DOI: https://doi.org/10.1016/j.aca.2021.338381

Untargeted metabolomics of prostate cancer zwitterionic and positively charged compounds in urine

Andrea Cerrato[1,‡], Carmen Bedia[2,‡], Anna Laura Capriotti[1,*], Chiara Cavaliere[1], Vincenzo Gentile[3], Martina Maggi[3], Carmela Maria Montone[1], Susy Piovesana[1], Alessandro Sciarra[3], Roma Tauler[2,**], Aldo Laganà[1,4,**]

[1]Dipartimento di Chimica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy

[2] Institute of Environmental Assessment and Water Research (IDAEA-CSIC), Barcelona, Spain

[3]Dipartimento di Scienze ginecologio-ostetriche e Scienze urologiche, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy

[4]CNR NANOTEC, Campus Ecotekne, University of Salento, Via Monteroni, 73100 Lecce, Italy

[‡] These authors equally contributed to this work

** These authors equally contributed to this work

***Corresponding author**: Anna Laura Capriotti,

Department of Chemistry, Università di Roma "La Sapienza"

Piazzale Aldo Moro 5

00185 Rome, Italy

E-mail: annalaura.capriotti@uniroma1.it

tel: +39 06 4991 3679

DOI: https://doi.org/10.1016/j.aca.2021.338381

Abstract

Prostate cancer, a leading cause of cancer-related deaths worldwide, principally occurs in over 50-year-old men. Nowadays there is urgency to discover biomarkers alternative to prostate-specific antigen, as it cannot discriminate patients with benign prostatic hyperplasia from clinically significant forms of prostatic cancer. In the present paper, 32 benign prostatic hyperplasia and 41 prostatic cancer urine samples were collected and analyzed. Polar and positively charged metabolites were therein investigated using an analytical platform comprising an up to 40-fold analyte enrichment step by graphitized carbon black solid-phase extraction, HILIC separation, and untargeted high-resolution mass spectrometry analysis. These classes of compounds are often neglected in common metabolomics experiments even though previous studies reported their significance in cancer biomarker discovery. The complex metabolomics big datasets, generated by the UHPLC-HRMS, were analyzed with the ROIMCR procedure, based on the selection of the MS regions of interest data and their analysis by the Multivariate Curve-Resolution Alternating Least Squares chemometrics method. This approach allowed the resolution and tentative identification of the metabolites differentially expressed by the two data sets. Among these, amino acids and carnitine derivatives were tentatively identified highlighting the importance of the proposed methodology for cancer biomarker research.

Keywords: untargeted metabolomics; positively charged compounds; chemometrics; polar compounds; ROI-MCR-ALS; LC-HRMS

DOI: https://doi.org/10.1016/j.aca.2021.338381

Introduction

Prostate cancer (PC), a leading cause of cancer-related deaths worldwide [1] and the second most diagnosed cancer in men, principally occurs in over 50-year-old men. Nowadays, early detection of PC is achieved by testing the prostate-specific antigen (PSA) level in blood and by digital rectal examination. Regrettably, PSA cannot discriminate benign prostatic hyperplasia (BPH) or prostatitis from clinically significant forms of PC due to its limited sensitivity and specificity; moreover, PSA levels may be affected by several other factors, such as age and urinary tract infections [2]. As such, there is an urgent need for validation of alternative biomarkers suitable for the early, non-invasive differential diagnosis of PC from other benign prostatic pathologies [3–5]. For these reasons, several PSA derivatives, in particular free forms, have been investigated; the circulating free PSA unbound to proteins had been initially reported as a promising biomarker but failed to show a significant predictive value [6]; precursor forms of PSA (pPSA) were also investigated, as they were detected at higher levels in PC tissues than in BPH tissues [7]. The truncated form of PSA precursor, [-2]pPSA, was estimated to be 25-95% of the free PSA in PC patients but only 6–19% in BPH cases [8]. However, data are too limited for implementation of these markers into routine programs for the early diagnosis of PC based on the European Urological Association (EAU) guidelines (https://uroweb.org/guideline/prostate-cancer).

Metabolomics is a valuable tool to discover disease-associated markers [9] because changes in the concentration of metabolites in bio-fluids reflect alterations in the physiological status of an individual [10]. This approach can be used to understand tumor metabolic pathways and identify metabolites markers suitable for PC early detection, prognostic stratification, and therapy response monitoring. To date, targeted and untargeted metabolomics are the two most common approaches applied to prostate tissue and bio-fluids. Metabolomics studies indicated a correlation between tumor aggressiveness and low levels of spermine and citrate in prostate tissue [11]. Changes in the metabolism of polyamines, tricarboxylic acid cycle, and amino acids, have been reported in

different studies on biofluids from PC patients [10]. Metabolomics of urine further supported the altered amino acid metabolism to be a significant factor underlying the pathogenesis of PC [12,13]. Serum citrate, sarcosine, glycine, alanine, and its derivatives have been found as capable of discriminating PC from BPH [14,15], with sarcosine possibly having more predictive value than PSA in differentiating PC patients from negative controls [16]. Decreased urinary levels of glycine, threonine, and alanine were also observed in another study [17]. A multiplatform untargeted metabolomics study revealed the possible role of urea, purine, and tricarboxylic acid metabolisms in the pathogenesis of PC [13]. These previous studies proved much evidence of the high potential of polar compounds as PC biomarkers, also for screening improvement [18].

Compared to tissues or serum, the advantages of urine include the possibility of a non-invasive sampling, large sample volumes, and a low protein content, which simplifies sample preparation procedures [19–21]. However, urine has drawbacks as well due to the high dilution and high complexity; several compound classes are found in urine, including amino acids, anionic conjugated metabolites, lipids, steroids, bile acids, and xenobiotics, and they can hinder the detection of low abundance metabolites [22]. Solid-phase extraction (SPE) is the technique of choice for sample clean-up and preconcentration [23,24], which are essential in targeted metabolomics [25–28] but quite uncommon in untargeted metabolomics. Nevertheless, some enrichment techniques were recently applied to polar, neutral, and ionic compounds also in sample preparations for untargeted metabolomics studies [24,29,30].

Given the above, in this study, a sample preparation method based on graphitized carbon black (GCB) SPE was applied to urine samples from PC patients and individuals diagnosed with BPH to purify low-abundance polar metabolites from ion suppressing salts and proteins. Samples were analyzed by untargeted ultra-high performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS). Hydrophilic interaction liquid chromatography (HILIC) was used for the efficient retention and separation of polar metabolites, to improve compound identification. A chemometric strategy, based on the combination of the region of interest (ROI) concept [31] and

DOI: https://doi.org/10.1016/j.aca.2021.338381

the multivariate curve resolution-alternating least squares (MCR-ALS) method [32], has been applied to filter, compress and resolve the complex metabolomics big data sets gathered by the UHPLC-HRMS. This platform was optimal to identify a set of polar metabolites that could contribute to a better understanding of the pathophysiological processes involved in the onset and progression of PC.

2. Material and Methods

2.1 Chemical Reagents

Optima LC-MS grade water, acetonitrile (ACN), and methanol (MeOH) were purchased from Thermo Fisher Scientific (Waltham, Massachusetts, USA). Trifluoroacetic acid (TFA) was supplied by Romil Ltd. (Cambridge). Formic acid and ammonium formate were purchased from Sigma-Aldrich (Germany). Dichloromethane (DCM) was provided by VWR International (Milan, Italy). Cartridges packed with 500 mg Carbograph 4 were supplied from Lara S.R.L (Lara S.r.l., Formello, RM, Italy).

2.2 Population and Sample Collection

This is an experimental observational research study in which patients have been managed and treated following the normal clinical practice and international guidelines. The protocol was carried out following the current International Conference for Good Clinical Practice and the principle of the Declaration of Helsinki. The study was approved by the University of Rome “La Sapienza” Ethics Committee (Protocol number 5742) and all the patients provided written informed consent. All patients were outpatients referred to the Department of Urology of the University of Rome “La Sapienza”. Patients were distinguished into two groups based on a histologically confirmed

diagnosis; Group 1 consisted of BPH cases; Group 2 consisted of PC cases. Patients were consecutively included in the study according to the following inclusion criteria. In Group 1 (BPH) inclusion criteria were: aged 45-80, histologically confirmed (prostatic biopsy) diagnosis of BPH, no clinical and pathological evidence of PC, prostate volume > 30 mL. In Group 2 (PC), the inclusion criteria were: aged 45-80, histologically confirmed (prostatic biopsy) diagnosis of PC, no clinical evidence of metastatic disease. For both groups, the following exclusion criteria were respected: no diagnosis of other malignancies, no previous prostate surgery or radiotherapy, no previous or concomitant medical therapies that potentially influenced prostatic metabolism and growth (i.e. androgen deprivation therapies, 5-alpha-reductase inhibitors, or chemotherapies), no concomitant inflammatory or metabolic diseases. In Group 1, the prostatic volume was measured by trans-rectal ultrasonography using the ellipsoid formula. In Group 2, the risk classes were assigned following the D'Amico and EAU classification.

All urine samples were collected after at least 30 days from diagnosis (biopsy) and no patient had undergone surgery. Before urine sampling, all patients were fasting (solid and liquid) from 10:00 p.m. of the previous day and no drugs were administrated. From each patient, a complete single-voided urine sample (minimum 100 mL) was collected in the morning (7:30 – 8:00 a.m.) at the Department of Urology, labeled with an identification code to allow their anonymous manipulation in all the testing phases, and stored at -80 °C until metabolomic analysis.

