



Research article

^1H -NMR-based urine metabolomics of prostate cancer and benign prostatic hyperplasia

Mohammed Zniber^a, Tarja Lamminen^b, Pekka Taimen^c, Peter J. Boström^b, Tan-Phat Huynh^{a,*}^a Laboratory of Molecular Science and Engineering, Åbo Akademi University, Turku, Finland^b Department of Urology, University of Turku and Turku University Hospital, Turku, Finland^c Institute of Biomedicine and FICAN West Cancer Centre, University of Turku and Department of Pathology, Turku University Hospital, Turku, Finland

ARTICLE INFO

Keywords:

NMR
Metabolomics
Urine
Prostate cancer
Chemometrics

ABSTRACT

Background: Prostate cancer (PCa) and benign prostatic hyperplasia (BPH) are prevalent conditions affecting a significant portion of the male population, particularly with advancing age. Traditional diagnostic methods, such as digital rectal examination (DRE) and prostate-specific antigen (PSA) tests, have limitations in specificity and sensitivity, leading to potential over-diagnosis and unnecessary biopsies.

Significance: This study explores the effectiveness of ^1H NMR urine metabolomics in distinguishing PCa from BPH and in differentiating various PCa grades, presenting a non-invasive diagnostic alternative with the potential to enhance early detection and patient-specific treatment strategies.

Results: The study demonstrated the capability of ^1H NMR urine metabolomics in detecting distinct metabolic profiles between PCa and BPH, as well as among different Gleason grade groups. Notably, this method surpassed the PSA test in distinguishing PCa from BPH. Untargeted metabolomics analysis also revealed several metabolites with varying relative concentrations between PCa and BPH cases, suggesting potential biomarkers for these conditions.

1. Introduction

Prostate cancer (PCa) is the fourth leading cause of death and the second most widely diagnosed cancer type among men. In 2020, PCa accounted for approximately 374 000 deaths and 1.42 million new cases [1] whereas benign prostatic hyperplasia (BPH) impacts 62% of men aged >50 years old [2–4]. PCa and BPH are major health problems and the incidences of both diseases are expected to increase due to aging of the population [5]. With similarities and frequent coexistence between these two pathological conditions, the question of their relation is raised. Based on epidemiological studies the incidence of PCa and BPH is tightly correlated with age. Both are also hormone dependent [6], for instance, androgens are essential for normal prostate growth as well as the development of PCa and BPH. In brief, 5- α reductase type 1 and 2, are two isozymes involved in converting testosterone to dihydrotestosterone (DHT), which binds to androgen receptors and promotes cellular differentiation and proliferation within the prostate. The abnormal growth of

* Corresponding author.

E-mail addresses: mohammed.zniber@abo.fi (M. Zniber), tarja.lamminen@tyks.fi (T. Lamminen), pepeta@utu.fi (P. Taimen), peter.bostrom@utu.fi (P.J. Boström), tan.huynh@abo.fi (T.-P. Huynh).<https://doi.org/10.1016/j.heliyon.2024.e28949>

Received 30 December 2023; Received in revised form 27 March 2024; Accepted 27 March 2024

Available online 31 March 2024

2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the prostate is considered to involve the disruption of DHT-supported homeostasis between cell death and cell proliferation resulting in an inhibition of apoptotic processes and predomination of proliferative processes [5,7–9].

The most common diagnostic methods of PCa include digital rectal examination (DRE), prostate-specific-antigen (PSA), transrectal ultrasound (TRUS), magnetic resonance imaging (MRI) and prostate biopsy. Among them, PSA test is widely used method for screening purposes to estimate the risk of PCa. However, it suffers from various disadvantages including high false positivity and poor specificity of PSA for detecting cancer which leads to unnecessary prostate cancer biopsies [10–12].

In contrast to the abovementioned diagnostic methods, nuclear magnetic resonance (NMR) spectroscopy-based urine metabolomics could be useful in discriminating PCa and BPH. For instance, NMR spectroscopy is non-invasive, non-destructive, and does not require chromatographic separation, sample treatment, or chemical derivatization. Moreover, NMR is highly automatable, rapid, and reproducible; thus large-scale metabolomics studies are more feasible with NMR compared to other metabolomics platforms. For instance, compared to NMR, GC-MS and LC-MS can only process 20 to 30 percent of samples over the course of a 24-h period. This is due to longer chromatographic runs, manual sample preparation, and lengthy computer processing [13]. Furthermore, NMR is capable of detecting compounds that are less detectable in LC-MS such as sugars, organic acids, alcohols, polyols and other highly polar compounds [14]. NMR can also be used for quantitative analysis and enables the identification of several metabolites in a single experiment [15]. Recent studies have revealed the potential of ^1H -NMR-based metabolomics in cancer diagnosis such as colorectal cancer [16–18], breast cancer [19,20] and other types of cancer [21–23]. The literature regarding PCa and BPH studies based on urine and NMR metabolomics is relatively scarce [12,24–28]. Additionally, most studies have focused on the analysis of tissue samples [29–34] or serum/plasma [27,28,35–39]. Despite its advantages, NMR faces significant limitations, including poor sensitivity that requires metabolite concentrations of at least $1\ \mu\text{M}$ for detection, limiting its ability to identify low-abundance metabolites essential for comprehensive metabolomic studies. Like MS, NMR cannot detect all metabolites, necessitating the combined use of NMR and MS to enhance metabolome coverage and metabolite identification accuracy. NMR is also prone to variability in spectral data due to changes in sample conditions like pH, temperature, ionic strength, and composition, complicating data interpretation and comparison. However, advancements in NMR technology are improving its accessibility, affordability, and clinical application compatibility [13, 40].

In this study, we have investigated the metabolic differences between PCa and BPH using two different approaches of NMR spectroscopy: binning or bucketing and relative quantifications. Binned spectra were explored using two unsupervised techniques namely principal component analysis (PCA) and unsupervised random forest (URF). A classification model was built based on supervised random forest (RF) for the discrimination of the two studied groups. On the other hand, peak identification was used to measure relative quantifications of metabolites and determine the most significantly different ones between PCa and BPH.

2. Materials and methods

2.1. Study design and participants

The study protocol was approved by the Ethics Committee of the Hospital District of Southwest Finland (3/1801/2013). The volunteers enrolled for the study due to a clinical suspicion of PCa based on either elevated PSA levels or physical examination (palpation). A total of 81 patients included in this study, were encoded, comprising 40 pathologically confirmed PCa patients and 41 BPH patients. Individual donors could not be identified from the samples. Clinical characteristics of the individuals included in the study are shown in Table 1. All the urine samples were obtained perioperatively from patients undergoing transurethral resection of the prostate (TURP) for BPH or robot-assisted laparoscopic prostatectomy (RALP) for PCa at Turku University Hospital (TYKS) between 2021 and 2022.

Explicitly, urine samples were taken using an indwelling catheter placed routinely for patients undergoing a urological operation. Urine collection did not include any additional instruments or surgical procedures. All collected samples were stored in a $-80\ ^\circ\text{C}$.

2.2. ^1H NMR spectroscopy

2.2.1. Reagents

Sodium (3-trimethylsilyl)-2,2,3,3-tetradeuteriopropionate (TSP), sodium chloride (NaCl), deuterium oxide (D_2O), sodium phosphate monobasic (NaH_2PO_4) and sodium phosphate dibasic (Na_2HPO_4) were purchased from Sigma-Aldrich.

Table 1
Clinical data of the individuals included in the study.

	BPH group (n = 41) (Median, range)	PCa group (n = 40) (Median, range)
Age (years)	74.4 (66.2–85.2)	66.2 (48.1–72.3)
BMI (kg m^{-2})	25.2 (18.7–30.7)	27.1 (21.3–31.3)
Preop PSA (ng mL^{-1})	3.4 (0.025–35)	7.9 (2.4–30)
Number of cores	N/A	14 (3–31)
Positive cores	N/A	0 (0–2)
Grade group	N/A	3 (2–5)

2.2.2. Measurement

Aliquots of 450 μL human urine were mixed with 50 μL of D_2O containing 1 mM of TSP, 0.2 M of Na_2HPO_4 | 0.04 M of NaH_2PO_4 (pH 7.4) and 0.8 w/v% of NaCl to a total volume of 500 μL in Eppendorf tubes. The pH adjustment of all urine samples needed to be done before conducting any measurements to ensure consistency and reliability in our NMR experiments. This step is crucial as the pH can significantly affect the outcome of NMR spectroscopy by influencing the chemical shift and, consequently, the interpretation of the spectra [41]. The samples were then centrifuged at 3000 rpm for 5 min and transferred to clean 5 mm NMR tubes using fine Pasteur pipettes. Finally, the tubes were capped and labelled prior to NMR acquisition. NMR experiments were carried out using a 600 MHz NMR spectrometer (AVANCE III, Bruker, Germany) equipped with a liquid nitrogen cooled Prodigy TCI (inverted CryoProbe) at 298 K.

