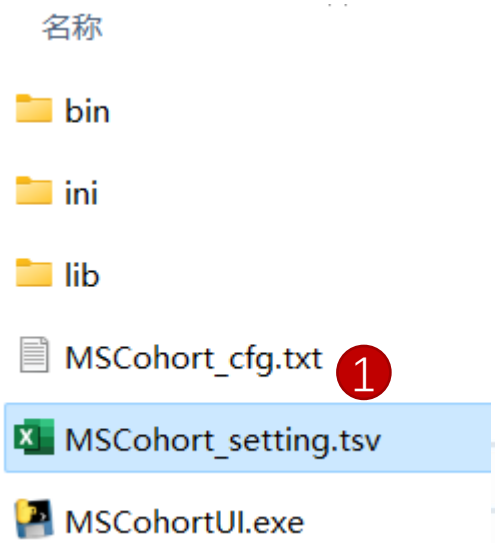
MSCohort_setting_SpectroDive.tsv

	A	B	C	D
1	Group Name	Raw Name	Experiment	Threshold
2	Group1	LXO-mix1.raw	LXO-mix1.raw	0
3	Group1	LXO-mix2.raw	LXO-mix2.raw	0
4	Group1	LXO-mix3.raw	LXO-mix3.raw	0
5	Group1	LXO-mix4.raw	LXO-mix4.raw	0
6	Group1	LXO-mix5.raw	LXO-mix5.raw	0
7	Group1	LXO-mix6.raw	LXO-mix6.raw	0
8	Group1	LXO-mix7.raw	LXO-mix7.raw	0
9	Group1	LXO-mix8.raw	LXO-mix8.raw	0
10	Group1	LXO-mix9.raw	LXO-mix9.raw	0
11	Group1	LXO-mix10.raw	LXO-mix10.raw	0
12	Group1	LXO-mix11.raw	LXO-mix11.raw	0
13	Group1	LXO-mix12.raw	LXO-mix12.raw	0
14	Group1	LXO-mix13.raw	LXO-mix13.raw	0

Taking the SpectroDive 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



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

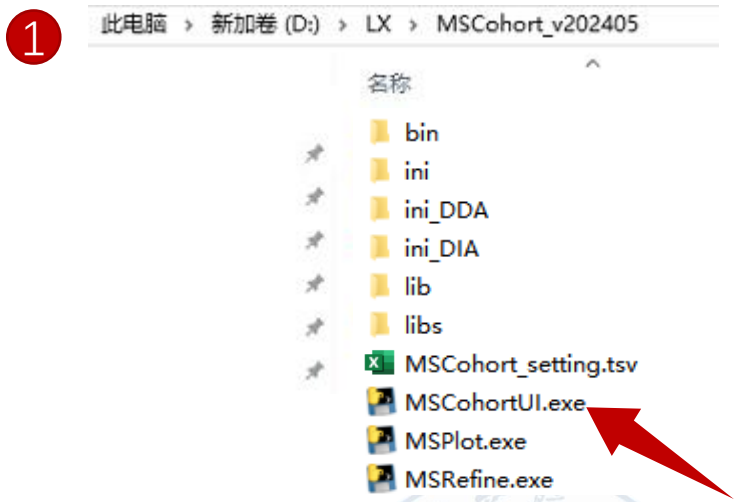
① Open MSCohort_setting.tsv with Excel;

