TARGETED AND UNTARGETED STUDY OF URINARY METABOLITES AS POTENTIAL CANCER MARKERS

W. Struck-Lewicka[1], D. Siluk[1], A. Yumba Mpanga[1], M. Markuszewski[2], R. Kaliszan[1], M. J. Markuszewski[1]

[1)] Department of Biopharmaceutics and Pharmacodynamics, Medical University of Gdańsk, Poland
[2)] Department of Urology, Medical University of Gdańsk, Poland

INTRODUCTION

Nowadays, in the prosperity of metabolomics, searching for putative biomarkers in cancer disease is rapidly developing. Thanks to the advanced analytical techniques involving mass spectrometry detection, the whole spectra of metabolites from biological samples can be determined and identified. Subsequently, by applying advanced bioinformatics methods, the obtained metabolic profiles can be compared and evaluated in terms of their usage as a prognostic markers in disease detection. Therefore, the aim of this study was a targeted as well as untargeted analysis of urine samples with the use of modern analytical techniques like LC-ESI-MS/MS as well as LC-ESI-TOF/MS from urogenital track cancer patients and healthy volunteers. Furthermore, the obtained data sets were statistically analyzed with the use of univariate (U-Mann Whitney test) and multivariate statistical analyses (PCA, HCA, PLS-DA, OPLS-DA, K-NN, SVM and logistic regression).

AIM of STUDY

1. Development and validation of reliable analytical procedure which bases on LC-MS/MS and LC-TOF-MS separation of urinary nucleosides.
2. Analysis of nucleosides in urine samples from healthy and cancer patients.
3. Univariate and multivariate analysis on the obtained data sets.

SPECTROMETRIC PARAMETERS

Optimized source parameters for LC –TOF /MS: Gas temperature 300°C, Drying Gas 12 L/min, Nebulizer 30 psig, Fragmentor 140 V, Skimmer 65 V, positive ion mode, Vcap 4000V.

Optimized source parameters for LC -MS /MS: Gas temperature 300°C, Drying Gas 12 L/min, Nebulizer 30 psig, positive polarity, Vcap 4000 V, fragmentor and collision energy are various for each analyte. Analysis was performed in dynamic MRM mode in min/max dwell time 38,17 ms/ 496 ms.

BIOINFORMATICS APPROACH

The obtained data sets were calculated and analyzed using Mass Profiler Professional software (Agilent Technologies) SIMCA-P+ 13 (Umetrics, Umea, Sweden) as well as Matlab (Mathworks, USA).

LC-ESI-QqQ-MS

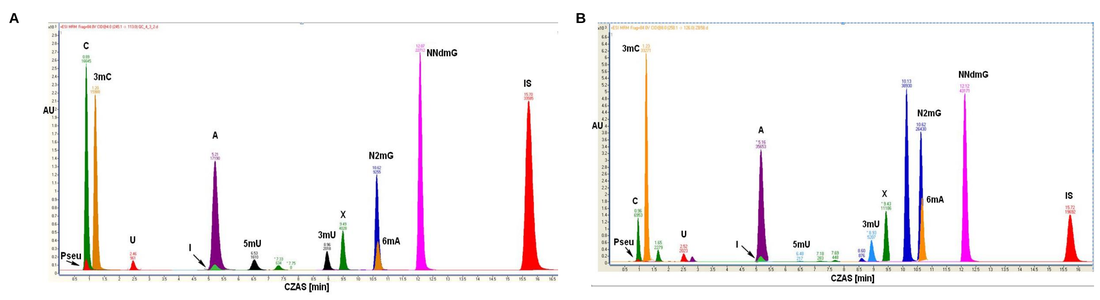
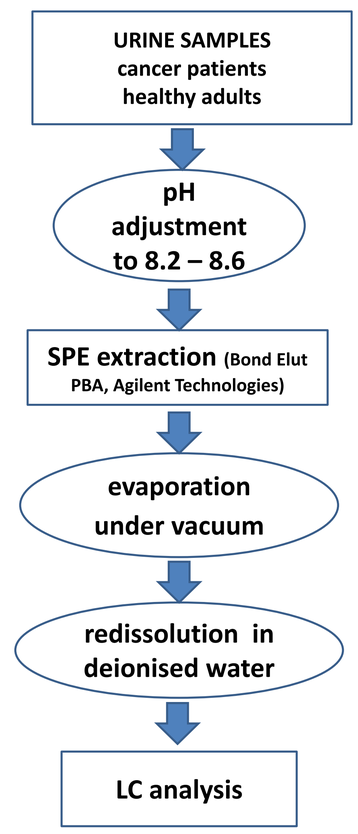


