[PRODUCT SOLUTION]

Discovery lipidomics study for colorectal cancer using a Xevo™ MRT Mass Spectrometer and Lipostar data processing workflow

Authors: Nyasha Munjoma[1], Paolo Tiberi[2], Laura Goracci[3] Lee A. Gethings[1], Jayne Kirk[1], Richard Lock[1]
Affiliations: [1]Discovery and Development, Waters Corporation, Wilmslow, SK9 4AX, United Kingdom
[2]Mass Analytica Ltd., Rambla del Celler, 113, local 08172 Sant Cugat del Vallés, Spain.
[3]University of Perugia, Department of Chemistry, Biology and Biotechnology, University of Perugia, Via Elce di Sotto 8, 06123 Perugia, Italy

INTRODUCTION

Despite developments in analytical technology the detection and identification of lipids remains a significant challenge. To streamline the process of data acquisition and analysis we have developed a workflow which combines the attributes of the Xevo MRT Mass Spectrometer, with data analysis powered by the Lipostar2 software (Mass Analytica, Barcelona, Spain). The benefits of this approach are demonstrated using the lipidomics data derived from the analysis of plasma samples from colorectal cancer (CRC) and healthy control plasma. We highlight the key features and benefits of the waters_connect™ Software Platfrom combined with Lipostar2 software, with data acquired in DIA mode. The resulting data showed dysregulation of the lipid metabolism pathways based on cancer type.

Data were acquired using the Xevo MRT MS (Figure 1) coupled to an ACQUITY™ Premier LC. The mobile phase A contained 10mM Ammonium Formate in Acetonitrile:Water and B contained Ammonium Formate in IPA:Acetonitrile. A CSH™ C_{18} (2.11 x 100 mm) Column and a 12-minute (50-99% B) reverse phase gradient. Detail can be found in the application note [A Robust and Reproducible Reversed-Phase Lipid Profiling Method for Large Sample Sets | Waters]

The data were collected collected using the waters_connect Software Platform and transferred to Lipostar2 software via the UNIFI™ application programming interface (API) and as mzML files. Peak picking, alignment, statistical analysis, identification, and pathway analysis were performed using Lipostar2 software.

Data Transfer from waters_connect platform to Lipostar2

The data generated using high-resolution instruments on waters_connect applications can be directly imported into the most recent version of Lipostar2 using the (API). However, many existing labs will have informatics and data pipelines already in place so for added flexibility, the user has options available to them to convert their data to mzML file format for use in popular third party software such as MS-Dial*, MZmine#, XCMS™ as well as Skyline. The conversion of format can either be triggered at the point of acquisition through the Acquisition Method Editor (AME) or post-acquisition, using DATA Convert Applications. The advantage of mzML is that it is a universal file format. This workflow is described in Figure 2.

Figure 1: The Xevo MRT Mass Spectrometer

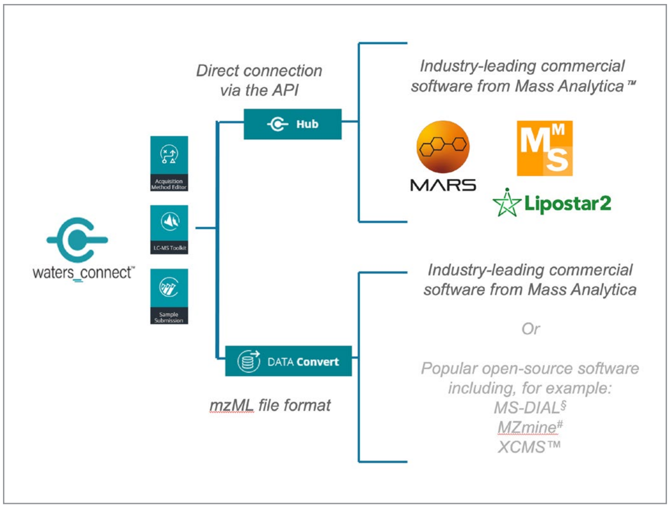


Figure 2: The data transfer workflow of data from waters_connect to Mass Analytica software and other third-party solutions

RESULTS AND DISCUSSION:

Validated Supervised models of Controls vs CRC samples with and without QC samples included. As well as stratification based on CRC type, colon, and rectum. The QCs are shown in red, they tightly cluster in the center of the model demonstrating good reproducibility over the analysis (Figure 3).

