Untargeted metabolomic analysis of urine samples in the diagnosis of some inherited metabolic disorders

Hana Janeckova[a,b], Alzbeta Kalivodova[b,c], Lukas Najdekr[b], David Friedecky[a,b], Karel Hron[c], Per Bruheim[d], Tomas Adam[a,b]

Background. Metabolomics is becoming an important tool in clinical research and the diagnosis of human diseases. It has been used in the diagnosis of inherited metabolic disorders with pronounced biochemical abnormalities. The aim of this study was to determine if it could be applied in the diagnosis of inherited metabolic disorders (IMDs) with less clear biochemical profiles from urine samples using an untargeted metabolomic approach.
Methods. A total of 14 control urine samples and 21 samples from infants with cystinuria, maple syrup urine disease, adenylosuccinate lyase deficiency and galactosemia were tested. Samples were analyzed by liquid chromatography on aminopropyl column in aqueous normal phase separation system using gradient elution of acetonitrile/ammonium acetate. Detection was performed by time-of-flight mass spectrometer fitted with electrospray ionisation in positive mode. The data were statistically processed using principal component analysis (PCA), principal component discriminant function analysis (PCA-DFA) and partial least squares (PLS) regression.
Results. All patient samples were first distinguished from controls using unsupervised PCA. Discrimination of the patient samples was then unambiguously verified using supervised PCA-DFA. Known markers of the diseases in question were successfully confirmed and a potential new marker emerged from the PLS regression.
Conclusion. This study showed that untargeted metabolomics can be applied in the diagnosis of mild IMDs with less clear biochemical profiles.

Keywords: inherited metabolic disorders, untargeted metabolomics, mass spectrometry

Received: April 22, 2014; Accepted: September 15, 2014; Available online: November 7, 2014
http://dx.doi.org/10.5507/bp.2014.048

[a]Laboratory for Inherited Metabolic Disorders, Department of Clinical Biochemistry, University Hospital Olomouc, Czech Republic
[b]Laboratory of Metabolomics, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University Olomouc
[c]Department of Mathematical Analysis and Applications of Mathematics, Faculty of Science, Palacky University Olomouc
[d]Department of Biotechnology, Norwegian University of Science and Technology, Sem Sælands vei 6/8, NO-7491 Trondheim, Norway
Corresponding author: Hana Janeckova, e-mail: janeckovah@gmail.com

INTRODUCTION

Metabolomics is an emerging science which studies the complex profile of low-molecular weight metabolites present in biological samples at a specific time. It has been applied in a number of fields (environmental chemistry, plant biochemistry, microbiology and nutritional studies). It has also become an important tool in clinical research and in the diagnosis of human diseases[1,2]. The first attempt to use metabolomic tools in diagnosing inherited metabolic disorders (IMDs) was conducted by Gary Siuzdak´s group[3]. These authors applied an untargeted metabolomic approach using reverse phase capillary liquid chromatography - time-of-flight mass spectrometry based on exact mass measurements and automatic data processing. The data were processed using nonlinear alignment XCMS software[4] and METLIN online database[5] (http://metlin.scripps.edu) to find and identify metabolites differently regulated in various diseases. The concept was successfully validated in the case of two clinically profound metabolic disturbances (methylmalonic acidemia and propionic acidemia) which are characterized by a prominent biochemical profile.

The aim of this study was to determine its validity in diagnosing IMDs in urine samples using an untargeted metabolomic approach by high performance liquid chromatography (HPLC) coupled with time-of-flight mass analyzer (Q-TOF). We tested four different IMDs in which the biochemical peculiarities are not so obvious – cystinuria (CYS), maple syrup urine disease (MSUD), adenylosuccinate lyase deficiency (ADSL), and galactosemia (GALT).

MATERIALS AND METHODS

Urine samples

The urine samples were taken from infants during routine diagnosic laboratory procedures. The collected urine samples were stored at -20 °C. Prior to preparation, the samples were allowed to thaw at room temperature.

Healthy control (n=14) and patient (n=21) urine samples including the 4 inherited metabolic disorders above (Table 1) were analyzed. The diagnoses had been confirmed by biochemical, enzyme or molecular-genetic analyses in all the patients (Table 2). Samples with the

Table 1. Summary of studied defects with their urine markers.

