A Multi-Platform Approach for Metabolomic Analysis of Human Liver Tissues

by Alessia Ferrarini[1], Cristina Di Poto[1], Rency S. Varghese[1], Mohammad R. Nezami Ranjbar[1], David E. Alonso[2], Joe Binkley[2], Habtom W. Ressom[1]*

[1] Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC

[2] LECO Corporation, St. Joseph, MI

**Corresponding author*

Metabolomics offers the opportunity to characterise complex diseases through a systematic evaluation of thousands of small molecules. Specifically, analysis of metabolites in liver tissues represents a suitable approach to investigate in-situ modifications and to highlight aberrant biochemical pathways directly at cellular level.

Considering that over 40,000 metabolites have been included so far in the Human Metabolome Database and many more remains unknown, the identification of the entire metabolome remains a great challenge. It is well known that a single analytical technique is not enough to characterise the high chemical diversity and the different concentration ranges of metabolites (e.g. spanning nearly over 11 orders of magnitude in serum) present in a biological compartment. Thus, a multi-platform approach has become a mandatory strategy to enhance metabolite coverage. In this article we present the use of three complementary platforms: GC-TOF-MS, GCxGC-TOF-MS and LC-QTOF-MS, that leads to an increment of the metabolite coverage for metabolomic characterisation of liver tissues.

Metabolomics refers to the study of all low molecular weight metabolites presents in biological systems [1]. Due to its capacity to reveal the relationship between multiple environmental stimuli and our genetic inheritance, metabolomics has become a useful tool in systems biology. The characterisation of complex disease phenotypes in fact could take us to the identification of new biomarkers for specific physiological responses and to the clarification of the pathophysiology of complex diseases such as cancer.

In the last decade, improvements in analytical technologies have increased the sensitivity, accuracy, and resolution of mass spectrometry-based platforms offering the possibility of detecting a greater number of chemical species. These advances, together with the establishment of different chemometric tools, have made possible the extrication of complex biological information hidden under the enormous amount of data generated. Even though a large part of the metabolome in a biological sample remains unknown, complementary platforms (e.g. GC-MS and LC-MS) are typically used to increase the metabolite coverage. GC-MS can be used for the detection of non-polar volatile compounds and medium to polar compounds after derivatisation, such as sugars, carboxylic acids, free fatty acids and other small lipids. In addition, GCxGC-MS offers a second dimension to the resolving power of chromatography and thus increasing the number of identified compounds. LC-MS is usually utilised with a reversed phase column for the analysis of different class of glycerolipids, glycerophospholipids, lysoglycerophospholipids, sphingolipids, carnitines, fatty acyls, amides, etc.

In previous studies, conducted by the same group, changes in metabolite levels in blood samples from patients with hepatocellular carcinoma (HCC) vs. cirrhotic controls have been extensively investigated, by using both GC-MS and LC-MS platforms [2,3,4,5]. HCC is a highly malignant form of liver cancer, with raising incidence and poor prognosis. Heterogeneous phenotypic and genotypic traits, along with a wide range of risks factors, make the identification of effective diagnostic markers and consequently the study of its pathophysiology complicated. The use of different platform helps characterise the alteration due to the cancer onset and guide on the choice of potential candidate biomarkers. In the present study the evaluation of three platforms (GC-TOF-MS, GCxGC-TOF-MS and LC-QTOF-MS) were used for characterisation of HCC based on untargeted metabolomic analysis of 15 human liver tissue samples.

Study Cohort

The 15 liver tissue samples were collected from 10 participants recruited at MedStar Georgetown University Hospital through a protocol approved by the Georgetown IRB. All subjects provided signed informed consent forms. The tissues represent 5 HCC cases (5 tumour and 5 adjacent cirrhotic tissues) and 5 patients with liver cirrhosis. Subjects in cases and controls were matched by gender, age, ethnicity and BMI.

Chemicals

Octafluoronaphthalene (OFN) was purchased as a custom standard from Ultra Scientific. Fatty acid methyl ester standards (FAMEs), C8, C9, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28; methoxyamine hydrochloride (MEOX), pyridine, debrisoquine, 4-nitrobenzoic acid and formic acid were purchased from Sigma Aldrich. MSTFA + 1% TMCS (N-Methyl-N-trimethylsilyltrifluoroacetamide + Chlorotrimethylsilane) was purchased from Thermo Scientific. HPLC grade 2-propanol, acetonitrile and water were used for metabolites extraction. Helium was purchased from Robert Oxygen and Air Gas.