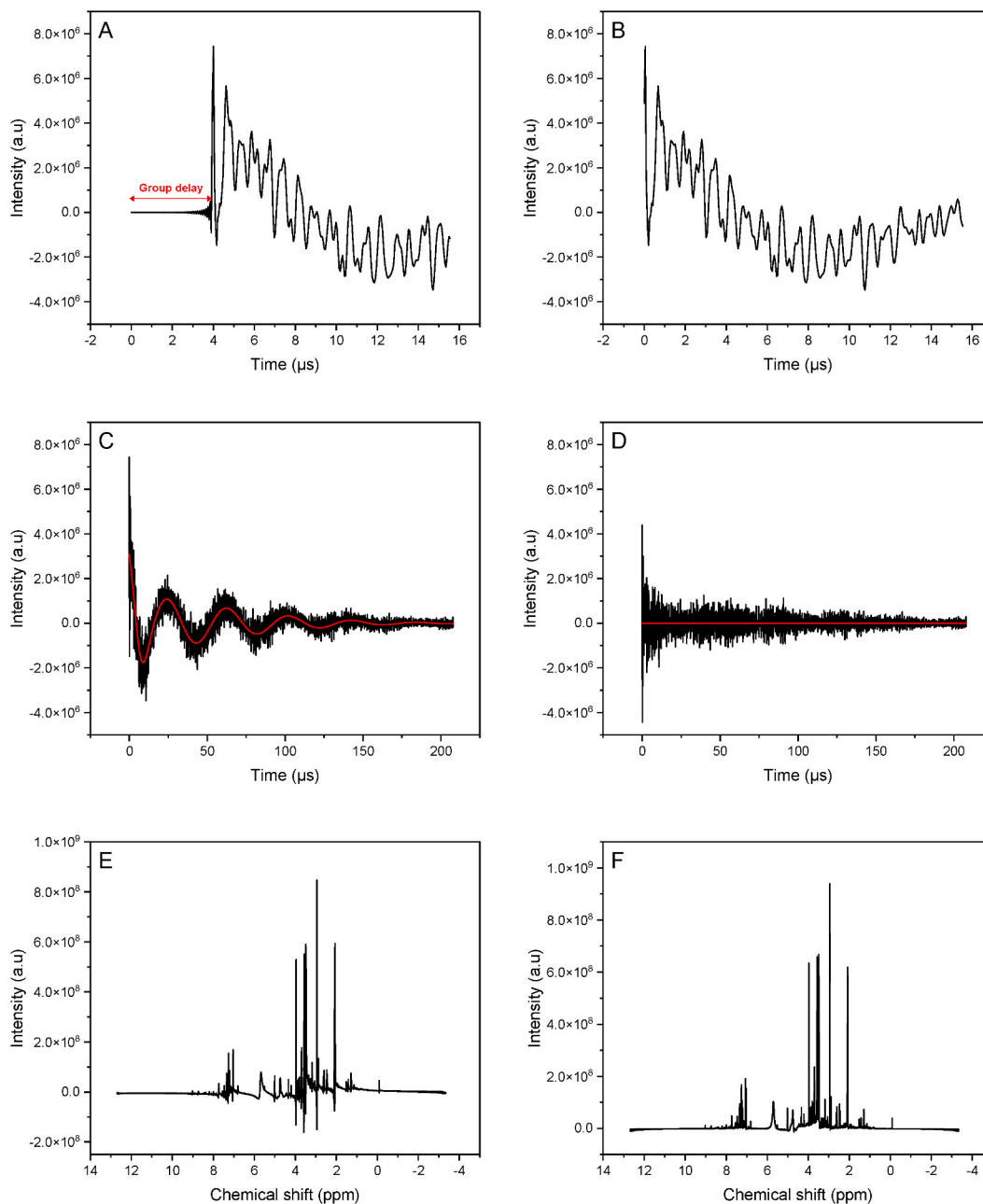


Fig. 1. Zoomed FID with group delay (A) and after group delay removal (B). FID spectrum with (C) and without solvent residuals signal (D). NMR spectrum after Fourier Transform (E) and after zero-order phase correction (F).

2.3. Data preprocessing and analysis

Raw FIDs (free induction decay) were preprocessed using the R statistical software package PepsNMR [42]. Subsequently, binned spectra were analyzed, employing techniques such as PCA and URF, to unveil clustering patterns and detect outliers within the two studied groups. Discriminant analysis was performed using RF to distinguish between the two studied groups. Metabolite identification and relative quantifications were carried out using ASICS [43]. Finally, statistical differences between PCa and BPH were assessed using the Kruskal-Wallis test.

3. Results

3.1. Data preprocessing

Initially, the raw Bruker files were imported into R resulting into two data frames: the complex FID signal matrix and the Bruker acquisition. Fig. S1 provides a visual representation of the real of the FID. Fig. 1A and B presents an NMR spectrum before and after correcting for group delay or death time. The FID indicates a death time of about 4 μ s, with negligible values before the start of the standard FID. Prior to group delay correction, the Fourier transform spectrum exhibits baseline fluctuations similar to the convolution with a sinc function centered in the spectrum. This can be explained by the Fourier transform's time shift theorem, which suggests that shifting the time domain signal by n points results in the frequency domain spectrum being the standard spectrum (unshifted FID) multiplied by $\exp(-i2\pi*w*n)$ [42]. Hence, these baseline wiggles due to a first-order phase shift introduced by the group delay can be rectified by subtracting the group delay from the FID. Next, the FIDs are processed using a Whittaker smoother to remove the variability of solvent (water) residuals in the spectrum, which can mask informative signals of other interesting compounds. Fig. 1C and D shows the estimated solvent residuals signal as well as the FID signal before and after its removal. Subsequently, the FIDs are converted into spectra in the frequency domain by Fourier Transform (Fig. 1E). This figure shows that the spectra present zero-order phase error of a certain angle, which is corrected by automatically finding an optimal ϕ_0 which shifts the spectrum into its pure absorptive mode (Fig. 1F).

Consequently, the NMR spectra are referenced with TSP (Fig. S2), a known reference compound whose chemical shift is set to 0 ppm by convention. A baseline removal was calculated and removed from the spectra as shown in Fig. 2A and B.

Negative values from prior steps are set to zero as they lack meaningful interpretation. Next, spectra are globally aligned to

