

# Development and Initial Validation of a Workflow for Small Metabolite Identifications

Jeffrey M. McGuire<sup>1</sup>, Nicholas J. Vaccaro<sup>2</sup>

(1) U.S. Army Edgewood Chemical Biological Center, (2) University of Maryland

## Introduction

Metabolomics is a core area of systems biology research that focuses on the study of low-molecular weight organic and inorganic metabolites. Metabolic profiling techniques have been applied to define metabolic changes related to genetic differences, environmental influences, and disease or drug perturbations. The untargeted nature of metabolic profiling and its focus on low-molecular weight metabolites can allow for new biomarkers of disease or toxic effect to be uncovered. Metabolomic analysis can be applied to the study of biological fluids collected in non-invasive or minimally-invasive ways such as blood or urine. Technological advances in analytical instrumentation and advances in data modeling are working in synergy to open up new perspectives and research agendas in metabolic research.

Traditionally, metabolic profiling has been performed using a range of analytical platforms in clinical, environmental, and toxicological studies. Of these, the various chromatography-mass spectrometry methodologies are perhaps the most widely applied techniques for the rapid observation of system perturbations in the metabolome. Gas chromatography-mass spectrometry (GC-MS) focuses on low-molecular-mass metabolites such as organic acids, amino acids, carbohydrates and some phosphorylated metabolites. In contrast, liquid chromatography-mass spectrometry (LC-MS) focuses on compounds such as bile acids, sterols, phospholipids and fatty acids.

To that end, we have developed a workflow for acquiring data related to the metabolites present in plasma samples. Initial studies have focused on using GC-MS to couple highly reproducible GC retention times with electron impact ionization mass spectra. Even though a single GC-MS analysis generates over 3000 molecular features, using software to compare and discriminate data sets via multivariate statistics, it is quite feasible to differentiate plasma samples from two different animal models, the Göttingen minipig and the New Zealand White rabbit.

## Sample Preparation

A 0.5 mL sample of plasma was first diluted (1:3) with acetonitrile (ACN) to "crash" any protein in the sample. Following vortex mixing, the sample was centrifuged (5 min at 14,000 x g) and an aliquot of the supernatant liquid (ca. 50 µL) removed for subsequent UPLC-MS/MS analysis. The remaining liquid was transferred to a 3.0 mL micro-reaction vessel and dried under vacuum for 30 min at 70 ° C. The dried residue was derivatized with 250 µL of *N,O*-bis(trimethylsilyl)trifluoroacetamide with trimethylchlorosilane and heated to 70 ° C for 30 min. After cooling the reaction vessel, 20 µL of internal standard solution (8 mg/mL of decane and 7 mg/mL of hexadecane) was added to the mixture.

## LC-MS/MS Analyses

A 5-µL aliquot of the supernatant liquid was injected onto an ACQUITY UPLC® BEH C<sub>18</sub> column (2.1 x 50 mm, 1.7 µm particle size) from Waters and chromatographed using a 0.1% formic acid (Solvent A) and methanol with 0.1% formic acid (Solvent B) mobile phase with a flow rate of 0.4 mL/min. A gradient was applied as shown in Table 2.

Quadrupole time-of-flight mass spectra were acquired using a Waters Xevo™ QToF MS (Milford, MA) operated in positive electrospray ionization mode. The mass spectrometer was operated in the MS<sup>E</sup> data acquisition mode using a desolvation temperature of 350 ° C, desolvation gas flow of 700 L/hr, source temperature of 120 ° C, and a capillary voltage of 1000 V.

Table 2. Time-gradient profile

min	% A	% B	Curve
0.0	100	0	-----
1.0	100	0	5
16.0	0	100	5
20.0	0	100	5
22.0	100	0	1

Data were acquired from 50 Da to 1000 Da using two functions. The first function applied a low collision energy (CE) of 4 eV for parent ion information, while the second function applied a high CE of 20 eV for fragment information and structural elucidation. The mass spectrometer was calibrated across the same mass range using a solution of sodium formate with a residual mass accuracy of 0.5 mDa. A third function acquired data for the external reference lock mass compound leucine enkephalin which was infused at 50 µL/min and generating a fragment ion at *m/z* 278.1141.