Sample preparation

Figure 1 depicts the sample preparation steps which were followed for untargeted metabolomic analysis of the 15 tissue samples. Liver tissues were homogenised and metabolite extraction was performed in one single step for both GC-MS and LC-MS analyses [2].

Specifically 20 mg of liver tissue were homogenised with 1 mL of pre-chilled Isopropanol:Acetonitrile:Water (3:3:2) in order to extract the metabolites and precipitate the proteins. Samples were then centrifuged at 14,500 g for 15 minutes at room temperature and the resulting supernatant was divided into two, 460µL each, and concentrated to dryness in speedvac (RC110B, Thermo Scientific, Waltham, Massachusetts, USA). The dried samples were kept at -20°C until subsequent steps.

Prior to GC-TOF-MS and GCxGC-TOF-MS analysis, one aliquot was reconstituted in 500 µL of water, of which 100 µL were dried-out, lyophilised (-50°C) and derivatised in order to protect functional groups and increase the volatility and the thermostability of the analytes. Dried samples were mixed with 20 µL of MEOX reagent (20 mg/mL in pyridine) and heated at 60°C for 1 hour. This was followed by reaction with 80 µL of MSTFA at 60°C for 1 hour. The MEOX reagent was spiked with octafluoronaphthalene (OFN) to monitor system performance and fatty acid methyl esters (C8, C9, and C10 – C28 even) for retention time reference. The reaction products were cooled down to room temperature, vortex-mixed and analysed.

Samples for LC-QTOF-MS analysis were directly reconstituted in 200 µL of mobile phase with spiked-in debrisoquine and 4-nitrobenzoic acid for quality assessment of positive and negative mode analysis, respectively.

Equipment and Parameters

Metabolites were analysed using Agilent 7890 GC with dual stage modulator, MPS2 autosampler and LECO Pegasus HT (Leco Corporation, St. Joseph, Michigan, USA), equipped with an electron ionisation source and TOF analyser. GC-TOF-MS data were acquired using both splitless and split 10:1 injection mode to compensate for very large or very low concentration range of tissue metabolites. Similarly, GCxGC-TOF-MS data were generated using two splits ratios (20:1 and 40:1). LC-MS analysis were carried out with a Waters ACQUITY UPLC coupled to XEVO G2 QTOF (Waters Corporation, Milford, Massachusetts, USA), operating in positive and negative polarity. The full method parameters employed for each platform are listed in Tables 1-3.

Data treatment and statistical analysis

The multi-platform data generated in this study were treated with different chemometric tools in order to extract the valuable information including biologically significant differences that are hidden by a multitude of unrelated ions, noise and contamination. The data treatment pipeline we used is depicted in Figure 2.

ChromaTOF GC software with True Signal Deconvolution package (Leco Corporation) was used for data pre-processing, including baseline calculation, peak finding, deconvolution and identification. Spectral similarity searches against the NIST 2011 databases and Fiehn Library (version 2013) resulted in matches ranging from 702 to 931 out of a possible score of maximum a 1000. The Statistical Compare software tool was used for alignment of the GC-MS data [2]. LC-QTOF-MS data were first converted into Network Common Data Format (NetCDF) using DataBridge Program from the MassLynx software (Waters) and then pre-processed and aligned using XCMS package (version 1.44, Scripps Center for Metabolomics, La Jolla, Ca, USA) [6]. XCMS performs first peak detection, then it matches the peaks across the samples to calculate the retention time deviations and relative intensity for ion comparison, and finally it performs missing values imputation. CAMERA (collection of annotation related methods for mass spectrometry data) package was used for ion annotation (process of grouping all adducts, cluster ions and charge states entities derived from same analyte) [7]. The data were filtered keeping only the ions present in at least 3 samples out of 5 within one biological group. The intensities of the missing ions in the GC-MS data were imputed with 1/6 of the minimum value in the corresponding biological group. In addition, possible contamination, artifacts, and column bleeding products were eliminated from the GC-MS data. For LC-MS data, XCMS performs the automatic missing values imputation by rereading the raw data files and integrating them in the region of the missing peaks.