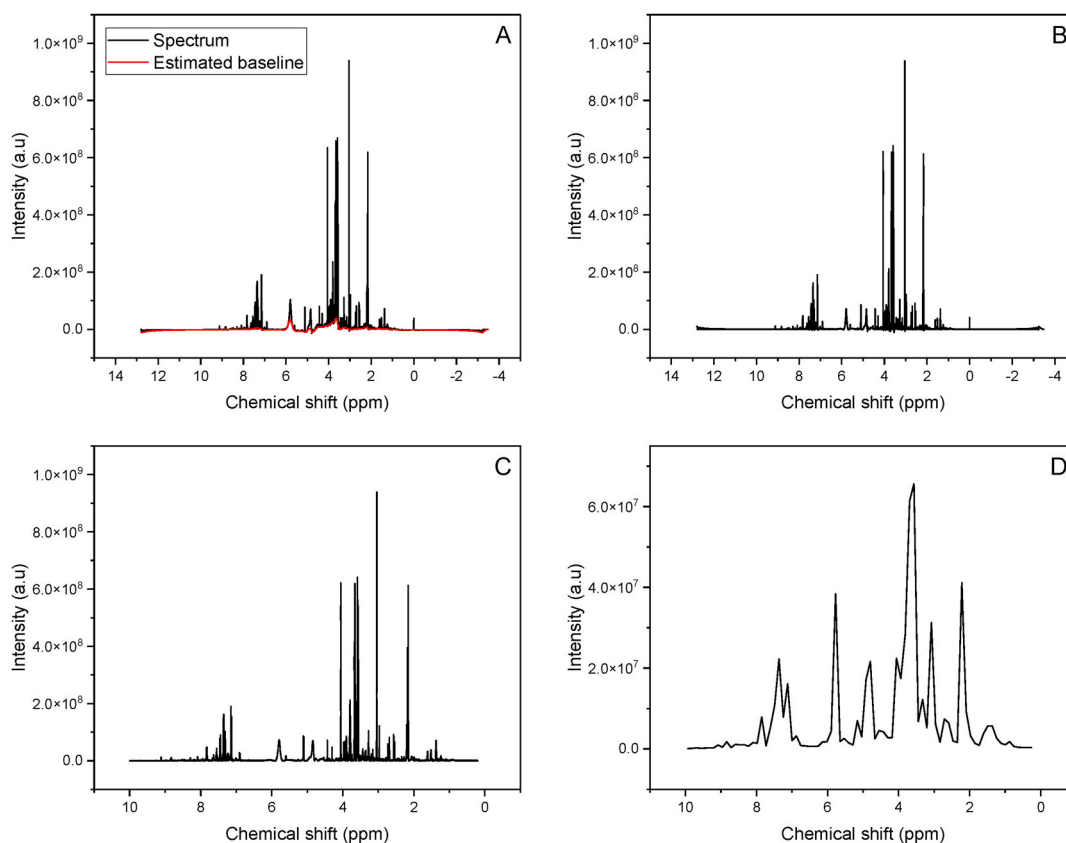


Fig. 2. NMR spectrum before (A) and after baseline correction (B). NMR spectrum before (C) and after bucketing (D).

minimize misalignment between identical peaks across different spectra. Negative values from prior steps are set to zero as they lack meaningful interpretation. Subsequently, spectra are globally aligned to minimize misalignment between identical peaks across different spectra. Finally, the high dimensions of the spectra is reduced using rectangular bucketing (Fig. 2C and D).

Region removal was applied to cut off the region (4.5–6.1 ppm) where the water resonance residuals, urea, and maleic acid are located. This region is influenced by the removal of the water resonance using presaturation. Finally, each spectrum was normalized by quantile normalization which attempts to achieve the same distribution of feature intensities across all spectra [44].

3.2. Analysis on buckets

3.2.1. Principal component analysis

The binned spectra were mean-centered, and any columns with all zeros were excluded. PCA was subsequently employed to explore potential associations and patterns in the data. Fig. 3A presents the first ten principal components (PCs), which collectively explain 90% of the total variance observed. In Fig. 3B, the score plot between PC1 and PC4 reveals that BPH samples tend to cluster with negative values of PC4, while PCa samples cluster with positive values of PC4. Additionally, PC2 exhibited slight separation of samples. This separation can be explained by the loadings illustrated in Fig. 3C and D, for instance, buckets located around 1.96, 2.09, 2.36, 3.01, 3.79, 4.05, 4.19, 8.50, and 8.63 ppm are related to the PCa patients while 2.23, 2.88, 3.14, 7.19, 7.45, and 7.71 ppm are related to BPH individuals. Additionally, loading of PC2 indicates that buckets around 7.58, 7.71, 7.84, 7.97, and 8.50 ppm are slightly related to the BPH group while 7.45, 7.32, 7.19, 3.66, 2.23, 2.09 ppm are slightly linked to the PCa group.

3.2.2. Random forest

The dataset was divided into training and testing sets (test size of 25%). RF classifier was then trained on the training dataset using repeated cross validation approach with a number of repetitions of one hundred and number of splits of five. Next, the testing set was used to predict the output classes and evaluate the performance of the model using the confusion matrix and receiver operating characteristic (ROC) curve.

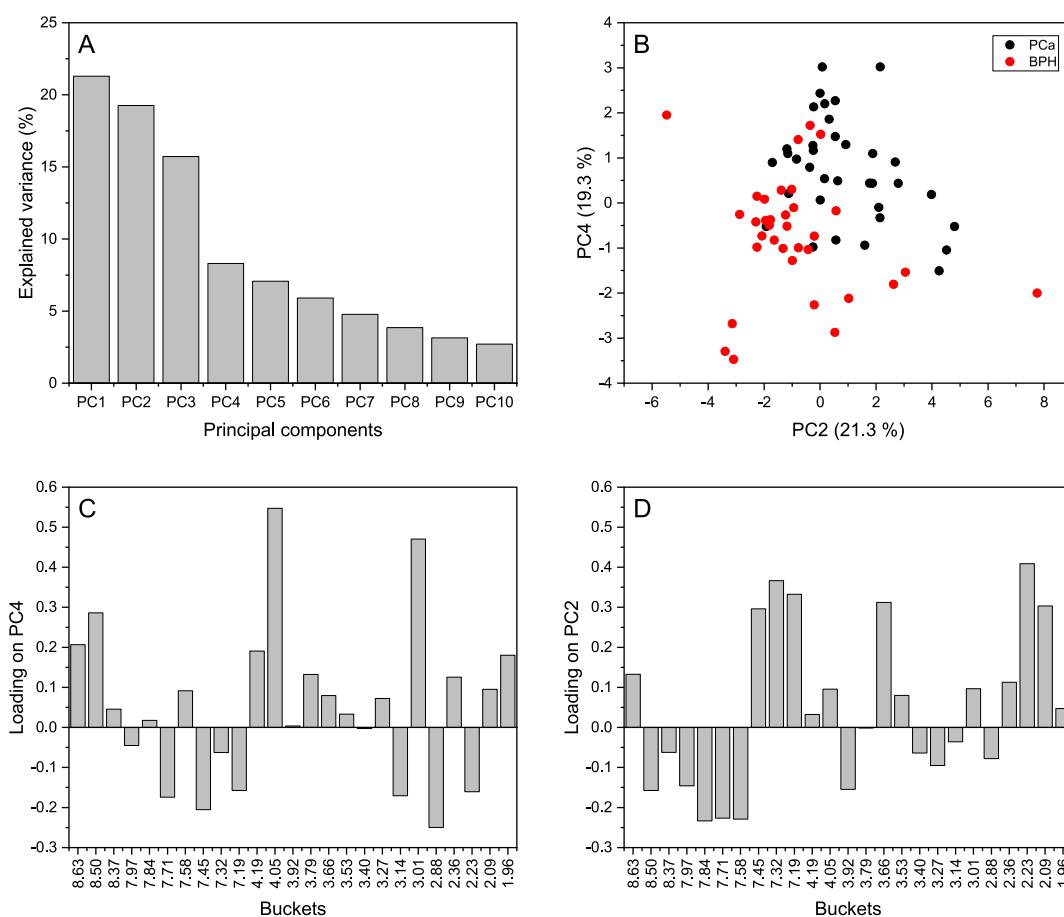


Fig. 3. Percentage of variance explained by PCs obtained from NMR binned spectra (A) and (B) scatter plot of PC2 versus PC4 (B); loading plots for PC4 (C) and PC2 (D)

Tables 2 and 3 display the average confusion matrix and classification report of the RF model. The misclassification rates were around 6.8% and 5% for cross validation (CV) and testing set, respectively, and were calculated as the number of all incorrect predictions divided by the total number of predictions. In addition, the average recalls were 93% and 95%, precisions were 96% and 93% and accuracies were 93% and 95% in CV and testing indicating that the model generalizes on unseen data.

3.2.3. Unsupervised random forest

In order to explain the efficiency of the RF classification model, the original data was visualized using a score plot based on PCA performed on the proximity matrix obtained from URF [45]. The proximity matrix symbolizes the similarities between observations, for instance, a large proximity value implies similarity while a small value indicates dissimilarity. Fig. 4A displays a score plot of PC1 versus PC2 calculated from the URF proximity matrix; two distinct groups appear corresponding the two pathological conditions BPH and PCa. The dissimilarities between the two groups are quite clear except for some samples that look a bit similar supporting the abovementioned results of RF in discriminating between BPH and PCa.

