The Human Serum Metabolome

Nick Psychogios[1], David Hau[1], Jun Peng[2], Igor Sinelnikov[1], Souhaila Bouatra[1], Rupasri Mandal[1], Ram Krishna Murthy[1], Jianguo Xia[4]
Fiona Bamforth[3], Janet McManus[5], Theresa Pedersen[6], Russ Greiner[1], Bruce McManus[5], John Newman[6] and David Wishart[1,4]

Depts. of Computing Science[1] , Chemistry, Laboratory Medicine and Pathology[3] and Biological Sciences[4], University of Alberta, Edmonton AB, Canada, Western Human Nutrition Research Center[6], USDA, ARS, 430 West Health Sciences Drive, Davis, CA, USA, iCAPTURE Centre for Cardiovascular and Pulmonary Research[5], St. Paul's Hospital, University of British Columbia, Vancouver, BC, Canada

Abstract: Continued advances in detection and separation analytical technologies in the field of metabolomics call for in-depth metabolic profiling of biofluids and underline the need for developing comprehensive reference resources. As part of our objective to systematically characterize the human metabolome and advance the fields of quantitative metabolomics, we currently present a global metabolic profiling of the human serum. For this holistic analysis, high-resolution NMR spectroscopy, GC–MS, LC-ESI-MS/MS for oxidized lipid mediators and LC/GC-FID (Flame-Ionization Detector) for lipidomics together with computer-aided literature mining were combined to identify and quantify the metabolites that can be commonly detected (with today's technology) in the human serum metabolome. Samples from healthy individuals and various patients were collected for the NMR, the LC-ESI-MS/MS analysis and the LC/GC-FID, whereas pooled blood serum samples were prepared for the GC–MS studies. Data acquisition and analysis was performed according to SOPs and in-house developed protocols. Our literature survey was facilitated by several computational tools developed for the Human Metabolome Database (http://www.hmdb.ca) and the in-house text-mining tool called PolySearch (http://wishart.biology.ualberta.ca/polysearch/). Our experimental results indicated that global metabolic profiling methods can routinely detect more than 4200 different compounds in serum. With NMR spectroscopy 44 compounds were detected and quantified, GC–MS methods could detect and quantify about 70 compounds, LC-ESI-MS/MS techniques were able to quantify 84 oxylipid mediators, whereas with LC/GC-FID more than 3100 lipids were quantified. Comprehensive, web-accessible tables containing the compounds, concentrations, spectra, protocols and links to disease associations that were revealed or identified from these combined experimental and literature mining efforts are presented, which are also freely available at: http://www.serummetabolome.ca.

Introduction

While it is relatively routine in genomics/transcriptomics and proteomics studies to identify and quantify 1000's of genes or proteins at a time, the same cannot be said of most metabolomics efforts. In an effort to further enhance the use of quantitative metabolomics, we (and others) have started to systematically determine the detectable metabolic composition of clinically important biofluids and tissue types. In order to facilitate future research into blood chemistry and blood metabolomics, it is crucial to establish a comprehensive, electronically accessible database of the detectable metabolites in human blood or human serum. Herein we present just such a database describing the detectable metabolites (including their concentrations and disease associations) that can be found in human serum.

Objectives

- Compare various metabolomic technologies, estimate their coverage and breath and suitability ofr studying blood serum
- Obtain values from literature
- Validate literature values
- Identify and expand the coverage of what is known in the serum metabolome and assess concentration ranges

Text mining/Bibliomics

With respect to the literature survey, more than 1000 references were studied, producing a list of 1087 metabolites purported to be present in human serum. Of the listed metabolites, 1021 of them have at least one reported normal concentration value and 476 have at least one concentration value linked to a specific illness.

Comprehensive tables containing the compounds concentrations, spectra, protocols and links to disease associations that were uncovered or identified from this work are freely available at: **http://www.serummetabolome.ca**

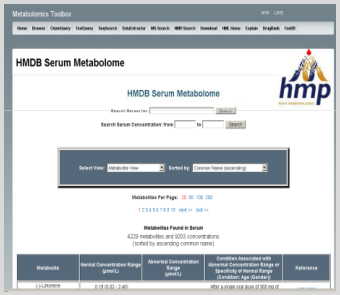

Figure 1. Human Serum Metabolome website **http://www.serummetabolome.ca**

NMR Spectroscopy

In total, 69 samples were analyzed: 15 from healthy subjects and 54 from heart transplant patients. 1D ^{1}H-NMR spectra were collected for each sample using the tnnoesy pulse sequence on a 500 MHz Inova VARIAN spectrometer. The spectra were processed using Chenomx NMR suite software v. 6.0 (Figure 2). Assigned compounds were confirmed with sample spiking. At least 95% of the area in each spectrum was accounted for.