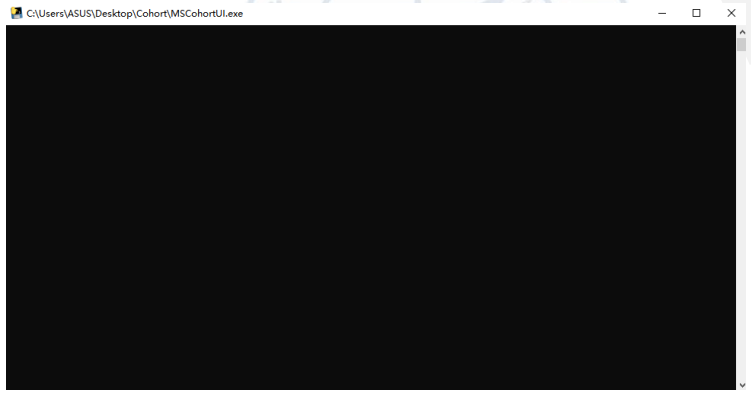
② Fill the columns as the example file (MSCohort_setting_SpectroDive.tsv) showed or alternatively, directly use the MSCohort_setting_SpectroDive.tsv file:

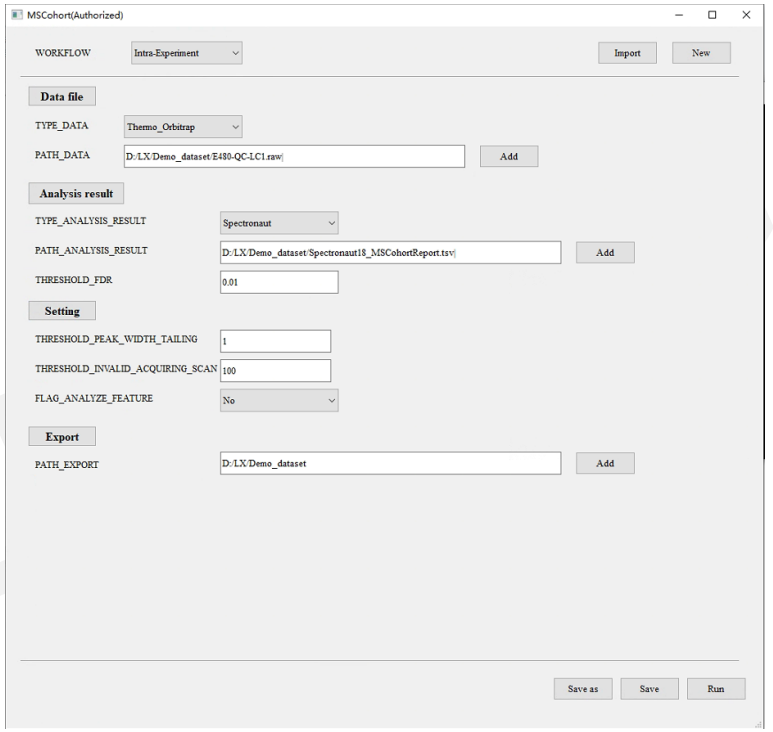
- The first column is the Group Name.
- The second column is the Raw Name, which is the same as R.FileName reported from SpectroDive.
- The third column is the Raw Name for short, which is convenient for display in large MSCohort report. This column could also be the same as Raw Name column.
- The fourth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