3.2.4. Prostate-specific-antigen test

PSA is a protein produced by the prostate gland and commonly measured as a screening test for prostate cancer. PSA levels are measured with a simple blood test, measured in nanograms of PSA per milliliter (ng mL^{-1}) of blood. A higher PSA level may indicate an increased risk of prostate cancer, but it is not a definitive diagnosis. Other factors, such as age, family history, and medical history, should also be considered when interpreting the results of a PSA test. Fig. 4B represents BPH and PCa samples and their PSA levels. A cutoff point of ng mL^{-1} or higher was used to decide if a person might have PCa and lower in the case of BPH. It can be seen that most of the PCa samples were classified correctly (above the cutoff) unlike the BPH samples where half of the samples were misclassified as PCa. Hence, high false positivity and poor specificity of PSA for discriminating PCa from BPH which results in unnecessary biopsies.

3.2.5. Comparison of ^1H NMR and PSA

Table 4 represents the confusion matrix calculated using the PSA cutoff of 4 ng/mL . It shows that the PCa samples were mostly predicted correctly unlike BPH samples. This can be seen from Table 5 by the higher and lower precision and recall obtained for PCa and BPH, respectively. By comparing the results obtained from ^1H NMR and PSA (Table 3 and Table 5), it can be clearly seen that NMR outperforms PSA in the prediction of PCa samples and especially in predicting BPH samples as PSA is not very sensitive to BPH. Fig. 5A illustrates the ROC curve for both PSA and RF model, along with their respective Area Under the Curve (AUC) values. The AUC of PSA is 0.84, which demonstrates its diagnostic potential. In comparison, the RF model achieved an even higher AUC of 0.98, highlighting its superior predictive ability. Meanwhile, the black dotted line represents the AUC of a random classifier, which is 0.5, indicating no discriminatory power between the two conditions.

3.2.6. Gleason score and grade group

The Gleason grading system, employed by pathologists to assess prostate cancer aggressiveness, assigns primary and secondary grade patterns scores (1–5) that are combined to generate Gleason scores (2–10). More recently, ISUP grade groups (GG1-GG5) have been introduced [46] and adopted, and they offer predictive value for recurrence-free progression in both biopsies and RP specimens following prostate cancer treatments [47]. The preprocessed spectra were employed to explore the potential of ^1H NMR's in separating between patients diagnosed with BPH and PCa based on their Gleason grade group. While ^1H NMR exhibited promise in this regard (Fig. 5B), further confirmation necessitates a larger sample set.

3.3. Analysis on relative concentrations

Untargeted metabolomics identified 183 metabolites using ASICS, while metabolites with over 50% missing values were excluded. Kruskal-Wallis test revealed that forty-seven metabolites showed significant differences (p -value <0.05) between the two studied groups (Table S1). Moreover, six metabolites illustrated in Fig. 6, exhibited high statistical significance (p -value <0.001). These six metabolites showed a significant increase in relative concentrations, while relative concentrations of hippuric acid (HA) decreased (from BPH to PCa). Similar studies [36,48] showed a decreasing level of HA in urine and serum. However, the levels of hippurate can be influenced by other factors including physical conditions, emotional states, dietary habits, psychiatric disorders, and the presence of certain diseases [49]. Organic acids including acetic acid, pantothenic, benzoic, formic, malonic, pyroglutamic, succinic, 3-hydroxyisovaleric, isocitric, and 2-aminoadipic acid were also different between PCa and BPH. For instance, acetic acid, isocitric acid were down-regulated in PCa compared to BPH and accords with [48] except that they investigated their differences in PCa and HC. Our

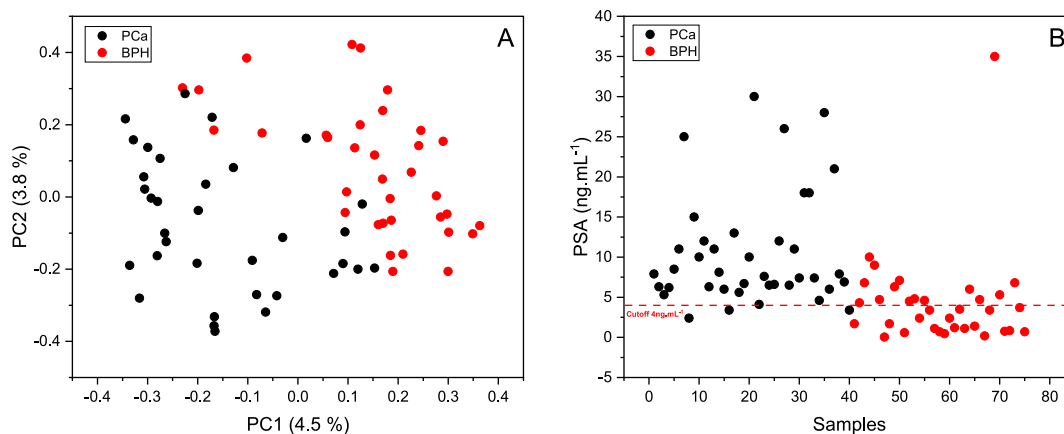
Table 2
Confusion matrix of RF obtained in cross validation.

	Actual	Predicted	
		BPH	PCa
Training set	BPH	22	1
	PCa	2	22
Testing set	BPH	10	0
	PCa	1	10

Table 3

Classification report calculated from Table 1.

		Precision	Recall	Accuracy	F1-score
Training set (Cross-validation)	BPH	0.92	0.93	0.93	0.93
	PCa	0.93	0.92	0.93	0.93
Testing set	BPH	1.00	0.90	0.95	0.95
	PCa	0.92	1.00	0.95	0.96

**Fig. 4.** PCA score plot obtained from URF proximity matrix (A); Scatter plot showing patients, their PSA level and the cutoff used to discriminate PCa and BPH (B)**Table 4**

Confusion matrix of PSA test.

Actual	Predicted	
	BPH	PCa
BPH	37	15
PCa	3	20

Table 5

Classification report calculated from Table 3.

	Precision	Recall	Accuracy	F1-score
BPH	0.87	0.57	0.76	0.80
PCa	0.71	0.93	0.76	0.69

results also revealed a significant increase of 3-hydroxyisovaleric acid in PCa which is in agreement with [50] in prostate cancer cells compared to normal cells. Similarly, succinic acid and malonic acid were also increased in PCa compared to BPH, but, succinic was down-regulated in PCa versus HC [48]. 2-amino adipic acid showed lower concentrations in PCa, unlike its elevated concentration in malignant tissue samples compared to non-malignant ones [51,52]. Benzoic acid was also decreased in patients diagnosed with PCa and accords with [48] which also found to be decreasing in ovarian cancer compared to ovarian benign tumor [53]. Similarly, pantothenic acid was also down-regulated in PCa, but it was found to be increased in PCa compared to BPH tissue extracts [52] and correlated with Gleason score which indicates its potential as a prostate cancer biomarker [54]. Pyroglutamic acid was decreased in PCa compared to BPH samples, however, it was found to be increasing in plasma of PCa samples [55]. Other organic acids were also significantly different such as 4-aminohippuric acid and glyoxylic acid, however, their role in prostate cancer have never been reported before. The amounts of lactose and citrulline were also down-regulated in PCa compared to BPH similar to Ref. [48] where the two metabolites were decreased in PCa compared with HC. Also, tryptophan decreased in PCa and are in accordance with [48,55] in urine and plasma, respectively. The decreased levels of glycine in urine of PCa compared to BPH was also observed and agrees with recent studies [48,56]. The levels of uracil were decreased in PCa, unlike other studies that reported its increased levels in BPH and PCa serum samples compared to HC group [37] as well as in urine samples of PCa and BPH [57]. A significant increase of creatinine in PCa was also observed and is in agreement with previous studies in filtered-serum of PCa compared with BPH and HC [37]. Indoxyl sulfate was also found to be decreased in urine samples of PCa and accords well with [58]. Moreover, this metabolites was also detected and

