Delivering the Right Results

GCxGC-TOFMS Data Interpretation of Metabolic Biomarkers from Diabetic and Non-diabetic Urine Utilizing Fisher Ratios Prior to Multivariate Analysis

John R. Heim, Scott Pugh, and Mark Libardoni • LECO Corporation, St. Joseph, Michigan

OVERVIEW

- Demonstrate the increased peak capacity available utilizing comprehensive multi-dimensional gas chromatography (GCxGC) for the analysis of complex small molecule metabolite profiles.
- Show the benefit of time-of-flight mass spectrometry (TOFMS) to provide the data density necessary for optimum characterization of complex GCxGC separations.
- Illustrate the application of the new features in LECO's ChromaTOF® software, Statistical Compare and Fisher Ratios, applied to a small molecule metabolite comparison study as a pre-statistical data analysis tool that is easily exported into multivariate analysis for a large data set representing diseased and non-diseased state classes.
- Demonstrate the comprehensive flow path used to simply export the data generated by Statistical Compare and Fisher Ratio calculations into multivariate analysis platforms.
- Show the graphical results from PCA and Clustering analysis of this small molecule metabolite profile study that indicate possible significant variance between disease and non-disease state classes.

INTRODUCTION

Metabolomics presents challenges for both the analytical methods used and the data reduction required to interpret the results. Comprehensive multi-dimensional gas chromatography time-of-flight mass spectrometry has emerged as an excellent instrumental option for the characterization of small metabolite profiles. GCxGC-TOFMS provides increased peak capacity and resolution for the chromatographic separation while fast TOFMS acquires the mass spectral data density necessary to characterize complex biological samples. This poster presents a data mining strategy from results obtained from a diabetic and non-diabetic Trimethylsilyl (TMS) derivatized urine study analyzed by GCxGC-TOFMS. A Fisher Ratio plot was also generated from grouped sample comparison results. A compound table based on the Fisher Ratio results was then applied to principal component analysis (PCA) whereby possible differences between non-diseased and diseased state groups were graphically represented.

The Initial experimentation focused on optimized method development for sample extraction, BSTFA derivatization, and GCxGC-TOFMS method optimization. Following method development, GCxGC-TOFMS analysis was conducted on six derivatized samples from each of four subjects, two diseased, and two non-diseased. The GCxGC-TOFMS data was refined through background elimination of erroneous peaks produced by column bleed and derivatization reagents before processing by statistical comparison and Fisher Ratio calculations.

GCxGC-TOFMS analysis was conducted on diabetic and non-diabetic urine samples that were first extracted with Methylene Chloride and then derivatized with BSTFA. A total of twenty-four samples were analyzed and data processed before applying a statistical comparison for the diseased and non-diseased state classes. Data alignment was carried out by the Statistical Compare feature contained in LECO's ChromaTOF® software. A Fisher Ratio plot was then calculated for each analyte of the Compound Table and utilized to identify metabolites with the highest variance. The high variance data was subsequently exported as a spreadsheet and applied to multivariate principal component and clustering analysis. Graphical representations of the multivariate analysis show significant analyte differences between diseased and non-diseased state sample classes.

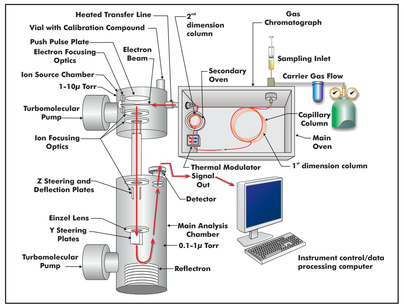

Figure 1. Pegasus® 4D GCxGC-TOFMS Schematic

EXPERIMENTAL METHODS

SAMPLE PREPARATION
4 urine samples
2 non-diabetic control subjects
1 type I diabetic
1 type II diabetic

Sample Extraction Procedure

10 mL aliquots were acidified to pH$_2$ with concentrated H_2SO_4.
6 samples from each subject were extracted with 2 mL methylene chloride.
Approximately 2 mg of sodium sulfate was added to each extract.

BSTFA Derivatization Procedure

200 uL extract was placed in a 2 mL amber glass autosampler vial containing 0.5 mg of sodium sulfate.
30 uL of pyridine was added to the vial.
100 uL of BSTFA was placed in the vial.
Samples were heated at 60°C for 1 hour.
Derivatized samples were then analyzed by GCxGC-TOFMS on the same day as prepared.