37 serum metabolites were identified in healthy subjects and 44 in heart-transplant patients. Serum pH together with 20 metabolites were found significantly different between the two groups (t-test, $p<0.05$) (Table 1). From the 37 compounds identified in the control serum group, 32 had corresponding concentration values in the literature, from which 12 had comparable concentrations, 12 significantly lower and 3 significantly higher concentrations compared with those determined by our group.

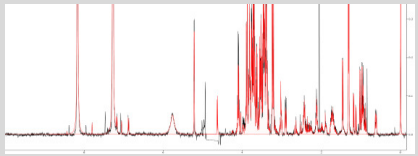

Figure 2. A 1D proton spectrum of serum. In black is the original spectrum. In red is the sum spectrum i.e. the combination of profiled compounds.

Metabolites	Concentration (mM) Controls (n=15)	Heart transplant (n=53)	p
3-Hydroxybutyrate	37	20	0.04
Alanine	251	194	0.03
Choline	7	6	0.02
Citrate	64	46	0.01
Creatinine	35	50	0.00
Formate	21	11	0.00
Glucose	3075	2139	0.02
Glutamine	299	215	0.00
Glycerol	384	77	0.00
Glycine	186	134	0.04
Histidine	75	26	0.00
Isoleucine	37	26	0.01
Leucine	58	43	0.01
Lysine	93	73	0.02
Phenylalanine	44	26	0.00
Propylene glycol	67	21	0.00
Threonine	73	47	0.00
Tyrosine	55	33	0.00
Urea	3471	1868	0.00
Valine	111	82	0.01
pH	7.34	7.14	0.00

Table 1. Comparison of average concentrations of different serum metabolites found in both normal and heart-transplant patients. P-values were calculated to determine the probability that the difference between the two averages was due to random chance.

GC-MS

Pooled blood serum samples were extracted separately for polar and lipid extracts and derivatized using N-methyl-N-trifluoroacetamide with 1% trimethylchlorosilane (TMS). Extracts were analyzed using an Agilent 7890-5975C GC-MS instrument operating in electron impact (EI) mode and the AMDIS software was used for the analysis of the EI-MS spectra. Metabolites were identified by matching the EI-MS spectral with NIST library (2008) and the experimental RI of each metabolite with an in-house compiled RI library for human metabolites.

Combining the metabolites identified in both polar extracts and lipid extracts, 70 metabolites were identified (Table 2).

Dicarboxylic acids	**Amino acids**	**Fatty acids**	**Alcohols & Polyols**
Oxalic acid	Threonine	Hexanoic acid	Tocopherol
Malonic acid	Aspartic acid	Nonanoic acid	Glycerol
Succinic acid	Pyroglutamic acid	Decanoic acid	Ribitol
Adipic acid	Alanine	n-pentadecanoic acid	Myo-inositol
Tricarboxylic acids	Glycine	Palmitic acid	Mannitol
Citrate	Isoleucine	Linoleic acid	**Carbohydrates**
Cis-aconitic acid	Leucine	Oleic acid	Erythritol
Hydroxy acids	Lysine	Stearic acid	Glucose
a-hydroxyvaleric acid	2-aminobutyric acid	Arachidonic acid	Fructose
Glyceric acid	Ornithine	Octanoic acid	**Amino alcohols**
2,3,4-trihydroxybutyric acid	Phenylalanine	Dodecanoic acid	Ethanolamine
2-Hydroxybutyrate	Proline	Tetradecanoic acid	**Amino ketones**
3-Hydroxybutyrate	Serine	Eicosanoic acid	Creatinine
Lactate	Tyrosine	Heptadecanoic acid	Urea
2,4-dihydroxybutyric acid	Valine	Palmitelaidic acid	**Acyl Phosphates**
Aromatic acids	Glutamate	**Steroids & Steroid Derivatives**	Glyceryl phosphate
Benzoic acid	Tryptophan	Cholesterol	**Alcohol Phosphates**
Ibuprofen	Cysteine	**Lipids**	Myo-inositol phosphate
Keto-acids	Aminomalonic acid	Oleamide	**Inorganic ions & gases**
Pyruvate	Aminoadipic acid	**Purines & Purine Derivatives**	Phosphate
	Misc	Uric acid	
	Hydroxamine	Hypoxanthine	