GC-MS and LC-MS data were normalised using the value of total proteins present in the sample, calculated using the BCA assay [8]. A logarithmic transformation was applied to smooth the data to a more normal distribution. Paired t-test ($p \leq 0.05$) between tumour and adjacent cirrhotic tissue (HCC vs. ADJ-CIRR) and unpaired t-test between tumour and cirrhotic tissue (HCC vs. CIRR) and adjacent cirrhotic tissue and cirrhotic tissue (ADJ-CIRR vs. CIRR) was applied to reveal significant changes in metabolite levels among the different groups. Finally, fold change and direction of change were calculated to establish relative quantification.

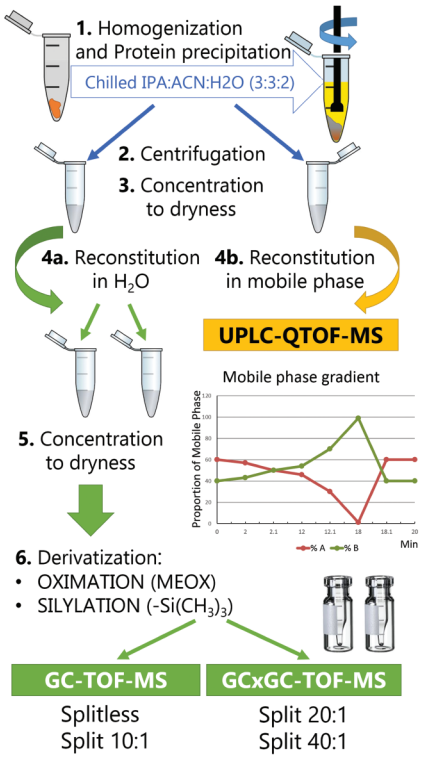


Figure 1: Overview of sample preparation for the three chromatographic techniques employed for the analysis of the liver samples.

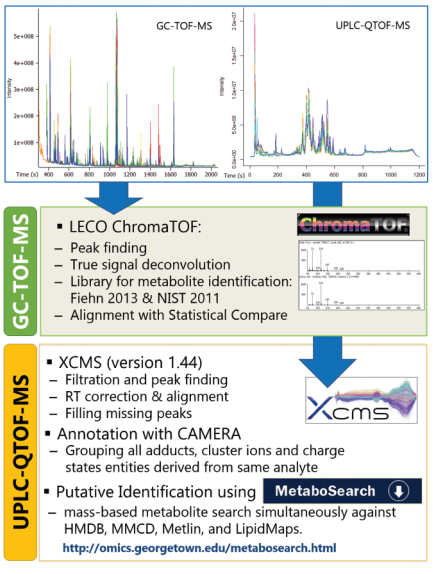


Figure 2: Data treatment pipeline involving multiple steps for GC-MS and LC-MS data analysis.

Table 1: GC-TOF-MS Conditions

Gas Chromatograph	Agilent 7890 with MPS2 Autosampler
Injection	1 µL, Splitless (& Split 10:1) at 280˚C
Carrier Gas	He at 1.0 mL/min, constant flow
Column	Rxi-5ms, 30 m × 0.25 mm i.d. × 0.25 µm coating (Restek, Bellefonte, PA, USA)
Oven program	70°C (4 min), to 300°C at 10°C/min (10 min)
Transfer Line	300°C
Total analysis time	34 min
Mass Spectrometer	LECO Pegasus® HT
Ion Source Temperature	250°C
Mass Range	30-600 m/z
Acquisition Rate	10 spectra/s

Table 2: GC×GC-TOF-MS Conditions

Gas Chromatograph	Agilent 7890 with MPS2 Autosampler
Injection	1 µL, split 20:1 (& Split 40:1) at 280°C
Carrier Gas	He at 1.0 mL/min, corrected constant flow
Column One	Rxi-5ms, 30 m × 0.25 mm i.d. × 0.25 µm coating (Restek)
Column Two	Rtx-200, 1.25 m × 0.25 mm i.d. × 0.25 µm coating (Restek)
Temperature Program	4 min at 70°C, ramped 5°C/min to 300°C, held 10 min - Secondary oven maintained +10°C relative to primary oven
Modulation	4 s with temperature maintained +15°C relative to secondary oven
Transfer Line	250°C
Total analysis time	60 min
Mass Spectrometer	Leco Pegasus® 4D
Ion Source Temperature	250°C
Mass Range	30-600 m/z
Acquisition Rate	200 spectra/s