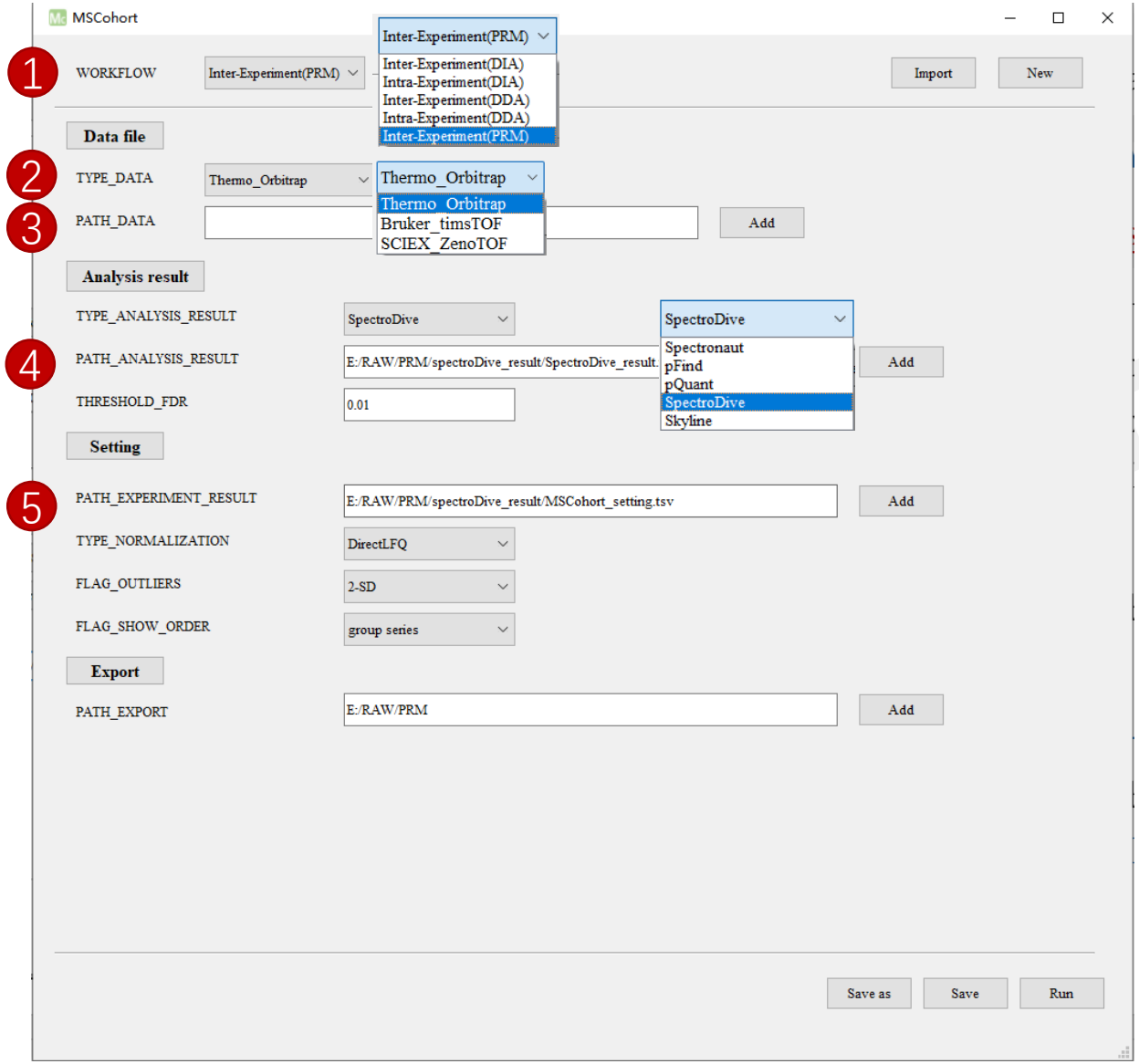


① Double-click MSCohortUI.exe;

② Wait for MSCohort program to load;

③ The MSCohortUI.exe settings screen is displayed.

2. Analyzing with MSCohort



① Set **WORKFLOW** as **Inter-experiment**;

