Data-Driven Soft Independent Modeling of Class Analogy in Paper Spray Ionization Mass Spectrometry-Based Metabolomics for Rapid Detection of Prostate Cancer

Article *in* Analytical Chemistry · January 2022

DOI: 10.1021/acs.analchem.1c04004

CITATION
1

READS
55

7 authors, including:

Frederico Garcia Pinto
Universidade Federal de Viçosa (UFV)
46 PUBLICATIONS **723** CITATIONS
SEE PROFILE

Iqbal Mahmud
University of Texas MD Anderson Cancer Center
60 PUBLICATIONS **307** CITATIONS
SEE PROFILE

Ademar Domingos Viagem Máquina
Universidade Federal de Uberlândia (UFU)
11 PUBLICATIONS **50** CITATIONS
SEE PROFILE

Anizia Fausta Furtado Durans
Universidade Federal de Uberlândia (UFU)
1 PUBLICATION **1** CITATION
SEE PROFILE

analytical chemistry

pubs.acs.org/ac

Technical Note

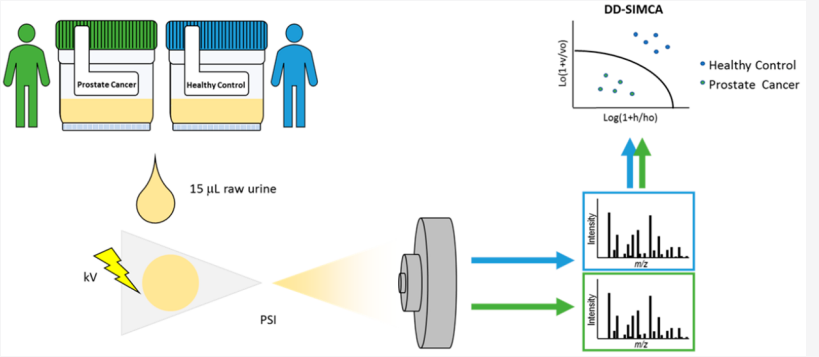
Data-Driven Soft Independent Modeling of Class Analogy in Paper Spray Ionization Mass Spectrometry-Based Metabolomics for Rapid Detection of Prostate Cancer

Frederico G. Pinto, Iqbal Mahmud, Vanessa Y. Rubio, Ademar Domingos Viagem Máquina, Anízia Fausta Furtado Durans, Waldomiro Borges Neto, and Timothy J. Garrett*

Cite This: https://doi.org/10.1021/acs.analchem.1c04004

Read Online

ACCESS | Metrics & More | Article Recommendations



ABSTRACT: Sensitive, rapid, and meaningful diagnostic tools for prostate cancer (PC) screening are urgently needed. Paper spray ionization mass spectrometry (PSI-MS) is an emerging rapid technology for detecting biomarker and disease diagnoses. Due to lack of chromatography and difficulties in employing tandem MS, PSI-MS-based untargeted metabolomics often suffers from increased ion suppression and subsequent feature detection, affecting chemometric methods for disease classification. This study first evaluated the data-driven soft independent modeling of class analogy (DD-SIMCA) model to analyze PSI-MS-based global metabolomics of a urine data matrix to classify PC. The efficiency of DD-SIMCA was analyzed based on the sensitivity and specificity parameters that showed 100% correct classification of the training set, based on only PC and test set samples, based on normal and PC. This analytical methodology is easy to interpret and efficient and does not require any prior information from the healthy individual. This new application of DD-SIMCA in PSI-MS-based metabolomics for PC disease classification could also be extended to other diseases and opens a rapid strategy to discriminate against health problems.

INTRODUCTION

Prostate cancer (PC) is the second most commonly diagnosed cancer in men worldwide, with an estimated 248,530 new cases and 34,130 new deaths in 2021.[1] The main barrier to reduce PC mortality is the early diagnosis of disease.[2] The current most used clinical tests applied for PC diagnosis are the prostate specific antigen (PSA) blood test and digital rectal examination (DRE).[2] However, these tests present numerous drawbacks.[2] The PSA blood test has low specificity, presenting a high rate of false positives or false negatives, and biopsies are an unpleasant procedure for patient which can produce several health complications such as infections, incontinence, and erectile dysfunction.[2,3] In addition, current biopsies frequently miss cancer detection due to tumor heterogeneity.[4] Thus, the development of efficient technology using a noninvasive approach for rapid and reproducible PC diagnosis is an urgent clinical need.