Table 2. List of polar and lipid compounds profiled by GC-MS

LC-ESI-MS/MS

Four human plasma sample replicates were extracted by Solid Phase Extraction (SPE) and analyzed by LC-ESI-MS/MS for 66 oxidized lipid mediators. Two internal standards and 11 deuterated surrogates were used to track instrument drift and extraction efficiency.

Samples were spin filtered at 0.22 micron and analyzed by LC-ESI-MS/MS (Waters UPLC-API 4000 QTRAP) for an expanded list of oxidized lipids. Data was analyzed with Analyst software (Applied Biosystems).

The 18-carbon oxygenated fatty acid profiles were dominated by diols in both the linoleates and alpha-linolenate series. However, the linoleates showed high mid-chain alcohol levels, which were not observed in the alpha-linoleates. The 20 carbon oxygenated fatty acid profiles were dominated by mid-chain alcohol levels and the DHA diols showed the highest concentrations.

In addition, 2 serum samples were run using a Phenomenex Luna C18 RP column connected to a Waters Quattro Premier triple quadrupole MS and analyzed with the MassLynx V4.0 software. With this method, 18 additional oxylipids were identified.

I8 samples were analyzed with the AbsoluteIDQ kit from Biocrates using a 4000 QTrap (MDS Sciex) mass spectrometer. The standard flow injection method of the AbsoluteIDQ kit comprising two 20 uL injections was applied for all measurements. Multiple reaction monitoring (MRM) detection was used for quantification.

In total, 162 metabolites were identified and quantified: 42 acylcarnitines, 14 amino acids, 91 glycerophospholipids and 15 sphingolipids

TLC/GC-FID-FAMEs

Two human plasma samples were used for the lipid extraction. Individual lipid classes were separated by preparative TLC and trans-esterified in 3 N methanolic-HCl. The resulting fatty acid methyl esters (FAMEs) were separated and quantified by capillary GC (6890 Hewlett-Packard) equipped with a 60-m DB-23 capillary column, a Flame Ionization Detector (FID), and were analyzed with the ChemStation software.

38 unique FA constituents were identified for 7 different lipid classes. Major lipid classes identified include cholesterol esters, free fatty acids, triacylgycerols, and phospholipids. Major fatty acid constituents detected include all C12 to C24 saturated and unsaturated chains (>25 fatty acids in total). Quantification of lipids and fatty acids was achieved using defined standards. Combinatorial lipid reconstruction (CLR) was used to "regenerate" triacylglycerols and phospholipids with >1 fatty acid chain. CLR uses fractional abundance of each fatty acid chain and total abundance of a given lipid class to estimate most probable and upper limit concentrations.

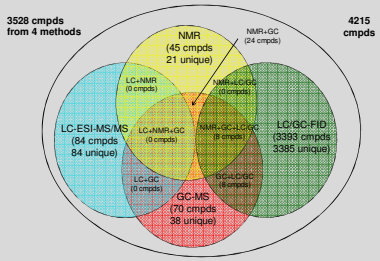

Figure 2. Venn diagram showing the overlap of blood serum metabolites detected by global NMR, GC–MS, LC–ESI-MS/MS and LC/GC-FID methods compared to the detectable serum metabolome.

Conclusions

Global metabolic profiling methods can routinely detect about 3528 different compounds in serum. In total, an average of 32 compounds were detected and quantified with NMR spectroscopy; 70 compounds with GC–MS; 84 oxylipids and 163 compounds with LC–ESI-MS/MS, 25 CEs, 30 LYPCs and 27 FFAs with TLC/GC-FID-FAMEs, whereas estimated concentration for 3393 lipids including TGs, DG, PCs and PEs were based on CLR (Figure 2).

Acknowledgments

Funding was provided by Genome Alberta, Genome Canada, the University of Alberta and the Pan-Alberta Metabolomics Platform (PANAMP)