Analytica Chimica Acta 1307 (2024) 342574

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

An untargeted analytical workflow based on Kendrick mass defect filtering reveals dysregulations in acylcarnitines in prostate cancer tissue

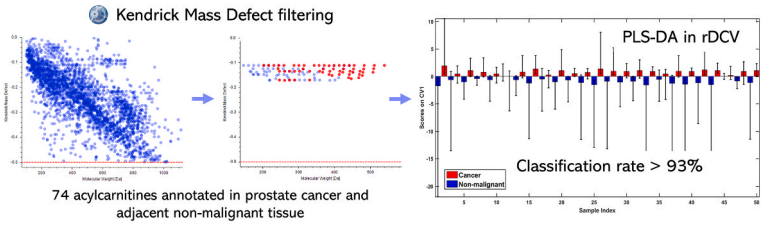
Andrea Cerrato [a], Sara Elsa Aita [a], Alessandra Biancolillo [b], Aldo Laganà [a], Federico Marini [a], Carmela Maria Montone [a], Davide Rosati [c], Stefano Salciccia [c], Alessandro Sciarra [c], Enrico Taglioni [a], Anna Laura Capriotti [a,*]

[a] Department of Chemistry, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy
[b] Department of Physical and Chemical Sciences, University of L'Aquila, Via Vetoio, 67100, L'Aquila, Coppito, Italy
[c] Department Maternal and Child and Urological Sciences, Sapienza University of Rome, Viale del Policlinico 155, 00161, Rome, Italy

HIGHLIGHTS

- Untargeted annotation of 74 ACs from PCa and non-malignant prostate tissue.
- Kendrick Mass Defect filtering allows straightforward ACs annotation.
- The ACs datasets allowed classification rates >93 % after PLS-DA in rDCV.
- Hydroxyhexanoylcarnitines were significantly upregulated in PCa tissue.
- Oxidative phenomena on short and medium ACs seem correlated to cancer metabolism.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Dr. L. Liang

Keywords:
Metabolomics
Retention time prediction
PLS-DA
Repeated double cross validation
Prostatic neoplasm
β-oxidation

ABSTRACT

Background: Metabolomics is nowadays considered one the most powerful analytical for the discovery of metabolic dysregulations associated with the insurgence of cancer, given the reprogramming of the cell metabolism to meet the bioenergetic and biosynthetic demands of the malignant cell. Notwithstanding, several challenges still exist regarding quality control, method standardization, data processing, and compound identification. Therefore, there is a need for effective and straightforward approaches for the untargeted analysis of structurally related classes of compounds, such as acylcarnitines, that have been widely investigated in prostate cancer research for their role in energy metabolism and transport and β-oxidation of fatty acids.
Results: In the present study, an innovative analytical platform was developed for the straightforward albeit comprehensive characterization of acylcarnitines based on high-resolution mass spectrometry, Kendrick mass defect filtering, and confirmation by prediction of their retention time in reversed-phase chromatography. In particular, a customized data processing workflow was set up on Compound Discoverer software to enable the Kendrick mass defect filtering, which allowed filtering out more than 90 % of the initial features resulting from the processing of 25 tumoral and adjacent non-malignant prostate tissues collected from patients undergoing radical prostatectomy. Later, a partial least square–discriminant analysis model validated by repeated double cross-validation was built on the dataset of 74 annotated acylcarnitines, with classification rates higher than 93 % for both groups, and univariate statistical analysis helped elucidate the individual role of the annotated metabolites.

* Corresponding author.
E-mail address: annalaura.capriotti@uniroma1.it (A.L. Capriotti).

https://doi.org/10.1016/j.aca.2024.342574
Received 14 December 2023; Received in revised form 25 March 2024; Accepted 1 April 2024
Available online 13 April 2024

Significance: Hydroxylation of short- and medium-chain minor acylcarnitines appeared to be a significant variable in describing tissue differences, suggesting the hypothesis that the neoplastic growth is linked to oxidation phenomena on selected metabolites and reinforcing the need for effective methods for the annotation of minor metabolites.

1. Introduction

Prostate cancer (PCa) is currently the second most diagnosed cancer worldwide and has the highest incidence in Western and high-income countries (https://www.iarc.who.int/). The development of PCa is influenced by several demographic factors, including aging, ethnicity, lifestyle, and family history [1]. If PCa is detected at early stages, the 10-year survival rate is more than 99 %; however, up to 40 % of men with PCa have no clinical signs, and the 5-year survival rate of highly metastatic patients drops to 30 % [2,3]. Currently, the early detection of PCa is based on the measure in blood of the prostate-specific antigen (PSA) [4], followed by magnetic multiparametric resonance (mMR) and targeted prostate biopsy for final diagnosis [5]. Unfortunately, PSA measurement has the significant drawback of limited specificity, causing incorrect PCa diagnoses that lead to unnecessary biopsies and over-treatment of indolent PCa [6]. As such, elevated PSA levels can be associated with benign conditions [7], including benign prostatic hyperplasia (BPH), or prostatitis, and, on the other hand, several PCa patients show low PSA levels [8]. For these reasons, there is an urgent need for alternative biomarkers to improve the detection and treatment of PCa. In this context, the omics sciences and, in particular, metabolomics have emerged in the latest years as powerful tools [3,9,10]. Given the considerable evidence of metabolic dysregulations in the development and growth of PCa [11] and differential metabolic activities of cancer cells [12], the study of the whole metabolome, i.e., the entire set of metabolites in a given biological matrix, has the potential to identify novel putative biomarkers [10]. In particular, untargeted metabolomics studies by either nuclear magnetic resonance (NMR) or high-resolution mass spectrometry (HRMS) coupled to gas- or liquid chromatography (LC) are currently the prime approaches in cancer research, allowing ideally the simultaneous analysis of the whole metabolome for highlighting any up- and downregulations [13]. Numerous studies have been conducted in recent years aimed at identifying novel putative biomarkers for PCa by means of untargeted metabolomics [3,9,10,14–17], highlighting several compound classes and metabolic pathways that were found to be dysregulated in PCa cells, such as amino acids [18–21], polyamines [22–24], tricarboxylic acid cycle metabolites [20,23], phospholipids [25], and acylcarnitines (ACs) [26–28]. Although untargeted global metabolomics has had a tremendous impact across different applications, several challenges still exist regarding quality control (QC), method standardization, computational methods for data processing, data analysis, and compound identification, which pose a much more complex challenge compared to targeted methods [29,30]. Untargeted suspect screening approaches, aimed at structurally-related classes of compounds, allow much more straightforward analytical workflows in terms of sample preparation and LC-HRMS conditions as well as for data processing and compound identification [31,32]. ACs are a broad structurally related class of metabolites that arise from the conjugation of fatty acids (FAs) with L-carnitine and play essential roles in energy metabolism, transport of long-chain FA across the mitochondrial membranes, and the β-oxidation of FAs [33]. Anti-inflammatory and antioxidant properties of carnitine and a stabilizing effect on mitochondrial membranes have been reported [34]. According to the nature of the acyl group, ACs can be classified into short-, medium-, long-, and very-long-chain, as well as saturated/unsaturated, hydroxy, dicarboxyl, and branched-chain ACs [35]. Given the large number of reported ACs, i.e., 1240 entries on the Human Metabolome Database (HMDB [36]), several papers have dealt with the setup of innovative and specific methods for a rapid and straightforward analysis of ACs. Feng et al. recently proposed a novel strategy for ACs identification in human plasma using a data-independent-acquisition-based retention time (RT) prediction modeling [37]. In 2019, an isotope labeling strategy followed by LC-MS was set up by Li et al. to identify 108 ACs in human urine [38]. Moreover, Tang et al. employed an integrated Tmt-PP derivatization-based LC-MS method for the absolute quantification of ACs in an untargeted fashion using a few analytical standards [39]. In the present paper, an untargeted HRMS approach based on Kendrick mass defect (KMD) filtering and RT prediction was set up for the identification of ACs in human PCa tissue and adjacent non-malignant tissue. KMD allows rapid visualization of compound homologs that differ only for specific molecular fragments (e.g., CH_2) [40] and has found several applications in environmental analyses [41]. The use of KMD filtering allowed straightforward data processing and identification without any derivatization or extensive sample preparation. Multivariate and univariate statistical tools were finally employed on the extracted data matrix for investigating dysregulations in the ACs linked to cancer cell metabolism.

2. Experimental section

2.1. Chemical and reagents

MS-grade absolute water, methanol (MeOH), ethanol (EtOH), acetonitrile (ACN), isopropanol (iPrOH), and formic acid (HCOOH) were purchased from Fisher Scientific (Waltham, MA, USA). Phosphate buffer saline (PBS), pH 7.4 was purchased by Sigma now Merck (Darmstadt, Germany). Acetyl-L-carnitine hydrochloride, propionyl-L-carnitine, oleoyl-L-carnitine, *trans*-2-tetradecenoyl-L-carnitine, *trans*-2-Hexadecenoyl-L-carnitine, malonyl-L-carnitine lithium salt, adipoyl-L-carnitine lithium salt, sebacoyl-L-carnitine, suberoyl-L-carnitine, 3-hydroxyisovaleryl-L-carnitine, [(3R)-3-hydroxydecanoyl]-L-carnitine, [(3R)-3-hydroxy-cis-tetradec-9-enoyl]-L-carnitine, [(3R)-3-hydroxyoctadecanoyl]-L-carnitine, [(3R)-3-hydroxy-cis-octadec-9-enoyl]-L-carnitine, acylcarnitines mix 2 (C4, iC4, C5, iC5, C6, C8, C10, C12, C14, C18) solution were purchased by Merck. ^{13}C caffeine (internal standard, IS) was also purchased by Merck. ACs standard stock solutions were prepared at 1 mg mL^{-1} in either H_2O or MeOH/H_2O 50:50 (*v/v*). The ACs working mix solution was prepared by appropriate dilution at 0.1 μg mL^{-1} in H_2O/ACN 90:10 (*v/v*). The IS solution was prepared at 5 μg mL^{-1} in H_2O/ACN 90:10 (*v/v*).