The emerging field of metabolomics such as liquid biopsies, child malnutrition, and COVID-19, in which a large number of small molecules (metabolites) from biological organisms are detected in both a qualitative and quantitative manner in a single step, has immense potential for early diagnosis of many diseases.[5−8] Metabolic profiling provides an overview of the cellular and physiological status about an individual being highly recommended in clinical diagnostics.[9] Metabolomics has been increasingly used for diagnosis of several diseases in humans, such as cancer,[10−13] COVID-19,[14,15] Alzheimer's,[16]

Received: September 14, 2021
Accepted: January 7, 2022

Parkinson's,[17,18] diabetes,[19] child malnutrition,[6] fatigue, depression, and pain.[20]

Liquid chromatography−mass spectrometry (LC−MS) is one of the fastest-growing techniques for metabolomics analysis. Although LC−MS is an excellent match, it presents some drawbacks for large-scale metabolomics projects such as time-consuming analysis (20−30 min per injection), laborious sample preparation, use of a large amount of solvent over the acquisition time, and potential high cost,[21] except for the microflow-based LC-MS system as an instance that minimizes a significant amount of sample and solvent consumption.[22] Thus, to overcome the drawbacks to apply the mass spectrometry-based metabolomics approach for large-scale studies and then establish the MS technique as a tool for clinical tests used for disease diagnosis, paper spray ionization-mass spectrometry (PSI-MS) appears to be a promising alternative.[5]

PSI is an ambient MS ionization method that produces gas-phase ions in the ambient air, removing chromatographic separation and minimizing prior sample preparation.[5] PSI-MS promotes a simple straightforward sample introduction and analysis with low cost and high speed. In this approach, the sample is placed directly onto a rigid sheet of chromatography paper, allowed to dry, and then analyzed by MS after application of a designated spray solvent for elution and a high voltage for ionization.[5] In addition, once the sample dries upon initial application onto the paper, the sample does not require any special handling or refrigeration. Clinical applications of PSI-MS are at an early stage but show promise as a technique for disease diagnosis.[5,23] PSI-MS has the potential to be used remotely as a small volume of a liquid sample can be applied onto a simple cellulose filter paper and then shipped for analysis to a centralized location equipped with an MS for rapid diagnosis of a given disease.[24] Storage of samples as dried spots improves sample stability at room temperature in many cases.[25] PSI-MS has been shown to be a high-throughput and suitable technique for measurement of many clinical analytes in biofluids in less than 1 min of acquisition time. This makes it an optimal tool for integration into a rapid metabolomic profiling assay.

Several studies highlight the use of chemometric methods in metabolomic analyses. Raja et al.[26] conducted a metabolomics study based on ^{1}H NMR for cancer targeting and metabolic engineering, using traditional chemometric methods such as principal component analysis (PC) and partial least-squares discriminant analysis (PLS-DA) to examine metabolic differences. Deev et al.[27] developed prostate cancer screening methodologies using chemometric processing of GC−MS profiles obtained in the headspace above urine samples, where they applied the PC for exploratory data analysis, the methods k-nearest neighbors (KNN), and the PLS-DA to make the models allowing attribution of a urine sample to the particular class. Drivelos et al.[28] developed methodologies for identifying the geographical origin and botanical type honey authentication through elemental metabolomics via chemometrics. However, the methods used by these authors have low efficiency when they are used to discriminate samples whose information is not incorporated in the calibration process, as demonstrated in detail by Rodionova et al., in 2016.[29] Thus, the applicability of these methods for noninvasive diagnostic of prostate cancer using urine liquid biopsy is reduced due to the existence of a variety of unknown samples.

As an alternative, the modeling of classes or classifying methods[30,31] are widely used for several purposes[32−35] due to the following advantages: for their calibration, only information from the samples of the class of interest is necessary; have good efficiency in the classification of unknown samples; and are easily interpreted. The data driven soft independent modeling of class analogy (DD-SIMCA) model is an example of these methods that is developed using only information from samples of the class of interest to later detect whether the information from a new sample resembles this class or not.[30,31] DD-SIMCA performs modeling of each class by using principal components (PCs) of the data for each class. For the unknown sample classifications, DD-SIMCA uses a collection of the individual class models where PLS-DA modeling is based on a single global model.[30,31] Moreover, DD-SIMCA decomposes the data by principal component analysis (PCA) to calculate the score distances and orthogonal distances for each object in order to establish two tolerance thresholds: the acceptance area for a given significance level and the outliers area.[30,31] Due to its functionality, DD-SIMCA is used as an authentication technique for various purposes, such as for the detection of melamine and sucrose as adulterants in powdered milk;[36] for monitoring the adulteration of the B10 blend of diesel and crambe biodiesel;[35] for authentication and identification of adulterants in virgin coconut oil;[37] and for detecting the presence of adulterants and confirming the provenance of edible bird's nests produced in Malaysia.[38]