GCxGC-TOFMS Analysis Parameters

Gas Chromatograph: Agilent 7890 equipped with a LECO quad jet dual stage cryogenic modulator and secondary oven
Primary Column: 30 m x 0.25 mm id. x 0.25 μm film thickness Rtx-5ms
(Restek Corp., Bellefonte, PA)
Secondary Column: 1.5 m x 0.18 mm id. x 0.20 μm film thickness Rtx-200
(Restek Corp., Bellefonte, PA)
Carrier Gas: Helium set @ 1.5 mL/min
Injection Mode: Splitless Injection volume: 3 uL
Inlet Temperature: 260 °C
Primary Column Temperature Program: Initial temperature set @ 40°C for 1 min ramped @ 6°C/min to 290°C final hold time 10 minutes
Secondary Column Temperature Program: Initial temperature set @ 50°C for 1 min ramped @ 6°C/min to 300°C final hold time 10 minutes
Total Run Time: 52.67 minutes

GCxGC parameters

Column Temperature offset: 10 °C
Modulator Temperature offset: 25 °C
Modulation Period: 5 s Hot Pulse Time: 0.8 s Cool Time Between Stages: 1.7 s

Mass Spectrometer: LECO Pegasus 4D

Acquisition Delay: 250 s
Mass Range: 45 - 800 μ
Acquisition Rate: 200 spectra/s
Ion source Temperature: 230 °C
Detector Voltage: 1750 V
Electron Energy: -70 eV

REFINED DATA RESULTS → STATISTICAL COMPARE → FISHER RATIO CALCULATION → EXPORT → MULTIVARIATE ANALYSIS

GCxGC-TOFMS RESULTS

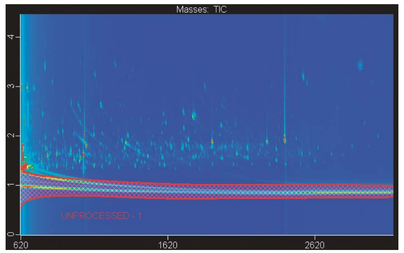

Figure 2. Normal Control non-diabetic subject: Contour Plot Total Ion Chromatogram of TMS-derivatized urine sample showing the small molecule metabolite profile.

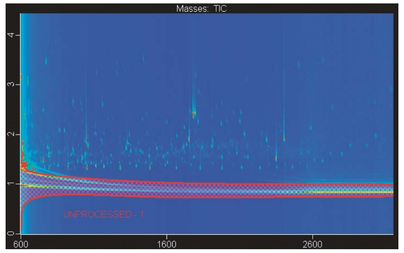

Figure 3. Disease state subject Type I diabetic: Contour Plot Total Ion Chromatogram of TMS-derivatized urine sample showing the small molecule metabolite profile.

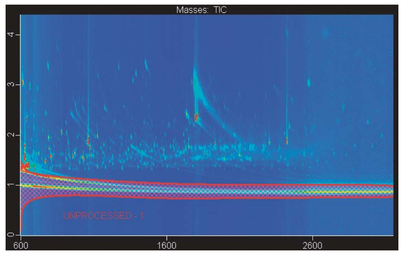

Figure 4. Disease state subject Type II diabetic: Contour Plot Total Ion chromatogram of TMS-derivatized urine sample showing the small molecule metabolite profile

Results of the Diabetic Profile Study between diabetic and non-diabetic subjects are shown in the figures above by the total ion chromatograms depicted as two dimensional contour plots. These chromatographic examples visually illustrate peak differences between sample types and highlight the benefits that GCxGC-TOFMS offer which include increased peak capacity, improved analyte detectability, and enhanced resolution. On average over 1000 peaks were found per sample with a signal to noise ratio of 100 in this study. The red cross hatched area in each contour plot is an unprocessed region developed in the Classifications feature of ChromaTOF® software which eliminates unwanted background peaks.

DATA MINING STRATEGY AND FISHER RATIO CALCULATION

Refined Peak Table Data

2-Azetidinone, 1-(t-butyldimethylsilyl)-4-[(t-butyldimethylsil	91631	1150	1.670	mainlib	93905	147	232.13
2-Butenoic acid, tert-butyldimethylsilyl ester	1383278	595	1.685	mainlib	91523	143	5704.4
2-Dimethyl(trimethylsilyl)silyloxytridecane)	57191	1595	1.590	mainlib	93984	147	149.43
2-Ethyl-1-dimethyl(isopropyl)silyloxyhexane)	43839	830	1.455	mainlib	36212	187	354.47
2-Ethyl-3-ketohexanoate, bis(O-trimethylsilyl)-	16174	1295	1.520	mainlib	34609	287	205.33
2-Ethyl-3-ketohexanoate, bis(O-trimethylsilyl)-	25553	1315	1.540	mainlib	34609	287	326.13
2-Ethyl-3-trimethylsilyloxy(trimethylsilyl)butyrate	49199	1255	1.510	mainlib	33690	147	124.30
2-Ethylhexanoic acid, trimethylsilyl ester	95570	945	1.595	mainlib	34173	201	656.97

Figure 4. The figure above represents a portion of a peak table from the diabetic versus non-diabetic GCxGC-TOFMS analysis. Erroneous background peaks were eliminated prior to processing the data with Statistical Compare in LECO's ChromaTOF® software.

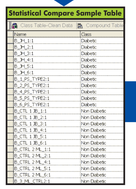

Figure 5. In Figure 5 above the sample data files for Diabetic and Non-diabetic Classes are shown in the grouped Sample Table.

