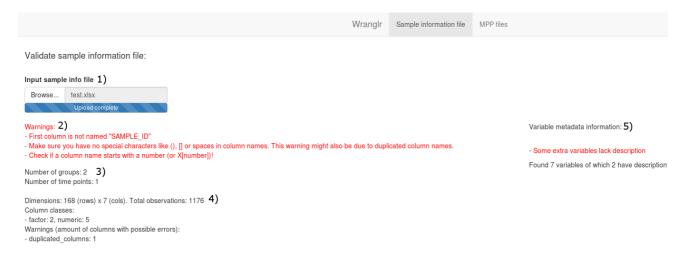# Wranglr instructions

Sample information file tab

Input sample information in a .xlsx file (see specification file). Wranglr will process the sample information file and produce files that are compatible with MPP and Worklist.

Wranglr provides useful information to check that your sample information file meets the requirements and that the file is read correctly.

1) Input Excel sheet as described in sample_form_specification.docx
2) Wranglr displays warnings if your file does not meet the requirements. Any warnings here will prevent processing and must be dealt with.
3) If column(s) named GROUP and/or TIME are found, the number of groups and/or time points found will be displayed.
4) Some extra information on the sample information sheet. Warnings here do not prevent processing
5) Information on the variable metadata. Warnings here do not prevent processing.

Randomize and add QC

Project title:

Demo

The complete path to project home folder.
Defaults to D:\MassHunter\Data\[project title]\

☐ Save project code Toggle this if you want Wranglr to remember the code given to this project.

Project home folder

D:\MassHunter\Data\Demo\

Sample run order

Random ▼

Choose between original order and randomized.

Sample positioning type

◉ 96-well plate
◯ 54-vial plate

Number of samples before every QC:

12 ▼

Choose between multiple intervals.
It is possible to run without any QC.

Modes

☑ HILIC neg ☑ HILIC pos ☑ RP neg ☑ RP pos Choose which modes will be run

Number of QC before beginning of mode:

HILIC neg	HILIC pos	RP neg	RP pos
10	10	10	10

QC sample position of mode: The same position will be given to every QC sample in the respective mode. Defaults to the last position on the second plate.

HILIC neg	HILIC pos	RP neg	RP pos
P2-H12	P2-H12	P2-H12	P2-H12

Second column for MPP

Group ▼

Choose the column used for grouping in MPP.
Defaults to GROUP column if found,
else defaults to QC.

Submit & Process

Note that the number of QC samples in the beginning of modes can change in course of analysis. In this case, instead of manually changing the worklist file, you can rerun the processing with corrected values. The randomization of run order calculates its seed number from the project title, so the run order will be identical if the sample information file of any individual project is reprocessed. Also, if you saved the project code, the same code will be used.

Example of a processed sample information file used in MPP and further analysis:

Preview:

◉ Processed sample information file for MPP
◯ Worklist file

Show 25 ▾ entries

DATAFILE	GROUP	RUN_ORDER	INTERNAL_SAMPLE_ID	SAMPLE_ID	QC
Demo_HILIC_neg_0001		1	QC_1		TRUE
Demo_HILIC_neg_0002		2	QC_2		TRUE
Demo_HILIC_neg_0003		3	QC_3		TRUE
Demo_HILIC_neg_0004		4	QC_4		TRUE
Demo_HILIC_neg_0005		5	QC_5		TRUE
Demo_HILIC_neg_0006		6	QC_6		TRUE
Demo_HILIC_neg_0007		7	QC_7		TRUE
Demo_HILIC_neg_0008		8	QC_8		TRUE
Demo_HILIC_neg_0009		9	QC_9		TRUE
Demo_HILIC_neg_0010		10	QC_10		TRUE
Demo_HILIC_neg_0011	C	11	A_1	115	FALSE
Demo_HILIC_neg_0012	B	12	A_2	42	FALSE
Demo_HILIC_neg_0013	A	13	A_3	68	FALSE
Demo_HILIC_neg_0014	B	14	A_4	56	FALSE
Demo_HILIC_neg_0015	C	15	A_5	75	FALSE
Demo_HILIC_neg_0016	B	16	A_6	74	FALSE
Demo_HILIC_neg_0017	B	17	A_7	5	FALSE
Demo_HILIC_neg_0018	B	18	A_8	58	FALSE

Wranglr generates four new columns: DATAFILE, RUN_ORDER, INTERNAL_SAMPLE_ID and QC. The internal sample IDs begin with project-specific code. The code is 1 or 2 letters from A to ZZ. If the project code is saved, the next project will be given the following code. The processed sample information file will also include all the columns of the original sample information file.