Previously, we reported discrimination of PC from normal by PLS-DA modeling based on their PSI spectra acquired by PSI-MS.[5] However, there are no reports in the literature that use the DD-SIMCA method for PSI-MS-based untargeted metabolomics. Herein, we report DD-SIMCA analysis of the PSI spectra obtained for PC versus normal urine samples. The DD-SIMCA classification model was evaluated using 80 urine samples from 40 PC and 40 healthy individuals. The efficiency of this methodology was analyzed based on the sensitivity and specificity parameters that showed 100% correct classification of the training and test set samples. Based on our understanding between the classification methods such as DD-SIMCA and PLS-DA, we suggest that DD-SIMCA can be an alternative approach of PLS-DA in the case of analyzing multiple sample classes including unknown samples.

EXPERIMENTAL SECTION

Materials. Acetonitrile (ACN), methanol (MeOH), water (H_2O), and formic acid (FA) solvents were Optima LC/MS grade solvents purchased from Thermo Fisher Scientific (Fairlawn, NJ).

Urine Samples. Urine samples were obtained from the Biospecimen Core of the SPORE network in Prostate Cancer at Northwestern University (P50 CA180995). Urine samples from 40 healthy individuals with no prior medical history of cancer (NPC) were obtained from the Life Study (University of Florida, Gainesville, FL). The institutional review board (IRB) of the Florida Hospital approved sample use.

Raw Urine Direct Analysis by PSI-HRMS. Raw urine samples were directly analyzed using PSI-HRMS as previously described.[5] Briefly, 15 μL of raw urine was deposited onto precut triangular paper Velox cartridges obtained from Prosolia, Inc. (Indianapolis, IN). A 3D-printed bracing device was used (Prosolia, Inc.) for reproducibility in sample dispensing and positioning. PSI cartridges were loaded into the Prosolia Velox 360 PSI automated unit for MS analysis. A

https://doi.org/10.1021/acs.analchem.1c04004
Anal. Chem. XXXX, XXX, XXX−XXX

volume of 80 μL of 3:7 MeOH:H_2O containing 0.1% FA (v/v/v) was used as the wetting and spray solvent. Pooled samples from each group were used to check reproducibility.

PSI-HRMS Analysis. PSI-HRMS mass spectra were obtained using a Thermo Scientific Q Exactive MS. Data acquisition was performed in full scan mode (positive ionization mode) with a mass range of 70−1,000 m/z at 140,000 mass resolution. Other PSI-HRMS instrument parameters for data acquisition were used as previously described.[5,39,40] Metabolomics data files were converted to mzXML file format using RawConverter.[41] Mzmine 2.0 was employed for peak picking and feature alignment.[42] Non-detected species with zero intensity were replaced with half of the minimum value of all features detected.[42] Data were filtered by blank feature filtration (BFF) to remove features with ≥10% signal intensity contribution from the background signal.

Chemometrics Analysis. The chemometric analyses involving DD-SIMCA for modeling paper spray ionization mass spectrometry data were performed with MATLAB software version R2017a (Mathworks, Inc.) using a graphical interface developed by Zontov et al.[31] To perform the multivariate procedures, the paper spray ionization mass spectra data were organized in an ordered array of rows and columns, where each row corresponds to one sample "i" and each column corresponds to one variable "j", in which i = 1,2,3,...,80 (The first 40 samples are normal, and the remaining 40 are PC samples.) and j = 1,2,3,...,783 (number of features after BFF). Data from 25 samples of interest class (PC) are extracted from this matrix to constitute the training set, called matrix $\mathbf{X}$. The data were preprocessed using sum normalization, that is, dividing each region of the spectrum (bin) by the respective sum total of the spectrum integration area, according to eq 1[43]

$$\mathbf{X}_{ij(\text{bin\_nor})} = \frac{\mathbf{X}_{ij}}{\sum_{j=1}^{j} |\mathbf{X}_{ij}|} \tag{1}$$

where $\mathbf{X}_{ij(\text{bin\_nor})}$ is bin normalized by the sum, $\mathbf{X}_{ij}$ is the original bin, and $\sum_{j=1}^{783}\mathbf{X}_{ij}$ is the sum of integration areas for each sample column.

