Abstract
Prostate cancer (PCa) is the second leading cause of mortality in men, the present diagnosis method being based on serum prostate-specific antigen (PSA) screening, with low specificity and overestimated values. A combined untargeted and targeted metabolomic study of metabolites from blood serum samples collected from healthy (n=11), hyperplasia (n=39) and prostate cancer (n=83) patients is presented, using the HPLC–(ESI+) QTOF-MS analysis. The profile of blood serum samples provided complementary information obtained from Base Peak Chromatograms and Dissect chromatograms. Based on dissect chromatograms, two different methods of statistical analysis were used, either based on the instrument software with automated alignments with/without normalization (Profile Analysis), or based on manual alignment followed by statistical analysis (Unscrambler10.X software). Both methods used the unsupervised Principal Component Analysis which discriminated between normal, hyperplasia and cancer patients. The second method allowed a better discrimination between groups, by qualitative and quantitative parameters (m/z values versus peak areas) and better possibilities to identify the molecules responsible for such discriminations.

Considering the retention time interval (6-17 min), four molecules to be considered as putative biomarkers for hyperplasia or prostate cancer were identified: Prostaglandins E2/G2, Pregnenolone/ethyltestosterone, Lysophosphatidylcholine18:2/0:0, Galactosylceramide (18:1/24:1).

By using larger patient cohorts and optimizing the data processing and chemometric analysis, more reliable biomarkers for prostate hyperplasia and cancer can be discovered and quantified. This preliminary study has had promising findings for the implementation and validation of metabolomic targeted analysis in clinical laboratories.

Keywords: untargeted/targeted metabolomics, prostate cancer, biomarkers, small metabolites, LC-MS

INTRODUCTION
Prostate cancer (PCa) is the second leading cause of cancer-related mortality in men (Cancer Facts & Figures 2013; Landis et al., 1998). The common diagnosis method is based on the serum prostate-specific antigen (PSA), digital rectal examination and ultrasound-guided prostate biopsy screening test.

Metabolome is a data-rich source of information concerning all the low-molecular-weight metabolites in a tissue, cell or biofluid, which can indicate biochemical changes in different pathways and possible perturbations in pathologic situations. The emerging field of metabolomics, in which a large number of small-molecule metabolites from body fluids or tissues are detec-
ted qualitatively or quantitatively, in a single step, promises great potential for early diagnosis, for therapy monitoring and for understanding the pathogenesis of many diseases (Park et al., 2012). Metabolic biomarkers are useful for routine screening of inborn errors of metabolism using either chromatography coupled with mass spectrometry or tandem mass spectrometry (Chace et al., 2002; Piraud et al., 2003).

However, global metabolite profiling of prostate cancer patients still remains in an early stage, and there is no biomarker panel currently in use for clinical testing (Trock, 2011).

Tissue sarcosine levels were proposed some years ago as a potential biomarker for the aggressive form of the disease in a metabolomic profiling study using both liquid- and gas-chromatography coupled to mass spectrometry (LC-MS and GC-MS). Its concentration in prostate-cancer related tissue specimens was highly increased during PCa progression to metastasis, but differences in urine were much less marked (Sreekumar et al., 2009). The controversial blood serum sarcosine does not discriminate between the three groups of men (control, localized prostate cancer and metastasized castration-resistant prostate cancer). Moreover, on an individual basis the sarcosine concentration was not relevant with respect to increased PSA concentrations and to prostate cancer progression. Serum sarcosine data in conjunction with the urinary sarcosine data (Sreekumar et al., 2009) shows that there are no major alterations in the extracellular concentrations of sarcosine, the assessment of sarcosine having a limited potential in the diagnostic algorithm of prostate cancer (Stuys et al., 2010).

The analysis of cancerous tissues by proton high-resolution magic-angle spinning nuclear magnetic resonance (NMR) spectroscopy has shown a decrease in the concentrations of citrate and polyamines and increases in choline, glycerophospholipids and lactate concentrations during prostate cancer proliferation (Swanson et al., 2003, 2006). Increased levels of cholesterol, as well as alterations in amino acid metabolism were detected in metastatic bone samples by GC-MS (Thysell et al., 2010). However, none of these studies included enough patients to offer strong leads on the metabolic alterations associated with prostate cancer.