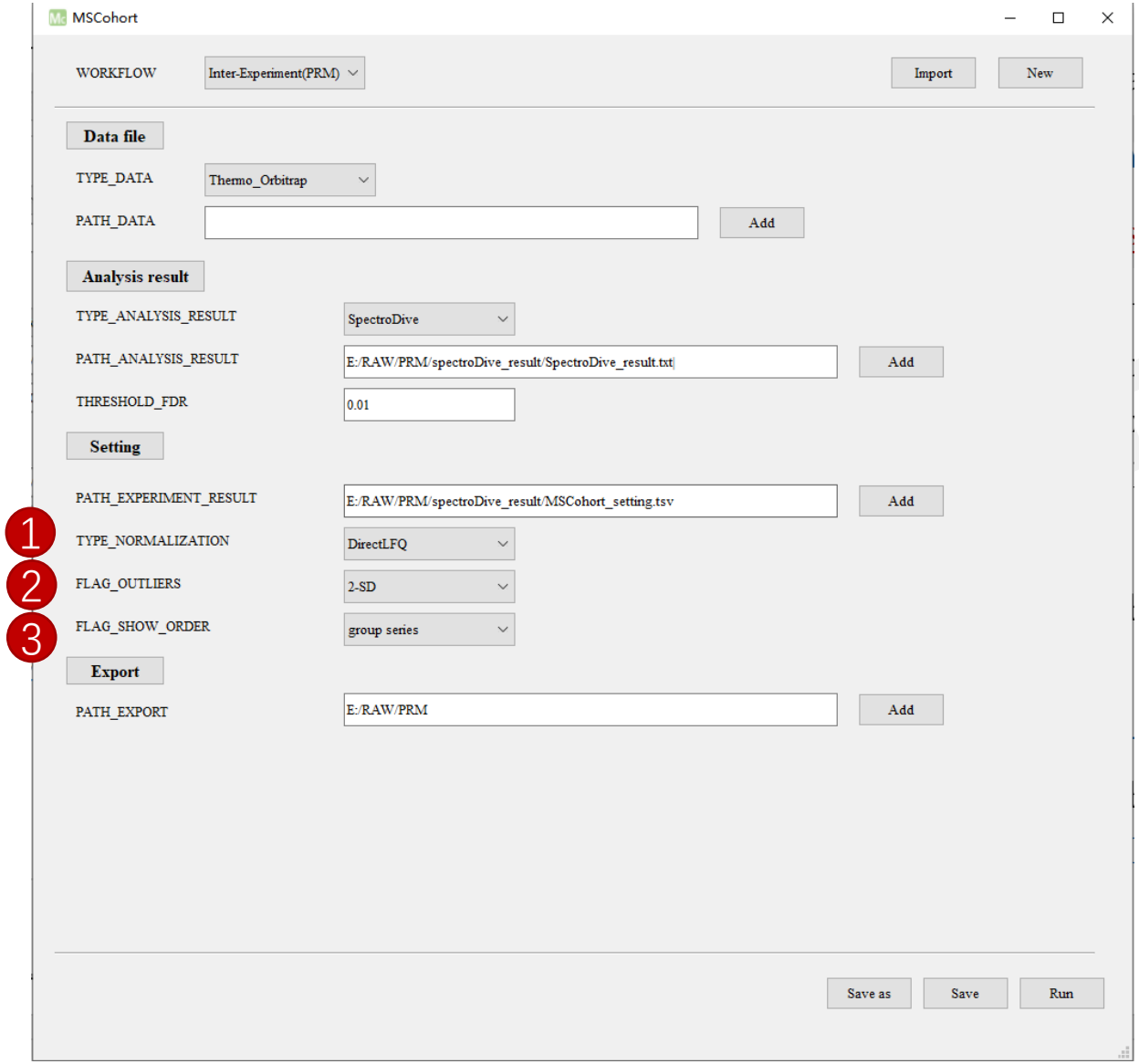
② Select **TYPE_DATA** according to the data type ;