2.3 Sample Treatments

2.3.1 Preparation of Urine Samples and GCB-SPE Enrichment

Samples were analyzed altogether about three months after their collection and conservation at -80 °C. Before analysis, the urine samples were thawed at room temperature, centrifuged for 10 minutes at 1000 × g. The creatinine concentration of each urine sample was measured by a colorimetric

DOI: https://doi.org/10.1016/j.aca.2021.338381

method based on the Jaffe reaction using the commercial Invitrogen Creatinine Urinary Detection kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). To avoid artifacts due to overloading, the amount of urine loaded on the GCB cartridge was calculated by normalization on the creatinine content. SPE and clean-up were carried out on cartridges packed with 500 mg Carbograph 4 using a procedure optimized in our previous work [33]. Briefly, cartridges were prepared by manually packing 500 mg of Carbograph 4 bulk material (130 m^2/g surface area, 20/400–120/200 mesh size) into 6 mL polypropylene tubes (Sigma-Aldrich); then, the cartridge was washed with 5 mL of DCM/MeOH, 80:20 (*v/v*) with 20 mmol L^{-1} TFA and 5 mL of MeOH with 20 mmol L^{-1} TFA. The material was activated by flushing 10 mL of 0.1 mol L^{-1} HCl and conditioned with 10 mL of 20 mmol L^{-1} TFA. Then, for every sample, the proper volume of urine was diluted with 20 mmol L^{-1} TFA, to reach the final volume of 10 mL, and loaded onto the cartridge, which was then sequentially washed with 2 mL of 20 mmol L^{-1} TFA and 0.5 mL MeOH. Finally, analytes were eluted in back-flushing with 10 mL of DCM/MeOH, 80:20 (*v/v*) with 20 mmol L^{-1} TFA. The eluate was evaporated at room temperature in a Speed-Vac SC250 Express (Thermo Savant, Holbrook, NY, USA), and the residue reconstituted in 200 µL of ACN/H_2O, 75:25 (*v/v*) for HILIC separation. Equal aliquots, from each extracted urine sample, were pooled to create multiple quality control (QC) samples.

2.3.2 Ultra-High Performance Liquid Chromatography-HRMS Analysis

A Vanquish binary pump H (Thermo Fisher Scientific, Bremen, Germany), equipped with a thermostated autosampler and column compartment, was used for the chromatographic separation. An iHILIC-Fusion UHPLC Column, SS (100 × 2.1 mm, 1.8 µm particle size, Hilicon, Umea, Sweden) was employed for HILIC separation. Flow rate, column temperature, and gradient parameters were used as reported in our previous work without any modification [34].

DOI: https://doi.org/10.1016/j.aca.2021.338381

The chromatographic system was coupled to a hybrid quadrupole-Orbitrap mass spectrometer Q Exactive (Thermo Fisher Scientific) using heated electrospray ionization (HESI) source. The HESI source was operated in the positive ionization mode and set up as previously reported [33]. Ions were monitored in the range *m*/*z* 150-750 with a resolution (full width at half maximum, FWHM, *m*/*z* 200) of 70,000. Top 5 data-dependent acquisition (DDA) higher-energy collisional dissociation (HCD) fragmentation was performed at 40% normalized collision energy at a resolution of 35,000 (FWHM, *m*/*z* 200).

Samples were run according to the following worklist, which included QC samples to account for the analytical variability. Moreover, solvent blank samples were run before and after each analytical section to check the UHPLC-HRMS system performance and stability, and evaluate the presence of carryover before analysis. The system was then conditioned by running 10 QCs samples, followed by a blank sample. This latter blank was used for background subtraction to remove contaminants from the mobile phases, from the UHPLC-HRMS system, and the compounds suffering from considerable carry-over (more than 25%), which are not suitable for statistical analysis. Then, 10 more QC samples were run for system conditioning. After this stage, randomized samples and controls were run in groups of five followed by a QC injection. All samples were acquired in full scan acquisition mode with exception of three QC samples run at the end of the sequence and acquired in top 5 DDA mode for structure elucidation and identification of biomarkers. The identification-only QCs were preprocessed together with the other samples. The .raw data files were acquired and converted in .cdf format using Xcalibur (version 3.1, Thermo Fisher Scientific).

2.4 Data Processing

Data files in .cdf format were imported into the MATLAB computational environment (release R2018a, The Mathworks Inc.), using the mzcdfread and mzcdf2peaks functions from the MATLAB

DOI: https://doi.org/10.1016/j.aca.2021.338381

Bioinformatics Toolbox. The different data preprocessing and processing steps are summarized in Figure 1.

Each one of the UHPLC-HRMS data files had an approximate size of 1.2 GB. The simultaneous analysis of the 88 samples (32 BPH + 41 PC + 15 QC) would have required ca. 100 GB of computer storage. Therefore, the raw data were conveniently reduced by decreasing the number of retention times used in each file and dividing the selected time range into three regions (i.e. time windows, later referred to as A, B, and C), which were selected considering the density of peaks observed in the total ion current chromatogram (TIC) and allowing a slight overlap between windows to avoid the loss of information on the borders. The mass spectra (*m*/*z* and intensity dyadic values), collected between minutes 4-14, were retained at every 5 elution times, resulting in a reduction from 2265 to 453 retention times (mass spectra) for each sample. The application of the three time windows provided three different time-compressed files containing 100, 115, and 252 retention times, respectively. These individual reduced sizes and time-windowed compressed files were concatenated column-wisely to build three augmented data files, containing the information of the 88 samples in each of the three pre-selected time windows (Figure 1). Therefore, the column-wise augmented data files A, B, and C had in total 9000, 10350, and 22680 retention times, respectively.

As shown in Figure 1, these files were then independently analyzed using the ROIMCR procedure. The ROIMCR strategy [31] consisted of two parts. The first one was the searching of the ROI in the mass spectra, a procedure that looks for the more significant features in the measured mass spectra using a set of specific parameters (see below). As a result of the application of the ROI procedure, a compressed squared data table or data matrix is obtained containing the *m*/*z* aligned mass spectra having the selected significant signal features in the set of samples analyzed simultaneously. The second part of the ROIMCR procedure performs the bilinear decomposition of the resulting ROI MS compressed data matrix by MCR-ALS [35], to resolve the set of elution profiles and mass spectra of the constituents present the simultaneously analyzed samples.

DOI: https://doi.org/10.1016/j.aca.2021.338381

As mentioned above and shown in Figure 1, the ROI procedure looks for those *m/z* values whose MS signals have significant intensities, in this work set higher than 0.5% of the maximum signal intensity. The choice of this value was based on a compromise between the number of resulting ROIs and the absence/presence of noisy peaks after visual inspection for tests at threshold values between 1 and 0.1% [31]. Lower signals were considered instrumental noise or signals belonging to sample constituents at very low concentrations. Every ROI was defined by two additional parameters. The first one was the *m/z* deviation (or error) within which the measured signal is considered to be from the same ion. This error depends on the MS instrument mass accuracy, and in this study, it was set at 0.005 mass units. The other parameter, needed to define every ROI, is the minimum number of consecutive signals that define a chromatographic peak in the UHPLC system used for analysis. Due to the high compression used in this work, this was set to 2 (for each sample, i.e., the signal had to appear 176 times along the augmented data matrices). The ROI procedure used in this work has been previously described in detail by Gorrochategui et al. [31]. The application of these three parameters and the visual inspection of the obtained ROIs, to detect possible signal artifacts, resulted in 135, 164, and 82 ROI values, for the A, B, and C time window files, respectively. This represented an important data size reduction. It converted the raw data files, having an irregular number of measured values (*m*/*z* - intensity dyad values) at each retention time, to new square data matrices, containing the ROI mass spectra at all selected retention times and samples. In the second part of the ROIMCR procedure [35], each of these three ROI data matrices was analyzed by the MCR-ALS method (Figure 1). Briefly, this method performs the bilinear decomposition of a data matrix into the product of two-factor matrices, one related to the elution profiles of the constituents in the analyzed samples and the other related to the spectra of these sample constituents. In the present work, the MATLAB MCR-ALS Toolbox [32] was used for this purpose. The constraints used during the ALS optimization of the elution and spectral profiles were the non-negativity for both, and the normalization of the spectra of the different sample constituents to have their maximum height equal to one. The ALS optimization needs the postulation of several

components and a set of initial estimates of the elution or spectra profiles of these components. The number of components should be large enough to include the number of sample constituents avoiding noisy ones and matrix rank deficiency problems. The initial estimates of the profiles were taken in this work from the 'purest' MS spectra (not noise) among the measured ones, following a procedure similar to that proposed in the literature [36].

