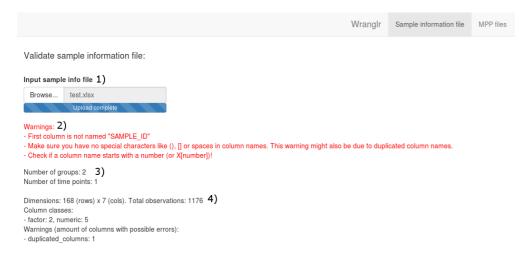
Wranglr instructions

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.



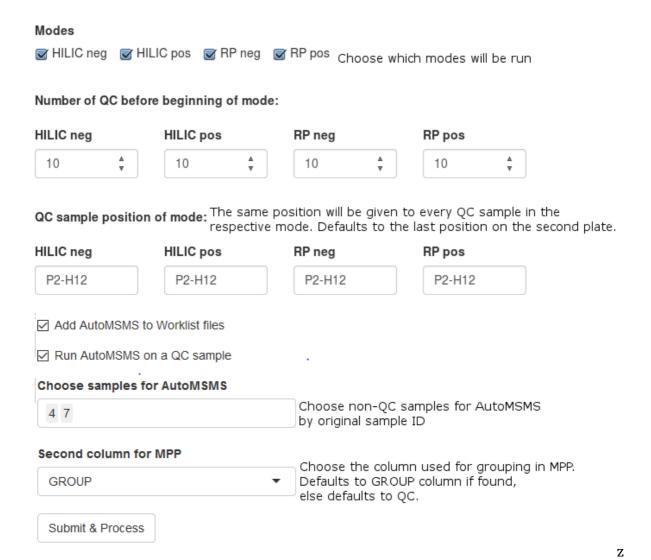
- 1) Input Excel sheet as described in sample form specification.docx
- 2) Wranglr displays warnings if your file does not meet the requirements. Any warnings here will prevent processing and must be dealt with.
- 3) If column(s) named GROUP and/or TIME are found, the number of groups and/or time points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

Randomize and add QC Project title: demo Project code Project code can be chosen freely (Metavuo will be used to generate project codes in the future.) В Wranglr will suggest a letter or combination of letters that it hasn't seen before Save project code Choose this to save the project code, so Wranglr doesn't suggest it to anyone else. Wrangl also remembers the code in the case you rerun the same project. Project home folder Project home folder, will be concatenated to the beginning of data file paths D:\MassHunter\Data\demo\ Sample run order Choose method for determining run order: original order, globally random or Random within group random within groups (e.g. sample types) Column containing groups for randomization If "Random within group" is chosen, a selection for the group GROUP column will appear If multiple samples originate from the same subject, they can be run sequently Run samples from same subject sequently to reduce variance between them Subject ID column SUBJECT_NO If the above checkbox is ticked, subject ID column can be chosen Sample positioning type 96-well plate 54-vial plate Number of plates available Choose how many plates are available to the machine. (the new Orbitrap can take 3 plates, while the older machine can only take 2) Number of samples before every QC: Choose the interval of the QC samples 12

If you have run the project with the same title before, Wranglr remembers the project code you used, ans suggests it as a default.

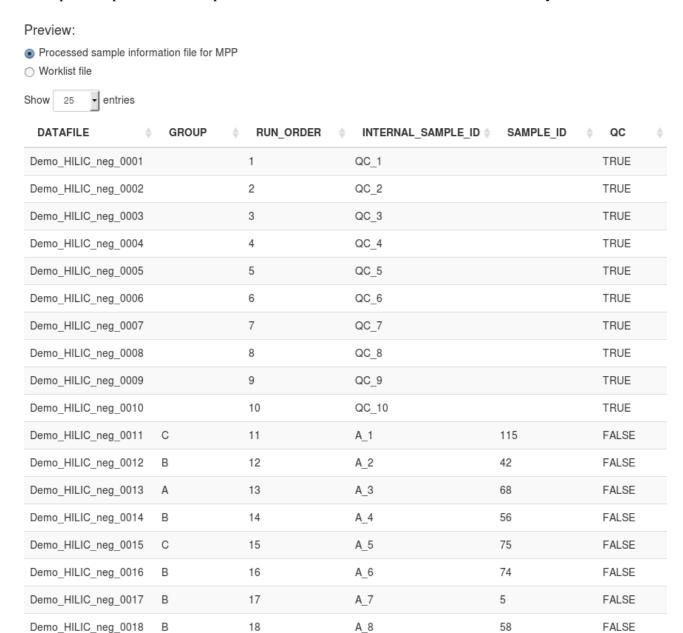
Running samples from the same subject sequently does not interfere with the randomization order. For example, if Sample run order is set to "Global Random" and the "Run samples from same subject sequently"-checkbox is ticked, the subject order will be randomized, but all samples from the same subject will be run sequently (in random order).

If Sample run order is set to "Random within group" as above and the "Run samples from same subject sequently"-checkbox is ticked, the subject order is randomized inside the groups and all samples from the same subject will be run sequently (in random order).