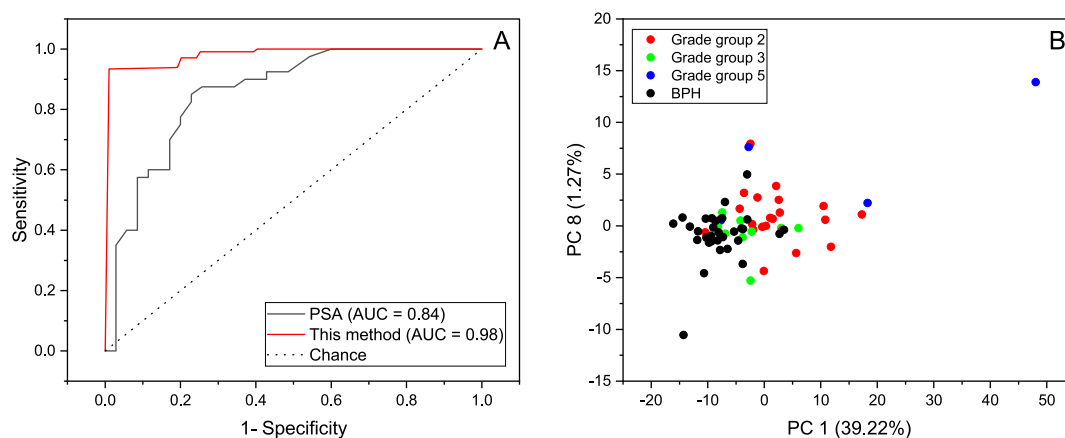


Fig. 5. ROC curve of PSA and the Random Forest model (A), PCA score plot of patients diagnosed with BPH and PCa (Grade Group 2, 3 and 5) (B)

quantified in serum samples, however, its concentration was higher in PCa compared to HC [59]. Metabolic pathway analysis conducted through MetaboAnalyst has identified significant alterations ($p < 0.05$) in six pathways in PCa, highlighting substantial metabolic changes that are intricately linked with the disease's development and progression. These pathways include glyoxylate and dicarboxylate metabolism, galactose metabolism, pantothenate and CoA biosynthesis, citrate cycle (TCA cycle), glutathione metabolism, glycine, serine, and threonine metabolism (Fig. 7).

Fig. 8A and B shows the score plot of the first two PCs obtained from the relative quantifications of the six highly different metabolites present in the two studied groups. It can be observed that the first component (PC1) is the main direction explaining the separation between PCa and BPH. This separation can be explained by the loading used to construct PC1 indicating that glycerol, acetone, levulinic acid, saccharic acid and alpha-hydroxyisobutyric acid are related to the PCa patients while hippuric acid is related to BPH individuals. Another PCA model was built using the whole statistically different metabolites between PCa and BPH. Fig. 8C and D displays score plots of PC1 vs PC2 and PC2 vs PC4 and indicate that PC2 is the most important principal component explaining the separation between the two studied groups. They also show that the model separates well the data compared to the one with only six compounds. This is due to the addition of other statistically different variables or compounds in the new model (Fig. S3 shows the loadings plot of PC1).

4. Discussion

The results presented in this study suggest that ^1H NMR has a strong discriminatory power in separating PCa from BPH. The study utilized various techniques, such as PCA and URF, to demonstrate clear separation between the two groups. These findings are supported by both the score plots, which reveal distinct clusters for PCa and BPH, and the loading plots, which indicate the metabolites contributing to this separation. They also showed that ^1H NMR outperforms the PSA test in terms of specificity, sensitivity, accuracy, and overall predictive ability, highlighting its potential as a more reliable diagnostic tool for distinguishing between PCa and BPH, potentially reducing unnecessary biopsies and improving patient care. This study also explored the potential of ^1H NMR in distinguishing patients diagnosed with PCa and BPH based on their Gleason grade group and showed promising results on the utility of ^1H NMR in grading PCa.

Untargeted metabolomics analysis identified several metabolites that exhibit significant differences between PCa and BPH. Notably, six metabolites— glycerol, acetone, levulinic acid, saccharic acid, alpha-Hydroxyisobutyric acid, and hippuric acid— emerged as highly statistically significant contributors to the differentiation between these two conditions. Furthermore, based on the relative concentrations of these six biomarkers, a PCA model was conducted, which further substantiated the ability to separate PCa from BPH.

The study points out the need for further research and validation, particularly with larger sample sizes. Additionally, exploring the metabolic pathways related to the identified metabolites and understanding their clinical relevance can provide valuable insights into the biology of PCa and BPH.

In conclusion, the findings of this study shed light on the potential of ^1H -NMR-based urine metabolomics as a powerful tool for distinguishing between PCa and BPH. This approach offers improved screening compared to the traditional PSA test; however, further validation is needed.

Ethical statement

The ethical statement is 3/1801/2013 given by Ethics Committee of the Hospital District of Southwest Finland. Participants are required to provide written informed consent.

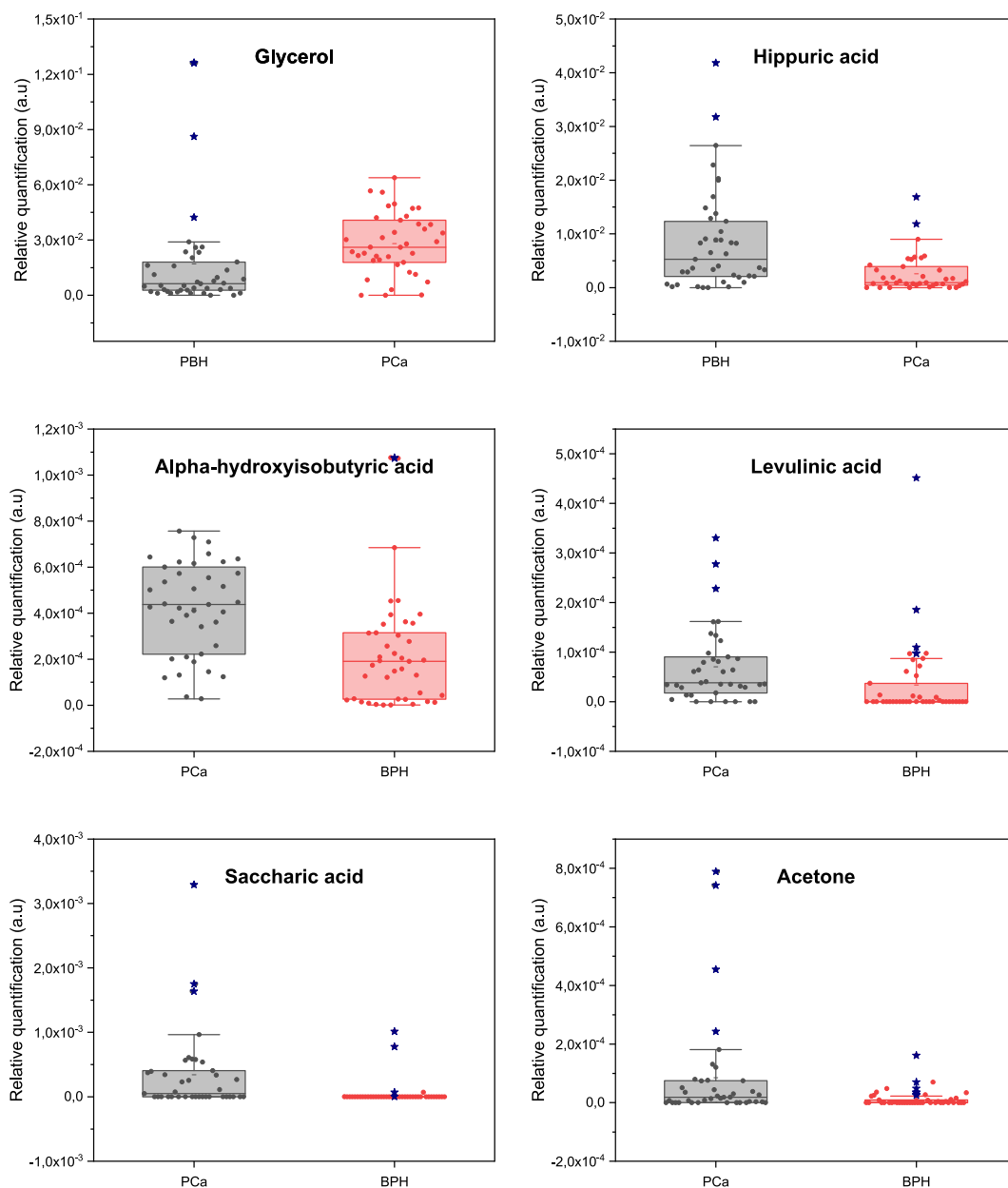


Fig. 6. Relative concentrations of the six high statistically significant metabolites: glycerol, hippuric Acid, alpha-hydroxyisobutyric acid, levulinic acid, saccharic acid, and acetone in PCa vs. BPH.

Funding

T-P H acknowledges the Liv och Hälsa Foundation, Academy of Finland (Grant No. 323240 and 331774), and the European Cooperation in Science and Technology (COST Action: CA17120) for financial support.

Data availability

The data will be provided upon request due to its large size.