As said above, the application of MCR-ALS to the three ROI data matrices resolve the elution profiles and spectra of the different sample constituents (an example of the resolved peaks in one sample for the different components is provided in Supplementary Fig S1). The relative amounts of these components in every sample can be estimated from the relative peak areas of the MCR-ALS resolved elution profiles of the same constituent in the different analyzed samples. From the results of the MCR-ALS analysis of the three A, B, and C matrices, the peak areas of those sample constituents giving significant differences ($p<0.05$ in Welch's t-test [37]) between BPH and PC samples were arranged in a new peak areas data matrix for further analysis by Principal Component Analysis (PCA) [38], and Partial Least-Squares Discriminant Analysis (PLS-DA) [39], (Figure 1) using the PLS-Toolbox (Eigenvector Inc.). PCA was applied directly to the newly arranged peak areas data matrix and used for the unsupervised pattern recognition and clustering of the analyzed samples. Projection of the sample scores in the two principal components (PC1 and PC2) displays graphically the similarities and differences of the analyzed samples in 2D plots. To further discriminate between samples and build the different PLS-DA models, two sample classes were considered, BPH vs PC samples. The model was validated using 12 randomly selected samples from the dataset and using them to test the predictive power of the model (see supplementary Figure S2). The Matthews Correlation Coefficient (MCC) [41] was calculated as a figure of merit of the discrimination model.

The elucidation of the urine metabolites, responsible for the differences between BPH and PC samples, was especially important in this study. Variable Importance in Projection (VIP) scores [42] of the PLS-DA models helped to identify the most influential variables (ROI *m*/*z* values) and were

set to > 1 for relevant metabolites able to discriminate between BPH from PC samples. To illustrate the diagnostic ability of the selected metabolites as potential biomarkers, the receiver operating characteristic (ROC) curves were calculated. For every selected metabolite, the ROC curves display the true-positive rate against the false-positive rate at different threshold settings. The area under the ROC curves (AUCROC) was estimated for all the selected metabolites and used for metabolite classification; AUCROC values equal to 1 indicated 100% sensitivity and specificity in the class assignation, whereas values equal to 0.5 indicated a random classification ability. ROC analysis was performed using the Metaboanalyst online platform (https://www.metaboanalyst.ca).

2.5 Metabolite Identification

Metabolites associated with the significant ROI *m/z* values were identified using their MS/MS data by searching existing databases of human metabolites, such as HMDB (https://hmdb.ca/) and KEGG (https://www.genome.jp/kegg/), and databases of structurally related classes of compounds, such as Lipid Maps (https://www.lipidmaps.org/). Manual MS/MS spectra interpretation was aided by mzCloud (https://www.mzcloud.org/), the largest database of Orbitrap-MS, and MS/MS HCD spectra.

2.6. Software and Computer Specifications

The calculations performed in the present work were carried out using MATLAB R2018a (The Mathworks Inc.), and PLSToolbox (Eigenvector Inc) running on a Fujitsu Celsius R940n workstation equipped with two Intel Xeon CPU E5-2620v3 processors and 128 Gb RAM using Microsoft Windows 7. MCR-ALS analysis was performed using the MCR-ALS toolbox freely available at http://www.mcrals.info/

3. Results and Discussions

DOI: https://doi.org/10.1016/j.aca.2021.338381

3.1. Clinical Characteristics of the Subjects Enrolled in the Study

A total of 32 BPH cases and 41 PC cases were enrolled in Group 1 and Group 2, respectively. Table 1 shows the clinical and pathological characteristics of the population in the two groups. As expected, the mean age was similar between the two groups (Group 1: 65.2 ± 7.3; range 52-79 years; Group 2: 65.9 ± 6.5; range 49-79 years). In Group 1, the mean prostatic volume was 62.4 ± 24.1 mL, range 33-124 mL. In Group 2, the cases were well distributed based on the risk classes (Low risk: 8 cases; Intermediate risk: 27 cases; High risk: 6 cases) and Gleason Score (GS).

3.2 Metabolomics Workflow

Although the collection of data without pre-existing knowledge is one major advantage of untargeted metabolomics, it is also accompanied by the caveat that sample preparation and analytical methods do have a direct impact on the qualitative results. Sample purification steps, separation methods, instrument platforms, and parameters influence the subset of detected metabolites depending upon the compound polarity, which can vary significantly due to the diverse composition of the metabolome [43]. For instance, metabolite extraction is complicated by the extreme concentration range and great physical diversity of metabolites often present in the metabolome; abundant metabolites may saturate the mass spectrometer detector, potentially masking important changes, or suppress the ionization of other molecules [44]. In this context, the clean-up of sample extracts and the development of extractions specific for polar metabolites could increase the metabolome coverage, although very little data on metabolite coverage is currently available on sample preparation in metabolomics, especially for untargeted metabolomics.
For the above reasons, in this work, the typical metabolomics workflow was modified with the introduction of a selective step for the enrichment and purification of polar positively charged compounds, which are of great importance in the discrimination of PC [45]. Sample preparation was considered the optimal strategy for obtaining a representative profile and comprehensive

coverage of a specific subset of compounds, which in turn would allow improving the detection of low abundance potential biomarkers [46]. GCB-SPE enrichment and purification were chosen based on our previous study [33], because of the strong interaction that the GCB sorbent can establish with positively charged and aromatic amino acids, due to the right combination of van der Waals contribution, and π–π interaction with the aromatic moieties. To date, only anion and cation exchange mixed-mode polymeric SPE cartridges were employed for global urine metabolomics [29]; this method allowed to identify of a wide range of compounds such as bile acids, lipids, organic acids, but also molecules connected with pharmaceutical treatment, diet, or lifestyle. Our methodology was more specifically dedicated to zwitterionic and positively charge compounds such as amino acids and carnitines. More recently, an advanced chemoselective metabolite enrichment by tagging and proteolytic release of highly polar metabolites (quaternary amines and carnitines) in breast cancer cells was developed [30].

The iHILIC column was selected for obtaining a more efficient separation for polar urinary compounds. The iHILIC-Fusion is a charge-modulated hydroxyethyl amide HILIC, in which the cationic ammonium site is at the terminal position separated from mixed sulfate and phosphate anionic sites by a linker with hydroxyethyl amide side chains. HILIC has grown in popularity due to its orthogonality to reversed-phase chromatography and the good compatibility with MS [47–50]. The use of the iHILIC Fusion column is still limited in the literature, although it has been recently demonstrated to provide better coverage of metabolite classes like amino acids than conventional HILIC columns [51].

When Q Exactive instrumentations are employed for untargeted analysis, DDA methods are usually preferred over data-independent acquisition (DIA) approaches, as the latter are highly time-consuming and showed poor performance on slow orbitrap-based instruments [52,53]. Since the numbers of points per peak drastically decrease when HCD fragmentation is performed, in this study samples and QCs were acquired in full-scan mode, which guaranteed high-quality peak shapes for both high- and low-abundance substances [54].

DOI: https://doi.org/10.1016/j.aca.2021.338381

3.3 Urinary Metabolic Profile of PC vs BPH

The ROIMCR data processing methodology [55–57] including the application of the MCR-ALS method [58–60], has been described and successfully applied in previous investigations of a variety of metabolomics systems using different LC-MS platforms.

The 88 urine samples (including 32 BPH, 41 PC, and 15 QC samples) investigated in this work were simultaneously processed by the ROIMCR method described in section 2.4 for untargeted analysis. A total number of 100, 85, and 30 MCR-ALS components were resolved in the analysis of the three augmented time window data matrices A, B, and C (see section 2.4), respectively, explaining more than 99.5% of their data variance in the three cases. From these resolved components, 60 showed statistically significant differences ($p<0.05$ in Welch's t-test) between BPH and PC, and they were selected to build the data table summarizing the changes in the peak areas of these 60 components in the 88 samples (size of the corresponding data matrix is 88x60).

PCA was applied to this peak areas data matrix to perform its unsupervised pattern recognition analysis (Figure 2). The PCA scores plot (PC1 vs PC2, Figure 2A) showed that BPH samples were clustered together and differentiated from the rest of the PC samples along the PC1 x-axis (30.1% of explained variance). The QC samples were tightly clustered in the small region of the PC1-PC2 vector space close to the axis origin, a position that confirmed the adequacy of both the experimental and data analysis procedures. PCA was then performed without the QC samples and the same distribution was obtained for PC and BPH samples, with only slight changes in the PC1 and PC2 explained variance (Figure 2B). As shown in Figure 2C and 2D, PC3 and PC4 did not offer any better separation of the samples. Interestingly, the distribution of PC samples in the scores plot was rather dispersed, which was an indication that the molecular profile of the urine of PC patients appreciably differed among them. However, PC samples still had features that enabled differentiating them from the molecular profile of BPH samples, which seemed to be much more

consistent. To investigate if PC samples (and patients) could be discriminated from BPH samples (and patients), these two types of samples were identified using a class vector which was then regressed against the peak areas of the 60 selected MCR-ALS components using the PLS-DA method [39]. BPH and PC samples were projected in the first two latent variables (LV1 vs LV2) scores plot of the PLS-DA model and displayed in Figure 2E. The PLS-DA model using the three first latent variables explained 90% of the Y-variance (class membership) and 42% of the X-variance (peak areas) (Figure 2D).

The reliability of the PLS-DA model was checked by calculating the MCC (see section 4.2) from the cross-validated model. The MCC was 0.945, which indicated that the molecular profile of the analyzed samples enabled good discrimination between BPH and PC samples. The sensitivities or true positive rates were 0.96 and 0.97 for BPH and PC, respectively, and the specificities or true negative rates (TNR) were 0.97 and 0.96 for BPH and PC, respectively. PLS-DA VIP scores revealed the most influential metabolites that enabled the discrimination between the two classes of samples. Only the MCR-ALS components, whose peak areas (variables) gave VIP > 1, were finally retained for metabolite identification and biochemical interpretation (Figure 2F).