Other metabolomic studies focused on the identification of relevant blood biomarkers for early diagnosis of prostate cancer. Zhou (2012) identified the lipid profile of 141 plasma lipids of patients with prostate cancer and 36 samples from healthy patients. The most significant lipid species as a biomarker candidate was lysophosphatidilcholine 22:6 (p=0.144). Higher concentrations of cholesteryl ester species with linoleic acid and phosphatidylycholine PC(34:2) were noticed. Other studies (Keshari, 2011; Zhou, 2011; Zaidi, 2013) identified significant modifications of phosphatidylethanolamine (PE), ether-phosphatidylethanolamine (ePE) and ether-phosphatidilcholine (ePC). Studies on larger cohorts are needed. More robust and reproducible metabolic models still need to be developed for improved understanding of disease progression and more reliable prostate cancer detection.

This paper reports an untargeted metabolomic procedure based LC-QTOF-(ESI+)-MS technique applied for polar molecules from blood serum of prostate cancer patients (histologically classified) compared with healthy controls. By multivariate analysis (Principal Component Analysis and Cluster Analysis), the discriminations between groups have been proved and possible biomarker candidates for diagnosis have been identified.

**MATERIALS AND METHODS**

**Blood sample collection**

The blood samples were collected from 37 prostate cancer patients before surgery (CO group) (mean age 62.9±5.7 years), 46 prostate cancer patients before biopsy (CB group) (mean age 66.7±7.6), 39 benign prostatic hyperplasia patients (H group) (mean age 63.8±7.7), and 11 healthy individuals (M group) (mean age 27±5.2). After securing the patients’ agreement on the protocol, blood samples were collected at Municipal Clinic Hospital Cluj-Napoca by venipuncture into vacutainer tubes without anticoagulant. Serum was obtained by centrifugation at 2000 rpm for 10 minutes and aliquots of 1 ml were frozen and stored at −20 °C for further use. The sample collection and storage procedures for prostate cancer patients and healthy individuals were identical. Histologic classification was obtained meanwhile and quantified by Gleason scores, as well the PSA values, for each patient.
For the metabolomic analysis, a stock sample of human blood serum was used to develop the serum sample preparation protocol and HPLC–(ESI+)QTOF-MS method. Serum samples were thawed on ice, and protein precipitation was performed by the addition of methanol to a volume of 200 μL serum in a 5:1 volume ratio. Samples were vortexed for 20s, sonicated for 5 minutes and centrifuged at 1500g for 15 minutes. The supernatant was collected, filtered through 0.2 µm filters and kept in the deep freezer until analysis.

**HPLC–ESI(+)–QTOF-MS analysis**

Aliquots of 5 µl serum were subjected to chromatography on a Thermo Scientific HPLC UltiMate 3000 system equipped with a quaternary pump system DionexUltiMate 3000 (UHPLC), a DionexUltimate 3000 photodiode array detector, column oven and autosampler.

Serum metabolites were separated using a Thermo Scientific Acclaim C-18 column (3μm, 2.1 X 50 mm) at 40°C. The mobile phase consisted of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The flow rate was set at 0.5 mL·min⁻¹. The gradient elution initial conditions were 1% B with linear gradient to 15% B from 0 to 3 min, followed by linear gradient to 50% B at 6 min, linear gradient to 95% B at 9 min, isocratic on 95% B for 6 min and then returned to initial conditions at 15 min and kept isocratic at 1% B for 5 min. The PDA detector was set at 270 nm. The eluent was introduced directly into the mass spectrometer by electrospray.

The mass spectrometry was performed on a BrukerDaltonicsMaxis Impact QTOF instrument, operating in positive ion mode (ESI+). The mass range was set between 50-1000 m/z. The nebulizing gas pressure was set at 2 bar, the drying gas flow at 8 L/min, the drying gas temperature at 180 °C. Before each chromatographic run, a calibrant solution of sodium formate was injected.