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

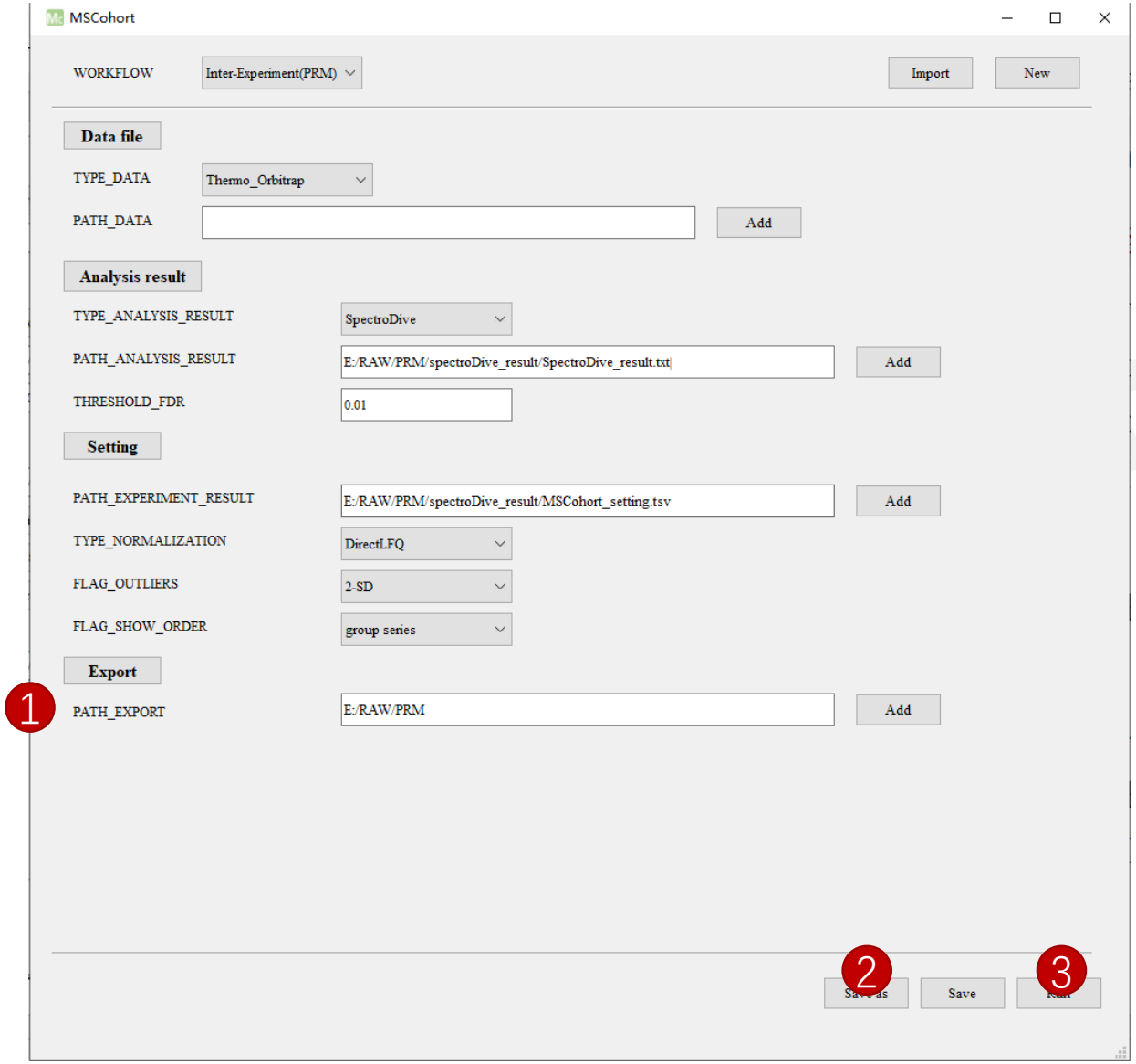


① Choose normalization strategies in **TYPE_NORMALIZATION**. MSCohort support directLFQ, MaxLFQ, and Quantile normalization (default is directLFQ);

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



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.

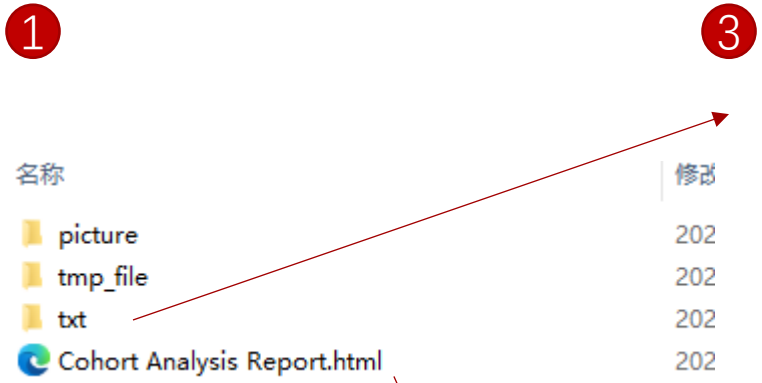
② Click **Save as** button to save the config file;