The most intense signals (12 retained signals) of the MS spectra of these selected ROIMCR components were analyzed for their possible metabolite assignation: 7 compounds were tentatively identified, 3 were unknown carnitine derivatives and 2 could not be annotated (Table 2, which summarizes the details about compound *m*/*z* value, retention time, and variance explained in the MCR-ALS models). The identified molecules were adenosine, hydroxyvaleroyl carnitine, mandelyl carnitine, methyl inosine, N-acetyl arginine, N-acetyl N-methyl arginine, and tryptophan. For each compound, the mean areas were calculated in the PC and BPH sample groups, and their ratio was used for relative quantitation.

To test the performance of the most effective identified biomarkers for BPH and PC class prediction, a new PLS-DA model was built considering the area values obtained for adenosine, hydroxyvaleroyl carnitine, mandelyl carnitine, methyl inosine, N-acetyl arginine, N-acetyl N-

methyl arginine, and tryptophan in all the samples. As a result, a very good discriminant model was obtained using only one latent variable that explained 70.4% of the Y-variance and 49% of the X-variance (Figure 3A). The MCC values, the sensitivity, and the specificity values were the same as in the model using 60 variables. In this reduced model, the highest VIP value was attributed to N-acetyl arginine, followed by hydroxyvaleroyl carnitine and tryptophan (see Figure 3B).

The AUCROC values were also reported as a measure of the potential of these molecules to be biomarkers of PC. All these molecules were significantly more abundant in the PC samples (up to a 10^2-fold increase) and they were almost absent in BPH samples, as observed for hydroxyvaleroyl carnitine, mandelyl carnitine, and N-acetyl arginine. The AUCROC values of these metabolites were between 0.78 and 0.925; the values indicated a good biomarker classification ability, with N-acetyl arginine as the best classifier (AUCROC=0.925), followed by tryptophan (AUCROC=0.892), and hydroxyvaleroyl carnitine (AUCROC=0.870). These results were coincident with those obtained in the PLS-DA model, reinforcing the idea that these three molecules were the best-found biomarkers that discriminate PC from BPH. The ROC curve plots for these 7 biomarkers are given in Figure 3C (plots for the unidentified molecules can be consulted in Supplementary Figure S3).

3.4 Biological Interpretation

The urinary metabolomics profiles in the PC samples of this study were rather dispersed; the result reflected the clinical heterogeneity of PC which is usually observed in clinical practice, as PC is a highly heterogeneous neoplasm with some men presenting indolent disease and others whose disease is rapidly progressive. A low GS (≤ 7 (3+4)) can be associated with a PC's low aggressiveness. On the contrary, clinically significant PC cases, and in particular those with a higher GS, show rapid growth and progression, probably sustained by a different metabolic profile. However, the present study did not confirm that the metabolic profile of PC cases was influenced by the disease aggressiveness or that the profile of low GS PC cases (6 (3+3) and 7 (3+4)) could be

closer to that of benign hyperplastic cases. No significant differences in the metabolic profiles were found by stratifying the PC cases based on the pathologic aggressiveness (GS); moreover, the same differences were found independently from the GS distribution when comparing PC to BPH cases (Supplementary Figure S4).

To our knowledge, there are no data in the literature on a defined role of the individual compounds characterizing the PC samples inside the metabolic mechanisms related to the prostate gland and PC, although previous studies linked some of these compounds to pro-inflammatory or hypoxia-related processes. Metabolic syndrome is considered a consistent risk factor for PC development, including chronic prostatic inflammation and high concentrations of inflammation-related markers able to enhance tumor growth [62]. However inflammatory processes have been similarly associated with BPH progression: BPH cases displaying prostate chronic inflammation have a higher risk of BPH progression and complications [63]. As in several neoplasms, hypoxia-induced angiogenesis is a crucial mechanism to induce PC progression and invasion through specific key-regulators of response [64]. In a recent meta-analysis, L-carnitine has been shown to significantly reduce inflammatory mediators possibly related to neoplastic transformation processes in different systems, such as tumor necrosis factor alpha (TNF) and interleukin-6 (IL-6) [65]. Coras et al. showed that hydroxivaleric carnitine is a pro-inflammatory metabolite related to rheumatoid arthritis pathogenesis [66]. The carnitine metabolic system has also been described as a gridlock to finely trigger the metabolic flexibility of cancer cells. The carnitine system has an enzymatic role as a crucial factor in cancer metabolic plasticity, so to survive in the face of adverse environmental conditions [67]. Sasso et al. showed that N-acelylarginine can reduce the activity of antioxidant enzymes such as catalase, increasing the oxidative stress possibly related to neoplastic induction in different systems [68]. An abundant supply of amino acids is important to sustain cancer proliferative drive and arginine can be an auxotrophy nutrient for different neoplastic cells [69]. Some evidence suggests that arginine derivatives can promote PC cell proliferation, migration, and invasion, by increasing nitric oxide expression, as well as pro-angiogenic growth factors [70].

DOI: https://doi.org/10.1016/j.aca.2021.338381

Moreover, experimental data on PC suggest a protective role of acetylcarnitine derivatives due to an anti-oxidant and anti-inflammatory activity in the prostate gland, which in turn hinders the production of pro-inflammatory cytokines and chemokines [71]. Despite this experimental evidence, no clinical research has been carried out so far and no data are present in the literature to clarify the metabolic role of the specific compounds extracted in our analysis from PC cases.

4. Conclusions

PC diagnosis, especially at early stages, is a critical issue for clinical research. Nowadays, untargeted metabolomics followed by statistical analysis and putative biomarker identification is one of the most promising techniques to reach this aim. In the present paper, an analytical workflow dedicated to polar and charged metabolite analysis was described for the first time, since such classes of metabolites are often neglected in common metabolomics workflows. In comparison to the common "dilute-and-shoot" strategy, this approach can afford a 10-40-fold pre-concentration of sample extracts. A chemometric approach based on ROIMCR chemometric method was applied to the raw datasets generated by the UHPLC-HRMS analysis. PCA and PLS-DA analysis of the relative concentrations of these metabolites confirmed the differences and discrimination between PC and BPH samples and allowed the identification of possible biomarkers of prostatic malignancy. Among these metabolites, adenosine, hydroxyvaleroyl carnitine, mandelyl carnitine, methyl inosine, N-acetyl arginine, N-acetyl N-methyl arginine, and tryptophan were identified. These metabolites belonged to the classes of carnitine and amino acid derivatives, which were significantly increased in the PC samples and almost absent in most BPH samples. This effect was particularly evident for hydroxyvaleroyl carnitine, mandelyl carnitine, and N-acetyl arginine.

This study confirmed the importance of polar and charged metabolites in cancer studies, even though dedicated methodologies for their analysis are rarely reported in metabolomics studies; the approach can be extended to study other malignancies, to improve diagnosis and patient prognosis.

DOI: https://doi.org/10.1016/j.aca.2021.338381

5. Acknowledgment

The work was supported by the PRIN project Prot. 2017Y2PAB8, entitled “Cutting Edge Analytical Chemistry Methodologies and Bio-Tools to Boost Precision Medicine in Hormone-Related Diseases”, provided by the Italian Ministry of Education, Universities and Research.

Moreover, this work was also supported by Generalitat de Catalunya (Suport a les activitats de Grups de Recerca, 2017 SGR753)

6. References

[1] A. Heidenreich, P.J. Bastian, J. Bellmunt, M. Bolla, S. Joniau, T. van der Kwast, M. Mason, V. Matveev, T. Wiegel, F. Zattoni, N. Mottet, EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent—Update 2013, Eur. Urol. 65 (2014) 124–137. https://doi.org/10.1016/j.eururo.2013.09.046.

[2] A.R. Lima, M. de L. Bastos, M. Carvalho, P. Guedes de Pinho, Biomarker discovery in human prostate cancer: An update in metabolomics studies, Transl. Oncol. (2016). https://doi.org/10.1016/j.tranon.2016.05.004.

[3] G.M. Busetto, F. Del Giudice, M. Maggi, F. De Marco, A. Porreca, I. Sperduti, F.M. Magliocca, S. Salciccia, B.I. Chung, E. De Berardinis, A. Sciarra, Prospective assessment of two-gene urinary test with multiparametric magnetic resonance imaging of the prostate for men undergoing primary prostate biopsy, World J. Urol. (2020). https://doi.org/10.1007/s00345-020-03359-w.

[4] C.M. de la Calle, V. Fasulo, J.E. Cowan, P.E. Lonergan, M. Maggi, A.J. Gadzinski, R.A. Yeung, A. Saita, M.R. Cooperberg, K. Shinohara, P.R. Carroll, H.G. Nguyen, Clinical Utility of 4Kscore ® , ExosomeDx™ and Magnetic Resonance Imaging for the Early Detection of

High Grade Prostate Cancer, J. Urol. 205 (2021) 452–460. https://doi.org/10.1097/JU.0000000000001361.