2.2. Population and sample collection

This is an experimental research study in which patients have been treated following normal clinical practice and international guidelines. The study was developed following the current Good Clinical Practice and the principles of the Declaration of Helsinki. The study was approved by our Ethic Committee of Sapienza University (Protocol 0919/2021–6511) and all patients provided informed consent. The population was represented by 25 consecutive unselected clinically significant PCa patients referred to the Department of Urology, Policlinico Umberto I and Sapienza University of Rome, for to RARP procedure. The inclusion criteria were: a histologic diagnosis of prostatic adenocarcinoma at biopsy; clinically significant, intermediate, or high-risk PCa according to the European Urological Association (EAU) classification; indication for RARP procedure according to the EAU guidelines. In all cases the following exclusion criteria were respected: no clinical evidence of metastatic disease; no previous or current surgical,

radiotherapeutic, hormone, or chemotherapeutic treatments for PCa; no concomitant other neoplastic diseases; no concomitant medical therapies that potentially influenced prostatic metabolism and growth; no concomitant inflammatory or metabolic disorders. Clinical and pathological characteristics of the population are summarized in Table 1. Immediately after removal of the entire prostate, a sample of non-malignant (BPH) prostatic tissue and a sample of neoplastic tissue were obtained in each patient. Each tissue sample measured approximately 5 mm in diameter and 1 g in weight. The site of sampling in the prostate was evaluated based on the localization performed on the previous mMR and the diagnosis of non-malignant BPH and prostate adenocarcinoma tissue was confirmed by pathologic examination in each sampling. Final pathologic examination with the definition of PCa pathologic staging, ISUP grading, and surgical margins were reported in all patients. All 50 tissue samples (1 non-malignant sample and 1 PCa sample for each patient) were immediately stored at −80 °C until metabolomics analysis.

2.3. Acylcarnitine extraction

Tissue samples were thawed at room temperature for 30 min, hashed with a scalpel, put in a LoBind tube (Eppendorf, Hamburg, Germany), and weighed. Subsequently, 1.5 mL of EtOH/H_2O 70:30 (*v/v*) with 10 mM PBS was added to each sample, which was sequentially vortexed for 10 min, sonicated for 5 min, and vortexed again for 10 min at 4 °C. The extraction mixture was then centrifuged at 14,000×*g* and 4 °C for 5 min. The supernatant was transferred to a test tube and the extraction was repeated once. The reunited supernatants were then dried out using a Speed-Vac SC 250 Express (Thermo 164 Avant, Holbrook, NY, USA). Subsequently, 500 μL of H_2O/ACN 90:10 (*v/v*) were added, and the extracts were filtered using Acrodisc™ MS syringe filters, 0.2 μm, 13 mm, wwPTFE (Pall Corporation, Port Washington, NY, USA) after vortexing and sonicating for 5 min. For each sample, 145 μL of the extract was transferred into an injection vial and 5 μL of the IS solution was added. Moreover, 50 μL of each extract was collected and mixed to obtain a matrix-matched pooled QC sample. Process blank samples were obtained following the described extraction procedure on a solvent sample.

Table 1
Clinical and pathological characteristics of the population in terms of the number of cases, age, body mass index (SMI), prostatic volume, total PSA value, European Association of Europe (EAU) risk, pathologic stage, pathological International Society of Urological Pathology (ISUP) grading, surgical margins, and perineural invasion (PNI).

Parameter	Number of cases (%) or mean ± SD (median) and range
Number of cases	25
Age (years)	63.92 ± 6.82 (64) 44-74
BMI	22.42 ± 2.90 (21.9) 17.40–29.06
Prostate volume (mL)	44.52 ± 11.70 (45) 25-68
PSA total ng mL^{-1}	25.40 ± 10.92 (25.4) 2.5–40.0
Tumor max diameter (mm)	12.76 ± 3.64 (11) 10-20
EAU Risk class	
-Intermediate	11 (44.0)
-High	14 (56.0)
Pathologic stage	
-pT2N0	13 (52.0)
-pT3aN0	10 (40.0)
-pT3bN0	2 (8.0)
Pathologic ISUP grading	
−2	18 (72.0)
−3	3 (12.0)
−4	2 (8.0)
−5	2 (8.0)
Surgical margins	
- negative	21 (84.0)
-positive	4 (16.0)
Perineural invasion (PNI)	
-no	12 (48.0)
-yes	13 (52.0)

2.4. UHPLC-HRMS and RT regression models

A Vanquish Binary Pump H system, equipped with a thermostated autosampler (4 °C) and column compartment (40 °C), was interfaced to a hybrid quadrupole-Orbitrap Q Exactive mass analyzer (Thermo Fisher Scientific, Bremen, Germany) with a heated electrospray ionization (HESI) source. The separation was achieved on a reversed-phase (RP) Accucore™ C8 column (150 × 2.1 mm I.D., 2.6 μm particle size, Thermo Fisher Scientific) using H_2O 0.1 % HCOOH (phase A) and ACN/iPrOH 80:20 (*v/v*) 0.1 % HCOOH at a constant flow rate of 0.4 mL min^{-1}. The gradient was as follows: 1 % phase B to 40 % phase B in 7 min; 40 % phase B to 99 % phase B in 21 min; 99 % phase B for 5 min (washing step); 99 % phase B to 1 % phase B in 2 min; 1 % phase B for 7 min (re-equilibration step). Samples were analyzed in the positive ion mode (ESI+) with the following HESI parameters: spray voltage at 3200 V, auxiliary gas heater temperature at 280 °C, sheath gas at 50 (arbitrary units), auxiliary gas at 25 (arbitrary units), sweep gas at 0 (arbitrary units), and S-Lens RF level was 50 (%). Samples and control were run in full-scan mode to ensure a sufficient number of points per peak for the precise measurement of the peak areas in the range 100–1000 *m/z* with a resolution of 70,000 (full width at half-maximum, FWHM, *m/z* 200). The automatic gain control (AGC) target was set at 1,000,000, with a maximum injection time of 50 ms. For the *identification-only* QC runs, the top 5 data-dependent acquisition (DDA) mode was performed at a resolution of 17,500 (FWHM, *m/z* 200), the AGC target at 100,000, maximum injection time at 50 ms, isolation window at 2.0 *m/z*, and normalized collision energy at 30. RT linear regression models were built by plotting the experimental RT of the 23 standard ACs *vs.* their calculated logP (ClogP, calculated using ChemDraw 14.0). Based on the retention behavior of the different ACs subclasses, three linear regression models were built: non-functionalized (acylated) ACs (R^2 = 0.9983), hydroxy ACs (R^2 = 0.9969), and dicarboxyl ACs (R^2 = 0.9903) (Fig. S1).

2.5. Acylcarnitine untargeted analysis

Untargeted data acquisition was performed following the recommendations of the metabolomics Quality Assurance and Quality Control Consortium (mQACC) [42]. Samples were injected in a randomized order and the chromatographic worklist is schematized in Supplementary Material Table S1. For system suitability testing, the column stability and performance were tested before and after each analytical section using solvent blank samples (H_2O/ACN, 90:10, *v/v*) and working mix standard solutions. System conditioning, consisting of ten consecutive pooled QC sample injections, preceded the process blank sample injection for background subtraction, which allowed to discard of the contaminants originating from the extraction solvents, mobile phases, and the HPLC-MS system, as well as the compounds subjected to high carry-over effects (more than 10 %), which alter peak areas and can result in biased statistical analysis. After further system reconditioning with ten more QC samples, randomized samples were run in groups of five, followed by a QC injection. Raw MS and MS/MS data were acquired by Xcalibur software (version 3.1, Thermo Fisher Scientific).

2.6. Data preprocessing and compound identification

The.raw data files obtained by the analysis of samples, controls, QCs, identification-only QCs, and the process blank were preprocessed using Compound Discoverer (Thermo Fisher Scientific) using a homemade workflow that was specifically designed for ACs. Feature alignment was obtained by the adaptive curve regression model; whenever the adaptive curve model failed, the linear model was automatically selected instead.

Features were aligned and filtered to remove the features whose areas in the process blank were at 10 % of the areas in the QCs, as they were attributed to either contaminants or carry-over artifacts. The tool "calculate mass defect" was enabled to automatically calculate the KMD of the extracted features to match the experimental KMD with the calculated ones of the ACs classes reported on HMDB: 0.111 (saturated ACs), −0.124 (unsaturated ACs, 1 double bond), −0.138 (unsaturated ACs, 2 double bonds), −0.151 (unsaturated ACs, 3 double bonds), −0.165 (unsaturated ACs, 4 double bonds), −0.134 (hydroxyl saturated ACs), −0.147 (hydroxyl unsaturated ACs, 1 double bond), −0.161 (hydroxyl unsaturated ACs, 2 double bonds), and −0.170 (dicarboxyl ACs). Later, features whose calculated KMD was different from the aforementioned ones were filtered out, and the remaining compounds were manually annotated based on the MS/MS spectra and RT, with 71 tentatively identified ACs. Furthermore, features whose areas in the QCs had a standard deviation higher than 25 % were also filtered out. QC-based normalization of the features was carried out based on the peak area variations over the time of the acquisition due to different instrumental fluctuations. For each feature, the peak area correction over time was performed by building a linear regression of the areas in the QC samples over time and eventually correcting each linear regression so that the slope is zero. Finally, the areas in each sample and control were corrected based on the time of data acquisition. The normalization filter resulted in the removal of 12 out of the 74 ACs and the obtention of a data matrix that was later employed for statistical analysis.