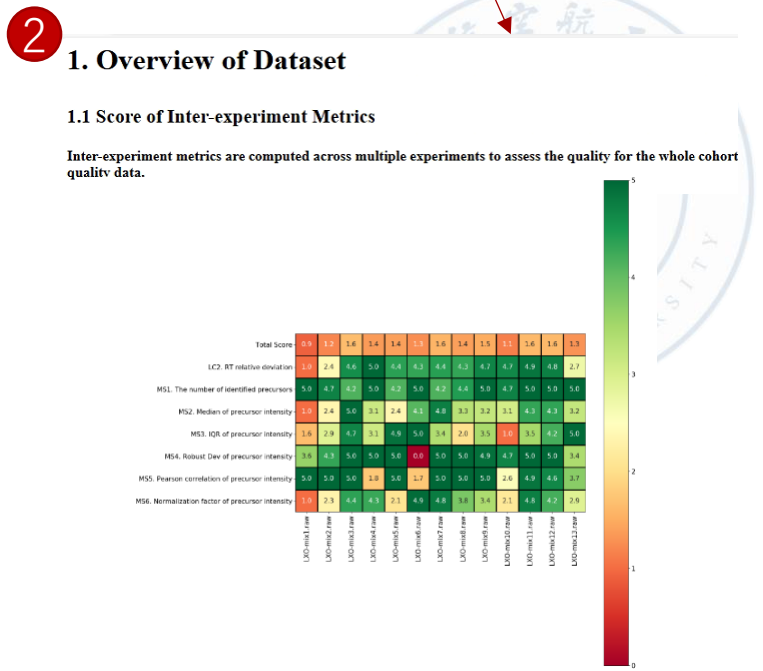
③ Click **Run** button to start the MSCohort, the progress information will be shown in the command-line interface.

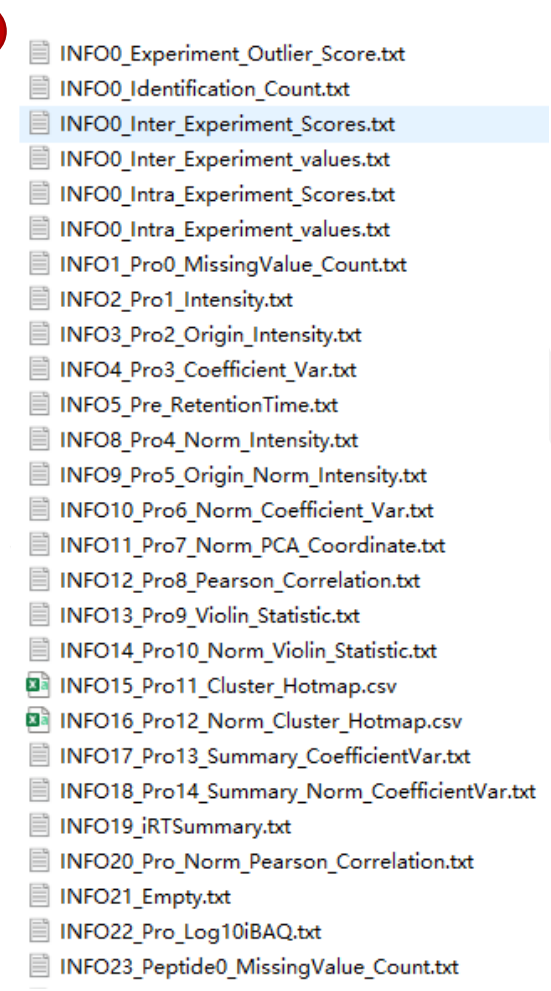
Note：

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 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.

Thanks!