Table 3: UPLC-QTOF-MS Conditions

Chromatograph	Waters ACQUITY UPLC
Injection	3µL
Column	Acquity UPLC CSH C18, 1.7 µm d.p., 2.1 mm i.d. × 100 mm (Waters)
Column Temperature	55°C
Mobile phase	A: 60:40 Acetonitrile/Water + 0.1% Formic acid + 10mM Ammonium formate B: 90:10 Isopropanol/ Acetonitrile + 0.1% Formic acid + 10mM Ammonium formate
Flow rate	0.4 mL/min
Total analysis time	20 min
Mass Spectrometer	XEVO G2 QTOF
Capillary voltage	3 and 1.5 kV for Positive and Negative mode
Sampling Cone	35 and 30 V for Positive and Negative mode
Ionisation Source Temperature	120°C
Desolvation Temperature	350°C
Desolvation gas flow	750 (L/Hr)
Cone gas flow	25 (L/Hr)
Mass Range	50-1200 m/z, acquired in centroid mode
Lock Spray	Leucine enkephalin ([M+H]+ = 556.2766 and [M-H]− = 554.2620) at rate of 20 µLmin-1

Putative metabolite identifications of selected significant ions detected by LC-MS were performed using MetaboSearch, a mass-based tool we developed previously to obtain putative IDs by combining information retrieved from Human Metabolome DataBase, Madison Metabolomics Consortium Database, Metlin, and LipidMaps [9].

Results and Discussion

A total of six different matrices of data were generated after analysis. The total number of ions detected in all 3 platforms were 720 and 427 for splitless and split 10:1 data in GC-MS analysis, 1408 and 1222 for split 20:1 and split 40:1 data in GCxGC-MS analysis and 6119 and 672 for positive and negative ion mode data in LC-MS analysis, respectively. Metabolites putatively identified in this study ranged around 55-49% in GC-MS analysis (splitless and split 10:1 data), 22-24% in GCxGC-MS analysis (split 20:1 and split 40:1 data) and 5-15% in LC-MS analysis (positive and negative polarity mode data), respectively. Undoubtedly the resolving power and the sensibility of GCxGC-MS and LC-MS are higher than GC-MS, with respect to the proportion of putatively identified metabolites against the ones detected confirming the difficulties in characterising and identifying unknown metabolites present in the human metabolome.

As illustrated by the Venn diagram shown in Figure 3, only a few compounds were obtained on two platforms and only one significant metabolite was detected in all three techniques. This demonstrates the need for using a multiplatform approach to characterise complex disease and specimens.

It should be noted that although the three approaches are complementary, the most significant amount of data was obtained using the liquid chromatographic approach; however the selection of just one analytical platform will clearly prevent the identification of potentially significant metabolites and possible candidate markers. It is well establish that utilising a RP column for LC separation would bias the analysis for the detection of medium-to-non-polar metabolites and preclude the identification of more polar compounds, found to be of great importance in cancer metabolism. For instance, metabolites implicated in the Krebs cycle and glycolysis usually elute in the front of the LC chromatogram, a region poorly reproducible. For these metabolites

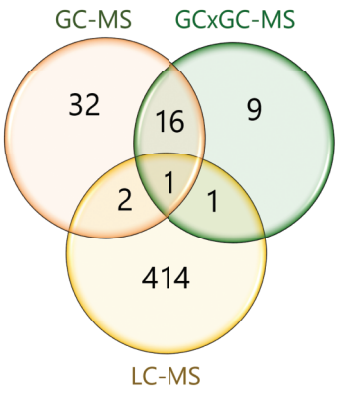


Figure 3: Putatively identified compounds resulted to be statistically significant and classified accordingly to the analytical platform. The LC-MS data shown in Figure 3 comprises of compounds detected in both positive and negative mode. The GC-MS and GCxGC-MS data from the same figure combines splitless and split 10:1 data, and split 20:1 and split 40:1 data, respectively.