CRedit authorship contribution statement

Mohammed Zniber: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation,

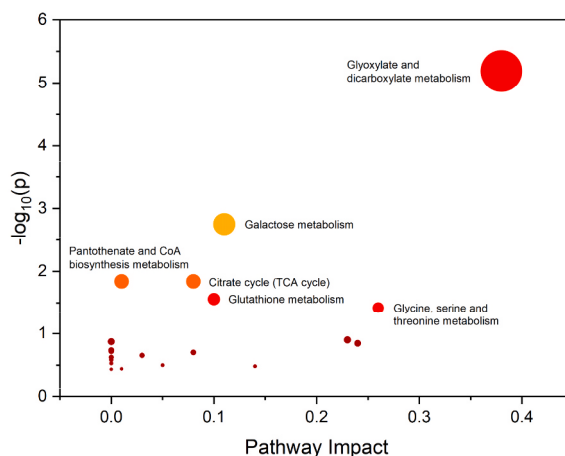


Fig. 7. Pathway analysis conducted with MetaboAnalyst 6.0 identified six significant metabolic pathways.

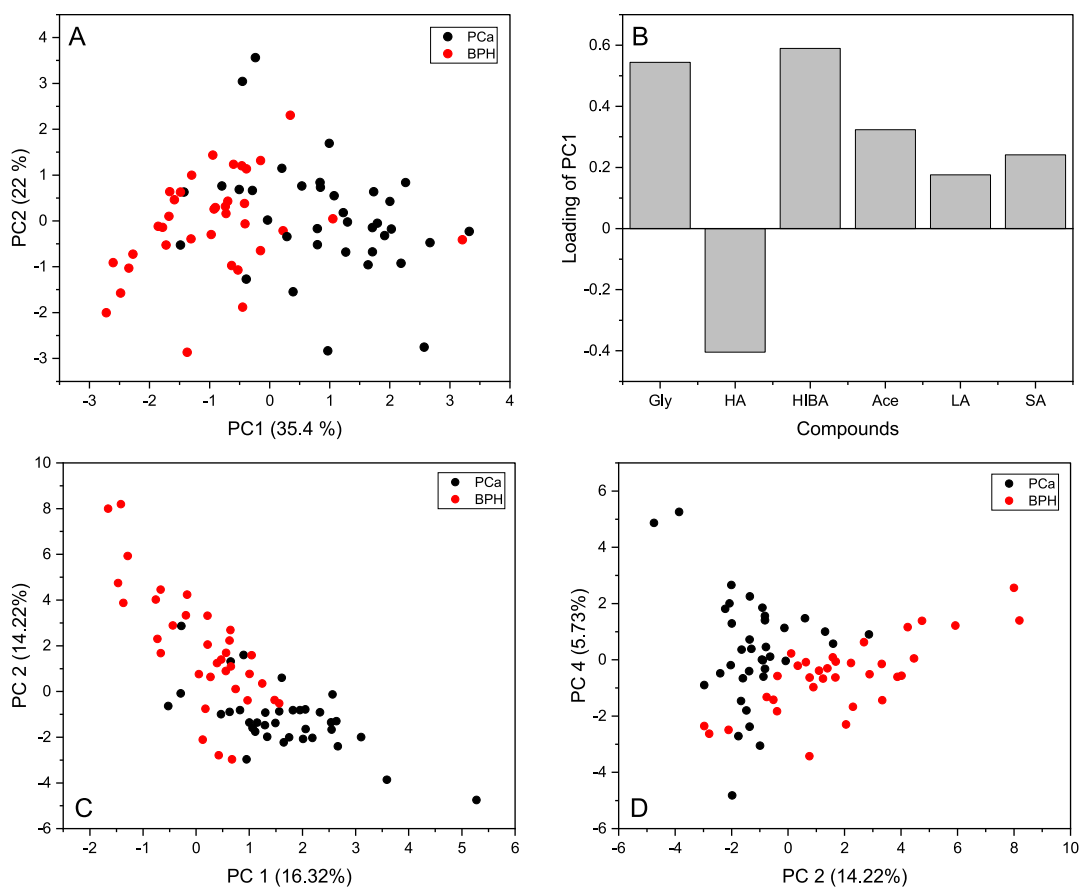


Fig. 8. Score plot of PC1 vs PC2 and loading of PC1 calculated using only six highly different metabolites (A and B); Score plots of PC1 vs PC2 and PC2 vs PC4 obtained from all different metabolites (C and D)

Conceptualization. **Tarja Lamminen**: Writing – review & editing, Resources, Methodology. **Pekka Taimen**: Writing – review & editing, Validation, Formal analysis. **Peter J. Boström**: Writing – review & editing, Resources, Formal analysis, Conceptualization. **Tan-Phat Huynh**: Writing – review & editing, Validation, Supervision, Resources, Funding acquisition, Formal analysis, Conceptualization.

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.

Acknowledgements

The authors express their gratitude towards the personnel of Turku University Hospital and Auria Biobank, as well as all the collaborators in Turku Prostate Cancer Consortium (TPCC). They also acknowledge Pauliina Toivonen for her technical support in sample collection and preparation. Additionally, the authors extend a special appreciation to all the patients who took part in this investigation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2024.e28949>.