Construction of the Data Driven Soft Independent Modeling of Class Analogy (DD-SIMCA) Model. The DD-SIMCA model construction procedures are well developed in the literature.[29,44−49] The procedure consists of two steps: First, the $(i \times j)$ data matrix $\mathbf{X}$ is decomposed by the principal component analysis (PCA) according to eq 2[31]

$$\mathbf{X} = \mathbf{TP^t} + \mathbf{E} \tag{2}$$

where $\mathbf{T} = \{\mathbf{t_{ia}}\}$ is the $[(i \times A)$, i.e., $(40 \times A)]$ score matrix, $\mathbf{P} = \{\mathbf{p_{ia}}\}$ is the $[(j \times A)$, i.e., $(783 \times A)]$ loadings matrix, $\mathbf{E} = \{\mathbf{e_{ij}}\}$ is the $[(i \times j)$, i.e., $(80 \times 783)]$ matrix of residuals, and A is the number of principal components.

In the second step, for each object I = 1,...,I from the training set, two distances are calculated according to eq 3. They are score distances (SD), h_i, and the orthogonal distance (OD), v_i[45]

$$h_i = t_i^t(\mathbf{T^tT})^{-1}t_i = \sum_{i=1}^{t}\frac{t_{ia}^2}{\lambda_a}, \qquad v_i = \sum_{j=1}^{J} e_{ij}^2 \tag{3}$$

where λ_a, a = 1,..,A are the diagonal elements of matrix $\mathbf{T^tT}$ = $\mathbf{\Lambda}$ = diag(λ_1,...,λ_A).

SD represents the position of a sample within the score space, while OD characterizes a sample distance to the score space. DD-SIMCA adds the possibility of estimation of the data-driven distribution parameters, and it was shown that distributions of both distances are well approximated by the scaled chi-squared distribution,[45] as follows

$$N_h\frac{h}{h_0}\alpha\chi^2(N_h) \quad N_v\frac{v}{v_0}\alpha\chi^2(N_v) \tag{4}$$

where v_0 and h_0 are the scaling factors, and N_h and N_v are the numbers of the degrees of freedom (DoF). These parameters are considered unknown and estimated using the method of moments (MM) as described previously.[45] The statistics c, called the total distance, is calculated by the following:[50]

$$c = N_h\frac{h}{h_0} + N_v\frac{v}{v_0} \propto \chi^2(N_h + N_v) \tag{5}$$

It is used to generate the decision rules on the acceptance area for a given value. Any decision rule is determined by an inequality eq 6. The third step defines the acceptance area or thresholds for the target class (samples of interest class, in this case PC). Given the type I error, α, the acceptance area is determined as the following[47]

$$c \le c_{crit}(\alpha) \tag{6}$$

where

$$c_{crit} = \chi^{-2}(1 - \alpha, N_h + N_v) \tag{7}$$

is the $(1-\alpha)$ quantile of the chi-squared distribution with N_h + N_v.[50] To calculate the type II error, β, we should assume that an alternative class is available, as follows[47]

$$\beta = Pr\left\{\chi'^2(N_h + N_v, s) < \frac{c_{crit}}{c_0'}\right\} \tag{8}$$

where c_{crit} is defined in eq 7, and χ'^2 is the noncentral chi-squared distribution. Parameters c_0' and s are found on the type II error in the SIMCA method.[47]

After this step, the model is ready for classification of new samples and can be represented by an acceptance area in the orthogonal (h/h_0) against the score distance (v/v_0) (acceptance plot) defined for the given α value. The α value specifies a type I error, i.e., a share of the false-negative decisions. Each sample from the training set is characterized by its position in the acceptance plot and has a status of either a "regular" sample, i.e., a sample attributed to the target class, or an "extreme" sample, which is located out of the acceptance area and is determined to be an alien (a nonmember). Besides that, a second cutoff level is determined to be used as the outlier border constructed for the given α value. This value specifies the probability that at least one regular object from the data set will be erroneously considered as an outlier. The outlier area depends on the size of a training set, I. For a specific α value, the greater I is, the farther the outlier area will be. For a moderate data set, a common value of α is equal to 0.05 or 0.01.[31,44]