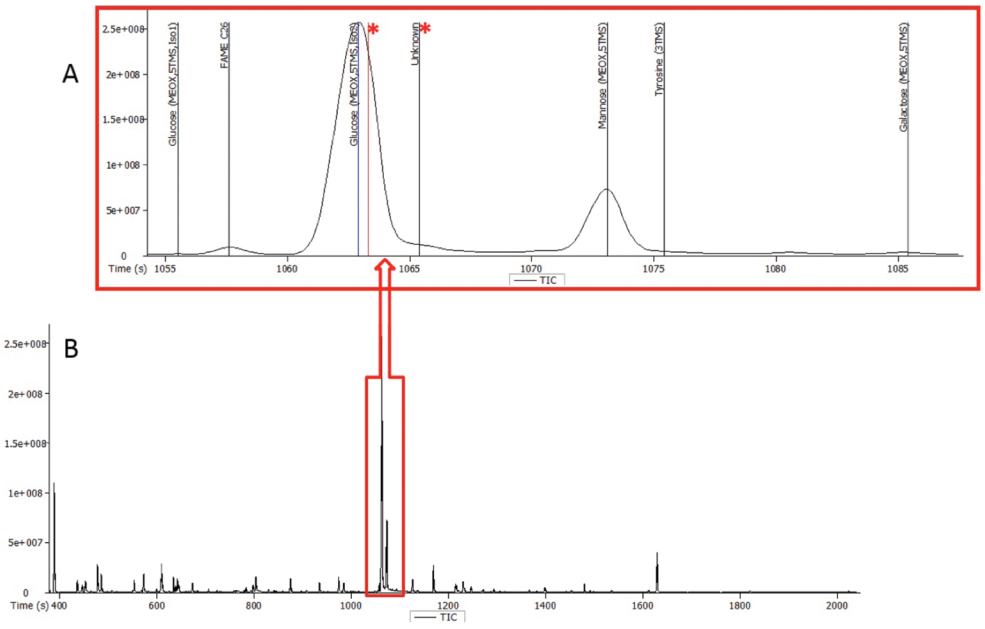


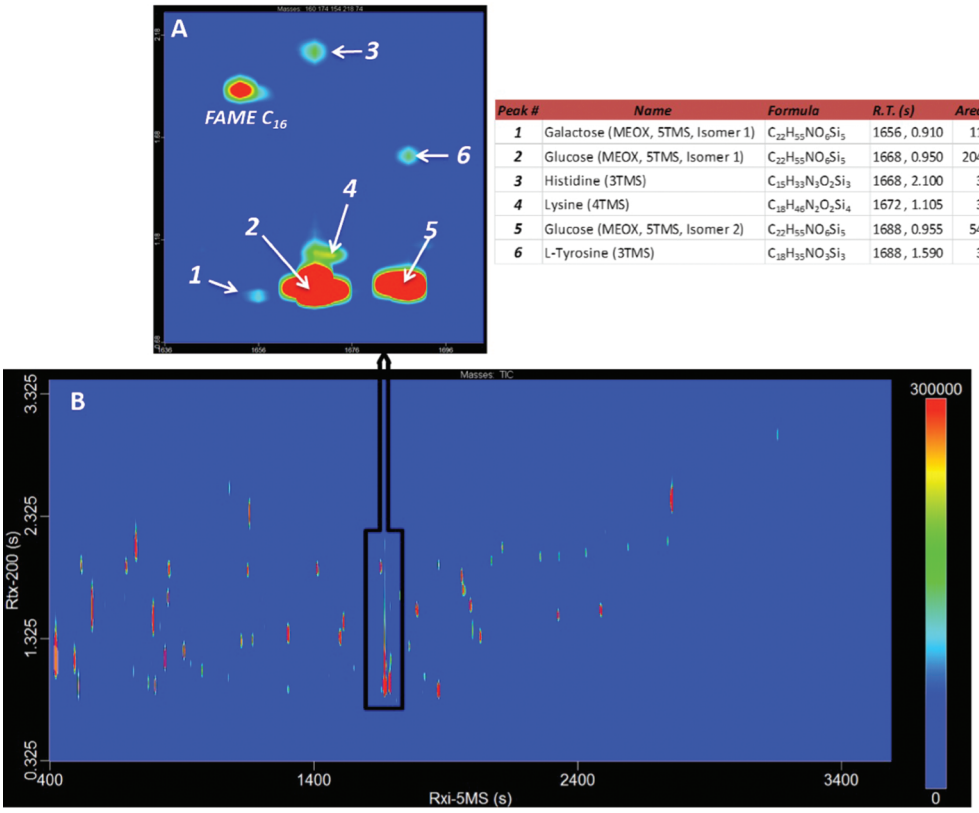
Figure 4. An expanded region (A) of the Total Ions Chromatograms (B) of a liver tissue sample acquired by GC-MS. The picture shows two unknowns (*) under the very large glucose (MEOX, 5TMs) signal.