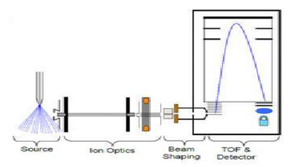
Fig. 1. Multiple reaction monitoring (MRM) mode for 12 nucleosides from A) reference sample and B) urine sample obtained using LC-QqQ-MS

SAMPLE PREPARATION



CHROMATOGRAPHIC PARAMETERS

Instrumentation:



LC Agilent 1200 (Agilent Technologies)
6224 TOF -MS (Agilent Technologies)
6430 Triple Quadrupole –MS (Agilent Technologies)

Column:

Zorbax Extend-C18, Rapid Resolution HT (2,1 x 50 mm; 1,8 µm)

Optimized separation conditions:

Mobile phase A : 0,05% Formic acid in water pH 3,

Mobile phase B: 0,05% Formic acid in methanol pH 3 .

Elution: gradient from 99:1 to 96.8:3.2 (A:B, v/v) from 0 till 3.5 min, then gradient elution from 96.8:3.2 to 88: 12 (A:B, v/v) from 3.5 till 7 min, then isocration elution till 12 min, subsequently gradient elution from 88:12 to 85:15 (A:B, v/v) from 12 till 13 minutes and isocratic elution till 17 minutes.

Flow rate: 0.3 ml/min, column temperature: 8°C

LC-ESI-TOF-MS

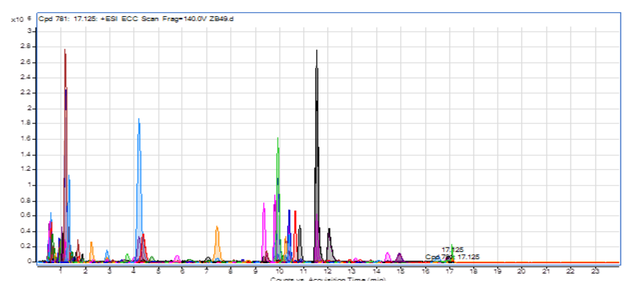


Fig.2
Metabolites extracted from urine sample using molecular feature extractor (MFE) algorithm.

VALIDATION RESULTS

LINEARITY RANGE: **02-200 µM/L**	CORRELATION COEFFICIENT: **R=0.983-0.999**
INTRA-DAY PRECISION OF PEAK AREA: **CV= 0.83 -7.25 %**	INTERDAY PRECISION OF PEAK AREA: **CV= 3.54 – 14.15 %**
LOD: **0.04 – 0.18 µg/ml** LOQ: **0.12 – 0.54 µg/ml**	ACCURACY: **85.89 – 108.21 %**

PCA

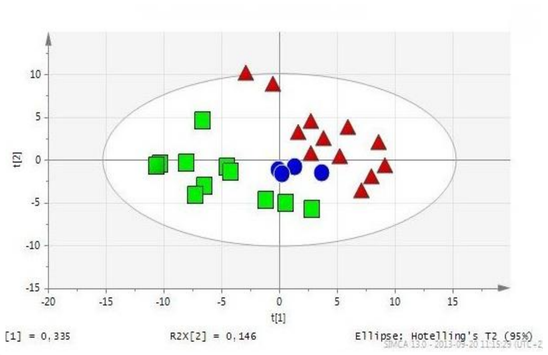


Fig. 3
Principal Component Analysis on data set from LC-TOF-MS analysis. Red triangles, green boxes and blue circles correspond to healthy controls, cancer patients, and quality controls samples, respectively.

HCA

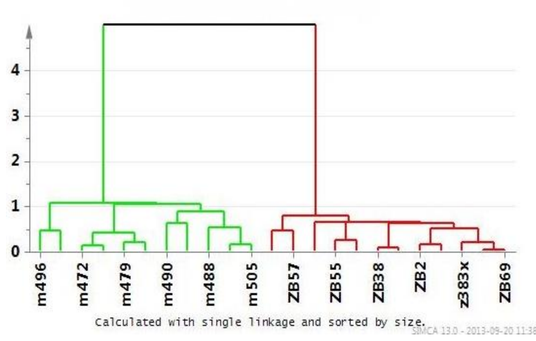


Fig. 4
Hierarchical Cluster Analysis on samples determined with the use of LC-TOF-MS analysis. Green and red lines correspond to cancer patients and healthy volunteers.