[5] G.M. Busetto, E. De Berardinis, A. Sciarra, V. Panebianco, R. Giovannone, S. Rosato, P. D'Errigo, F. Di Silverio, V. Gentile, S. Salciccia, Prostate Cancer Gene 3 and Multiparametric Magnetic Resonance Can Reduce Unnecessary Biopsies: Decision Curve Analysis to Evaluate Predictive Models, Urology. 82 (2013) 1355–1362. https://doi.org/10.1016/j.urology.2013.06.078.

[6] H. Vasarainen, J. Salman, H. Salminen, R. Valdagni, T. Pickles, C. Bangma, M.J. Roobol, A. Rannikko, Predictive role of free prostate-specific antigen in a prospective active surveillance program (PRIAS), World J. Urol. (2015). https://doi.org/10.1007/s00345-015-1542-3.

[7] S.D. Mikolajczyk, L.S. Millar, T.J. Wang, H.G. Rittenhouse, L.S. Marks, W. Song, T.M. Wheeler, K.M. Slawin, A precursor form of prostate-specific antigen is more highly elevated in prostate cancer compared with benign transition zone prostate tissue, Cancer Res. (2000).

[8] S.D. Mikolajczyk, K.M. Marker, L.S. Millar, A. Kumar, M.S. Saedi, J.K. Payne, C.L. Evans, C.L. Gasior, H.J. Linton, P. Carpenter, H.G. Rittenhouse, A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer, Cancer Res. (2001).

[9] X. Zhang, Q. Li, Z. Xu, J. Dou, Mass spectrometry-based metabolomics in health and medical science: a systematic review, RSC Adv. 10 (2020) 3092–3104. https://doi.org/10.1039/C9RA08985C.

[10] N. Gómez-Cebrián, A. Rojas-Benedicto, A. Albors-Vaquer, J. López-Guerrero, A. Pineda-Lucena, L. Puchades-Carrasco, Metabolomics Contributions to the Discovery of Prostate Cancer Biomarkers, Metabolites. 9 (2019) 48. https://doi.org/10.3390/metabo9030048.

[11] G.F. Giskeødegård, H. Bertilsson, K.M. Selnæs, A.J. Wright, T.F. Bathen, T. Viset, J. Halgunset, A. Angelsen, I.S. Gribbestad, M.B. Tessem, Spermine and Citrate as Metabolic Biomarkers for Assessing Prostate Cancer Aggressiveness, PLoS One. 8 (2013) 1–9. https://doi.org/10.1371/journal.pone.0062375.

[12] P. Dereziński, A. Klupczynska, W. Sawicki, J.A. Pałka, Z.J. Kokot, Amino Acid Profiles of Serum and Urine in Search for Prostate Cancer Biomarkers: a Pilot Study, Int. J. Med. Sci. 14 (2017) 1–12. https://doi.org/10.7150/ijms.15783.

[13] W. Struck-Lewicka, M. Kordalewska, R. Bujak, A. Yumba Mpanga, M. Markuszewski, J. Jacyna, M. Matuszewski, R. Kaliszan, M.J. Markuszewski, Urine metabolic fingerprinting using LC–MS and GC–MS reveals metabolite changes in prostate cancer: A pilot study, J. Pharm. Biomed. Anal. 111 (2015) 351–361. https://doi.org/10.1016/j.jpba.2014.12.026.

[14] C. Pérez-Rambla, L. Puchades-Carrasco, M. García-Flores, J. Rubio-Briones, J.A. López-Guerrero, A. Pineda-Lucena, Non-invasive urinary metabolomic profiling discriminates prostate cancer from benign prostatic hyperplasia, Metabolomics. 13 (2017) 52. https://doi.org/10.1007/s11306-017-1194-y.

[15] D. Kumar, A. Gupta, A. Mandhani, S.N. Sankhwar, NMR spectroscopy of filtered serum of prostate cancer: A new frontier in metabolomics, Prostate. 76 (2016) 1106–1119. https://doi.org/10.1002/pros.23198.

[16] N. Melling, E. Thomsen, M.C. Tsourlakis, M. Kluth, C. Hube-Magg, S. Minner, C. Koop, M. Graefen, H. Heinzer, C. Wittmer, G. Sauter, W. Wilczak, H. Huland, R. Simon, T. Schlomm, S. Steurer, T. Krech, Overexpression of enhancer of zeste homolog 2 (EZH2) characterizes an aggressive subset of prostate cancers and predicts patient prognosis independently from pre- and postoperatively assessed clinicopathological parameters, Carcinogenesis. 36 (2015) 1333–1340. https://doi.org/10.1093/carcin/bgv137.

[17] S.M. Lloyd, J. Arnold, A. Sreekumar, Metabolomic profiling of hormone-dependent cancers: a bird's eye view, Trends Endocrinol. Metab. 26 (2015) 477–485. https://doi.org/10.1016/j.tem.2015.07.001.

[18] S.S. Dinges, A. Hohm, L.A. Vandergrift, J. Nowak, P. Habbel, I.A. Kaltashov, L.L. Cheng, Cancer metabolomic markers in urine: evidence, techniques and recommendations, Nat. Rev. Urol. 16 (2019) 339–362. https://doi.org/10.1038/s41585-019-0185-3.

[19] C. Burton, Y. Ma, Current Trends in Cancer Biomarker Discovery Using Urinary Metabolomics: Achievements and New Challenges, Curr. Med. Chem. 26 (2019) 5–28. https://doi.org/10.2174/0929867324666170914102236.

[20] M. Clos-Garcia, A. Loizaga-Iriarte, P. Zuñiga-Garcia, P. Sánchez-Mosquera, A. Rosa Cortazar, E. González, V. Torrano, C. Alonso, M. Pérez-Cormenzana, A. Ugalde-Olano, I. Lacasa-Viscasillas, A. Castro, F. Royo, M. Unda, A. Carracedo, J.M. Falcón-Pérez, Metabolic alterations in urine extracellular vesicles are associated to prostate cancer pathogenesis and progression, J. Extracell. Vesicles. 7 (2018) 1470442. https://doi.org/10.1080/20013078.2018.1470442.

[21] G.N. Gowda, S. Zhang, H. Gu, V. Asiago, N. Shanaiah, D. Raftery, Metabolomics-based methods for early disease diagnostics, Expert Rev. Mol. Diagn. 8 (2008) 617–633. https://doi.org/10.1586/14737159.8.5.617.

[22] E.J. Want, I.D. Wilson, H. Gika, G. Theodoridis, R.S. Plumb, J. Shockcor, E. Holmes, J.K. Nicholson, Global metabolic profiling procedures for urine using UPLC–MS, Nat. Protoc. 5 (2010) 1005–1018. https://doi.org/10.1038/nprot.2010.50.

[23] M.A. Fernández-Peralbo, M.D. Luque de Castro, Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis, TrAC Trends Anal. Chem.

41 (2012) 75–85. https://doi.org/10.1016/j.trac.2012.08.011.

[24] F. Michopoulos, H. Gika, D. Palachanis, G. Theodoridis, I.D. Wilson, Solid phase extraction methodology for UPLC-MS based metabolic profiling of urine samples, Electrophoresis. 36 (2015) 2170–2178. https://doi.org/10.1002/elps.201500101.

[25] A. Oikawa, N. Fujita, R. Horie, K. Saito, K. Tawaraya, Solid-phase extraction for metabolomic analysis of high-salinity samples by capillary electrophoresis-mass spectrometry, J. Sep. Sci. 34 (2011) 1063–1068. https://doi.org/10.1002/jssc.201000890.

[26] F. Michopoulos, L. Lai, H. Gika, G. Theodoridis, I. Wilson, UPLC-MS-Based Analysis of Human Plasma for Metabonomics Using Solvent Precipitation or Solid Phase Extraction, J. Proteome Res. 8 (2009) 2114–2121. https://doi.org/10.1021/pr801045q.

[27] A.D. de Jager, J. V. Warner, M. Henman, W. Ferguson, A. Hall, LC–MS/MS method for the quantitation of metabolites of eight commonly-used synthetic cannabinoids in human urine – An Australian perspective, J. Chromatogr. B. 897 (2012) 22–31. https://doi.org/10.1016/j.jchromb.2012.04.002.

[28] R. Fan, R. Ramage, D. Wang, J. Zhou, J. She, Determination of ten monohydroxylated polycyclic aromatic hydrocarbons by liquid–liquid extraction and liquid chromatography/tandem mass spectrometry, Talanta. 93 (2012) 383–391. https://doi.org/10.1016/j.talanta.2012.02.059.

[29] A.J. Chetwynd, A. Abdul-Sada, E.M. Hill, Solid-Phase Extraction and Nanoflow Liquid Chromatography-Nanoelectrospray Ionization Mass Spectrometry for Improved Global Urine Metabolomics, Anal. Chem. 87 (2015) 1158–1165. https://doi.org/10.1021/ac503769q.

[30] E.E. Carlson, B.F. Cravatt, Enrichment tags for enhanced-resolution profiling of the polar metabolome, J. Am. Chem. Soc. (2007). https://doi.org/10.1021/ja0779506.

[31] E. Gorrochategui, J. Jaumot, R. Tauler, ROIMCR: A powerful analysis strategy for LC-MS metabolomic datasets, BMC Bioinformatics. (2019). https://doi.org/10.1186/s12859-019-2848-8.

[32] J. Jaumot, A. de Juan, R. Tauler, MCR-ALS GUI 2.0: New features and applications, Chemom. Intell. Lab. Syst. 140 (2015) 1–12. https://doi.org/10.1016/j.chemolab.2014.10.003.