References

- [1] J. Ferlay, M. Colombet, I. Soerjomataram, D.M. Parkin, M. Piñeros, A. Znaor, F. Bray, Cancer statistics for the year 2020: an overview, *Int. J. Cancer* 149 (2021) 778–789, <https://doi.org/10.1002/IJC.33588>.
- [2] N.M. Gades, D.J. Jacobson, C.J. Girman, R.O. Roberts, M.M. Lieber, S.J. Jacobsen, Prevalence of conditions potentially associated with lower urinary tract symptoms in men, *BJU Int.* 95 (2005) 549–553, <https://doi.org/10.1111/J.1464-410X.2005.05337.X>.
- [3] J. Ferlay, P. Autier, M. Boniol, M. Heanue, M. Colombet, P. Boyle, Estimates of the cancer incidence and mortality in Europe in 2006, *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 18 (2007) 581–592, <https://doi.org/10.1093/ANNONC/MDL498>.
- [4] M. Colonna, A. Danzon, P. Delafosse, N. Mitton, S. Bara, A.M. Bouvier, O. Ganry, A.V. Guizard, G. Launoy, F. Molinie, E.A. Sauleau, C. Schwartz, M. Velten, P. Grosclaude, B. Tretarre, Cancer prevalence in France: time trend, situation in 2002 and extrapolation to 2012, *Eur. J. Cancer* 44 (2008) 115–122, <https://doi.org/10.1016/J.EJCA.2007.10.022>.
- [5] A. Alcaraz, P. Hammerer, A. Tubaro, F.H. Schröder, R. Castro, Is There Evidence of a relationship between benign prostatic hyperplasia and prostate cancer? Findings of a literature review, *Eur. Urol.* 55 (2009) 864–875, <https://doi.org/10.1016/J.EURURO.2008.11.011>.
- [6] D.G. Bostwick, H.B. Burke, D. Djakiew, S. Euling, S.M. Ho, J. Landolph, H. Morrison, B. Sonawane, T. Shifflett, D.J. Waters, B. Timms, Human prostate cancer risk factors, *Cancer* 101 (2004) 2371–2490, <https://doi.org/10.1002/CNCR.20408>.
- [7] P.J.M. Hendriksen, N.F.J. Dits, K. Kokame, A. Veldhoven, W.M. Van Weerden, C.H. Bangma, J. Trapman, G. Jenster, Evolution of the androgen receptor pathway during progression of prostate cancer, *Cancer Res.* 66 (2006) 5012–5020, <https://doi.org/10.1158/0008-5472.CAN-05-3082>.
- [8] D.J. Tindall, R.S. Rittmaster, The Rationale for inhibiting 5 α -reductase Isoenzymes in the Prevention and treatment of prostate cancer, *J. Urol.* 179 (2008) 1235–1242, <https://doi.org/10.1016/J.JURO.2007.11.033>.
- [9] C. Carson, R. Rittmaster, The role of dihydrotestosterone in benign prostatic hyperplasia, *Urology* 61 (2003) 2–7, [https://doi.org/10.1016/S0090-4295\(03\)00045-1](https://doi.org/10.1016/S0090-4295(03)00045-1).
- [10] K. Chandra, S. Al-Harathi, F. Almulhim, A.-H. Emwas, Ł. Jaremko, M. Jaremko, The robust {NMR} toolbox for metabolomics, *Mol. Omi.* 17 (2021) 719–724, <https://doi.org/10.1039/d1mo00118c>.
- [11] O.T. Okotie, K.A. Roehl, M. Han, S. Loeb, S.N. Gashti, W.J. Catalona, Characteristics of prostate cancer detected by digital rectal examination only, *Urology* 70 (2007) 1117–1120, <https://doi.org/10.1016/J.UROLOGY.2007.07.019>.
- [12] 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) 1–12, <https://doi.org/10.1007/s11306-017-1194-y>.
- [13] D.S. Wishart, L.L. Cheng, V. Copié, A.S. Edison, H.R. Eghbalnia, J.C. Hoch, G.J. Gouveia, W. Pathmasiri, R. Powers, T.B. Schock, L.W. Sumner, M. Uchimiya, NMR and metabolomics—a Roadmap for the Future, *Metab.* 12 (12) (2022) 678, <https://doi.org/10.3390/METABO12080678>, 2022) 678.
- [14] A.H. Emwas, R. Roy, R.T. McKay, L. Tenori, E. Saccenti, G.A. Nagana Gowda, D. Raftery, F. Alahmari, L. Jaremko, M. Jaremko, D.S. Wishart, NMR spectroscopy for metabolomics research, *Metab* 9 (123 9) (2019) 123, <https://doi.org/10.3390/METABO9070123> (2019).
- [15] A. Zhang, H. Sun, X. Wang, Power of metabolomics in biomarker discovery and mining mechanisms of obesity, *Obes. Rev.* 14 (2013) 344–349, <https://doi.org/10.1111/OBR.12011>.
- [16] A. Zhang, H. Sun, G. Yan, P. Wang, Y. Han, X. Wang, Metabolomics in diagnosis and biomarker discovery of colorectal cancer, *Cancer Lett.* 345 (2014) 17–20, <https://doi.org/10.1016/j.canlet.2013.11.011>.
- [17] Y. Lin, C. Ma, C. Liu, Z. Wang, J. Yang, X. Liu, Z. Shen, R. Wu, NMR-based fecal metabolomics fingerprinting as predictors of earlier diagnosis in patients with colorectal cancer, *Oncotarget* 7 (2016) 29454, <https://doi.org/10.18632/oncotarget.8762>.
- [18] E.R. Kim, H.N. Kwon, H. Nam, J.J. Kim, S. Park, Y.H. Kim, Urine-NMR metabolomics for screening of advanced colorectal adenoma and early stage colorectal cancer, *Sci. Reports* 91 (9) (2019) 1–10, <https://doi.org/10.1038/s41598-019-41216-y>, 2019.
- [19] U. Sharma, N.R. Jagannathan, Breast cancer metabolomics using NMR, *Methods Mol. Biol.* 2037 (2019) 195–213, https://doi.org/10.1007/978-1-4939-9690-2_12.
- [20] L. Lécuyer, A.V. Bala, M. Deschasaux, N. Bouchemal, M.N. Triba, M.P. Vasson, A. Rossary, A. Demidov, P. Galan, S. Hercberg, V. Partula, L. Le Moyec, B. Srour, T. Fiolet, P. Latino-Martel, E. Kesse-Guyot, P. Savarin, M. Touvier, NMR metabolomic signatures reveal predictive plasma metabolites associated with long-term risk of developing breast cancer, *Int. J. Epidemiol.* 47 (2018) 484–494, <https://doi.org/10.1093/ije/dyx271>.
- [21] S. Tiziani, V. Lopes, U.L. Günther, Early stage diagnosis of oral cancer using 1H NMR-based metabolomics, *Neoplasia* 11 (2009), <https://doi.org/10.1593/neo.81396>, 269-IN10.
- [22] L. Michálková, Š. Horník, J. Sýkora, L. Habartová, V. Setnická, Diagnosis of pancreatic cancer via 1 H NMR metabolomics of human plasma, *Analyst* 143 (2018) 5974–5978, <https://doi.org/10.1039/C8AN01310A>.
- [23] M. Cao, L. Zhao, H. Chen, W. Xue, D. Lin, NMR-Based metabolomic analysis of human Bladder cancer, *Anal. Sci.* 28 (2012) 451–456, <https://doi.org/10.2116/analsci.28.451>.
- [24] C. Bruzzzone, A. Loizaga-Iriarte, A. Loizaga-Iriarte, P. Sánchez-Mosquera, R. Gil-Redondo, I. Astobiza, I. Astobiza, T. Diercks, A.R. Cortazar, A.R. Cortazar, A. Ugalde-Olano, A. Ugalde-Olano, H. Schäfer, F.J. Blanco, F.J. Blanco, M. Unda, M. Unda, C. Cannet, M. Spraul, J.M. Mato, N. Embade, A. Carracedo, A. Carracedo, A. Carracedo, O. Millet, 1H NMR-based urine metabolomics reveals Signs of enhanced Carbon and nitrogen Recycling in prostate cancer, *J. Proteome Res.* 19 (2020) 2419–2428, <https://doi.org/10.1021/acs.jproteome.0c00091>.