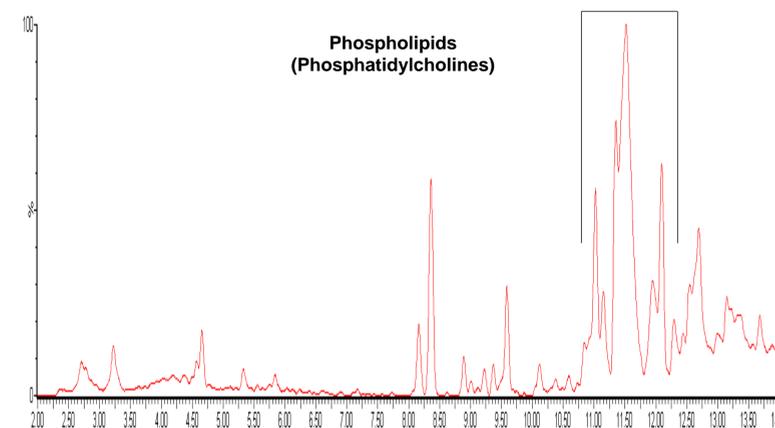
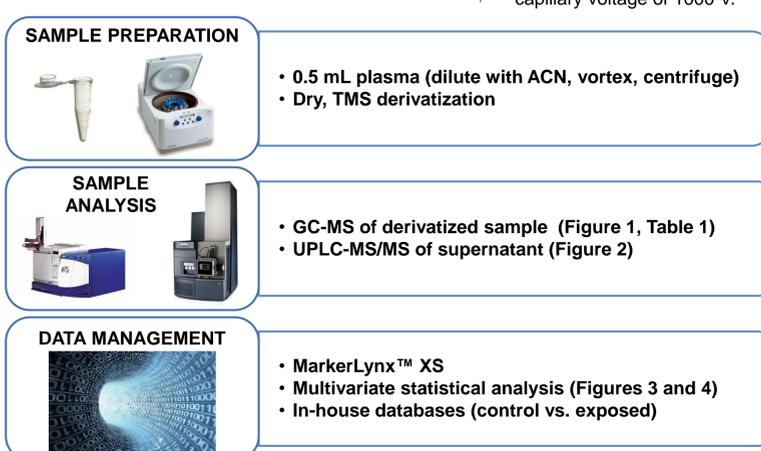


Figure 2. Total ion chromatogram from UPLC-MS/MS of supernatant from minipig plasma



## GC-MS Analyses

Analyses were performed on a Waters Quattro micro GC Mass Spectrometer. Gas chromatographic separations were achieved using a Restek Rtx®-5 column, 30m x 0.25mm i.d. with a 0.25µm film thickness. The carrier gas was helium with a flow rate of 1mL/min. Injections of 1.0 µL were made by autoinjector into a split injector port (20:1) at a temperature of 280 ° C. The initial oven temperature of 70 ° C was held for 5 min, then ramped at 5 ° C/min to 320 ° C. Samples were ionized by 70-e.v. electron impact (EI). Mass spectra were acquired while scanning from 50 Da to 650 Da following calibration and tuning with heptacosafuorotributylamine. Putative identification of metabolites detected by GC-MS was accomplished using the NIST library.

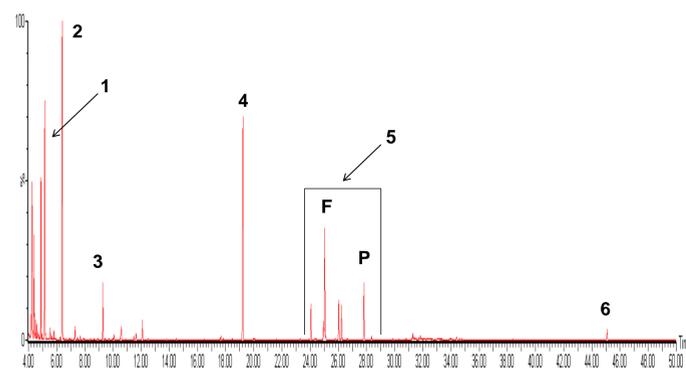


Figure 1. Total ion chromatogram from GC-MS analysis of minipig plasma

Table 1. Peak identifications from GC-MS analysis

Peak No.	Identification
1	n-Decane (IS)
2	Lactic Acid
3	Urea
4	n-Hexadecane (IS)
5	Sugars (Furanose/Pyranose)
6	Cholesterol