2.7. Chemometric strategies for data processing

All the chemometric calculations were run in MATLAB (R2015b; The Mathworks, Natick, MA) through custom routines developed in-house.

2.7.1. Classification: Partial Least Squares Discriminant Analysis

Partial Least Squares Discriminant Analysis (PLS-DA) [43] is a discriminant classification method that takes advantage of a regression approach (PLS) to address classification problems. This method is made feasible by the employment of a *dummy* response matrix $\boldsymbol{Y}$, which encodes the information about the class-belonging of the investigated samples [44]. In the context of a two-class problem, such as the one examined in this study, where discrimination between cases and controls is achieved, the Y-dummy is a binary vector $\mathbf{y}$ of dimensions N × 1 (with N equal to the number of analyzed samples). The true class of each sample is coded either as one for the cases, $y_{case} = 1$, or zero for the controls by $y_{control} = 0$ [44].

After having suitably coded the response $\mathbf{y}$, a PLS-DA calibration model is established by solving Eq. (1):

$$\mathbf{y} = \boldsymbol{Xb} + \boldsymbol{e} \quad (1)$$

This enables the estimation of the vectors of regression coefficients and residuals ($\boldsymbol{b}$ and $\boldsymbol{e}$, respectively). Then, since the predicted response is real-valued and not binary as its target values a classification rule is essential to assign the samples to one of the investigated classes. In this study, this goal is achieved through the approach proposed by Perez et al. [45].

Eventually, Variance Importance in Projection (VIP) analysis [46] was employed to identify the most significant analytes from a classification standpoint. This strategy allows associating each variable with an index that ranks its contribution to the solution of the problem.

2.7.2. Validation

Validation of the model's predictive ability and the consistency and reliability of the identified candidate markers was carried out through a Repeated Double Cross-Validation (rDCV) procedure [47]. This involves 8 cancellation groups in the inner and 10 in the outer loops, with 50 runs. For each outer loop cancelation group, the optimal model in terms of the number of latent variables (LVs) is selected based on the mean classification error on the inner loop samples. Moreover, to rule out the possibility of chance correlations and overoptimistic results, the classification figures of merit are compared to their null distributions (and corresponding p-values are calculated) non-parametrically estimated by permutation tests [48] with 1000 randomizations.

2.7.3. Univariate statistical analysis

Metaboanalyst 5.0 was employed for univariate statistical analysis on the ACs data matrix [49]. The data matrix was submitted as a text file that was prepared according to the furnished by the developers. The interquartile range (IQR) was selected for data filtering, whereas the autoscaling algorithm was selected for data scaling. A volcano plot analysis was performed to evaluate the individual contribution of the annotated ACs to discriminate between the two sets of samples. Moreover, a correlation heatmap was obtained to display the correlation among selected ACs based on the trends in tumor and adjacent non-malignant tissue samples.

3. Results and discussion

3.1. Acylcarnitine separation and retention time prediction

Aiming at setting up a method for the separation of ACs, three main goals were set: (i) obtaining the retention and separation of ACs in a wide range of polarity, (ii) allowing the discrimination of positional isomers (linear vs branched chain ACs), and (iii) obtaining a linear regression between the hydrophobicity of ACs and their RT. The separation of ACs and their RT prediction has been achieved by RP separation systems given that the acyl and fatty acyl chains are responsible for differentiating the ACs subclasses [37,50]. It is known, in fact, that RP separates lipids based on their fatty acyl chains, whereas hydrophilic interaction liquid chromatography (HILIC) separates lipids according to their polar heads [51]. Preliminary results on an Acquity UPLC® BEH Amide HILIC column (100 × 2.1 mm, 1.7 μm I.D., Waters, Milford, MA, USA) confirmed the need for RP by showing no separation between linear and branched-chain AC 4:0 and 5:0 (Fig. S2). Subsequently, four different RP columns with different characteristics were compared: fully-porous Luna Omega Polar C18, solid-core Accucore™ C30 and C8, and Kinetex® F5. For all columns, the ACs standard mix was separated using a linear gradient (details are reported in the Supplementary Material). The F5 column allowed the maximum retention of the most polar ACs, but, at the same time, showed the poorest baseline separation of the isomeric butyryl- and isobutyrylcarnitine (AC 4:0, Fig. S3d) and valeryl- and isovaleryl-carnitine (AC 5:0, Fig. S4d). The fully-porous C18 column, which is specifically designed for polar compounds, did not allow the baseline separation of the AC 4:0 isomers (Fig. S3a) as well as showing no retention of acetylcarnitine and poor peak shape of propionylcarnitine (Figure S5). As expected, the Accucore™ C8 gave similar performance to the C30 in terms of the isomer separation but allowed higher retention of the polar species possibly due to stronger interactions with the silica gel due to the lower carbon load of the C8 column compared to the C30 (Figs. S3b–S3c). The experimental RT of the standard ACs were then plotted against their ClogP to evaluate the trends in their RT. As shown in Fig. S6, the four tested columns furnished good linearity, with the R^2 values ranging from 0.9863 (C30) and 0.9879 (F5). It was then noticed that if non-functionalized ACs, hydroxy ACs, and dicarboxyl ACs were plotted separately the R^2 values exceeded 0.99 (Figs. S7–S10). As demonstrated by Yu [50], subclasses of ACs are more efficiently plotted against ClogP, thus allowing better RT predictions. Based on the previous considerations, the Accucore™ C8 was then selected for further gradient optimization and the final set-up of the LC-HRMS method. As such, based on the general trends observed in Fig. S6b, in which the experimental RT of short-chain ACs (C2–C6) were consistently slightly lower than predicted and those of longer ACs (C8–C18) were often higher than predicted, a two-slope gradient was set up. In the linear gradient, the percentage of phase B reached around 25 % in 7 min. Therefore, three different gradients were tested on the

Accucore™ C8 column, reaching 30 %, 40 %, and 50 % phase B in 7 min, respectively, the second of which gave the best results in terms of R^2 (Fig. S1). The final conditions were subsequently employed for metabolomics analysis of prostate tissue extracts following the guidelines of the mQACC.

3.2. Data processing and acylcarnitine identification

In this work, fifty prostate tissue samples (tumor and non-malignant BPH adjacent tissue) from 25 PCa patients undergoing robotic radical prostatectomy (RARP) were analyzed using untargeted MS-based metabolomics. Data preprocessing represents a critical step in metabolomics workflows, as it allows the obtention of a data matrix from the acquired raw data files through a series of algorithms that can greatly affect the outcome of the data analysis [52,53]. A major advantage of proper data preprocessing is the possibility of filtering out most of the unwanted features, including contaminations, polymers, redundancies, and features suffering from carry-over or other issues that hinder their proper peak area measurement [54]. The removal of undesired mass spectral features has been obtained by Kendrick Mass Filter [55], which groups compound homologs that differ from a specific repeating unit. Whenever the compounds of interest are characterized by specific repeating units, the KMD can be employed to filter out all other compounds to simplify the data [56,57]. Aiming at setting up an efficient data processing method for ACs, an untargeted metabolomics workflow was customized on Compound Discoverer for the purpose. In particular, three different strategies were tested and compared. The first workflow (W1, Fig. S11) was a standard untargeted metabolomics workflow with QC-based normalization and gap filling. In the second approach (W2, Fig. S12), the tool "Filter by Mass Defect" was enabled in the filter direction "keep" by inserting the composition of the five most abundant subclasses of ACs. Finally, the tool "Calculate Mass Defect" was enabled in the third method (W3, Fig. S13). The latter methods differ significantly based on the order of the algorithms employed by the software. Whereas in W2 the mass defect filter operates right after the alignment step and before all the other algorithms, including detection and grouping of MS adduct, prediction of the molecular formulas, gap filling, area normalization, and database search, the customized tool in W3 operates at the end of data processing and the KMD filters must be manually enabled by the user. As a result, W2 significantly reduced the processing time, since most features were automatically discarded before the application of most algorithms. The main drawback of W2 is represented by the possibility of inserting only a maximum of five compositions, thus reducing the number of ACs subclasses that can be analyzed at once. In Fig. 1, the Kendrick diagrams resulting from the three workflows are shown. The standard approach resulted in 4608 features after the background removal, i.e., the features that were present in the processing blank with a peak area that was higher than 10 % of the average area in the QC samples (Fig. 1a). Conversely, W2 and W3 resulted in 301 and 154 features, respectively (Fig. 1b–c), a higher than