GC-MS offers better separation and identification capability. Considering the other way around, the sole use of the GC-MS platform would prevent the detection of a variety of lipid classes that are becoming more and more important in the study of different diseases.

An example of the importance of combining different platform is given in Figures 4 and 5. Overlapping signals of low-level and high concentrated analytes in GC-MS analysis can be resolved when the GCxGC-MS platform is employed (Figure 5).

After data filtration and missing values imputation, statistical analysis revealed several compounds that are significantly altered in the different disease groups, as depicted in Figure 6. These compounds belong to various biochemical categories including acids, di-acids, amino acids, bases, sugars, phosphorylated sugars, sugar alcohols, fatty acids, nucleosides, nucleotides, mono- di- & tri-acylglycerides, lysophosphatidylglycerol lipids, glycerolipids, sphingolipids, and carnitines. Among them, amino acids and phospholipids are up-regulated in HCC versus cirrhosis, whereas carboxylic acids, bile acids and long chain carnitines are down-regulated.

As illustrated in Figure 6, we found more number of tissue metabolites significantly altered in HCC vs. CIRR and ADJ-CIRR vs. CIRR compared to those found in comparing HCC vs. adjacent cirrhotic tissues. This could be due to that fact that the samples in the latter comparison were derived from the same set of patients, thus presenting relatively smaller inter-individual variability. In addition, we observed that LC-MS-based analysis of tissues led to the greatest number of statistically significant compounds among the three comparisons, followed by GCxGC-MS. The largest number



Peak #	Name	Formula	R.T. (s)	Area	Similarity
1	Galactose (MEOX, 5TMS, Isomer 1)	$C_{22}H_{55}NO_6Si_5$	1656 , 0.910	1196476	919
2	Glucose (MEOX, 5TMS, Isomer 1)	$C_{22}H_{55}NO_6Si_5$	1668 , 0.950	20429945	929
3	Histidine (3TMS)	$C_{15}H_{33}N_3O_2Si_3$	1668 , 2.100	370428	789
4	Lysine (4TMS)	$C_{18}H_{46}N_2O_2Si_4$	1672 , 1.105	326836	719
5	Glucose (MEOX, 5TMS, Isomer 2)	$C_{22}H_{55}NO_6Si_5$	1688 , 0.955	5454926	945
6	L-Tyrosine (3TMS)	$C_{18}H_{35}NO_3Si_3$	1688 , 1.590	326956	829

Figure 5. Extracted Ion Chromatograms (masses = 74, 154, 160, 174, 215) of an expanded region (A) of the contour plot (B) showing signals for several analytes which were chromatographically resolved in the 2nd dimension. Lysine (4TMS) and histidine (3TMS) are two amino acids not identified in the 1D analysis, but clearly resolved in the 4D contour plot as shown in the extracted ion chromatogram contour plot expansion.