[33] S. Piovesana, A.L. Capriotti, A. Cerrato, C. Crescenzi, G. La Barbera, A. Laganà, C.M. Montone, C. Cavaliere, Graphitized Carbon Black Enrichment and UHPLC-MS/MS Allow to Meet the Challenge of Small Chain Peptidomics in Urine, Anal. Chem. 91 (2019) 11474–11481. https://doi.org/10.1021/acs.analchem.9b03034.

[34] A. Cerrato, S.E. Aita, A.L. Capriotti, C. Cavaliere, C.M. Montone, A. Laganà, S. Piovesana, A new opening for the tricky untargeted investigation of natural and modified short peptides, Talanta. 219 (2020) 121262. https://doi.org/10.1016/j.talanta.2020.121262.

[35] R. Tauler, Multivariate curve resolution applied to second order data, Chemom. Intell. Lab. Syst. 30 (1995) 133–146. https://doi.org/10.1016/0169-7439(95)00047-X.

[36] W. Windig, D.A. Stephenson, Self-modeling mixture analysis of second-derivative near-infrared spectral data using the SIMPLISMA approach, Anal. Chem. 64 (1992) 2735–2742. https://doi.org/10.1021/ac00046a015.

[37] B.L. WELCH, THE GENERALIZATION OF ‘STUDENT’S’ PROBLEM WHEN SEVERAL DIFFERENT POPULATION VARLANCES ARE INVOLVED, Biometrika. 34 (1947) 28–35. https://doi.org/10.1093/biomet/34.1-2.28.

[38] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods. 6 (2014) 2812–2831. https://doi.org/10.1039/C3AY41907J.

[39] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemom. 17 (2003) 166–173. https://doi.org/10.1002/cem.785.

[40] M. Bevilacqua, R. Bro, Can We Trust Score Plots?, Metabolites. 10 (2020) 278. https://doi.org/10.3390/metabo10070278.

[41] B.W. Matthews, Comparison of the predicted and observed secondary structure of T4 phage lysozyme, Biochim. Biophys. Acta - Protein Struct. 405 (1975) 442–451. https://doi.org/10.1016/0005-2795(75)90109-9.

[42] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.

[43] R.J. Raterink, P.W. Lindenburg, R.J. Vreeken, R. Ramautar, T. Hankemeier, Recent developments in sample-pretreatment techniques for mass spectrometry-based metabolomics, TrAC - Trends Anal. Chem. (2014). https://doi.org/10.1016/j.trac.2014.06.003.

[44] J.S. Kirkwood, C. Maier, J.F. Stevens, Simultaneous, untargeted metabolic profiling of polar and nonpolar metabolites by LC-Q-TOF mass spectrometry, Curr. Protoc. Toxicol. (2013). https://doi.org/10.1002/0471140856.tx0439s56.

[45] A.R. Lima, M. de L. Bastos, M. Carvalho, P. Guedes de Pinho, Biomarker Discovery in Human Prostate Cancer: an Update in Metabolomics Studies, Transl. Oncol. 9 (2016) 357–370. https://doi.org/10.1016/j.tranon.2016.05.004.

[46] J. Ivanisevic, E.J. Want, From Samples to Insights into Metabolism: Uncovering Biologically Relevant Information in LC-HRMS Metabolomics Data, Metabolites. 9 (2019) 308. https://doi.org/10.3390/metabo9120308.

[47] C. Vosse, C. Wienken, C. Cadenas, H. Hayen, Separation and identification of phospholipids

by hydrophilic interaction liquid chromatography coupled to tandem high resolution mass spectrometry with focus on isomeric phosphatidylglycerol and bis(monoacylglycero)phosphate, J. Chromatogr. A. 1565 (2018) 105–113. https://doi.org/10.1016/j.chroma.2018.06.039.

[48] D. K Trivedi, H. Jones, A. Shah, R. K Iles, Development of Zwitterionic Hydrophilic Liquid Chromatography (ZIC®HILIC-MS) Metabolomics Method for Shotgun Analysis of Human Urine, J. Chromatogr. Sep. Tech. 03 (2012). https://doi.org/10.4172/2157-7064.1000144.

[49] D.-Q. Tang, L. Zou, X.-X. Yin, C.N. Ong, HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS, Mass Spectrom. Rev. 35 (2016) 574–600. https://doi.org/10.1002/mas.21445.

[50] R. Pérez-Míguez, M. Castro-Puyana, E. Sánchez-López, M. Plaza, M.L. Marina, Untargeted HILIC-MS-Based Metabolomics Approach to Evaluate Coffee Roasting Process: Contributing to an Integrated Metabolomics Multiplatform, Molecules. 25 (2020) 887. https://doi.org/10.3390/molecules25040887.

[51] N. Sillner, A. Walker, E.-M. Harrieder, P. Schmitt-Kopplin, M. Witting, Development and application of a HILIC UHPLC-MS method for polar fecal metabolome profiling, J. Chromatogr. B. 1109 (2019) 142–148. https://doi.org/10.1016/j.jchromb.2019.01.016.

[52] R. Wang, Y. Yin, Z.J. Zhu, Advancing untargeted metabolomics using data-independent acquisition mass spectrometry technology, Anal. Bioanal. Chem. 411 (2019) 4349–4357. https://doi.org/10.1007/s00216-019-01709-1.

[53] J.F. Xiao, B. Zhou, H.W. Ressom, Metabolite identification and quantitation in LC-MS/MS-based metabolomics, TrAC - Trends Anal. Chem. 32 (2012) 1–14. https://doi.org/10.1016/j.trac.2011.08.009.

[54] A.C. Schrimpe-Rutledge, S.G. Codreanu, S.D. Sherrod, J.A. McLean, Untargeted Metabolomics Strategies—Challenges and Emerging Directions, J. Am. Soc. Mass Spectrom. 27 (2016) 1897–1905. https://doi.org/10.1007/s13361-016-1469-y.

[55] M. Navarro-Reig, J. Jaumot, A. Baglai, G. Vivó-Truyols, P.J. Schoenmakers, R. Tauler, Untargeted Comprehensive Two-Dimensional Liquid Chromatography Coupled with High-Resolution Mass Spectrometry Analysis of Rice Metabolome Using Multivariate Curve Resolution, Anal. Chem. 89 (2017) 7675–7683. https://doi.org/10.1021/acs.analchem.7b01648.

[56] N. Dalmau, N. Andrieu-Abadie, R. Tauler, C. Bedia, Untargeted lipidomic analysis of primary human epidermal melanocytes acutely and chronically exposed to UV radiation, Mol. Omi. 14 (2018) 170–180. https://doi.org/10.1039/C8MO00060C.

[57] C. Bedia, M. Badia, L. Muixí, T. Levade, R. Tauler, A. Sierra, GM2-GM3 gangliosides ratio is dependent on GRP94 through down-regulation of GM2-AP cofactor in brain metastasis cells, Sci. Rep. 9 (2019) 14241. https://doi.org/10.1038/s41598-019-50761-5.

[58] M. Farrés, B. Piña, R. Tauler, Chemometric evaluation of Saccharomyces cerevisiae metabolic profiles using LC–MS, Metabolomics. 11 (2015) 210–224. https://doi.org/10.1007/s11306-014-0689-z.

[59] C. Bedia, N. Dalmau, J. Jaumot, R. Tauler, Phenotypic malignant changes and untargeted lipidomic analysis of long-term exposed prostate cancer cells to endocrine disruptors, Environ. Res. 140 (2015) 18–31. https://doi.org/10.1016/j.envres.2015.03.014.

[60] E. Ortiz-Villanueva, F. Benavente, B. Piña, V. Sanz-Nebot, R. Tauler, J. Jaumot, Knowledge integration strategies for untargeted metabolomics based on MCR-ALS analysis of CE-MS and LC-MS data, Anal. Chim. Acta. 978 (2017) 10–23.

https://doi.org/10.1016/j.aca.2017.04.049.

[61] M. Logozzi, D.F. Angelini, A. Giuliani, D. Mizzoni, R. Di Raimo, M. Maggi, A. Gentilucci, V. Marzio, S. Salciccia, G. Borsellino, L. Battistini, A. Sciarra, S. Fais, Increased Plasmatic Levels of PSA-Expressing Exosomes Distinguish Prostate Cancer Patients from Benign Prostatic Hyperplasia: A Prospective Study, Cancers (Basel). 11 (2019) 1449. https://doi.org/10.3390/cancers11101449.

[62] A. Sciarra, A. Gentilucci, S. Salciccia, F. Pierella, F. Del Bianco, V. Gentile, I. Silvestri, S. Cattarino, Prognostic value of inflammation in prostate cancer progression and response to therapeutic: a critical review, J. Inflamm. 13 (2016) 35. https://doi.org/10.1186/s12950-016-0143-2.

[63] A. Sciarra, F. Di Silverio, S. Salciccia, A.M. Autran Gomez, A. Gentilucci, V. Gentile, Inflammation and Chronic Prostatic Diseases: Evidence for a Link?, Eur. Urol. 52 (2007) 964–972. https://doi.org/10.1016/j.eururo.2007.06.038.