Disease	Urine markers
ADSL	↑ SAICAr, SAdo
GALT	↑ galactose, galactitol
CYS	↑Orn, Lys, Arg, cystin, cystein-homocystein disulfide
MSUD	↑3-hydroxybutyric acid, 2-oxoisovaleric acid, 2-hydroxyisovaleric acid, 2-oxo-3-methylvaleric acid, 2-oxoisocaproic acid, 2-hydroxy-3-methylvaleric acid, 2-hydroxyisocaproic acid, xLeu, Val

SAdo, succinyladenosine; SAICAr, succinylaminoimidazole carboxamide riboside, Orn, ornithine; Lys, lysine; Arg, arginine; xLeu, the sum of leucine, isoleucine and alloisoleucine; Val, valine.

Table 2. Summary of analyzed urine samples from patients with IMDs.

Patient*	Gender	Age (year)	Previous biochemical findings in urine sample (μmol/mmolcreatinine)
ADSL 1	M	20	SAICAr 1,7; SAdo 13,8
ADSL 2a	F	21	SAICAr 9,3; SAdo 27,8
ADSL 2b	F	22	SAICAr 10,5; SAdo 30,3
ADSL 2c	F	22	SAICAr 6,2; SAdo 17,0
CYS 1a	M	unk	Cys 225, Orn 75, Lys 554, Arg 231
CYS 1b	M	13	unk
CYS 1c	M	unk	unk
CYS 2a	F	unk	Arg 1190, cystine 352, Orn 420, Lys 1060
CYS 2b	F	unk	Arg 1120, cystine 760, Orn 710, Lys 2110
CYS 2c	F	unk	Arg 750, cystine 315, Orn 290, Lys 800
CYS 3a	M	unk	cystine 73, Gly 480, Lys 400, Arg 15
CYS 3b	M	14	cystine 58, Gly 480, Lys 235
GALT 1a	M	16	galactitol 84
GALT 1b	M	16	hypoaminoaciduria
GALT 1c	M	15	galactitol 100
GALT 2a	M	32	galactitol below LOD
GALT 2b	M	32	galactitol below LOD
MSUD 1	F	8	Leu 42, Ile 12
MSUD 2a	F	18	Val 13, Ile 10, Leu 35
MSUD 2b	F	16	unk
MSUD 2c	F	17	unk

*The patient samples with the same number were taken from the same patients in different time (assigned by small letter).
F, female; M, male; unk, unknown; Gly, glycine; LOD, limit of detection; Leu, leucine; Ile, isoleucine.

same number were from the same patients taken at a different time (assigned small letters).

Untargeted metabolomic analysis

Urine samples were diluted to a creatinine concentration of 0.5 mmol/L with mobile phase and 2 μL of the diluted urine were injected and analyzed by liquid chromatography coupled with mass spectrometry. For separation, a modification of a published method was used[6,7]. Separations were performed on aminopropyl column (Luna 3 μm NH2, 2 x 150 mm, Phenomenex, Torrance, CA, UHPLC Agilent 1200 Series). Gradient elution at flow rate of 0.25 mL/min was set using a program with the mobile phase A (20 mmol/L ammonium acetate, pH 9.45) and mobile phase B (acetonitrile) as follows: 0–15 min, 85% B to 15% B; 15–25 min, 15% B; 25–25.1 min, 15% B to 85% B. Between runs, the system was equilibrated with 85% B for 9 min. Total analysis time was 35 min. Detection was performed with Agilent G6520A Q-TOF fitted with electrospray ionisation in positive mode. Electrospray ionisation source settings were: V Cap of 3000 V, skimmer of 65 V and fragmentor of 100 V. The nebulizer was set at 45 psi and the nitrogen drying gas was set at a flow rate of 8 L/min. Drying gas temperature was maintained at 325 °C. The data were acquired in continuum mode at a scan rate of 1.0 Hz in a mass range of *m/z* 70-1100 and with mass resolution of 20.000.

Data processing and statistical analysis

Data were extracted from "mzdata" raw files using the XCMS package in R software[8] (www.r-project.org). After extraction, the data were baseline corrected, normalized by the total ion count, mean centered on zero and scaled to unit standard deviation (auto-scaled). Data were statistically processed by principal component analysis (PCA) and principal components discriminant function analysis (PCA-DFA) using R software. Features most important in discrimination were found from the partial least squares (PLS) regression. Metabolites were identified using automatic workflows PUTMEDID_LCMS (ref.[9]), METLIN,