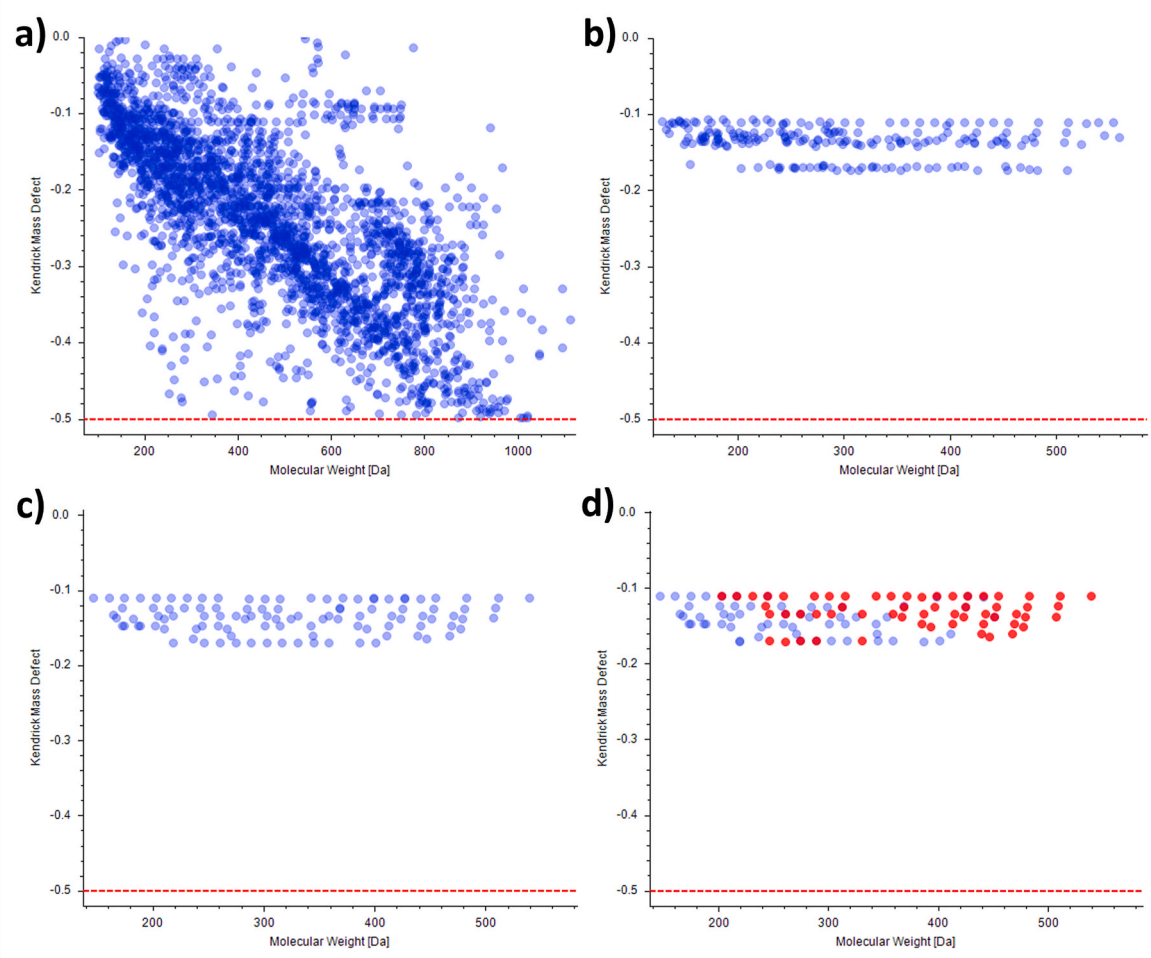


Fig. 1. Kendrick diagrams obtained after data processing of the metabolomics raw data on Compound Discoverer using a standard metabolomics workflow (W1, a), the workflow employing the "Filter by Mass Defect" tool (W2, b), the workflow employing the "Calculate Mass Defect" tool (W3, c), and W3 after the identification of ACs (manually annotated ACs are marked in red, d). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

90 % decrease compared to M1. Once the list of putative features was obtained from M2 and M3, manual MS/MS spectra inspection was performed to tentatively identify the ACs. The MS/MS spectra of these compounds usually contain three main peaks, i.e., $C_4H_5O_2^+$ (*m/z* 85.0284), trimethylamine ion ($C_3H_{10}N^+$, *m/z* 60.0808), and trimethylamine neutral loss from the protonated precursor (C_3H_9N, mass 59.0735). A fourth and less abundant peak is represented by dehydrated carnitine arising from the neutral loss of the fatty acyl moiety (*m/z* 144.1019). Moreover, hydroxylated ACs display a diagnostic peak that depends on the site of hydroxylation (*m/z* 145.0495 in the common third position of the acyl chain) [58]. Based on the structural properties of ACs, the MS/MS spectra, albeit diagnostic, fail to furnish information on the acyl chains.

The combination of KMD filtering, MS/MS spectra inspection, and RT confirmation was therefore needed to corroborate the untargeted identification of ACs. A total of 63 and 74 ACs (Fig. 1d) were annotated following W2 and W3, respectively. Not unexpectedly, W2 resulted in a lower number of identifications since only 5 subclasses could be included in the method. A possible solution to this limitation could be the repetition of the pre-processing workflow by including other subclasses. Considering the lower number of filtered features, however, W3 was still preferred, and the "Filter by Mass Defect" tool would probably be the best option for classes of compounds that show a more limited range of variations. Table S2 reports the list of the 74 annotated ACs following W3, comprising their molecular formula, ClogP, experimental and predicted RT (with the related ΔRT), experimental and calculated *m/z* (with the related ΔMass), KMD, HMDB entry, main diagnostic product ions, and confidence level according to Schymansky et al. [59]. The ClogP of the annotated ACs ranged from −7.24 (malonylcarnitine) and 5.86 (hexacosylcarnitine), demonstrating the wide range of polarity that could be simultaneously separated by the LC method. As regards the ΔRT, a maximum tolerance of 10 % was considered, with 6 exceptions. Acetylcarnitine (ΔRT 14.4 %) was identified by matching with the analytical standard, and the relatively high error is surely due to the very low absolute value of its RT (1 min), considering the injection peak at 0.8 min which corresponds to a retention factor of 1.25. Low-molecular-weight dicarboxyl ACs exhibited the highest ΔRT. Among these three compounds, however, malonylcarnitine (ΔRT -19.7 %) was again identified by matching with the analytical standards. Similar to acetylcarnitine, these compounds present low absolute values of RT, and their RT prediction is also affected by a lower value of R^2 (0.9903 vs 0.9984). Luckily, dicarboxyl ACs present typical fragmentation patterns [58] that permitted the identification of succinyl- and methylmalonylcarnitine. Finally, two of the three isomers of 3-hydroxybutyrylcarnitine resulted in ΔRT higher than 10 %. In this particular case, since HMDB reports a single compound corresponding to this molecular formula (3-hydroxybutytylcarnitine, HMDB0013127), the ΔRT is due to the attribution of the same ClogP to the three isomers. It is important to point out that 3-hydroxyisobutyrylcarnitine has been previously reported in the literature [60], and its ClogP (−7.34) would result in a much lower ΔRT (4.3 %). Following the manual spectra annotation, the identified ACs were filtered to remove those that had a higher-than-25 % standard deviation in the QC samples over time and that could be normalized. The filtered data matrix of 62 annotated and normalized ACs was then subject to statistical analysis.

3.3. Chemometric evaluation of the dysregulation of ACs associated with PCa

PCa cells often use FAs metabolism as the main source of energy and the *de novo* lipid biogenesis and β-oxidation are the most altered pathways in PCa metabolism, thus generating dysregulations in the ACs system [61]. In the present study, a series of PLS-DA models were built on the ACs data after autoscaling and validated through a rDCV procedure. On average, the prediction ability on the rDCV outer loop samples, which mimic an external test set was found to be 93.8 ± 4.6 % and 93.3 ± 4.2 % for cases and controls, respectively. In Fig. 2a, the average scores of the outer loop samples along the only canonical variate of the models are shown together with their confidence intervals. In the figure, red and blue bars represent cases and controls, respectively, whereas black whiskers delimit the confidence intervals. The plot reveals a distinct pattern where cases exhibit positive values on the CV, while controls have negative scores. On the other hand, the weights of the analytes on the canonical variate are shown in Fig. 2b. By comparing Fig. 2a and b it can be affirmed that the variables with a positive weight are present at higher concentrations in cancer tissues, while in non-malignant specimens there is a higher amount of those with a negative weight. It is, therefore, possible to affirm that some compounds, such as malonylcarnitine (AC 3:1; O2), were present at higher concentrations in cancer tissues, whereas others, such as decanoylcarnitine (AC 10:0), were more abundant in the non-malignant samples.

To have a more refined interpretation of the variables characteristic of the discrimination between the two groups of samples, VIP analysis was employed to rank the most significant features, shedding light on the crucial factors that contributed to the model's efficacy. This has revealed that there were a number of variables that significantly contributed to the discrimination between cancer and non-malignant tissues, i.e., propionylcarnitine, 2-methylbutyroylcarnitine (AC 5:0_1), isovalerylcarnitine (AC 5:0_2), malonylcarnitine, 3-hydroxyhexanoylcarnitine (AC 3-OH 6:0_1), 3-hydroxyisohexanoylcarnitine (AC 3-OH 6:0_2), decanoylcarnitine, tetradecadienoylcarnitine (AC 14:2), 3-hydroxyhexadecanoylcarnitine (AC 3-OH 16:0), octadecadienoylcarnitine (AC 18:2), nonadecanoylcarnitine (AC 19:0), eicosatetraenoylcarnitine (AC 20:4), eicosatrienoylcarnitine (AC 20:3), and lignoceroylcarnitine (AC 24:0), confirming the indication obtained by the inspection of the CV1 weights. The interpretation of the multivariate results based on the AC subclasses is not straightforward, but some interesting trends can be observed by the CV1 weights shown in Fig. 2b. A closer look at the data for short- and medium-chain ACs (C2–C12), highlights generally negative values of CV1 for non-oxidized compounds, such as acetylcarnitine, isobutyrylcarnitine, isovalerylcarnitine, hexanoylcarnitine (AC 6:0), and decanoylcarnitine, with the notable exception of propionylcarnitine. On the other hand, several oxidized ACs display positive values, including malonylcarnitine, 3-hydroxybutyrylcarnitine (AC 3-OH 4:0_2), 3-hydroxyhexanoylcarnitine, and 3-hydroxyisohexanoylcarnitine. Interestingly, most of these compounds are among the variables that were selected by VIP analysis. A previous study by Albini et al. found that some medium-chain ACs, such as decanoylcarnitine and octanoylcarnitine, were significantly decreased in serum samples from PCa patients when compared to those from BPH cases, suggesting a protective role of these ACs against PCa progression through an anti-angiogenetic and angiopreventive effect in two microenvironment settings: hypoxia and inflammation [62]. Moreover, the higher abundances of propionylcarnitine in cancer tissues were in agreement with a case-control study on the progression of PCa [61]. In opposition to short ACs, long- and very long-chain ACs had often positive CV1 values, such as oleylcarnitine (AC 18:1_1), nonadecanoylcarnitine, and eicosatrienoylcarnitine. Previous studies have highlighted opposite trends of short- and long-chain ACs in association with tumors. In the case of hepatocellular carcinoma, evidence suggested a significant decrease in the serum levels of short- and medium-chain ACs compared to healthy controls, while measuring a simultaneous increase in long-chain ACs [33,63]. A multi-center case--control study on 1077 PCa cases and 1077 controls showed that plasmatic oleylcarnitine and propionylcarnitine had an opposite association with the risk of overall PCa [64]. In correlation with opposite trends for short- and long-chain ACs, the ratio (AC 16:0 + 18:1)/AC 2:0 is an index of the activity of carnitine palmitoyltransferase II, an enzyme that oxidizes long-chain FAs in the mitochondria and whose alteration has been associated with tumors [65]. The area under receiver operative characteristic (AUC) curve, a performance measurement tool for the