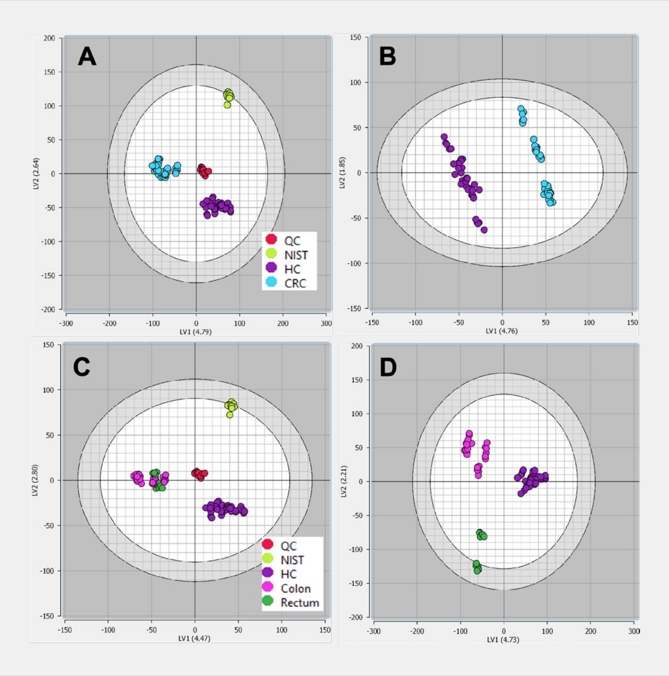
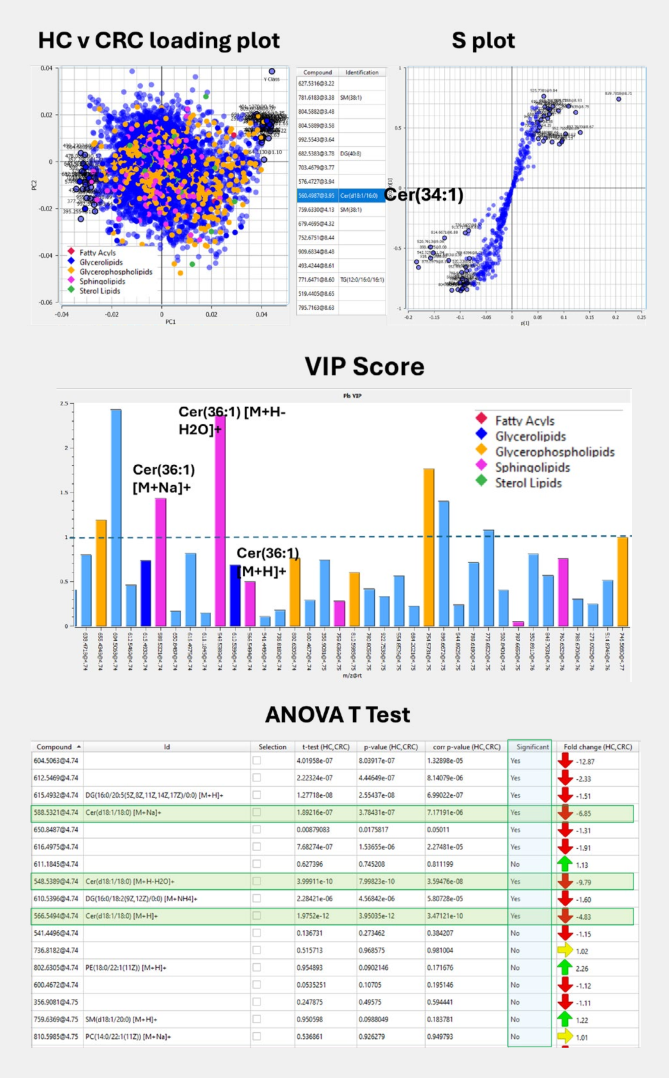