A special extreme plot is also created to demonstrate the dependence of the observed number of the extreme samples versus the theoretically expected values, calculated as $n = \alpha I$.[50] This dependence is shown in the plot together with the tolerance limits calculated according to eq 9.[31] The extreme

https://doi.org/10.1021/acs.analchem.1c04004
Anal. Chem. XXXX, XXX, XXX−XXX

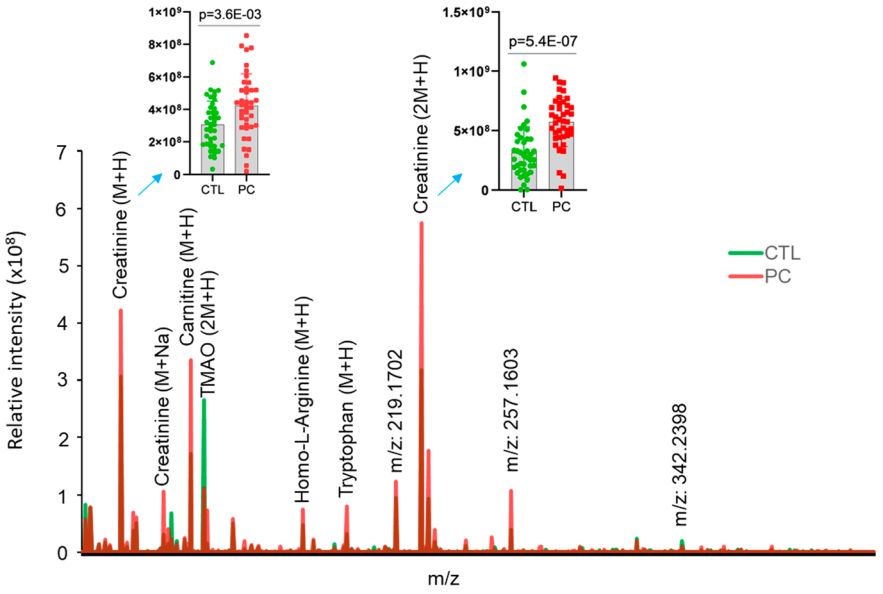


Figure 1. Stacked group averaged spectra for prostate cancer (PC) and healthy individuals as control (CTL). We show m/z 70−400 where the most intense and populated peak differences were observed between PC and CTL. The red line indicates PC average spectra (n = 40), and the green line indicates CTL spectra (n = 40). Annotated metabolites/features and the corresponding adduct are shown along with the respective peak. A bar graph with a whisker plot shows the creatinine level as an example of a metabolite across the scans as shown in the insets. The p-value was calculated using Student's t test and two-tailed unequal variance.

plot helps to analyze the quality of the classification model for the chosen number of principal components:[50]

$$t_\alpha = n \pm 2\sqrt{\alpha(1-\alpha)I} = n \pm 2\sqrt{n(1-\alpha/I)} \quad (9)$$

Evaluation of the DD-SIMCA Model. The efficiency of the model is analyzed based on the sensitivity and specificity parameters. Generally, sensitivity is determined for the training data set by eq 10 but can also be determined for the test data set if this set is built only with new target class samples. Specificity is determined for the test data set by eq 11, when this set is constructed with samples of the target class '*regular*' and alien class '*external objects*'.[31] Thus, samples correctly classified in their respective classes are considered true positive (TP) samples.

$$\text{Sensitivity} = \frac{\text{Samples} - \text{Extremes}}{\text{Samples}} \times 100\% \quad (10)$$

$$\text{Specificity} = \frac{\text{TP samples}}{\text{Test Samples}} \times 100\% \quad (11)$$

RESULTS AND DISCUSSION

PSI-HRMS Spectra of Urine Samples Show Difference between Normal and PC Individuals. Typical PSI-HRMS spectra of urine samples from normal and PC individuals acquired are presented in Figure 1. The typical PSI-HRMS molecular profiling between normal and PC urine groups showed a considerable difference. Figure 1 displays the average spectral comparison (m/z 70−400 region) with the presence of peaks related to low-molecular-weight species that are biologically relevant. Some highly differentiated metabolites or features such as creatinine ($[M + H]^+$ adduct, m/z 114.0667) or ($[2M + H]^+$ adduct, m/z 227.1249), carnitine ($[M + H]^+$ adduct, m/z 144.1018), TMAO ($[2M + H]^+$ adduct, m/z 151.1441), and tryptophan ($[M + H]^+$ adduct, m/z 205.0977) are shown in the urine spectra (Figure 1). Metabolites were putatively identified based on m/z exact mass (molecular weight tolerance of 5 ppm) matching to the HMDB database (https://hmdb.ca/spectra/ms/search)[51] in accordance with the metabolomics standards initiative.[52]