- [25] R.R. Reddy, N.R. Jagannathan, Potential of nuclear magnetic resonance metabolomics in the study of prostate cancer, *Indian J. Urol* 38 (2022) 99–109, <https://doi.org/10.4103/iju.iju.416.21>.
- [26] A.R. Lima, J. Pinto, D. Barros-Silva, C. Jerónimo, R. Henrique, M. de L. Bastos, M. Carvalho, P. Guedes Pinho, New findings on urinary prostate cancer metabolome through combined GC-MS and 1H NMR analytical platforms, *Metabolomics* 16 (2020) 1–9, <https://doi.org/10.1007/s11306-020-01691-1>.
- [27] N. Gómez-Cebrián, M. García-Flores, J. Rubio-Briones, J. López-Guerrero, A. Pineda-Lucena, L. Puchades-Carrasco, Targeted metabolomics analyses reveal specific metabolic alterations in high-grade prostate cancer patients, *J. Proteome Res.* 19 (2020) 4082–4092, <https://doi.org/10.1002/pros.23198>.
- [28] H. Zheng, A. Cai, Q. Zhou, P. Xu, L. Zhao, C. Li, B. Dong, H. Gao, Optimal preprocessing of serum and urine metabolomic data fusion for staging prostate cancer through design of experiment, *Anal. Chim. Acta* 991 (2017) 68–75, <https://doi.org/10.1016/J.ACA.2017.09.019>.
- [29] H. Zheng, Y. Zhu, X. Shao, A. Cai, B. Dong, W. Xue, H. Gao, Distinct metabolic signatures of hormone-sensitive and Castration-Resistant prostate cancer revealed by a 1H NMR-based metabolomics of biopsy tissue, *J. Proteome Res.* 19 (2020) 3741–3749, <https://doi.org/10.1021/acs.jproteome.0c00282>.
- [30] 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. Reports* 81 (8) (2018) 1–12, <https://doi.org/10.1038/s41598-018-23177-w>, 2018.
- [31] P.R. Braadland, G. Giskeødegård, E. Sandmark, H. Bertilsson, L.R. Euceda, A.F. Hansen, I.J. Guldvik, K.M. Selnaes, 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>.
- [32] G.F. Giskeødegård, H. Bertilsson, K.M. Selnaes, 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) e62375, <https://doi.org/10.1371/journal.pone.0062375>.
- [33] B. Madhu, G.L. Shaw, A.Y. Warren, D.E. Neal, J.R. Griffiths, Response of degarelix treatment in human prostate cancer monitored by HR-MAS 1H NMR spectroscopy, *Metabolomics* 12 (2016) 1–11, <https://doi.org/10.1007/s11306-016-1055-0>.
- [34] A.F. Hansen, E. Sandmark, M.B. Rye, A.J. Wright, H. Bertilsson, E. Richardsen, T. Viset, A.M. Bofin, A. Angelsen, K.M. Selnaes, T.F. Bathen, M.B. Tessem, Presence of TMPRSS2-ERG is associated with alterations of the metabolic profile in human prostate cancer, *Oncotarget* 7 (2016) 42071, <https://doi.org/10.18632/oncotarget.9817>.
- [35] C.S. Clendinning, D.A. Gaul, M.E. Monge, R.S. Arnold, A.S. Edison, J.A. Petros, F.M. Fernández, Preoperative metabolic signatures of prostate cancer recurrence following radical prostatectomy, *J. Proteome Res.* 18 (2019) 1316–1327, <https://doi.org/10.1021/acs.jproteome.8b00926>.
- [36] D. Kumar, A. Gupta, A. Mandhani, S.N. Sankhwar, Metabolomics-derived prostate cancer biomarkers: Fact or fiction? *J. Proteome Res.* 14 (2015) 1455–1464, <https://doi.org/10.1021/pr5011108>.
- [37] 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>.
- [38] G.F. Giskeødegård, A.F. Hansen, H. Bertilsson, S.V. Gonzalez, K.A. Kristiansen, P. Bruheim, S.A. Mjøs, A. Angelsen, T.F. Bathen, M.B. Tessem, Metabolic markers in blood can separate prostate cancer from benign prostatic hyperplasia, *Br. J. Cancer* 11312 (113) (2015) 1712–1719, <https://doi.org/10.1038/bjc.2015.411> (2015).
- [39] M.J. Roberts, R.S. Richards, C.W.K. Chow, M. Buck, J. Xaxley, M.F. Lavin, H.J. Schirra, R.A. Gardiner, Seminal plasma enables selection and monitoring of active surveillance candidates using nuclear magnetic resonance-based metabolomics: a preliminary investigation, *Prostate Int* 5 (2017) 149–157, <https://doi.org/10.1016/J.PRNIL.2017.03.005>.
- [40] M.J. Jeppesen, R. Powers, Multiplatform untargeted metabolomics, *Magn. Reson. Chem.* 61 (2023) 628–653, <https://doi.org/10.1002/MRC.5350>.
- [41] A.H. Emwas, C. Luchinat, P. Turano, L. Tenori, R. Roy, R.M. Salek, D. Ryan, J.S. Merzaban, R. Kaddurah-Daouk, A.C. Zeri, G.A. Nagana Gowda, D. Raftery, Y. Wang, L. Brennan, D.S. Wishart, Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review, *Metabolomics* 11 (2015) 872–894, <https://doi.org/10.1007/S11306-014-0746-7>.
- [42] M. Martin, B. Legat, J. Leenders, J. Vanwinsberghe, R. Rousseau, B. Boulanger, P.H.C. Eilers, P. De Tullio, B. Govaerts, PepsNMR for 1H NMR metabolomic data pre-processing, *Anal. Chim. Acta* 1019 (2018) 1–13, <https://doi.org/10.1016/J.ACA.2018.02.067>.
- [43] G. Lefort, L. Liaubet, C. Canlet, P. Tardivel, M.C. Pèrè, H. Quesnel, A. Paris, N. Iannuccelli, N. Vialaneix, R. Servien, J. Kelso, ASICS: an R package for a whole analysis workflow of 1D 1H NMR spectra, *Bioinformatics* 35 (2019) 4356–4363, <https://doi.org/10.1093/BIOINFORMATICS/BTZ248>.
- [44] S.M. Kohl, M.S. Klein, J. Hochrein, P.J. Oefner, R. Spang, W. Gronwald, State-of-the-art data normalization methods improve NMR-based metabolomic analysis, *Metabolomics* 8 (2012) 146, <https://doi.org/10.1007/S11306-011-0350-Z>.
- [45] L. Plantier, A. Smolinska, R. Fijten, M. Flamant, J. Dallinga, J.J. Mercadier, D. Pachen, M.P. d'Ortho, F.J. van Schooten, B. Crestani, A.W. Boots, The use of exhaled air analysis in discriminating interstitial lung diseases: a pilot study, *Respir. Res.* 23 (2022), <https://doi.org/10.1186/S12931-021-01923-5>.
- [46] P.M. Pierorazio, P.C. Walsh, A.W. Partin, J.I. Epstein, Prognostic Gleason grade grouping: data based on the modified Gleason scoring system, *BJU Int.* 111 (2013) 753–760, <https://doi.org/10.1111/j.1464-410X.2012.11611.x>.
- [47] J.I. Epstein, M.J. Zelefsky, D.D. Sjoberg, J.B. Nelson, L. Egevad, C. Magi-Galluzzi, A.J. Vickers, A.V. Parwani, V.E. Reuter, S.W. Fine, J.A. Eastham, P. Wiklund, M. Han, C.A. Reddy, J.P. Ciezki, T. Nyberg, E.A. Klein, A Contemporary prostate cancer grading system: a validated alternative to the gleason score, *Eur. Urol.* 69 (2016) 428–435, <https://doi.org/10.1016/j.eururo.2015.06.046>.
- [48] W. Struck-Lewicka, M. Kordalewska, R. Bujak, A. Yumba Mpanga, M. Markuszewski, J. Jacyna, M. Matuszewski, R. Kalisz, 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>.
- [49] Fourier transform infrared for noninvasive optical diagnosis of oral, oropharyngeal, and laryngeal cancer | Elsevier Enhanced Reader, (n.d.), <https://doi.org/10.1016/j.trsl.2013.09.006>.
- [50] A.R. Lima, A.M. Araújo, J. Pinto, C. Jerónimo, R. Henrique, M.D.L. Bastos, M. Carvalho, P.G. De Pinho, GC-MS-Based Endometabolome analysis differentiates prostate cancer from normal prostate cells, *Metab* 8 (2018) 23, <https://doi.org/10.3390/metabo8010023>, 23 8 (2018).
- [51] K. Jung, R. Reszka, B. Kamlage, B. Bethan, C. Stephan, M. Lein, G. Kristiansen, Tissue metabolite profiling identifies differentiating and prognostic biomarkers for prostate carcinoma, *Int. J. Cancer* 133 (2013) 2914–2924, <https://doi.org/10.1002/ijc.28303>.
- [52] 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>.
- [53] E.C. Eroglu, U. Kucukgoz Gulec, M.A. Vardar, S. Paydas, GC-MS based metabolite fingerprinting of serous ovarian carcinoma and benign ovarian tumor, *Eur. J. Mass Spectrom.* 28 (2022) 12–24, <https://doi.org/10.1177/14690667221098520>.
- [54] S. Meller, H.A. Meyer, B. Bethan, D. Dietrich, S.G. Maldonado, M. Lein, M. Montani, R. Reszka, P. Schatz, E. Peter, C. Stephan, K. Jung, B. Kamlage, G. Kristiansen, Integration of tissue metabolomics, transcriptomics and immunohistochemistry reveals ERG- and gleason score-specific metabolomic alterations in prostate cancer, *Oncotarget* 7 (2016) 1421, <https://doi.org/10.18632/oncotarget.6370>.
- [55] P.A. Markin, A. Brito, N. Moskaleva, E.V. Lartsova, Y.V. Shpot, Y.V. Lerner, V.Y. Mikhajlov, N.V. Potoldykova, D.V. Enikeev, M.R. La Frano, S.A. Appolonova, Plasma metabolomic profile in prostatic intraepithelial neoplasia and prostate cancer and associations with the prostate-specific antigen and the Gleason score, *Metabolomics* 16 (2020) 1–14, <https://doi.org/10.1007/s11306-020-01694-y>.
- [56] H. Wu, T. Liu, C. Ma, R. Xue, C. Deng, H. Zeng, X. Shen, GC/MS-based metabolomic approach to validate the role of urinary sarcosine and target biomarkers for human prostate cancer by microwave-assisted derivatization, *Anal. Bioanal. Chem.* 401 (2011) 635–646, <https://doi.org/10.1007/s00216-011-5098-9>.

- [57] 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>.
- [58] Pradeep Kumar, Rajeev Kumar, Virendra Kumar, Senthil S. Kumaran, Sanjay Thulkar, Siddhartha D. Gupta, Naranamangalam R. Jagannathan, Study of altered urinary metabolism in prostate cancer patients using NMR metabolomics, *Int. Soc. Magn. Reson. Med.* (2018). <https://archive.ismrm.org/2018/4008.html>. (Accessed 30 August 2022).
- [59] X. Zang, C.M. Jones, T.Q. Long, M.E. Monge, M. Zhou, L.D. Walker, R. Mezencev, A. Gray, J.F. McDonald, F.M. Fernández, Feasibility of detecting prostate cancer by ultraperformance liquid chromatography-mass spectrometry serum metabolomics, *J. Proteome Res.* 13 (2014) 3444–3454, <https://doi.org/10.1021/pr500409q>.