Figure 3: PLS-DA models of Controls vs CRC patients including NIST and study reference QCs (A) and with QCs removed from model (B). PLS-DA models of Controls vs Colon vs Rectum patients including NIST and study reference QCs (C) and with QCs removed from model (D)

When the QC samples (including NIST plasma) were removed clear separation of healthy controls and CRC cohorts is observed. Colon cancer plasma could be separated from Rectum cancer and healthy controls if SR QC samples are excluded from the model generation. However, the colon and rectum cancer samples cluster together if QC groups are included because PLS-DA seeks to find the maximum difference between groups. The resulting Multivariate Statistical Analysis (MVA) data were interrogated to identify the most significant lipid using the loading plot, S-Plot and VIP scores (Figure 4).



Compound ▲	Id	Selection	t-test (HC,CRC)	p-value (HC,CRC)	corr p-value (HC,CRC)	Significant	Fold change (HC,CRC)
604.5063@4.74		☐	4.01958e-07	8.03917e-07	1.32898e-05	Yes	↓ -12.87
612.5469@4.74		☐	2.22324e-07	4.44649e-07	8.14079e-06	Yes	↓ -2.33
615.4932@4.74	DG(16:0/20:5(5Z,8Z,11Z,14Z,17Z)/0:0) [M+H]+	☐	1.27718e-08	2.55437e-08	6.99022e-07	Yes	↓ -1.51
588.5321@4.74	Cer(d18:1/18:0) [M+Na]+	☐	1.89216e-07	3.78431e-07	7.17191e-06	Yes	↓ -6.85
650.8487@4.74		☐	0.00879083	0.0175817	0.05011	Yes	↓ -1.31
616.4975@4.74		☐	7.68274e-07	1.53655e-06	2.27481e-05	Yes	↓ -1.91
611.1845@4.74		☐	0.627396	0.745208	0.811199	No	↑ 1.13
548.5389@4.74	Cer(d18:1/18:0) [M+H-H2O]+	☐	3.99911e-10	7.99823e-10	3.59476e-08	Yes	↓ -9.79
610.5396@4.74	DG(16:0/18:2(9Z,12Z)/0:0) [M+NH4]+	☐	2.28421e-06	4.56842e-06	5.80728e-05	Yes	↓ -1.60
566.5494@4.74	Cer(d18:1/18:0) [M+H]+	☐	1.9752e-12	3.95035e-12	3.47121e-10	Yes	↓ -4.83
541.4496@4.74		☐	0.136731	0.273462	0.384207	No	↓ -1.15
736.8182@4.74		☐	0.515713	0.968575	0.981004	No	→ 1.02
802.6305@4.74	PE(18:0/22:1(11Z)) [M+H]+	☐	0.954893	0.0902146	0.171676	No	↑ 2.26
600.4672@4.74		☐	0.0535251	0.10705	0.195146	No	↓ -1.12
356.9081@4.75		☐	0.247875	0.49575	0.594441	No	↓ -1.11
759.6369@4.75	SM(d18:1/20:0) [M+H]+	☐	0.950598	0.0988049	0.183781	No	↑ 1.22
810.5985@4.75	PC(14:0/22:1(11Z)) [M+Na]+	☐	0.536861	0.926279	0.949793	No	→ 1.01

Figure 4: Various tool available in Lipostar to interrogate the most relevant features found to drive clustering of sample cohort groups

The variables were color coded based on the class identifications. The 50 most relevant variables resulting from the separation observed in the PLS-DA model, are highlighted in the loading plot and can be displayed as a list of mass/charge – retention time pair feature (*m/z*@tR). The associated identification information can be applied when an approved assignments Compounds selected a relevant will also appear in S plots have VIP scores tables as well. For VIP score plot, those greater than one are typically regarded as being important to the separation of groups. Further advanced tools for the visualization and interpretation of the statistical analysis are available as shown Anova T Test table with p-values/fold changes. A t-test is used to determine whether there is a statistically significant difference between the means of two groups. Figure 3 highlights how Cer(36:1) adduct mean is a significantly between control and CRC samples.