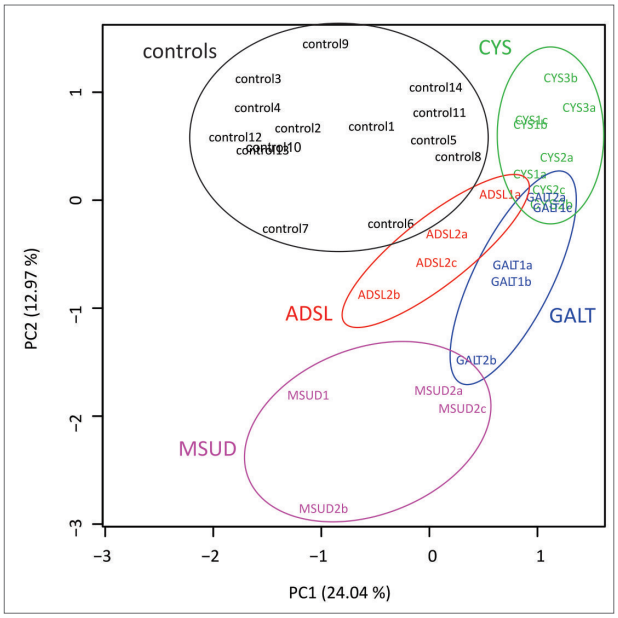

Fig. 1. Principal component analysis of urine samples.

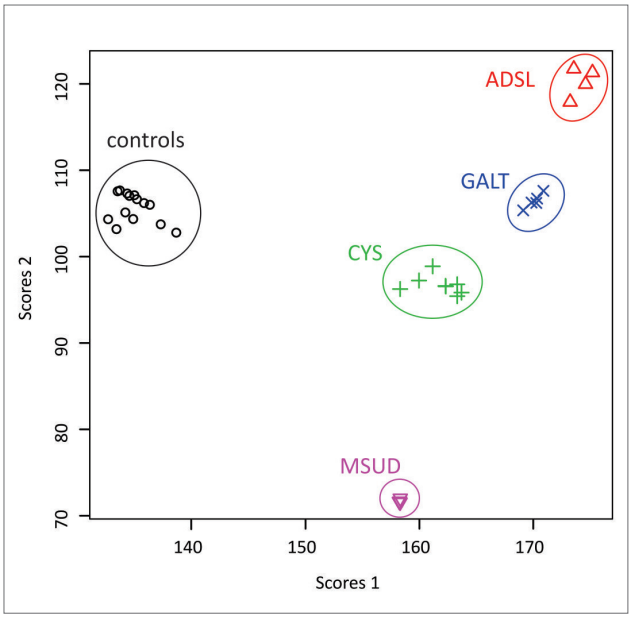

Fig. 2. Principal components discriminant function analysis of urine samples.

and HMDB (http://www.hmdb.ca) databases. Identity of metabolites was confirmed by comparison with corresponding standards.

RESULTS AND DISCUSSION

The total dataset after XCMS processing comprised a total of 1492 features. First, data were statistically processed using unsupervised PCA. All the patient samples were distinguished from the controls in the PCA analysis (Fig. 1). Second, PCA-DFA, a supervised method, was applied. Distinction of all five groups (controls and patients with CYS, MSUD, ADSL and GALT) was unambiguously confirmed (Fig. 2).

A list of twenty most significant features differentiating the diseases from normal urine samples was created for each disorder based on the PLS regression. The most unambiguous results were obtained for patients with cystinuria. This defect is characterized by several biochemical markers (Table 1), which can be easily analyzed by the method described. The majority of these markers were confirmed (Table 3). A feature with retention time in dead volume (measured *m/z* of 445.2397) listed first in Table 3 was not successfully identified. A feature with *m/z* of 196.0790 was identified as $[M+NH_4]^+$ of cysteinylglycine (Cys-Gly) based on exact mass. Identification of this compound was subsequently confirmed by analyzing the appropriate standard.

One of two known markers - SAdo (Table 1) - was confirmed in patients with an ADSL deficiency. This was markedly increased in patient samples compared to control samples. Features corresponding to SAdo were among the three most discriminatory compounds. Specifically, they were identified as SAdo (molecular ion with *m/z* of 384.1156), its isotope (*m/z* of 385.1175), and its fragment (*m/z* of 340.1329). This was observed at the same retention time of 1050 s, consistent with the appropriate

Table 3. List of ten most significant features differentiating patients with cystinuria from normal urine samples which were identified using partial least squares regression.

Order of significance	Rt (s)	m/z measured	m/z theoretic	Identification
1	105	445.2397	-	unk
2	652	147.1133	147.1128	Lys $[M+H]^+$
3	604	175.1199	175.1190	Arg $[M+H]^+$
4	667	133.0972	133.0972	Orn $[M+H]^+$
5	652	84.0812	84.0809	Lys fragment $[M+H]^+$
6	651	169.0973	169.0947	Lys $[M+Na]^+$
7	603	176.1213	176.1224	Arg isotope $[M+H]^+$
8	651	130.0865	130.0861	Lys fragment $[M+H]^+$
9	819	196.0790	196.0750	Cys-Gly $[M+NH_4]^+$
10	920	241.0311	241.0311	Cystine $[M+H]^+$

Rt, retention time.

standard. The second marker of this disease, SAICAr, was not identified in the samples. One reason for this may have been be its instability in the ion source[10].

