THE USE OF FRAGMENT ION AND COLLISION CROSS SECTION FOR CONFIDENT IDENTIFICATION FROM LC-ION MOBILITY-MS METABOLOMICS DATA

Giorgis Isaac[1]; Giuseppe Astarita[1]; Steven Lai[1]; Adam Ladak[1]; James Langridge[3]; John P. Shockcor[1]; Andy Borthwick[2]
[1]Waters Corporation, Milford, MA; [2]Nonlinear Dynamics, Newcastle upon Tyne, UK ;[3]Waters Corporation, Manchester, UK

OVERVIEW

A novel and confident identification of metabolomics and lipidomics UPLC-ion mobility MS datasets based on fragment ion and collision cross section.

INTRODUCTION

- MS based technologies often couple with UPLC-Ion Mobility-MS (UPLC-IM-MS) to measure the level and variation of metabolites in biofluids. Data generated through metabolomics studies can yield important metabolic insights into disease onset and progression.
- UPLC-IM-MS based metabolomics generate large and complex datasets and data analysis and interpretation is the slowest step in any metabolomics workflow.
- In this study, we present a simple and novel informatics solution that has been specifically developed for the large scale analysis of UPLC-Ion Mobility-MS data from metabolomics and lipidomics datasets.
- A dataset consisting three sample mixtures was created to illustrate the processing features of Progenesis QI. A control (C) urine sample with out spiking, low dose (LD) urine sample spiked with 10 LD drugs and high dose (HD) urine sample spiked with 10 HD drugs (Table 1). According to the experimental design, Progenesis QI should:
  1. Identify all the spiked standards by means of database search.
  2. Classify three distinct groups and identify the marker ions responsible for the group separation.
  3. High level of relative abundance of the spiked standards for the LD and HD groups compared to control.

METHODS

Sample Preparation

Urine samples were collected from a single healthy individual. The urine was centrifuged and 600 µL of the urine supernatant was taken and diluted with 1800 µL of water. The urine was divided into three groups: control (C), low dosed (LD) and high dosed (HD). To create a working differential urine samples 10 different drugs were spiked at LD and HD as shown in Table 1.

UPLC Conditions

Instrument	Waters ACQUITY UPLC System
Column	ACQUITYHSS T3 C_{18}, 1.7µm, 2.1 x 100 mm
Column temp	40 °C
Mobile phase	A) 100% water (0.1% formic acid)
	B) 100% ACN (0.1% formic acid)
Injection	1.0 µL; Flow rate 0.5 mL/min
Gradient:	time (min) %B
	0.0 1
	1.0 1
	3.0 15
	6.0 95
	10.0 95
	10.1 1

MS Conditions

MS System	Waters Synapt™ G2-S
Mode of operation	TOF MS^E and TOF $HDMS^E$
Ionization	ESI +ve
Capillary voltage	0.5 KV; Cone voltage 30.0 V
Transfer CE	Ramp 20-40 V (Function 2)
Source temp.	120.0
Desolvation temp.	500.0
Desolvation gas	800.0 L/hr (N_2)
Mobility gas	90.0 mL/min (N_2)
Acquisition range	100-1000

No.	Compound	Molecular Formula	Monoisotopic mass (Da)	Control (ng/µL)	LD (ng/µL)	HD (ng/µL)
1	Acetaminophen	$C_8H_9NO_2$	151.0633	0	0.5	5
2	Acetanilide	C_8H_9NO	135.0684	0	0.5	5
3	Caffeine	$C_8H_{10}N_4O_2$	194.0804	0	0.5	5
4	Phenancetin	$C_{10}H_{13}O_2N$	179.0946	0	0.5	5
5	Ranitidine	$C_{13}H_{22}N_4O_3S$	314.1413	0	0.1	1
6	Alprazolam	$C_{17}H_{13}ClN_4$	308.0829	0	0.1	1
7	Tolbutamide	$C_{12}H_{18}N_2O_3S$	270.1038	0	0.1	1
8	Warfarin	$C_{19}H_{16}O_4$	308.1049	0	0.1	1
9	Lidocaine	$C_{14}H_{22}N_2O$	234.1732	0	0.03	0.3
10	Plavix	$C_{16}H_{16}ClNO_2S$	321.0590	0	0.03	0.3

Table 1. Preparation of control, low dose and high dose drug mixtures spiked to urine matrix.

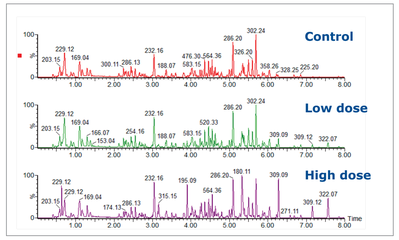

Figure 1. Representative chromatogram of control, low dose and high dose spiked urine sample.

RESULTS

Intuitive Step by Step Workflow

The Progenesis QI software adopts an intuitive step by step workflow to perform comparative UPLC-IM-MS metabolomics data analysis. The key to data processing and analysis is the ability of the software to distinguish biological variation and metabolic changes from analytical interferences. It is crucial that each sample is randomized and injected a minimum of three times to ensure that the data analysis is statistically valid. For this study the biological groups were randomized and injected six times with a set of QC pooled sample runs.

Progenesis QI Imports Multi Vendor Raw Data

The workflow starts with UPLC-IM-MS raw data file importing. Details of data file format and a list of expected adducts are entered to facilitate the handling of data import followed by automatic retention time alignment. The imported data is presented as Ion Intensity Map of m/z versus retention time enabling the visual check of the chromatography. Progenesis QI supports importing LC- data independent (MS^E), mobility data independent ($HDMS^E$), DDA, profile and centroid multi vendor raw MS data.