[64] G. Deep, R. Kumar, D.K. Nambiar, A.K. Jain, A.M. Ramteke, N.J. Serkova, C. Agarwal, R. Agarwal, Silibinin inhibits hypoxia-induced HIF-1α-mediated signaling, angiogenesis and lipogenesis in prostate cancer cells: In vitro evidence and in vivo functional imaging and metabolomics, Mol. Carcinog. 56 (2017) 833–848. https://doi.org/10.1002/mc.22537.

[65] F. Haghighatdoost, M. Jabbari, M. Hariri, The effect of L-carnitine on inflammatory mediators: a systematic review and meta-analysis of randomized clinical trials, Eur. J. Clin. Pharmacol. 75 (2019) 1037–1046. https://doi.org/10.1007/s00228-019-02666-5.

[66] R. Coras, J. Murillo-Saich, M. Guma, Circulating Pro- and Anti-Inflammatory Metabolites and Its Potential Role in Rheumatoid Arthritis Pathogenesis, Cells. 9 (2020) 827. https://doi.org/10.3390/cells9040827.

[67] M.A.B. Melone, A. Valentino, S. Margarucci, U. Galderisi, A. Giordano, G. Peluso, The carnitine system and cancer metabolic plasticity, Cell Death Dis. 9 (2018) 228. https://doi.org/10.1038/s41419-018-0313-7.

[68] S. Sasso, L. Dalmedico, D. Delwing-Dal Magro, A.T.S. Wyse, D. Delwing-de Lima, Effect of N -acetylarginine, a metabolite accumulated in hyperargininemia, on parameters of oxidative stress in rats: protective role of vitamins and L-NAME, Cell Biochem. Funct. 32 (2014) 511–519. https://doi.org/10.1002/cbf.3045.

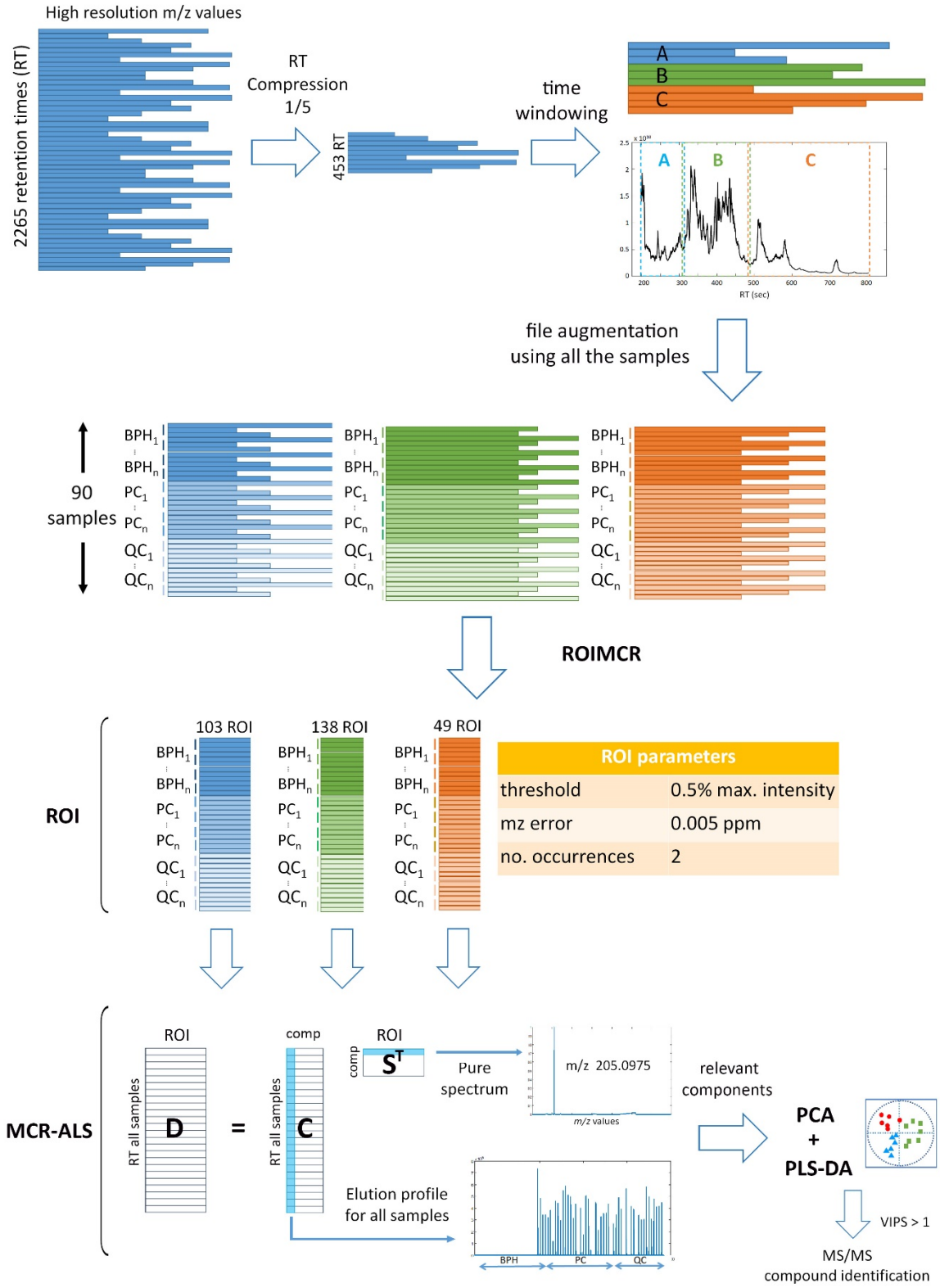
[69] L.G. Feun, M.T. Kuo, N. Savaraj, Arginine deprivation in cancer therapy, Curr. Opin. Clin. Nutr. Metab. Care. 18 (2015) 78–82. https://doi.org/10.1097/MCO.0000000000000122.

[70] K.R.K. Reddy, C. Dasari, D. Duscharla, B. Supriya, N.S. Ram, M. V. Surekha, J.M. Kumar, R. Ummanni, Dimethylarginine dimethylaminohydrolase-1 (DDAH1) is frequently upregulated in prostate cancer, and its overexpression conveys tumor growth and angiogenesis by metabolizing asymmetric dimethylarginine (ADMA), Angiogenesis. 21 (2018) 79–94. https://doi.org/10.1007/s10456-017-9587-0.

[71] D. Baci, A. Bruno, C. Cascini, M. Gallazzi, L. Mortara, F. Sessa, G. Pelosi, A. Albini, D.M. Noonan, Acetyl-L-Carnitine downregulates invasion (CXCR4/CXCL12, MMP-9) and angiogenesis (VEGF, CXCL8) pathways in prostate cancer cells: rationale for prevention and interception strategies, J. Exp. Clin. Cancer Res. 38 (2019) 464. https://doi.org/10.1186/s13046-019-1461-z.

DOI: https://doi.org/10.1016/j.aca.2021.338381

Figures



DOI: https://doi.org/10.1016/j.aca.2021.338381

Figure 1. Schematic representation of the data pre-processing and processing steps in the untargeted LC-MS analysis of the urine samples. First, the size of the raw LC-MS data from every sample was reduced taking one of every five retention times and divided into three-time windows (A, B, and C). Every one of these data windows for all 88 samples was then vertically concatenated before their simultaneous analysis by ROIMCR. In the first ROI compression step, three regular square data matrices (one for every time window, A, B, and C) were obtained. In the second step, the three ROI-compressed column-wise augmented data matrices were analyzed by the MCR-ALS method, associating the elution profiles and mass spectra of the metabolites present in the different urine and control samples. The data table of the peak areas of the elution profiles of the MCR-ALS resolved components with the most relevant changes was analyzed by PCA and PLS-DA. The most influential metabolites discriminated by PLS-DA were then identified by their MS/MS analysis and further used for the interpretation of their biological role.