Galactosemia is characterized by increased galactose and galactitol in urine (Table 1). In this study, we analyzed urine samples from treated adolescent and adult patients (Table 2). In treated patients, galactose and galactitol were reduced to normal or slightly above normal levels[11]. The intention of including GALT patients was to elucidate novel biomarkers of the disease. This was however not the case and patients were not distinguished by traditional markers.

Urine samples from patients with MSUD are characterized mainly by increased levels of various organic acids (Table 1). Measurements proceeded in positive mode. Hence their analyses were not sufficiently sensitive. Nevertheless, leucine and valine are also important markers of this disease. A fragment of leucine (*m/z* of 86.0954) was found among ten most significant features. It was substantially increased in the patient samples in comparison with controls. Its retention time fitted the appropriate standard.

CONCLUSION

This study tested the method of untargeted metabolomics in conjunction with unsupervised (PCA) and supervised (PCA-DFA) data processing for detecting patients with specific IMDs. Reported data on this method used samples from patients with diseases with an extreme biochemical profile to confirm the validity of this concept. We selected four diseases with mild biochemical abnormalities – cystinuria, maple syrup urine disease, adenylosuccinate lyase deficiency, and galactosemia. The approach was successful: patient samples were discriminated from controls by appropriate metabolites.

ABBREVIATIONS

IMDs, inherited metabolic disorders; PCA, principal component analysis; PCA-DFA, principal components discriminant function analysis; PLS, partial least squares; HPLC, high performance liquid chromatography; TOF, time-of-flight mass analyzer; CYS, cystinuria; MSUD, maple syrup urine disease; ADSL, adenylosuccinate lyase deficiency; GALT, galactosemia.

Acknowledgement: We thank Prof. Royston Goodacre, School of Chemistry and Manchester Interdisciplinary Biocentre (University of Manchester, UK) for help with data processing.

This work was supported by grants IGA MZČR NT12218, IGUP Olomouc LF_2014_011, NPÚ LO1304 and CZ.1.07/2.4.00/17.0015. The infrastructural part of this project (Institute of Molecular and Translational Medicine) was supported from the Operational programme Research and Development for Innovations (project CZ.1.05/2.1.00/01.0030).

Author contributions: HJ: data collection, data interpretation, manuscript writing, figures; AK, KH: statistical analysis; LN: data processing; DF: data collection, data interpretation; PB: data collection; TA: study design, data interpretation.

Conflict of interest statement: None declared.

RERERENCES

1. Roux A, Lison D, Junot C, Heilier JF. Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review. Clin Biochem 2011;44(1):119-35.
2. Becker S, Kortz L, Helmschrodt C, Thiery J, Ceglarek U. LC-MS-based metabolomics in the clinical laboratory. J Chromatogr B Analyt Technol Biomed Life Sci. 2012;883-884:68-75.
3. Wikoff WR, Gangoiti JA, Barshop BA, Siuzdak G. Metabolomics identifies perturbations in human disorders of propionate metabolism. Clin Chem 2007;53(12):2169-76.
4. Smith CA, Want EJ, 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.
5. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G. METLIN: a metabolite mass spectral database. Ther Drug Monit 2005;27(6):747-51.
6. Bajad SU, Lu W, Kimball EH, Yuan J, Peterson C, Rabinowitz JD. Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. J Chromatogr A 2006;1125(1):76-88.
7. Bajad S, Shulaev V. LC-MS-based metabolomics. Methods Mol Biol 2011;708:213-28.
8. R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.
9. Brown M, Wedge DC, Goodacre R, Kell DB, Baker PN, Kenny LC, Mamas MA, Neyses L, Dunn WB. Automated workflows for accurate mass-based putative metabolite identification in LC/MS-derived metabolomic datasets. Bioinformatics 2011;27(8):1108-12.
10. Hartmann S, Okun JG, Schmidt C, Langhans CD, Garbade SF, Burgard P, Haas D, Sass JO, Nyhan WL, Hoffmann GF. Comprehensive detection of disorders of purine and pyrimidine metabolism by HPLC with electrospray ionization tandem mass spectrometry. Clin Chem 2006;52(6):1127-37.
11. Bosch AM. Classical galactosaemia revisited. J Inherit Metab Dis 2006;29(4):516-25.