Example of a worklist file:

Show [25 ▾] entries   Search: [   ]

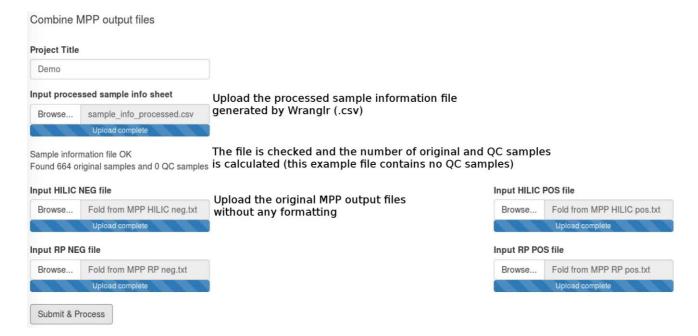| Sample Name	Sample Position	Data File
QC_1	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0001.d
QC_2	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0002.d
QC_3	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0003.d
QC_4	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0004.d
QC_5	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0005.d
QC_6	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0006.d
QC_7	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0007.d
QC_8	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0008.d
QC_9	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0009.d
QC_10	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0010.d
A_1	P1-A1	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0011.d
A_2	P1-A2	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0012.d
A_3	P1-A3	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0013.d
A_4	P1-A4	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0014.d
A_5	P1-A5	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0015.d
A_6	P1-A6	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0016.d
A_7	P1-A7	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0017.d
A_8	P1-A8	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0018.d
A_9	P1-A9	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0019.d
A_10	P1-A10	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0020.d
A_11	P1-A11	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0021.d
A_12	P1-A12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0022.d
QC_11	P2-H12	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0023.d
A_13	P1-B2	D:\MassHunter\Data\Demo\Demo_HILIC_neg\Demo_HILIC_neg_0024.d

The Sample Name column is identical to INTERNAL_SAMPLE_ID column in the processed sample information file. Sample position is generated automatically. The Data File column includes the complete file paths. Wranglr will generate a subfolder for each mode.

An individual worklist file will be generated for each mode. All the files generated by Wranglr can be downloaded as .csv files.

MPP files tab

In the MPP tab you can combine MPP output files with sample information. The MPP output files should include the raw abundance values and the following columns: Compound, Mass, Retention Time, CompositeSpectrum, Frequency, CompoundAlgo and ionization mode.

The combined data matrix is ready for further data analysis. Other information of the compounds is stored in separate annotation file. Both files can be downloaded as .csv files.

Combine MPP output files

Project Title

Demo

Input processed sample info sheet   Upload the processed sample information file
Browse... sample_info_processed.csv   generated by Wranglr (.csv)
Upload complete

Sample information file OK   The file is checked and the number of original and QC samples
Found 664 original samples and 0 QC samples   is calculated (this example file contains no QC samples)

Input HILIC NEG file   Upload the original MPP output files   Input HILIC POS file
Browse... Fold from MPP HILIC neg.txt   without any formatting   Browse... Fold from MPP HILIC pos.txt
Upload complete   Upload complete

Input RP NEG file   Input RP POS file
Browse... Fold from MPP RP neg.txt   Browse... Fold from MPP RP pos.txt
Upload complete   Upload complete

Submit & Process

Example of the combined data matrix:

Preview:
◉ Data matrix
○ Annotation file

Show 25 entries

RUN_ORDER	INTERNAL_SAMPLE_ID	SAMPLE_ID	GROUP	QC	COMPOUND_835_5739a0_60039145_HILIC_neg	COMPOUND_880_6041a0_6022438_HILIC_neg
1	A_1	86	1	FALSE	36876	29488
2	A_2	59	2	FALSE	21472	26513
3	A_3	47	2	FALSE	37850	18807
4	A_4	12	2	FALSE	35907	26220
5	A_5	8	1	FALSE	39034	22130
6	A_6	60	1	FALSE	41822	27177
7	A_7	46	1	FALSE	31022	35585
8	A_8	25	3	FALSE	37226	31912
9	A_9	77	2	FALSE	64101	34478
10	A_10	100	3	FALSE	41231	24626
11	A_11	131	2	FALSE	42003	32209

The data matrix has one row per sample. The columns from the processed sample sheet are combined with abundance values of all compounds found in all modes. The compound names are modified to better suit data analysis pipeline.

Example of an annotation file:

Preview:
○ Data matrix
● Annotation file

Show 25 ▾ entries   Search: [   ]

Compound	Mass	Retention.Time	CompositeSpectrum	Frequency	CompoundAlgo	Ionization.mode	chromatography
COMPOUND_835_5739a0_60039145_HILIC_neg	835.5739	0.6003915	(834.561, 618.15)(835.5636, 325.98)(836.5809, 2383.65)(837.5818, 1204.03)(838.5966, 6840.1)(839.6005, 3720.55)(840.5932, 2400.19)	161	FindByFormula	-	hilic
COMPOUND_880_6041a0_6022438_HILIC_neg	880.6041	0.6022438	(879.5964, 4105.43)(880.6052, 4976.04)(881.611, 2238.33)(882.6193, 1087.96)(883.6216, 516.88)	175	FindByFormula	-	hilic
COMPOUND_851_5713a0_61231047_HILIC_neg	851.5713	0.6123105	(850.5641, 40071.87)(851.5659, 20185.1)(852.5773, 24523.87)(853.5808, 11196.83)(854.5927, 18349.58)(855.5972, 8665.93)(856.6024, 4311.33)	181	FindByFormula	-	hilic
COMPOUND_827_5705a0_6225396_HILIC_neg	827.5705	0.6225396	(826.5634, 48164.44)(827.5662, 22517.98)(828.5764, 27366.94)(829.5799, 11596.97)(830.5916, 14789.17)(831.5959, 6514.86)(832.602, 3081.05)(833.6086, 1033.93)	181	FindByFormula	-	hilic
COMPOUND_1588_1497a0_63386375_HILIC_neg	1588.1497	0.6338637	(1587.1439, 1736.1)(1588.1519, 2181.12)(1589.1565, 1379.81)(1590.1635, 1079.31)(1591.1703, 599.48)	156	FindByFormula	-	hilic
COMPOUND_1583_1154a0_6332007_HILIC_neg	1583.1154	0.6332007	(1582.1046, 841.17)(1583.1107, 779.23)(1584.126, 2245.74)(1585.1256, 1989.31)(1586.1373, 2390.57)(1587.1411, 1810.44)(1588.1488, 1675.66)	165	FindByFormula	-	hilic
COMPOUND_803_5716a0_6422843_HILIC_neg	803.5716	0.6422843	(802.5643, 74265.74)(803.5665, 34034.41)(804.5776, 46775.65)(805.581, 19167.41)(806.5839, 5319.22)(807.5871, 1110.69)(808.5786, 257.74)	181	FindByFormula	-	hilic
COMPOUND_799_5405a0_6426328_HILIC_neg	799.5405	0.6426328	(798.5369, 502.62)(799.5325, 199.74)(800.5466, 7602.79)(801.5496, 3623.57)(1598.1224, 445.69)	179	FindByFormula	-	hilic
COMPOUND_743_5465a0_6734437_HILIC_neg	743.5465	0.6734437	(742.5393, 1332.7)(743.543, 587.2)(744.5544, 510.12)(745.5594, 150.68)	161	FindByFormula	-	hilic

The Chromatography column is added. Compound names correspond to those in the data matrix.