As previously stated, the accurate identification of putative lipids is extremely challenging, with databases being used extensively. A description of the Lipostar2 identification workflow is discussed in detail by Goracci *et al*, 2017. Lipostar2 uses a rule-based approach to generate theoretical fragments for lipid structures. For example, a fragmentation rule can be labelled as "mandatory" as in the case of the choline head group fragment with *m/z* 184.074 for the protonated phosphocholines or *m/z* 264.270 for protonated ceramides. Adducts of a particular compound can also be clustered based on retention time and *m/z* values. Identified lipids are linked to the various pathways they have been associated within peer reviewed literature. The example shown in Figure 5 is of a Cer 42:1 which was assigned green because it has a low mass accuracy error high fragment score, three associated adduct at the same retention time of 7.92 minutes.

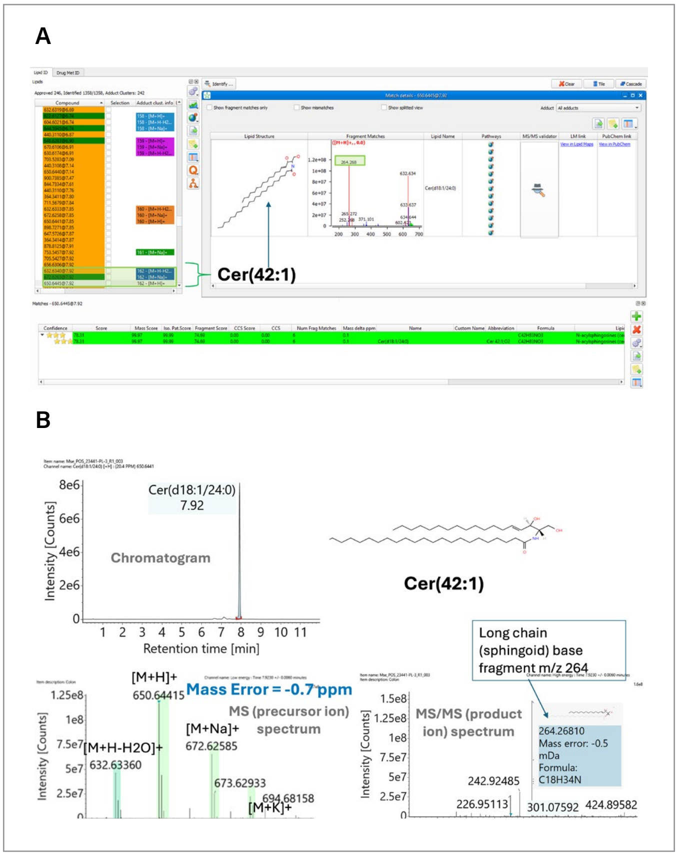


Figure 5: Example Cer(42:1) identification at 7.92 min as the in Lipostar2 (A) and in waters_connect Platform (B)

The Lipostar2 identification can be confirmed using waters_connect Platform as show on in Figure 4B. The extracted chromatogram at RT 7.92 minutes is identified as Cer(42:1) with the long chain base fragment observed with low mass error.

The box and whisker plots of ceramide and lysophoscholine lipids that have been reported to be involved with CRC dysregulation according to various peer reviewed publications as well as our findings from the MVA of samples are shown in Figure 6.

Some ceramide species are increased in CRC patients compared to healthy control similar to the pancreatic cancer pathway discussed previously. However, when the whole class it taken into consideration there is no significant difference between the two sample cohorts. However, the LPCs are downregulated in CRC samples compared to healthy controls on a class basis. Larger sample cohorts and targeted lipidomic approaches could be applied for further investigation.