Comparisons							
Platform	Mode	HCC vs. ADJ-CIRR (1)	Common between 1 and 2	HCC vs. CIRR (2)	Common between 2 and 3	ADJ-CIRR vs. CIRR (3)	Common between 1 and 3
GC-MS	Splitless	16	2	23	5	41	3
	Split 10:1	20	0	9	4	15	4
GCxGC-MS	Split 20:1	18	2	28	2	26	0
	Split 40:1	16	3	27	7	34	3
LC-MS	LC +	67	9	173	44	115	15
	LC -	20	8	35	2	38	4

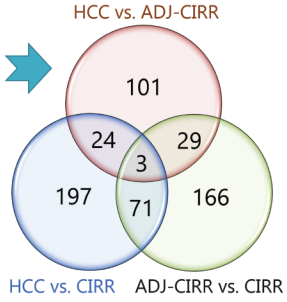


Figure 6: Summary of all statistically significant metabolites (putatively identified and unknown), listed accordingly to the group comparison and the different platform in which they were detected.

of compounds found by GC-MS was in the splitless mode when comparing ADJ-CIRR vs. CIRR. Furthermore, we found three compounds with statistically significant changes in all three comparisons. Finally, only one compound that belongs to the class of the sugar alcohols was identified. This compound was detected by GC-MS in the split ratio 10:1.

Conclusion

In this study, the combination of different platforms and the enhanced chromatographic resolution achieved by GCxGC-TOF-MS and UPLC-QTOF-MS analysis allowed the detection of a wide variety of class of compounds in liver tissue samples from HCC and cirrhotic patients. The study demonstrates the capability of the multi-platform approach for the improvement of the metabolome coverage and identification of changes in metabolite levels in tissues from distinct biological groups.

Acknowledgements

This work was supported by NIH/NCI U01CA185188 and NIH/NCATS UL1TR000101 grants.

GC-MS and GCxGC-MS analyses were performed by LECO Corporation.

LC-MS analysis was performed at the Proteomics & Metabolomics Shared Resource at Georgetown University, partially supported by NIH/NCI grant P30 CA051008.

References

1. Fiehn, O., Metabolomics--the link between genotypes and phenotypes. Plant Mol Biol 2002, 48 (1-2), 155-71.

2. Nezami Ranjbar, M. R.; Luo, Y.; Di Poto, C.; Varghese, R. S.; Ferrarini, A.; Zhang, C.; Sarhan, N. I.; Soliman, H.; Tadesse, M. G.; Ziada, D. H.; Roy, R.; Ressom, H. W., GC-MS Based Plasma Metabolomics for Identification of Candidate Biomarkers for Hepatocellular Carcinoma in Egyptian Cohort. PLoS One 2015, 10 (6), e0127299.

3. Ressom, H. W.; Xiao, J. F.; Tuli, L.; Varghese, R. S.; Zhou, B.; Tsai, T. H.; Ranjbar, M. R.; Zhao, Y.; Wang, J.; Di Poto, C.; Cheema, A. K.; Tadesse, M. G.; Goldman, R.; Shetty, K., Utilization of metabolomics to identify serum biomarkers for hepatocellular carcinoma in patients with liver cirrhosis. Anal Chim Acta 2012, 743, 90-100.

4. Xiao, J. F.; Varghese, R. S.; Zhou, B.; Nezami Ranjbar, M. R.; Zhao, Y.; Tsai, T. H.; Di Poto, C.; Wang, J.; Goerlitz, D.; Luo, Y.; Cheema, A. K.; Sarhan, N.; Soliman, H.; Tadesse, M. G.; Ziada, D. H.; Ressom, H. W., LC-MS based serum metabolomics for identification of hepatocellular carcinoma biomarkers in Egyptian cohort. J Proteome Res 2012, 11 (12), 5914-23.

5. Xiao, J.; Zhao, Y.; Varghese, R. S.; Zhou, B.; Di Poto, C.; Zhang, L.; Tadesse, M. G.; Ziada, D. H.; Shetty, K.; Ressom, H. W., Evaluation of metabolite biomarkers for hepatocellular carcinoma through stratified analysis by gender, race, and alcoholic cirrhosis. Cancer Epidemiol Biomarkers Prev 2014, 23 (1), 64-72.

6. Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G., XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 2006, 78 (3), 779-87.

7. Kuhl, C.; Tautenhahn, R.; Bottcher, C.; Larson, T. R.; Neumann, S., CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Anal Chem 2012, 84 (1), 283-9.

8. Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C., Measurement of protein using bicinchoninic acid. Anal Biochem 1985, 150 (1), 76-85.

9. Zhou, B.; Wang, J.; Ressom, H. W., MetaboSearch: tool for mass-based metabolite identification using multiple databases. PLoS One 2012, 7 (6), e40096.