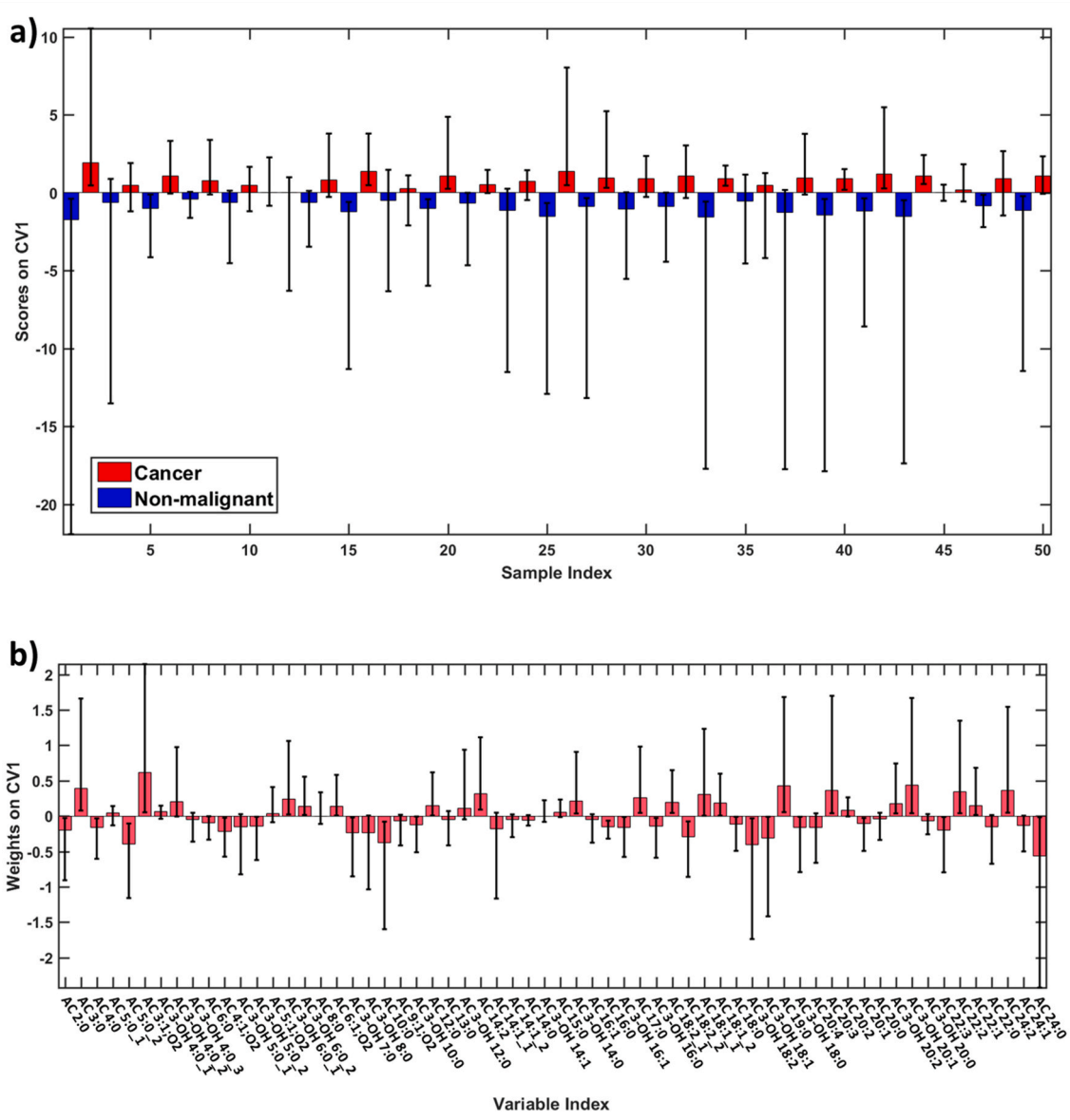


Fig. 2. (a) Mean scores of the outer loop samples along the only canonical variate of the model, together with their 95 % confidence intervals (black whiskers). Blue bars: Non-Malignant; Red bars: Cancer; (b) Weights of the measured variables for the definition of the only canonical variate of the model (pink bars) together with their 95 % confidence intervals (black whiskers). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

classification, however, demonstrated that the index was not affected (Fig. S14), and other compounds or parameters must be taken into account. To evaluate the individual contribution of the investigated ACs, univariate statistical analysis on the data matrix was performed by MetaboAnalyst. Fig. 3a displays the results of the volcano plot analysis, a scatter plot that plots significance (p-values) versus the fold-change of each of the compounds in the ACs data matrix. Among the investigated ACs, the two isomers of 3-hydroxyhexanoylcarnitine stood out among the others. These two metabolites were among the ones highlighted by the multivariate model that had consistently higher abundances in the cancer tissue samples, as shown in the exemplary box and whiskers plot in Fig. 3b for 3-hydroxyhexanoylcarnitine (AUC were around 0.65 as shown in Figs. S15–16). The correlation heatmap is a graphical tool that displays the correlation between multiple variables (on a scale from −1, negative, to 1, each variable with itself) and is therefore useful to visualize the existence of clusters of metabolites whose variability is correlated [66].

The Pearson correlation heatmap in Fig. 3c shows distinct clusters of ACs with significant correlation. The cluster with the highest Pearson correlation values is visible right at the center of the heatmap and is constituted by hydroxylated and non-hydroxylated long-chain ACs (C18 and C20) with up to two unsaturation. This cluster had also a good correlation with two other clusters of longer- (C22–C24) and shorter-chain (C14–C17) ACs. The latter cluster had also positive Pearson correlation values with a large cluster of compounds that is visible at the bottom right corner of the diagram and is constituted by medium- and long-chain (C6–C16) ACs, as well as two highly unsaturated compounds (AC 20:3 and 20:4) and propionylcarnitine. In contrast, at the top left corner of the heatmap, several ACs with few highly positive Pearson correlation values are shown, including acetylcarnitine and the dicarboxyl ACs. In this area of the diagram, there are also pairs of unsurprisingly highly correlated isomers (AC 5:0 and AC 3-OH 5:0), as well a small but interesting cluster of five short-chain hydroxylated ACs, i.e., the two isomers of 3- hydroxyhexanoylcarnitine and the three isomers of 3-hydroxybutyrylcarnitine. The existence of such a cluster supports the previously stated hypothesis of an increase in the abundance of short-