DOI: https://doi.org/10.1016/j.aca.2021.338381

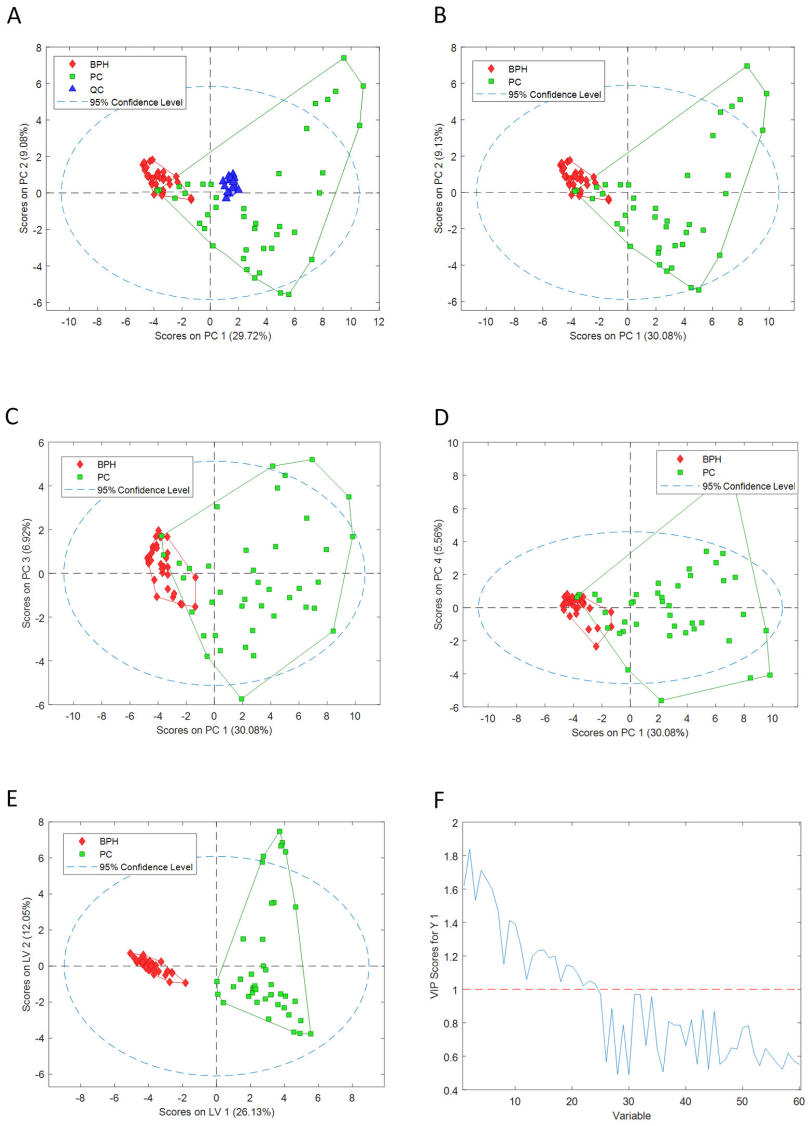


Figure 2. Unsupervised PCA and supervised PLS-DA of the peak areas of the elution profiles of the components resolved by MCR-ALS in the simultaneous analysis of all samples. A) PCA scores plot of BPH, PC, and QC samples projected in the PC1-PC2 vector space. B) PCA scores without QC samples in the PC1-PC2; C) PCA scores in the PC1-PC3 vector space; D) PCA scores in the PC1-PC4 vector space; E) PLS-DA scores plot of BPH vs PC samples, projected in the latent variables LV1 LV2 vector space; F) PLS-DA VIPS scores representation of the 60 variables used in this model.

DOI: https://doi.org/10.1016/j.aca.2021.338381

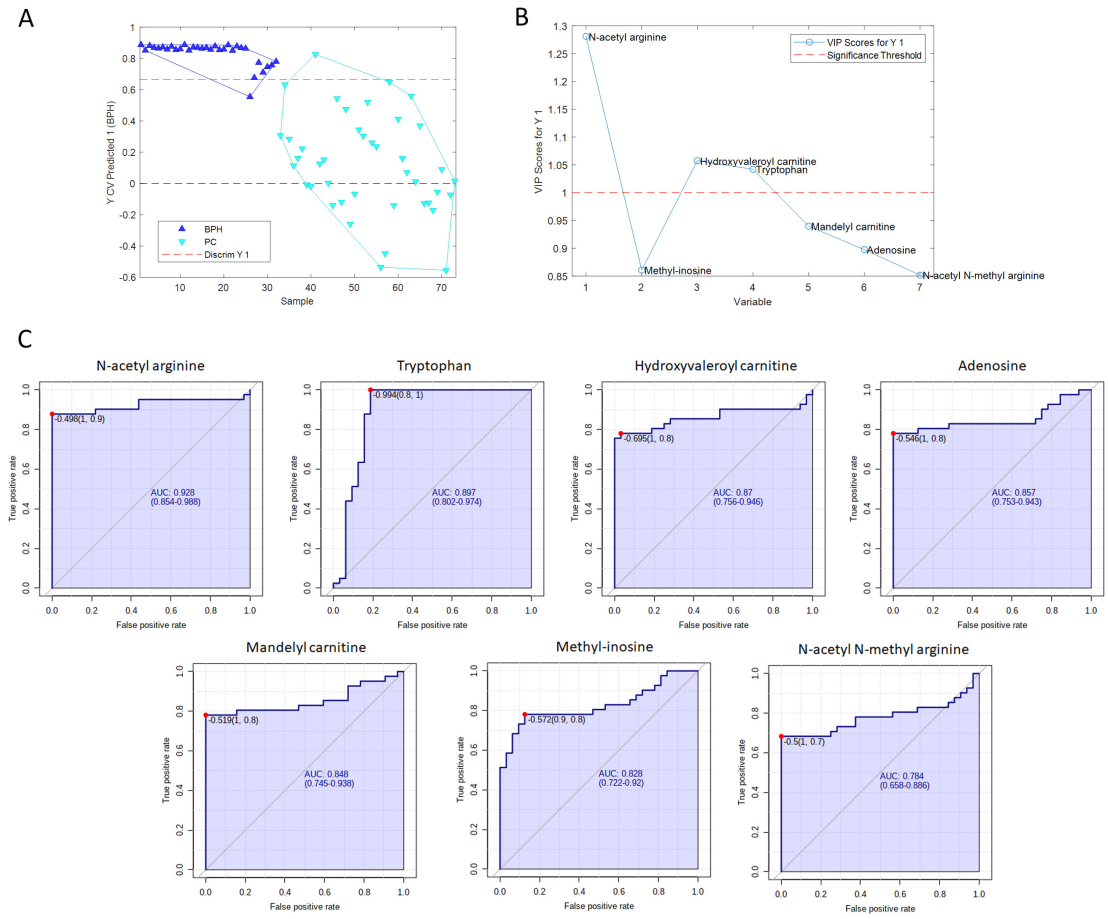


Figure 3. PLS-DA model using the 7 identified biomarkers and ROC curves. A) PLS-DA model scores plot of the LV1, B) PLS-DA VIPS scores representation of the 7 variables used in the model, C) ROC curves of the 7 identified biomarkers that discriminate PC from BPH.

DOI: https://doi.org/10.1016/j.aca.2021.338381

Tables

Table 1. Clinical and pathological characteristics of the population in the two groups (Group 1: BPH; Group 2: PC) in terms of the number of cases, age, prostatic volume, total PSA value, risk, GS, pathologic stage.

Parameter	Group 1 (BPH)	Group 2 (PC)
Number of cases	32	41
Age (years)		
mean ± SD	65.2 ± 7.3	65.9 ± 6.5
median (range)	64 (52-79)	65 (49-79)
Prostatic volume (mL)		
mean ± SD	62.4 ± 24.1	49.6 ± 15.3
median (range)	50 (33-124)	47 (30-85)
Total PSA (ng mL^{-1})		
mean ± SD	2.8 ± 2.1	7.2 ± 3.6
median (range)	2.4 (0.4-8.3)	6.7 (2.5-18)
Risk classes, n (%)	_	
Low		8 (19.5)
Intermediate		27 (65.9)
High		6 (14.6)
GS, n (%)	_	
6 (3+3)		10 (24.4)
7 (3+4)		11 (26.8)
7 (4+3)		14 (34.2)
8 – 10		6 (14.6)
Pathologic stage, n (%)	_	
T2 N0 M0		26 (63.4)
T3 N0 M0		15 (36.6)

DOI: https://doi.org/10.1016/j.aca.2021.338381

Table 2. List of 12 compounds with VIP > 1 relevant to differentiate PC from BPH patient samples. For each compound, the list gives its tentative identification, R2 (variance explained) in the MCR-ALS model, *m*/*z* value, retention time, molecular weight, mean area calculated for BPH and PC samples, the PC vs BPH area ratio, the AUCROC, and its t-test p-value.

Identification	R^2	*m/z*	Rt (min)	MW	Mean Area BPH	Mean Area PC	Ratio PC/BPH	AUC ROC curve	t-Test
Adenosine	0.00016	268.1046	3.3	267.0973	5.26E+06	3.94E+08	7.48E+01	0.853	1.8E-06
Hydroxyvaleroyl carnitine	3.84E-03	262.1653	6.5	261.1580	1.69E+06	2.64E+09	1.57E+03	0.861	5.3E-09
Mandelyl carnitine	0.01639	296.1500	5.9	295.1427	9.49E+05	5.35E+09	5.64E+03	0.853	4.6E-07
Methyl-inosine	0.00026	283.1042	3.6	282.0969	2.09E+07	3.97E+08	1.90E+01	0.827	5.5E-06
N-acetyl arginine	0.04105	217.1299	7.1	216.1227	1.00E+06	1.06E+10	1.06E+04	0.925	2.2E-14
N-acetyl N-methyl arginine	0.0012	231.1455	6.9	230.1382	2.20E+06	1.49E+09	6.77E+02	0.780	7.1E-06
Tryptophan	0.07636	205.0975	5.5	204.0902	2.74E+09	1.26E+10	4.60E+00	0.892	8.7E-10
Carnitine derivative	0.00172	346.2231	5.4	345.2158	3.73E+06	1.64E+09	4.40E+02	0.902	2.31E-06
Carnitine derivative	8.47E-05	348.2386	6.0	347.2313	5.28E+06	2.81E+08	5.32E+01	0.841	6.12E-06
Carnitine derivative	0.00026	368.1556	6.5	367.1483	8.23E+06	6.27E+08	7.62E+01	0.918	3.53E-05
unknown compound	0.00013	170.0609	5.5	169.0536	1.37E+07	2.68E+08	1.96E+01	0.866	1.88E-06
unknown compound	0.26998	202.1189	5.6	201.1117	3.09E+06	3.50E+10	1.13E+04	0.905	7.82E-18