The control of the instrument and data processing were done using ToFControl 3.2 and Data Analysis 4.1 (BrukerDaltonics).

**Statistical Analysis**

The unsupervised analysis was done by Principal Component Analysis (PCA) method to discriminate between normal and pathologic samples, using the instrument algorithm and Profile Analysis software (Bruker, Daltonics) (statistical method 1), as well as manual alignment of retention times and PCA analysis by the Unscrambler 10X.1 (statistical method 2).

Statistical method 1 included the alignment parameters, with or without normalization, in the whole separation interval, from min 1 to 17. The limits of bucket filter were either >50% (considering the profiles which are common for at least 50% of molecules) or > 90% (considering the profiles which are common for at least 90% of molecules). The S/N thresholds were 5.0 for chromatograms and 1.0 for MS peak finders. The coefficient of correlation was established to be 0.7 without smoothing and the search range m/z was 0.5.

Statistical method 2 included the manual alignment of the m/z values in the separation interval, from min 6 to 17. Such alignment was necessary since the same molecules had shifted values of RT, of about 0.1-0.3 min.

The PCA scores plots (by both methods) indicated the spectral similarity between samples and allowed the classification of normal or pathological samples. The loadings were based either on m/z values (to identify qualitative differences between samples) or peak areas (to identify semi-quantitative differences). The PCA analysis was used also to verify the metabolite homogeneity of each group (based on retention times, peak areas and m/z values).

**RESULTS AND DISCUSSION**

1. **Untargeted fingerprint of blood serum by HPLC–(ESI+)–QTOF-MS analysis**

The metabolomic serum profile by HPLC–(ESI+)–QTOF-MS analysis is a rapid and sensitive method which acquires robust results and allows the simultaneous identification of known and unknown compounds.

Fig. 1 (A- D) shows the different patterns of base peak chromatograms (BPC) of serum samples from patients with benign prostatic hyperplasia (H), CB, CO groups, compared with controls (M). There were 7-9 main peaks separated and identified in BPC chromatograms of the four groups H, CB, CO and M. Such peaks represent the major components, representing especially the phospholipids.

A more detailed composition is offered by the Dissect chromatograms, processed by the instrument Data analysis software, as shown in Fig.2 (A-D) for the same groups of patients.
Dissect chromatograms contain 68-85 peaks showing a large diversity of major and minor molecules. For each molecule, the m/z and peak area values were recorded.

2. Statistical analysis (method 1)

To perform an adequate interpretation of untargeted data, the unsupervised PCA method was applied, the alignments being made automatically by Profile Analysis software attached to Bruker Data Analysis which built specific bucket matrices.

Fig. 1. Base peak chromatograms recorded for serum samples from the four groups of patients: H (A), CB (B), CO (C) and M (D).

Fig. 2. Dissect chromatograms recorded for serum samples from the four groups of patients: H (A), CB (B), CO (C) and M (D).
Different parameters were applied to compare the principal components with the best variance percentages, to discriminate groups M-CB, M-H, M-CO, H-CB, H-CO, in the whole separation interval, from minute 1 to 17, as mentioned in Materials and Methods.

Fig. 3 A-D represents the scores and loadings plots which compare the M group (triangle) with CB (circles) group of patients, without normalization (A) or with normalization, considering the largest bucket values (B and D) or the sum of bucket values (C).

The patterns of these plots indicate differences between the non-normalized and normalized data, diagrams B and D were similar, while diagram C, which considered the sum of bucket values, discriminated groups M and CB more significantly.

3. Statistical analysis (method 2)

By manual alignment, the common m/z values were considered as found in each group between RT=6-17 min. minute 6 to 17, eliminating the interval 1 - 6 min, where insignificant, minor molecules were separated.

From the m/z values obtained, it was possible to establish the homogeneity of each group and that between groups, based on the relation RT-area and RT-m/z using Principal Component Analysis realized with Bruker-Profile Analysis and Unscrambler 10x software.