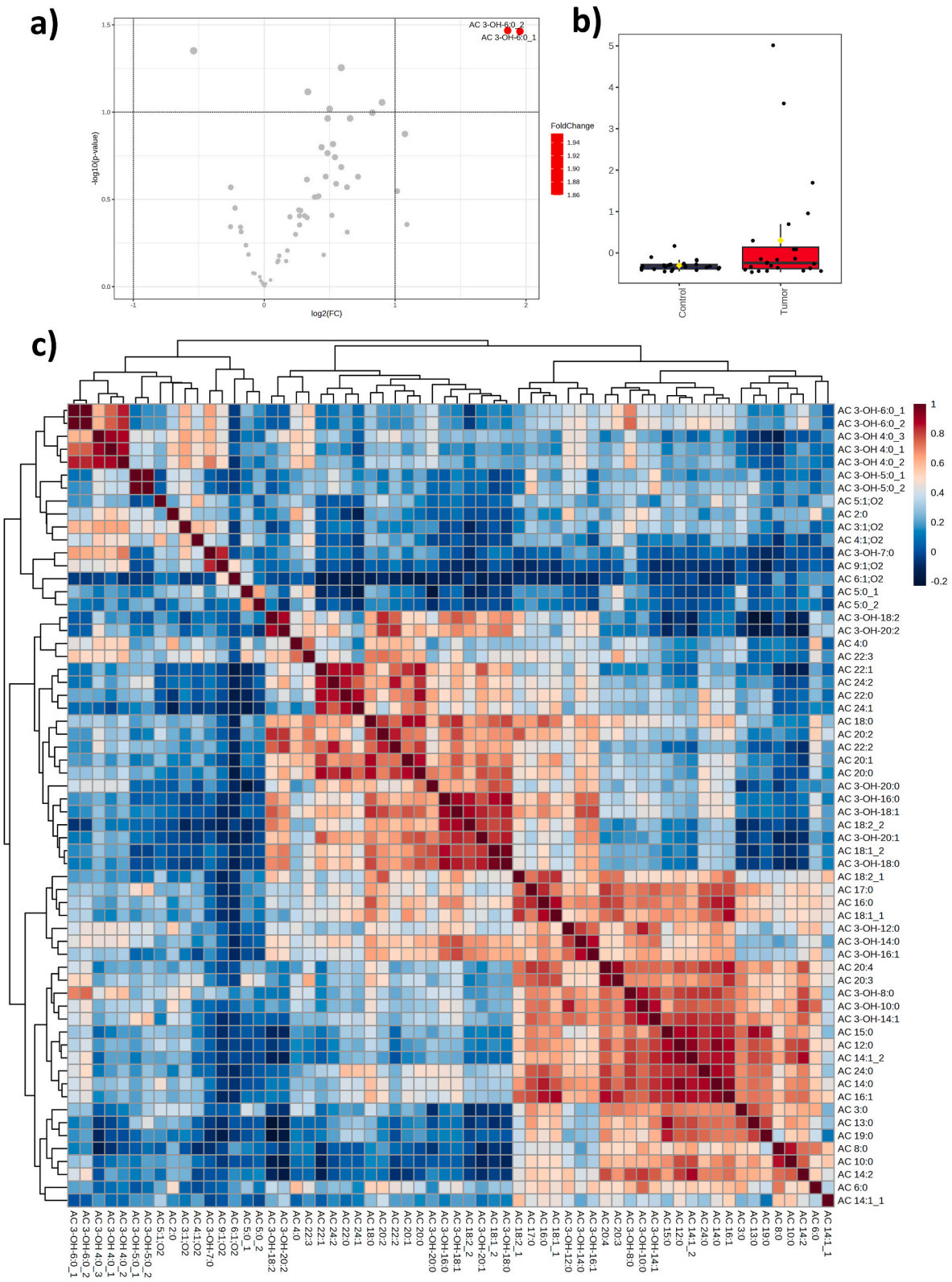


Fig. 3. Volcano plot (significance vs fold-change) analysis of the ACs data matrix from tumor and adjacent non-malignant tissue samples (a); box and whiskers plot of 3-hydroxyhexanoylcarnitine in the two sets of samples (b); correlation heatmap of the annotated ACs based on their trends in tumor and adjacent non-malignant tissue samples.

chain hydroxylated ACs and a contemporary decrease in their non-hydroxylated counterparts based on the CV1 values resulting from the PLS-DA multivariate model. Interestingly, a previous study by Puhka [27] reported significantly lower abundances of isobutyrylcarnitine (AC 4:0) in the urines of patients before surgical removal of prostate cancer compared to their urine after the operation and healthy controls. Moreover, Ren [26] measured a significant increase in AC OH-4:0 in prostate cancer tissue compared to the adjacent non-malignant tissue.

By considering the list of ACs with Pearson correlation values higher than 0.5 with the two isomers of 3-hydroxyhexanoylcarnitine selected by the volcano plot, additional confirmation of the hypothesis is obtained. Other than the three isomers of 3-hydroxybutyrylcarnitine (0.61–0.84), there were only three other compounds, i.e., 3-hydroxyoctanoylcarnitine (0.69–0.74), 3-hydroxyheptanoylcarnitine (0.62–0.65), and malonylcarnitine (0.57–0.60). By comparison, the Pearson correlation values of hydroxylated ACs with 10–16 carbon atoms were around 0.3–0.4 and fell around zero for hydroxylated longer-chain ACs. The insurgence of PCa appears therefore to be possibly linked to oxidation phenomena on short- and medium-chain ACs which do not occur on longer-chain compounds, in agreement with the previous hypothesis of anti-angiogenetic and angiopreventive effects of these compounds.

4. Conclusions

Metabolomics is a possible missing link between phenotype and genotype, and it reflects changes arising from the insurgence of pathologies, such as tumors, which are known to extensively alter the cell metabolism. In the present study, an innovative analytical platform was developed for a straightforward albeit comprehensive characterization of ACs based on untargeted HRMS, KMD filtering, and confirmation by prediction of their RT. As a result, a large number of ACs was identified from non-malignant BPH and malignant PCa tissue samples from the same patient in a cohort of patients undergoing RARP. This type of analysis, in which each patient is represented in the sample and control groups, allowed reducing the effect of comorbidities and other intra- and inter-group differences. Later, a PLS-DA model validated by rDCV was built on the ACs dataset, with classification rates higher than 93 % for both groups, and univariate statistical analysis helped elucidating the individual role of the annotated ACs. Hydroxylation of short- and medium-chain ACs appeared to be a significant variable in describing tissue differences. In particular, the fact that in PCa tissue there was an increase in short-chain hydroxylated ACs and a contemporary decrease in their non-hydroxylated counterparts, suggests the hypothesis that the neoplastic growth is linked to oxidation phenomena on selected ACs. Several previous studies based on targeted MS focused solely on selected high abundance ACs, overlooking the role of minor compounds, i.e., most of the hydroxylated ones. It is possible to speculate that these differences in the metabolic profiles of ACs could help in the early identification of PCa if significantly correlated to the analysis in biological fluids or by intervening on the hydroxylation mechanism of ACs to develop a strategy for preventive or therapeutic purposes. Starting from the specific results obtained from the tissue analysis, further studies are needed to compare and extended the analysis and models on biological fluids, to verify which biological fluid can best represent prostatic activity and how the radical removal of the prostate gland may cause significant changes in the following months.

CRediT authorship contribution statement

Andrea Cerrato: Writing – original draft, Methodology, Funding acquisition. **Sara Elsa Aita:** Investigation, Conceptualization. **Alessandra Biancolillo:** Writing – original draft, Formal analysis. **Aldo Laganà:** Project administration, Conceptualization. **Federico Marini:** Formal analysis. **Carmela Maria Montone:** Visualization, Investigation. **Davide Rosati:** Resources. **Stefano Salciccia:** Resources. **Alessandro Sciarra:** Conceptualization, Resources, Writing – review & editing. **Enrico Taglioni:** Investigation. **Anna Laura Capriotti:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by Avvio alla Ricerca 2022 project “Identification of key prognostic biomarkers in biological fluids of patients with prostatic carcinoma before and after radical prostatectomy by metabolomics and lipidomics” funded by La Sapienza University of Rome, protocol number: AR2221811A3F85C3.

Andrea Cerrato was supported by Fondazione Umberto Veronesi through Post-doctoral Fellowships 2024 - 2 Year Grant.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2024.342574.

References

[1] J.L. Beebe-Dimmer, A.L. Kapron, A.M. Fraser, K.R. Smith, K.A. Cooney, Risk of prostate cancer associated with familial and hereditary cancer syndromes, J. Clin. Oncol. 38 (2020) 1807–1813, https://doi.org/10.1200/JCO.19.02808.

[2] G. Wang, D. Zhao, D.J. Spring, R.A. DePinho, Genetics and biology of prostate cancer, Genes Dev. 32 (2018) 1105–1140, https://doi.org/10.1101/gad.315739.118.

[3] N. Gholami, A. Haghparast, I. Alipourfard, M. Nazari, Prostate cancer in omics era, Cancer Cell Int. 22 (2022) 274, https://doi.org/10.1186/s12935-022-02691-y.

[4] S.W.D. Merriel, L. Pocock, E. Gilbert, S. Creavin, F.M. Walter, A. Spencer, W. Hamilton, Systematic review and meta-analysis of the diagnostic accuracy of prostate-specific antigen (PSA) for the detection of prostate cancer in symptomatic patients, BMC Med. 20 (2022) 54, https://doi.org/10.1186/s12916-021-02230-y.

[5] N. Hübner, S. Shariat, M. Remzi, Prostate biopsy, Curr. Opin. Urol. 28 (2018) 354–359, https://doi.org/10.1097/MOU.0000000000000510.

[6] R.J. Hendriks, I.M. van Oort, J.A. Schalken, Blood-based and urinary prostate cancer biomarkers: a review and comparison of novel biomarkers for detection and treatment decisions, Prostate Cancer Prostatic Dis. 20 (2017) 12–19, https://doi.org/10.1038/pcan.2016.59.

[7] M. Bo, M. Ventura, R. Marinello, S. Capello, G. Casetta, F. Fabris, Relationship between prostatic specific antigen (PSA) and volume of the prostate in the benign prostatic hyperplasia in the elderly, Crit. Rev. Oncol. Hematol. 47 (2003) 207–211, https://doi.org/10.1016/S1040-8428(03)00094-5.

[8] I.M. Thompson, D.K. Pauler, P.J. Goodman, C.M. Tangen, M.S. Lucia, H.L. Parnes, L.M. Minasian, L.G. Ford, S.M. Lippman, E.D. Crawford, J.J. Crowley, C. A. Coltman, Prevalence of prostate cancer among men with a prostate-specific antigen level ≤4.0 ng per milliliter, N. Engl. J. Med. 350 (2004) 2239–2246, https://doi.org/10.1056/NEJMoa031918.

[9] S. Krishnan, S. Kanthaje, D.R. Punchappady, M. Mujeeburahiman, C. K. Ratnacaram, Circulating metabolite biomarkers: a game changer in the human prostate cancer diagnosis, J. Cancer Res. Clin. Oncol. 149 (2023) 951–967, https://doi.org/10.1007/s00432-022-04113-y.

[10] S. Salciccia, A.L. Capriotti, A. Laganà, S. Fais, M. Logozzi, E. De Berardinis, G. M. Busetto, G.B. Di Pierro, G.P. Ricciuti, F. Del Giudice, A. Sciarra, P.R. Carroll, M. R. Cooperberg, B. Sciarra, M. Maggi, Biomarkers in prostate cancer diagnosis: from current knowledge to the role of metabolomics and exosomes, Int. J. Mol. Sci. 22 (2021) 4367, https://doi.org/10.3390/ijms22094367.

[11] F. Giunchi, M. Fiorentino, M. Loda, The metabolic landscape of prostate cancer, Eur. Urol. Oncol. 2 (2019) 28–36, https://doi.org/10.1016/j.euo.2018.06.010.

[12] R.J. DeBerardinis, N.S. Chandel, Fundamentals of cancer metabolism, Sci. Adv. 2 (2016), https://doi.org/10.1126/sciadv.1600200.

[13] F. Danzi, R. Pacchiana, A. Mafficini, M.T. Scupoli, A. Scarpa, M. Donadelli, A. Fiore, To metabolomics and beyond: a technological portfolio to investigate cancer metabolism, Signal Transduct. Targeted Ther. 8 (2023) 137, https://doi.org/10.1038/s41392-023-01380-0.