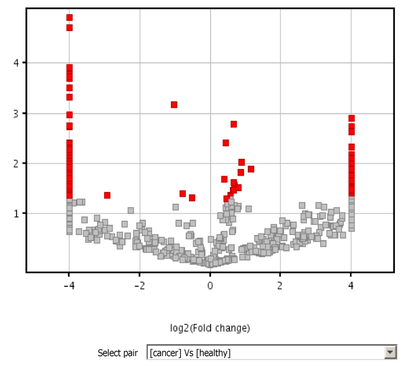


Fig. 5
Volcano plot on data set after alignment and filtering. The red boxes indicate massess of analytes (n=99) that are statistically significant ($p<0.05$)

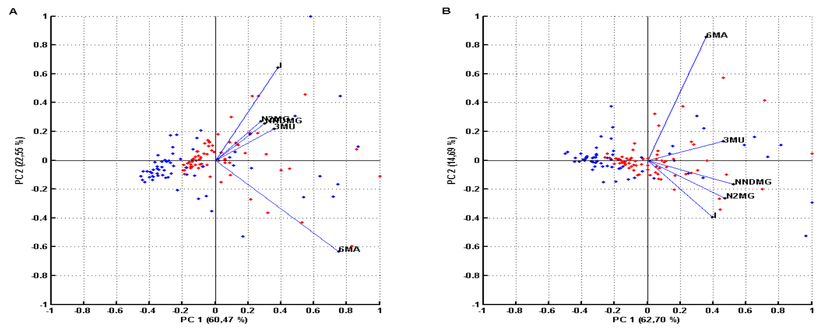


Fig. 6. Principal Component Analysis with biplot on statistically significant nucleosides (6mA, I, 3mU, N2mG, NNdmG) after U Mann-Whitney test ($p<0.05$) determined using LC-QqQ-MS . Blue and red stars correspond to healthy volunteers and cancer patients, respectively. A- data after autoscaling, B- data after level scaling.

Table 1. Overview of diagnostic outcomes of supervised pattern recognition analysis: K-nearest neighbours (k-nn) , Partial least square discriminat analysis (PLS-DA) , Support vector machine (SVM) and logistic regression for statistically significant nucleosides (n=5) Ccr – correct classification rate, cal - calibration set, test - test set.

Discrimination model	Parameter for LC-QqQ-MS data			
	ccr_cal	ccr_test	sensitivity	specificity
K-nn	65,56 %	61,54 %	71,43 %	50,00 %
PLS-DA	68,89 %	64,10 %	88,89 %	42,86 %
SVM	-	71,79 %	77,78 %	66,67 %
Logistic regression	64,44 %	61,54 %	72,22 %	52,38 %

OPLS-DA

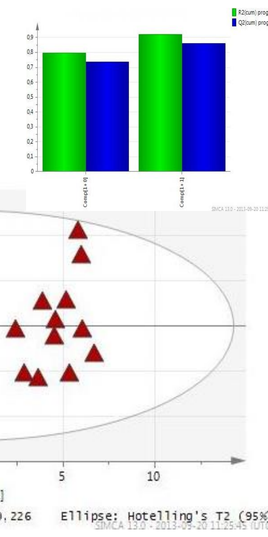


Fig. 7
Orthogonal projection to latent structures with discriminant analysis on data set obtained from LC-TOF-MS after UV scaling. Red triangles and green boxes correspond to healthy volunteers and cancer patients, respectively. **R^2 = 0.921, Q^2 =0.861.**

RESULTS:

1. Liquid chromatography with tripple quadrupole allows for determination of 12 nucleosides from urine samples in relatively short time (17 minutes) with good linearity and precision. Low values of quantification limit enable determination of even small nucleosides' levels.
2. Both techniques enable the analysis of nucleosides and other metabolites that could be statistically significant in cancerogenesis.
3. The presented results prove the usefulness of the metabolomic approach in studying urinary nucleoside profiles with high diagnostic potency in urogenital cancer diseases.
4. The sensitivity and specificity for data obtained from LC-MS/MS ranged from 71.43 to 88.89% as well as from 42.86 to 66.7%, respectively.
5. The results of the OPLS-DA analysis on LC-TOF-MS data showed clear separation of two groups. The high prediction ability was confirmed by the predicted variance Q^2=0.861.

Acknowledgments:

This research work was supported by the system project "InnoDoktorant – Scholarships for PhD students, IVth edition". Project is co-financed by the European Union in the frame of the European Social Fund.

The authors would like to thank Renata Bujak for her invaluable support in Simca P+13 calculations.