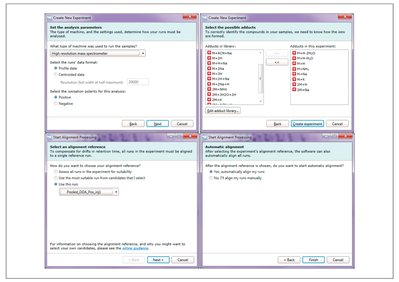

Figure 2. Data importing and retention time alignment. Progenesis supports multi vendor MS data.

Review Retention Time Alignment

Metabolomics experiments involve large amount of sample runs that may result shift in retention time. Progenesis aligns sample ions automatically to compensate for drift in retention time between analytical runs.

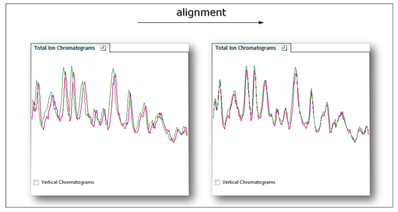

Figure 3. Total ion chromatogram displaying before (left) and after retention time alignment.

Experiment Design Setup

Multiple comparisons of biological conditions can be defined at the Experimental Design Setup stage or before and after a completed analysis. Progenesis QI provides an easy to follow and flexible "between-subjects" and "within-subjects" experimental design.

Peak picking and Normalization

Alignment of the data runs allows a common pattern of compound ion detection (peak picking) to be performed across all the runs in the experiment.

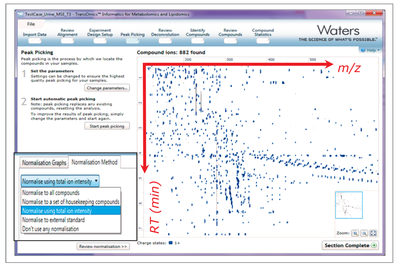

Figure 4. Peak picking with different data normalization options.

Review Deconvolution

Adducts of the same compound are automatically grouped during deconvolution. This uses the list of defined adducts provided when importing the raw data. The example in figure 5 shows the deconvolution of the compound eluting at 7.15 min and neutral mass of 308.1043 where three adducts (M+H, M+Na and M+K) are shown to be coincident with respect to Mass and Retention time. The Peak shows the actual abundance of the three adducts of this compound.

Figure 5. Deconvolution of the adducts from the same compound enables the neutral mass to be determined for more specific database identification and more accurate quantitation.

Identify Compounds

Using the integral Metascope search engine and a compound database all the 10 spiked compounds were identified in the complex sample set using defined tolerances for precursor exact mass, retention time, collision cross section (CCS) and fragment ion information.

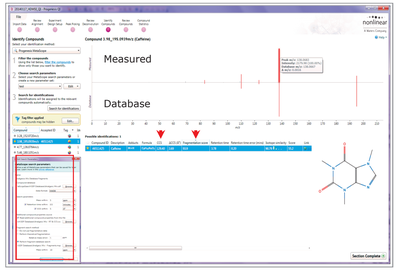

Figure 6. Identification based on precursor exact mass, retention time, isotopic distribution, CCS and fragment ion information provides confident metabolite ID.

Review Compound

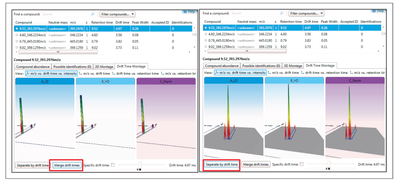

Figure 7. Co-eluting isobaric metabolites (left) are resolved by ion mobility (Right).

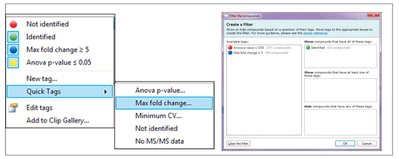

Figure 8. Progenesis QI allows to create a short list compound of interest that can be tagged for further review (left). Compounds can be shown or hidden based on a selection of their tags (right).

Compound Statistics

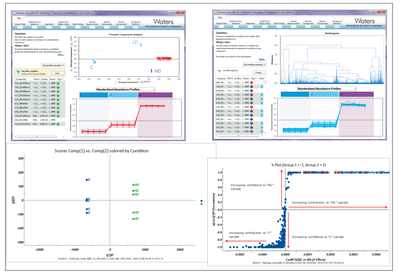

Figure 9. PCA, trend plot, correlation analysis, OPLS DA and S-plot generated for all compounds. From the S-plot the selected significant contributing features can be further investigated.

CONCLUSION

- **Progenesis QI workflow provides a easy-to-use, scalable system for analysis of metabolomic data.**
- **Accurate peak alignment and peak picking.**
- **Precise quantification of compounds, including deconvolution of the compound ions.**
- **LC-$HDMS^E$ (LC-DIA-IM-MS) provides both qualitative and quantitative information in a single experiment increasing the resolving power as the complexity of the sample increases.**
- **Compound identification based on precursor exact mass, isotopic distribution, retention time, fragmentation and collision cross section (CCS) provides confidence in metabolite identification.**
- **All 10 spiked standards were identified**
- **Three sample group clusters (C, LD & HD) were produced and the spiked standards were found the features that differentiate the most between the three groups. Based on the q-value, the standards are ranked at the top 20 most contributing features.**
- **The trend plot showed an increase in the LD and HD groups for the spiked standards compared to control.**

TO DOWNLOAD A COPY OF THIS POSTER, VISIT WWW.WATERS.COM/POSTERS