[14] A.R. Lima, J. Pinto, F. Amaro, M. de L. Bastos, M. Carvalho, P. Guedes de Pinho, Advances and perspectives in prostate cancer biomarker discovery in the last 5 Years through tissue and urine metabolomics, Metabolites 11 (2021) 181, https://doi.org/10.3390/metabo11030181.

[15] L.A. Vandergrift, E.A. Decelle, J. Kurth, S. Wu, T.L. Fuss, E.M. DeFeo, E.F. Halpern, M. Taupitz, W.S. McDougal, A.F. Olumi, C.-L. Wu, L.L. Cheng, Metabolomic prediction of human prostate cancer aggressiveness: magnetic resonance spectroscopy of histologically benign tissue, Sci. Rep. 8 (2018) 4997, https://doi.org/10.1038/s41598-018-23177-w.

[16] 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.

[17] N. Bansal, M. Kumar, S.N. Sankhwar, A. Gupta, Relevance of emerging metabolomics-based biomarkers of prostate cancer: a systematic review, Expet Rev. Mol. Med. 24 (2022) e25, https://doi.org/10.1017/erm.2022.20.
[18] A. Sreekumar, L.M. Poisson, T.M. Rajendiran, A.P. Khan, Q. Cao, J. Yu, B. Laxman, R. Mehra, R.J. Lonigro, Y. Li, M.K. Nyati, A. Ahsan, S. Kalyana-Sundaram, B. Han, X. Cao, J. Byun, G.S. Omenn, D. Ghosh, S. Pennathur, D.C. Alexander, A. Berger, J. R. Shuster, J.T. Wei, S. Varambally, C. Beecher, A.M. Chinnaiyan, Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression, Nature 457 (2009) 910–914, https://doi.org/10.1038/nature07762.
[19] M.A. Fernández-Peralbo, E. Gómez-Gómez, M. Calderón-Santiago, J. Carrasco-Valiente, J. Ruiz-García, M.J. Requena-Tapia, M.D. Luque de Castro, F. Priego-Capote, Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC-QTOF: upstream information on other omics, Sci. Rep. 6 (2016) 38243, https://doi.org/10.1038/srep38243.
[20] A. Franko, Y. Shao, M. Heni, J. Hennenlotter, M. Hoene, C. Hu, X. Liu, X. Zhao, Q. Wang, A.L. Birkenfeld, T. Todenhöfer, A. Stenzl, A. Peter, H.-U. Häring, R. Lehmann, G. Xu, S.Z. Lutz, Human prostate cancer is characterized by an increase in urea cycle metabolites, Cancers 12 (2020) 1814, https://doi.org/10.3390/cancers12071814.
[21] X. Zhang, B. Xia, H. Zheng, J. Ning, Y. Zhu, X. Shao, B. Liu, B. Dong, H. Gao, Identification of characteristic metabolic panels for different stages of prostate cancer by 1H NMR-based metabolomics analysis, J. Transl. Med. 20 (2022) 275, https://doi.org/10.1186/s12967-022-03478-5.
[22] P.R. Braadland, G. Giskeødegård, E. Sandsmark, H. Bertilsson, L.R. Euceda, A. F. Hansen, I.J. Guldvik, K.M. Selnæs, H.H. Grytli, B. Katz, A. Svindland, T. F. Bathen, L.M. Eri, S. Nygård, V. Berge, K.A. Taskén, M.-B. Tessem, Ex vivo metabolic fingerprinting identifies biomarkers predictive of prostate cancer recurrence following radical prostatectomy, Br. J. Cancer 117 (2017) 1656–1664, https://doi.org/10.1038/bjc.2017.346.
[23] A.F. Hansen, E. Sandsmark, M.B. Rye, A.J. Wright, H. Bertilsson, E. Richardsen, T. Viset, A.M. Bofin, A. Angelsen, K.M. Selnæs, T.F. Bathen, M.-B. Tessem, Presence of TMPRSS2-ERG is associated with alterations of the metabolic profile in human prostate cancer, Oncotarget. 7 (2106) 42071–42085. https://doi.org/10.18632/oncotarget.9817..
[24] E. Amante, A. Cerrato, E. Alladio, A.L. Capriotti, C. Cavaliere, F. Marini, C. M. Montone, S. Piovesana, A. Laganà, M. Vincenti, Comprehensive biomarker profiles and chemometric filtering of urinary metabolomics for effective discrimination of prostate carcinoma from benign hyperplasia, Sci. Rep. 12 (2022) 4361, https://doi.org/10.1038/s41598-022-08435-2.
[25] L.M. Butler, C.Y. Mah, J. Machiels, A.D. Vincent, S. Irani, S.M. Mutuku, X. Spotbeen, M. Bagadi, D. Waltregny, M. Moldovan, J. Dehairs, F. Vanderhoydonc, K. Bloch, R. Das, J. Stahl, J.G. Kench, T. Gevaert, R. Derua, E. Waelkens, Z. D. Nassar, L.A. Selth, P.J. Trim, M.F. Snel, D.J. Lynn, W.D. Tilley, L.G. Horvath, M. M. Centenera, J.V. Swinnen, Lipidomic profiling of clinical prostate cancer reveals targetable alterations in membrane lipid composition, Cancer Res. 81 (2021) 4981–4993, https://doi.org/10.1158/0008-5472.CAN-20-3863.
[26] S. Ren, Y. Shao, X. Zhao, C.S. Hong, F. Wang, X. Lu, J. Li, G. Ye, M. Yan, Z. Zhuang, C. Xu, G. Xu, Y. Sun, Integration of metabolomics and transcriptomics reveals major metabolic pathways and potential biomarker involved in prostate cancer, Mol. Cell. Proteomics 15 (2016) 154–163, https://doi.org/10.1074/mcp.M115.052381.
[27] M. Puhka, M. Takatalo, M.-E. Nordberg, S. Valkonen, J. Nandania, M. Aatonen, M. Yliperttula, S. Laitinen, V. Velagapudi, T. Mirtti, O. Kallioniemi, A. Rannikko, P. R.-M. Siljander, T.M. af Hällström, Metabolomic profiling of extracellular vesicles and alternative normalization methods reveal enriched metabolites and strategies to study prostate cancer-related changes, Theranostics 7 (2017) 3824–3841, https://doi.org/10.7150/thno.19890.
[28] A. Cerrato, C. Bedia, A.L. Capriotti, C. Cavaliere, V. Gentile, M. Maggi, C. M. Montone, S. Piovesana, A. Sciarra, R. Tauler, A. Laganà, Untargeted metabolomics of prostate cancer zwitterionic and positively charged compounds in urine, Anal. Chim. Acta 1158 (2021) 338381, https://doi.org/10.1016/j.aca.2021.338381.
[29] D. Broadhurst, R. Goodacre, S.N. Reinke, J. Kuligowski, I.D. Wilson, M.R. Lewis, W. B. Dunn, Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies, Metabolomics 14 (2018) 72, https://doi.org/10.1007/s11306-018-1367-3.
[30] E. Rampler, Y. El Abiead, H. Schoeny, M. Rusz, F. Hildebrand, V. Fitz, G. Koellensperger, Recurrent topics in mass spectrometry-based metabolomics and lipidomics—standardization, coverage, and throughput, Anal. Chem. 93 (2021) 519–545, https://doi.org/10.1021/acs.analchem.0c04698.
[31] M. Pourchet, L. Debrauwer, J. Klanova, E.J. Price, A. Covaci, N. Caballero-Casero, H. Oberacher, M. Lamoree, A. Damont, F. Fenaille, J. Vlaanderen, J. Meijer, M. Krauss, D. Sarigiannis, R. Barouki, B. Le Bizec, J.-P. Antignac, Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: from promises to challenges and harmonisation issues, Environ. Int. 139 (2020) 105545, https://doi.org/10.1016/j.envint.2020.105545.
[32] 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.
[33] S. Li, D. Gao, Y. Jiang, Function, detection and alteration of acylcarnitine metabolism in hepatocellular carcinoma, Metabolites 9 (2019) 36, https://doi.org/10.3390/metabo9020036.
[34] D. Baci, A. Bruno, B. Bassani, M. Tramacere, L. Mortara, A. Albini, D.M. Noonan, Acetyl- l -carnitine is an anti-angiogenic agent targeting the VEGFR2 and CXCR4 pathways, Cancer Lett. 429 (2018) 100–116, https://doi.org/10.1016/j.canlet.2018.04.018.
[35] M. Dambrova, M. Makrecka-Kuka, J. Kuka, R. Vilskersts, D. Nordberg, M. M. Attwood, S. Smesny, Z.D. Sen, A.C. Guo, E. Oler, S. Tian, J. Zheng, D.S. Wishart, E. Liepinsh, H.B. Schiöth, Acylcarnitines: nomenclature, biomarkers, therapeutic potential, drug targets, and clinical trials, Pharmacol. Rev. 74 (2022) 506–551, https://doi.org/10.1124/pharmrev.121.000408.
[36] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D.D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G. E. Duggan, G.D. MacInnis, A.M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B.D. Sykes, H.J. Vogel, L. Querengesser, HMDB: the human metabolome database, Nucleic Acids Res. 35 (2007), https://doi.org/10.1093/nar/gkl923.
[37] C. Feng, L. Xue, D. Lu, Y. Jin, X. Qiu, F.J. Gonzalez, G. Wang, Z. Zhou, Novel strategy for mining and identification of acylcarnitines using data-independent-acquisition-based retention time prediction modeling and pseudo-characteristic fragmentation ion matching, J. Proteome Res. 20 (2021) 1602–1611, https://doi.org/10.1021/acs.jproteome.0c00810.
[38] S. Li, D. Gao, C. Song, C. Tan, Y. Jiang, Isotope labeling strategies for acylcarnitines profile in biological samples by liquid chromatography–mass spectrometry, Anal. Chem. 91 (2019) 1701–1705, https://doi.org/10.1021/acs.analchem.8b05120.
[39] T. Tang, P. Zhang, S. Li, D. Xu, W. Li, Y. Tian, Y. Jiao, Z. Zhang, F. Xu, Absolute quantification of acylcarnitines using integrated tmt-PP derivatization-based LC–MS/MS and quantitative analysis of multi-components by a single marker strategy, Anal. Chem. 93 (2021) 12973–12980, https://doi.org/10.1021/acs.analchem.1c02606.
[40] C.A. Hughey, C.L. Hendrickson, R.P. Rodgers, A.G. Marshall, K. Qian, Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra, Anal. Chem. 73 (2001) 4676–4681, https://doi.org/10.1021/ac010560w.
[41] S. Merel, Critical assessment of the Kendrick mass defect analysis as an innovative approach to process high resolution mass spectrometry data for environmental applications, Chemosphere 313 (2023) 137443, https://doi.org/10.1016/j.chemosphere.2022.137443.
[42] J.A. Kirwan, H. Gika, R.D. Beger, D. Bearden, W.B. Dunn, R. Goodacre, G. Theodoridis, M. Witting, L.-R. Yu, I.D. Wilson, Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management, Metabolomics 18 (2022) 70, https://doi.org/10.1007/s11306-022-01926-3.
[43] L. Ståhle, S. Wold, Partial least squares analysis with cross-validation for the two-class problem: a Monte Carlo study, J. Chemom. 1 (1987) 185–196, https://doi.org/10.1002/cem.1180010306.
[44] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemom. 17 (2003) 166–173, https://doi.org/10.1002/cem.785.
[45] N.F. Pérez, J. Ferré, R. Boqué, Calculation of the reliability of classification in discriminant partial least-squares binary classification, Chemometr. Intell. Lab. Syst. 95 (2009) 122–128, https://doi.org/10.1016/j.chemolab.2008.09.005.
[46] S. Wold, E. Johansson, M. Cocchi, Pls - partial least-squares projections to latent structures, in: 3D QSAR Drug Des., 1993.
[47] P. Filzmoser, B. Liebmann, K. Varmuza, Repeated double cross validation, J. Chemom. 23 (2009) 160–171, https://doi.org/10.1002/cem.1225.
[48] J.A. Westerhuis, H.C.J. Hoefsloot, S. Smit, D.J. Vis, A.K. Smilde, E.J.J. van Velzen, J.P.M. van Duijnhoven, F.A. van Dorsten, Assessment of PLSDA cross validation, Metabolomics 4 (2008) 81–89, https://doi.org/10.1007/s11306-007-0099-6.
[49] J. Xia, D.S. Wishart, Metabolomic data processing, analysis, and interpretation using MetaboAnalyst, Curr. Protoc. Bioinf. 34 (2011), https://doi.org/10.1002/0471250953.bi1410s34.
[50] D. Yu, L. Zhou, Q. Xuan, L. Wang, X. Zhao, X. Lu, G. Xu, Strategy for comprehensive identification of acylcarnitines based on liquid chromatography–high-resolution mass spectrometry, Anal. Chem. 90 (2018) 5712–5718, https://doi.org/10.1021/acs.analchem.7b05471.
[51] G. Ventura, C.D. Calvano, V. Porcelli, L. Palmieri, A. De Giacomo, Y. Xu, R. Goodacre, F. Palmisano, T.R.I. Cataldi, Phospholipidomics of peripheral blood mononuclear cells (PBMCs): the tricky case of children with autism spectrum disorder (ASD) and their healthy siblings, Anal. Bioanal. Chem. 412 (2020) 6859–6874, https://doi.org/10.1007/s00216-020-02817-z.
[52] X. Wei, X. Shi, S. Kim, L. Zhang, J.S. Patrick, J. Binkley, C. McClain, X. Zhang, Data preprocessing method for liquid chromatography–mass spectrometry based metabolomics, Anal. Chem. 84 (2012) 7963–7971, https://doi.org/10.1021/ac3016856.
[53] I. Karaman, Preprocessing and Pretreatment of Metabolomics Data for Statistical Analysis, 2017, pp. 145–161, https://doi.org/10.1007/978-3-319-47656-8_6.
[54] C. Schiffman, L. Petrick, K. Perttula, Y. Yano, H. Carlsson, T. Whitehead, C. Metayer, J. Hayes, S. Rappaport, S. Dudoit, Filtering procedures for untargeted LC-MS metabolomics data, BMC Bioinf. 20 (2019) 334, https://doi.org/10.1186/s12859-019-2871-9.
[55] R.R. da Silva, F. Vargas, M. Ernst, N.H. Nguyen, S. Bolleddu, K.K. del Rosario, S. M. Tsunoda, P.C. Dorrestein, A.K. Jarmusch, Computational removal of undesired mass spectral features possessing repeat units via a Kendrick mass filter, J. Am. Soc. Mass Spectrom. 30 (2019) 268–277, https://doi.org/10.1007/s13361-018-2069-9.
[56] L. Renai, M. Del Bubba, S. Samanipour, R. Stafford, A.F.G. Gargano, Development of a comprehensive two-dimensional liquid chromatographic mass spectrometric