Fig.4 shows the results of PCA analysis (scores and loadings plots) made by Unscrambler software, after the manual alignment of molecules, considering the RT range from 6 to 17 min. Such analysis took into account the relation RT-Peak area (up) or RT-m/z values (down) for the four groups of patients (M, H, CB, CO). While the RT-Peak area relationship (Fig.4. up) represents a quantitative discrimination, the RT-m/z values (down) identify the specific molecules which discriminate the groups qualitatively.

In the score plots of RT-peak area correlations which are important for quantitative evaluation, one can see that RT values of 9.8, 10.1, 11.6, 12.6-12.7, 14.7, 16.8 min can explain the discrimination among groups. Loadings plots show a good separation of M group, but they do not indicate good discrimination between H, CB and CO groups.

In the score plots of RT-m/z correlations which are important for qualitative evaluation, one can see that compounds having RT values of 9.8, 10.1, 12.3-12.8, 14.8, 16.7 and 17 min can explain the discrimination among groups. In this case, loadings plots show a better separation (in positive versus negative axis) of M group, but they do not indicate good discrimination between H, CB and CO groups.

Finally, we can conclude that similar RTs explain qualitative and quantitative differences between the groups M-H-CB-CO, and that the manual alignment gave superior information of biological significance.

Fig. 3. Profile analysis of untargeted data with alignments, comparing the M (triangle) and CB (circles) groups, without normalization (A) or with normalization, considering the largest bucket values (B and D) or the sum of bucket values (C). The whole 1-17 RT interval was considered.
4. Targeted metabolomic analysis for different RT intervals

Based on the untargeted analysis and PCA statistics (methods 1 and 2), the targeted metabolomic analysis took into consideration the most significant retention times, mentioned above: RT= 8.1-8.2 min corresponding to pregnenolone/ethytestosterone with m/z= 317.13 (prostate-related hormone), RT= 11.6-11.9 min for prostaglandins PGE2/G2 with m/z= 353.26, RT= 9.8 min corresponding to LPC (18:2/0:0), m/z=520.34 and RT=14.1-14.3 min corresponding to GlcCer (d18:1/18:0), m/z=728.59. For these targeted molecules, the mean peak area values and standard deviation were calculated and represented (data not shown). The identification was made using Human Metabolome Database (www.hmdb.ca) and Lipidomics Gateway (www.lipidmaps.org). Other RT values of 11.0, 12.4-12.5, 13.8-14.0 and 14.2 were considered, where diacylglycerols and monoacylglycerols were identified.

Increases of prostaglandins peak areas were gradually noticed, from M group to CO group suggesting that PGE2/PGD2 may represent good biomarkers for prostate cancer. Nevertheless, these data need validation, since standard deviations were high. Considering the RT= 8.1-8.2 min peak areas, corresponding to pregnenolone/ethytestosterone, a slight decrease from M to H/CO/CB groups was observed, while for LPC (18:2/0:0) gradual increases from M to H and CO/CB groups were noticed. Concerning the Galactosylceramide (d18:1/18:0), peaks separated at RT=14.1-14.3 min, the highest increases were noticed for H group against controls (M) and slight increases in CO/CB groups. One can suggest that Galactosylceramide may represent a putative biomarker for Hyperplasia. In addition to these data, we noticed modifications in the peak areas corresponding to some free fatty acids (linoleic, stearic), amino acids and nucleosides.