Loadings Bi Plot Comp[1] vs. Comp[2] colored by 5.43\_51.0464

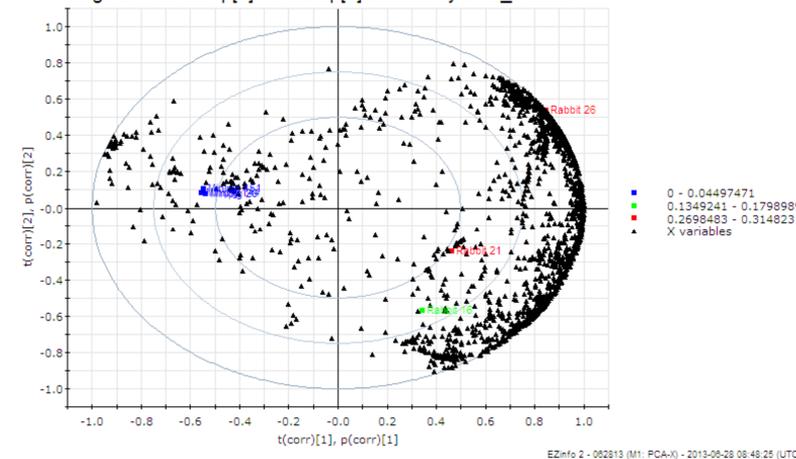


Figure 3. Loadings Plot

## Data Management

The initial validation of this workflow consisted of the GC-MS analyses of three different Göttingen minipig plasmas and three different New Zealand White rabbit plasmas. Figure 3 shows a plot of how the X variables (observations such as GC retention time, and mass-to-charge ratio and intensity) correlate with each other. The three minipig samples all cluster on the left side of the plot while the three rabbit samples cluster on the right side. This is clearly illustrated in Figure 4 which shows a 3-D plot of the  $t[1]$  and  $t[2]$  scores, the weighted averages of the original data.

Future studies will focus on creating custom in-house databases of small metabolites from both control and chemical warfare nerve agents exposed animals. By applying multivariate analysis software, such as MarkerLynx™ XS, subsequent analysis and mining of these complex datasets may provide some insight into potential biomarkers of nerve agent exposure.

Scores Num vs. Comp[1] vs. Comp[2], colored by 5.43\_51.0464

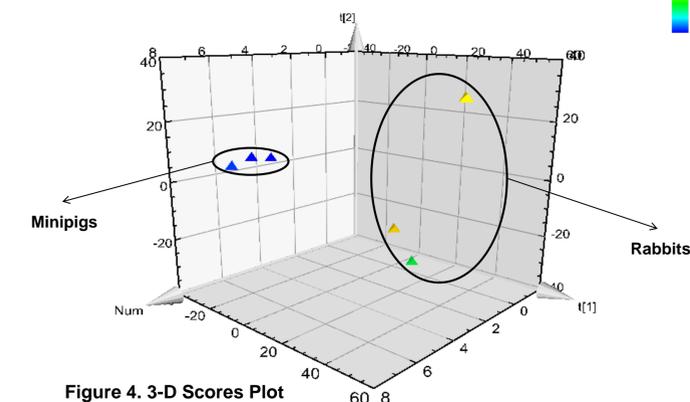


Figure 4. 3-D Scores Plot

Acknowledgements: The authors thank the Defense Threat Reduction Agency/Joint Science and Technology Office for their assistance and funding of this work under Project CB3883 and the Science and Engineering Apprenticeship Program (SEAP). The views expressed in this presentation are those of the authors and do not necessarily reflect official policy or the position of the Department of Defense or the U.S. Government.