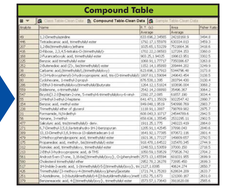

Figure 6. Figure 6 above shows the aligned Diabetic and Non-diabetic data from 24 samples along with the Fisher Ratio for each analyte.

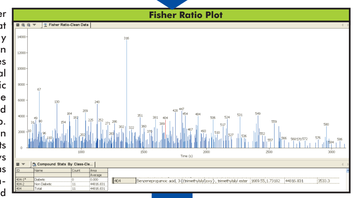

Figure 7. The Fisher Ratio plot shown at right graphically represents unknown chemical differences between the normal control non-diabetic sample group and the diabetic diseased state sample group. Peak 404 revealed in the Compound Stats table above shows that the analyte was found in 11 non-diabetic samples and 0 diabetic samples.

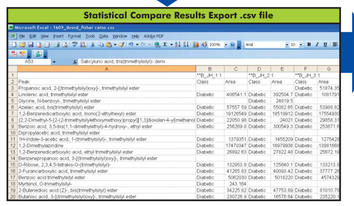

Figure 8. Illustrated above is an exported .csv file Partial Excel spreadsheet of the Statistical Compare results for the diabetic and non-diabetic urine GCxGC-TOFMS study. The compounds of highest variance by their Fisher Ratio were then imported into multivariate analysis programs.

MULTIVARIATE ANALYSIS

Multivariate analysis is based on multivariate statistics, which involves observation and analysis of more than one statistical variable at a time. The technique is used to perform studies across multiple dimensions while taking into account the effects of all variables on the responses of interest. This study applied ChromaTOF's® Statistical Compare and Fisher Ratios to a data set of twenty-four samples from diseased and non-diseased state subjects that determined the analytes with highest variation across the entire sample population. The Statistical Compare results generated a Compound Table which was exported as a .csv in Excel format and applied to several multivariate analysis platforms. PCA analysis was conducted on the variables of analyte identification, class, (diseased or non-diseased), and analyte peak area. Following PCA analysis, K-means clustering was applied. The three-dimensional graph shown below in figure 10 was developed in the commercially available Miner3D software. The graph shows clear differences as well as similarities in the small molecule metabolites found in both diabetic and non-diabetic TMS derivatized urine analyzed by GCxGC-TOFMS.

PCA PLOT

Miner 3D Analysis

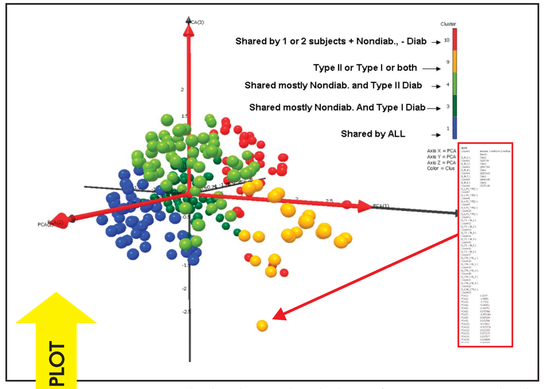

Figure 10. The three-dimensional graph above in figure 10 was generated in the multivariate analysis program Miner3D. The plot used PCA 1, PCA 2, and PCA 3 as the x, y, and z axis. K-means clustering was used as the fourth variable, color, which categorized the spheres into discrete analyte groups. The colors of the spheres differentiate discrete analyte groups as described above in the legend. The inset illustrates one of the analytes shown by the arrow indicating the yellow sphere. The chart inset describes the compound, Benzene, 1-methyl-4-(1-methylenethenyl)- and shows that it was found only in one diabetic subject.

Figure 9. Figure 9 above shows the Miner3D workspace with the .csv file data loaded from the exported ChromaTOF® Statistical Compare results.

CONCLUSIONS

The results presented from this study demonstrate that significantly increased analytical performance is achieved by utilizing comprehensive multi-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GCxGC-TOFMS) for the characterization of small molecule metabolite profiles. TOFMS provides non-skewed mass spectra and fast acquisition needed to deconvolute complex overlapping peaks as well as the data density required to characterize GCxGC analysis. Several new ChromaTOF software features were applied in this metabolomic study. Statistical Compare and Fisher Ratio calculation were applied in this study. These features allow the user to find significant unknown chemical differences among known classes of complex samples. The new features available in LECO's ChromaTOF software were used to align a large set of data and define the highest variance for analytes between diseased and non-diseased state subjects. Furthermore, it was demonstrated that the results from Statistical Compare and Fisher Ratio calculations can be exported quite simply into multivariate analysis programs whereby PCA and Clustering analysis can be conducted. A relatively simple comprehensive single flow path from sample preparation, GCxGC-TOFMS analysis, statistical comparison that targets high variance data, to multivariate PCA and Clustering analysis was demonstrated. This exploratory research presents an optimized GCxGC-TOFMS analysis followed by a data mining strategy using preliminary statistical methods prior to multivariate analysis that establishes a viable strategy which can identify significant metabolite variation in complex biological samples representing diseased and non-diseased state classes.