method for the non-targeted identification of poly- and perfluoroalkyl substances in aqueous film-forming foams, Anal. Chim. Acta 1232 (2022) 340485, https://doi.org/10.1016/j.aca.2022.340485.
[57] L.T. Richardson, E.K. Neumann, R.M. Caprioli, J.M. Spraggins, T. Solouki, Referenced Kendrick mass defect annotation and class-based filtering of imaging MS lipidomics experiments, Anal. Chem. 94 (2022) 5504–5513, https://doi.org/10.1021/acs.analchem.1c03715.
[58] X. Yan, S.P. Markey, R. Marupaka, Q. Dong, B.T. Cooper, Y.A. Mirokhin, W. E. Wallace, S.E. Stein, Mass spectral library of acylcarnitines derived from human urine, Anal. Chem. 92 (2020) 6521–6528, https://doi.org/10.1021/acs.analchem.0c00129.
[59] E.L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H.P. Singer, J. Hollender, Identifying small molecules via high resolution mass spectrometry: communicating confidence, Environ. Sci. Technol. 48 (2014) 2097–2098, https://doi.org/10.1021/es5002105.
[60] J. Wang, Z. Liu, M. Xu, X. Han, C. Ren, X. Yang, C. Zhang, F. Fang, Cinical, metabolic, and genetic analysis and follow-up of eight patients with HIBCH mutations presenting with leigh/leigh-like syndrome, Front. Pharmacol. 12 (2021), https://doi.org/10.3389/fphar.2021.605803.
[61] E. Zoni, M. Minoli, C. Bovet, A. Wehrhan, S. Piscuoglio, C.K.Y. Ng, P.C. Gray, M. Spahn, G.N. Thalmann, M. Kruithof-de Julio, Preoperative plasma fatty acid metabolites inform risk of prostate cancer progression and may be used for personalized patient stratification, BMC Cancer 19 (2019) 1216, https://doi.org/10.1186/s12885-019-6418-2.
[62] A. Albini, D. Briga, M. Conti, A. Bruno, D. Farioli, S. Canali, I. Sogno, G. D'Ambrosio, P. Consonni, D.M. Noonan, SANIST: a rapid mass spectrometric SACI/ESI data acquisition and elaboration platform for verifying potential candidate biomarkers, Rapid Commun. Mass Spectrom. 29 (2015) 1703–1710, https://doi.org/10.1002/rcm.7270.
[63] X. Lu, H. Nie, Y. Li, C. Zhan, X. Liu, X. Shi, M. Shi, Y. Zhang, Y. Li, Comprehensive characterization and evaluation of hepatocellular carcinoma by LC–MS based serum metabolomics, Metabolomics 11 (2015) 1381–1393, https://doi.org/10.1007/s11306-015-0797-4.
[64] J.A. Schmidt, G.K. Fensom, S. Rinaldi, A. Scalbert, P.N. Appleby, D. Achaintre, A. Gicquiau, M.J. Gunter, P. Ferrari, R. Kaaks, T. Kühn, A. Floegel, H. Boeing, A. Trichopoulou, P. Lagiou, E. Anifantis, C. Agnoli, D. Palli, M. Trevisan, R. Tumino, H.B. Bueno-de-Mesquita, A. Agudo, N. Larrañaga, D. Redondo-Sánchez, A. Barricarte, J.M. Huerta, J.R. Quirós, N. Wareham, K.-T. Khaw, A. Perez-Cornago, M. Johansson, A.J. Cross, K.K. Tsilidis, E. Riboli, T.J. Key, R.C. Travis, Pre-diagnostic metabolite concentrations and prostate cancer risk in 1077 cases and 1077 matched controls in the European Prospective Investigation into Cancer and Nutrition, BMC Med. 15 (2017) 122, https://doi.org/10.1186/s12916-017-0885-6.
[65] K. Enooku, H. Nakagawa, N. Fujiwara, M. Kondo, T. Minami, Y. Hoshida, J. Shibahara, R. Tateishi, K. Koike, Altered serum acylcarnitine profile is associated with the status of nonalcoholic fatty liver disease (NAFLD) and NAFLD-related hepatocellular carcinoma, Sci. Rep. 9 (2019) 10663, https://doi.org/10.1038/s41598-019-47216-2.
[66] H.-J. Lee, D.M. Kremer, P. Sajjakulnukit, L. Zhang, C.A. Lyssiotis, A large-scale analysis of targeted metabolomics data from heterogeneous biological samples provides insights into metabolite dynamics, Metabolomics 15 (2019) 103, https://doi.org/10.1007/s11306-019-1564-8.