Our results are in good agreement with literature data which show that blood prostaglandins increase in prostate cancer, PGE2 being highly expressed in the prostate, associated with prostate cancer progression (Jain, 2008; Terada, 2014). Such bioactive lipids are potent mediators of several signal transduction pathways that modulate cellular adhesion, growth and differentiation (Badawi, 2000). Previous findings have also shown abnormalities in fatty acids

---

**Fig.4.** PCA analysis (scores and loadings) determined after the manual alignment of molecules, at RT range from 6 to 17 min. Such analysis took into account the relation RT-Peak area (up) or RT-m/z values (down) for the four groups of patients (M, H, CB, CO).
(Liu, 2006), and amino acids (Sreekumar et al., 2009; Ebenezar et al., 2012; Fu et al., 2010) metabolism in prostate cancer patients. Such data have been confirmed recently by our results as well (Romanciuc et al, unpublished data). Overall, patients with prostate cancer (CO and CB groups) had increased serum polar lipid concentrations (excepting phosphatidic acid, a precursor significantly lower in patients with prostate cancer). This imbalance in plasma concentrations among lipid classes could be caused by increased synthesis of phospholipids in prostate cancer cells resulting in over-consumption of the precursor (Athenstaedt and Daum, 1999). In general, lysophospholipids can bind and activate specific cell-surface G protein-coupled receptors (GPCRs) that initiate cell growth, proliferation and survival pathways, and show altered concentrations in cancer cells, and thus in combination with other specific markers have the potential to contribute to early diagnosis (Murph et al., 2007). Among the different lysophospholipids identified that may play a role in cell signalling, lysoPC(18:2) and lysoPC(18:0) have been reported as biomarkers for prostate cancer detection within a panel of plasma lipids (Peyruchaud, 2009; Zhou et al., 2012). Ceramides, belonging to sphingolipid class, act as second messengers in apoptosis induction in many cells, including prostate cancer cells, being identified as potential chemotherapeutic agents for prostate cancer. Hormones such as androgens are crucial for the normal development of the prostate gland and in maintaining its functional state in the adult. The prolonged presence of androgens may be a risk factor in the development of prostate cancer (Armstrong, 1982; Wilding, 1995). Pregnenolone and ethyltestosterone derivatives are included in the steroid hormone pathway which supplies androgens, such as testosterone and 5α-dihydrotestosterone, which stimulate the androgen-dependent prostate cancer cells proliferation, as shown by many authors (Kliman et al., 1978; Zumoff et al., 1982; Castagnetta and Carruba, 1995; Signorello et al., 1997; Titus et al., 2005; Knudsen and Scher, 2009; Zang et al., 2014).

CONCLUSION

The results of this preliminary study show that the untargeted metabolomic analysis, using the chromatographic fingerprint HPLC–(ESI+)QTOF-MS and profile of blood serum samples offered complementary information, obtained either from Base Peak Chromatograms or from Dissect chromatograms. Based on disect chromatograms, two different methods of statistical analysis were used: one based on the instrument software with automated alignments with/without normalization (Profile Analysis), and another one based on manual alignment followed by statistical analysis using the specialized Unscrambler software.

Both methods used the unsupervised Principal Component Analysis (PCA) to discriminate between normal, hyperplasia and cancer patients, establishing meanwhile the homogeneity of each sample group and the clustering of groups. Comparing these two methods, we conclude that the second method allowed a better discrimination between groups, by both qualitative and quantitative parameters (m/z values versus peak areas), and better possibilities to identify the molecules responsible for such discriminations.

Considering the retention time interval from 6 to 17 minutes, by targeted metabolomic analysis four molecules were identified which can be considered good candidates as putative biomarkers for hyperplasia or prostate cancer: Prostaglandins PGE2/G2, Pregnenolone/ethyltestosterone, lysophosphatidylcholine18:2/0:0 and Galactosylceramide (18:1/24:1). Also, other molecular types were identified (e.g. fatty acids, aminoacids, nucleosides, etc) which will be presented in future studies.

Finally, we consider that using larger patient cohorts and optimizing the data processing and chemometric analysis, more reliable biomarkers for prostate hyperplasia and cancer can be discovered and quantified. This preliminary study showed good promise towards the implementation and validation of metabolomic targeted analysis in clinical laboratories.

Acknowledgments. This paper was published under the frame of EU Social Fund-Human Resources Development Operational Program, project POSDRU/159/1.5/S/132765

Bulletin UASVM Animal Science and Biotechnologies 72(2) / 2015
REFERENCES