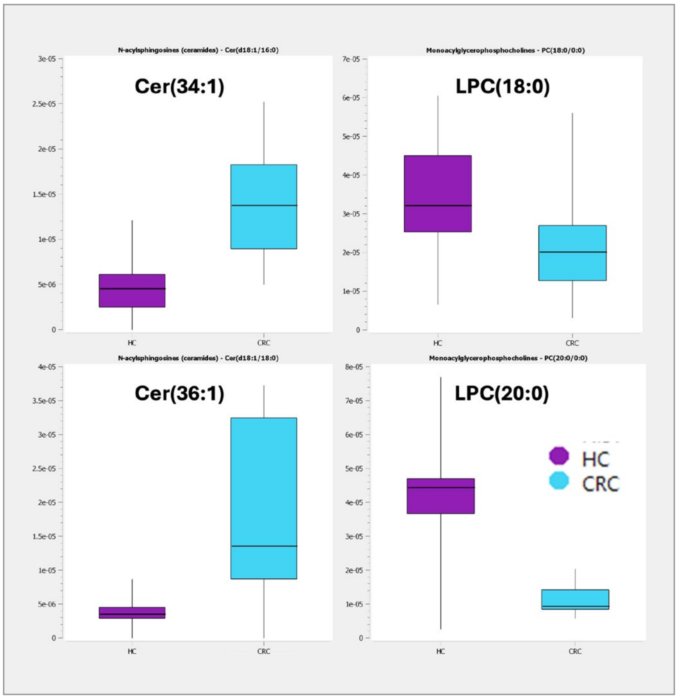
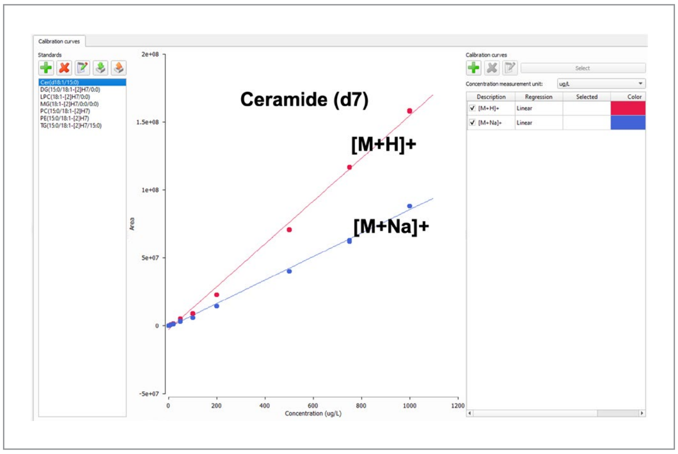


Figure 6: Example box and whisker plots of ceramides species and LPC species from the CRC sample cohort

Quantification of lipids can also be performed in Lipostar2 using calibration curves of standards spiked in plasma. The Xevo MRT MS has a linear dynamic range over five orders of magnitude. More researchers are moving towards quoting concentration values rather that relative responses when transferring data from different labs and geographies as well as that collected different instruments. Figure 7 show the ceramide standard calibration curve that can be assigned to lipids of the sample class and adduct to generate concentration values.



Concentration measurement unit: ug/L

Description	Regression	Selected	Color
☑ [M+H]+	Linear		🟥
☑ [M+Na]+	Linear		🟦

Figure 7: Example Calibration curves of Ceramide standard spiked in human plasma that can be used to calculate concentrations of lipid species if the same class

CONCLUSIONS:

The combination of reversed-phase UPLC, Xevo MRT MS and intelligent workflow driven software allows for simple, robust and rapid accurate lipidomic analysis of biological samples. The Lipostar2 software enables accurate identification and quantification of compounds within data sets acquired in DDA or DIA mode. The user-friendly, easy-to-use workflow described above allowed for the rapid processing of lipidomic data such as that generated for the CRC study described here. The data generated from the pilot study was subjected to extensive statistical analysis using Lipostar2 software to enable biological interpretation.

*https://mzio.io/#mzmine

#http://prime.psc.riken.jp/compms/msdial/main.html

Waters Corporation
34 Maple Street
Milford, MA 01757 U.S.A.
T: 1 508 478 2000
F: 1 508 872 1990
waters.com