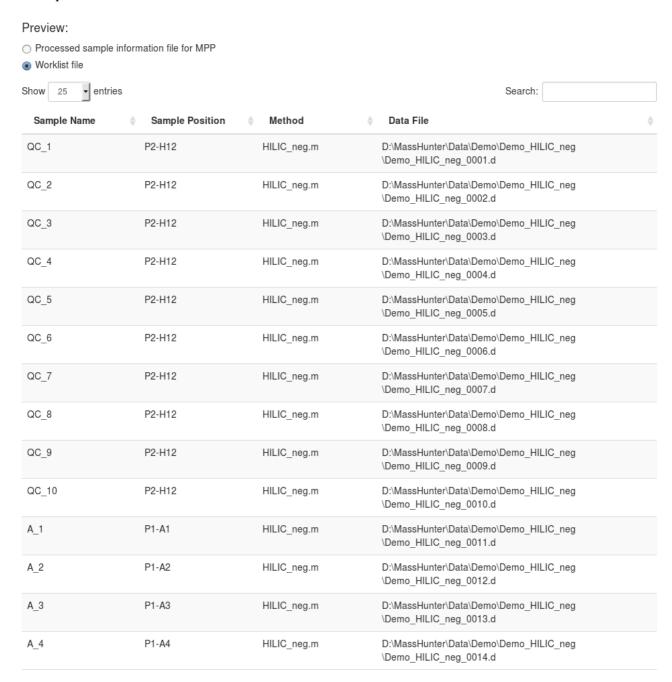
Note that the number of QC samples in the beginning of modes can change in course of analysis. In this case, instead of manually changing the worklist file, you can rerun the processing with corrected values. The randomization of run order calculates its seed number from the project title, so the run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

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



Wranglr generates four new columns: DATAFILE, RUN_ORDER, INTERNAL_SAMPLE_ID and QC. The internal sample IDs begin with project-specific code. The processed sample information file will also include all the columns of the original sample information file.

Example of a worklist file:



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

A_165	P2-F9	HILIC_neg.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0188.d
A_166	P2-F10	HILIC_neg.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0189.d
A_67_auto_msms_10V	P1-F7	HILIC_neg_AutoMSMS_10V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0190.d
A_67_auto_msms_20V	P1-F7	HILIC_neg_AutoMSMS_20V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0191.d
A_67_auto_msms_40V	P1-F7	HILIC_neg_AutoMSMS_40V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0192.d
A_80_auto_msms_10V	P1-G8	HILIC_neg_AutoMSMS_10V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0193.d
A_80_auto_msms_20V	P1-G8	HILIC_neg_AutoMSMS_20V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0194.d
A_80_auto_msms_40V	P1-G8	HILIC_neg_AutoMSMS_40V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0195.d
QC_1_auto_msms_10V	P2-H12	HILIC_neg_AutoMSMS_10V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0196.d
QC_1_auto_msms_20V	P2-H12	HILIC_neg_AutoMSMS_20V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0197.d
QC_1_auto_msms_40V	P2-H12	HILIC_neg_AutoMSMS_40V.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0198.d
STOP	Vial 1	STOP.m	D:\MassHunter\Data\Demo\Demo_HILIC_neg \Demo_HILIC_neg_0199.d

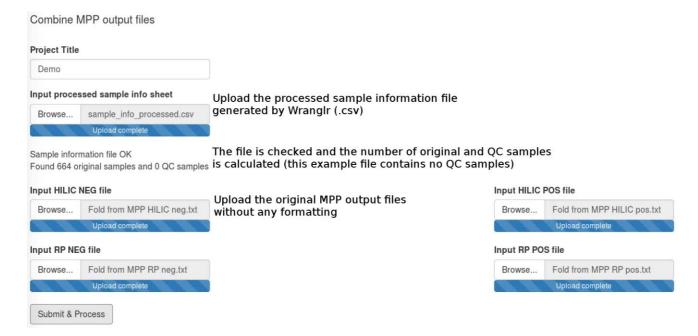
The AutoMSMS runs and the STOP run will be included in the end of the Worklist files. The position of STOP is always put to "Vial 1"

An individual worklist file will be generated for each mode. All the files generated by Wranglr can be downloaded as .csv files (with Windows line endings).

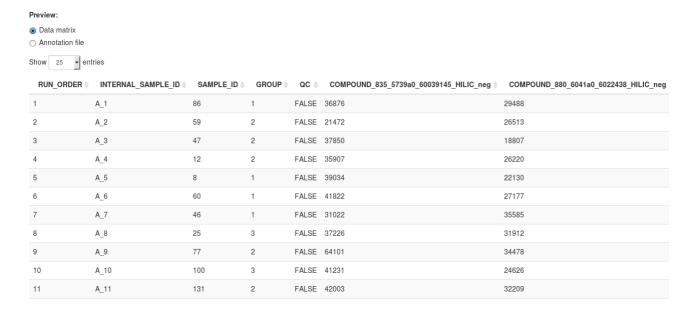
MPP files tab

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.



Example of the combined data matrix:



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:



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