To better understand the variability of a metabolite or feature during data collection, we showed a most differentiated metabolite such as creatinine as an example and observed that the level of creatinine is significantly intense in PC samples which is inconsistent with the average spectral comparison (Figure 1). These observations suggest that PSI-MS-based global metabolomics may delineate disease differentiation.

DD-SIMCA Modeling with the Training Set. We observed hundreds of metabolic features in PSI-MS-based global urine metabolomics which included both known and unknown molecules. However, the challenges in PSI-MS-based metabolomics are to properly classify groups associated with disease. Here, we leveraged the DD-SIMCA method for the first time in PSI-MS-based metabolomics to demonstrate correct classification of PC alone or from normal individuals. The DD-SIMCA model was developed using 4 principal components. This number of principal components showed the best results, explaining 93.96% of the total variance of training data with $\propto$ = 0.01, as shown in Table 1. The sensitivity of the model in training was 1, which is equivalent to saying that no sample of the target class (PC) was classified as extreme or as outliers. The sum of the normalized score distance and orthogonal distance defined the type of acceptance area in ≪chi-squared≫, i.e., a triangular area

Table 1. DD-SIMCA Modeling Result for the Training Set of Samples

alpha	principal component	explained variance	samples	extremes	outlier	sensitivity
0.01	4	93.96%	25	0	0	1

https://doi.org/10.1021/acs.analchem.1c04004
Anal. Chem. XXXX, XXX, XXX−XXX

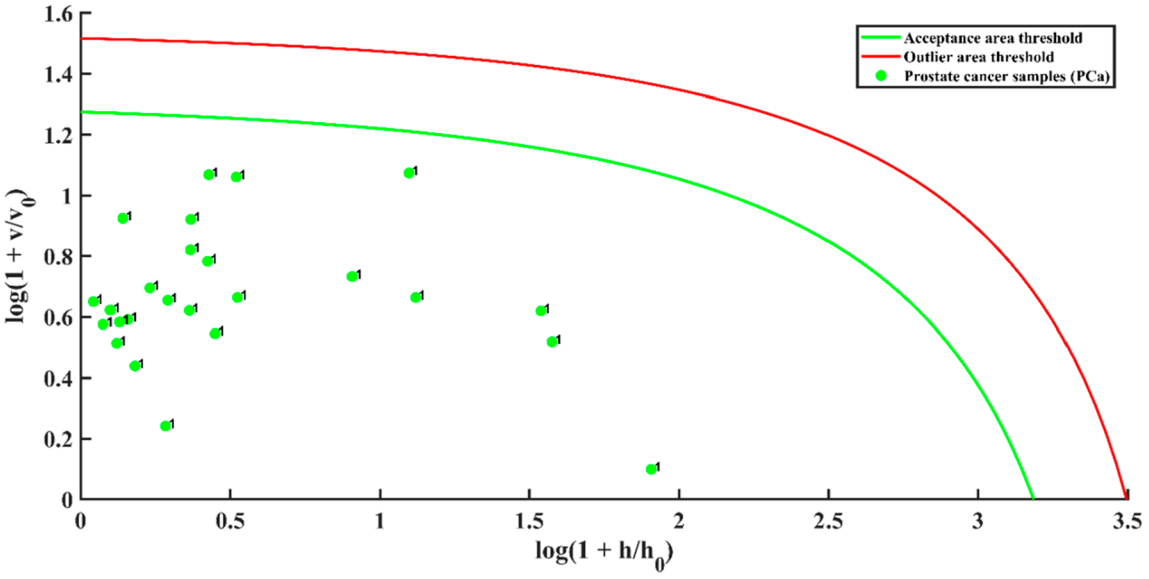


Figure 2. Acceptance plot of samples with prostate cancer by the DD-SIMCA model in the training process. The acceptance area is the threshold for $\propto = 0.05$, and the outlier delimitation is the limit for $\propto = 0.01$.

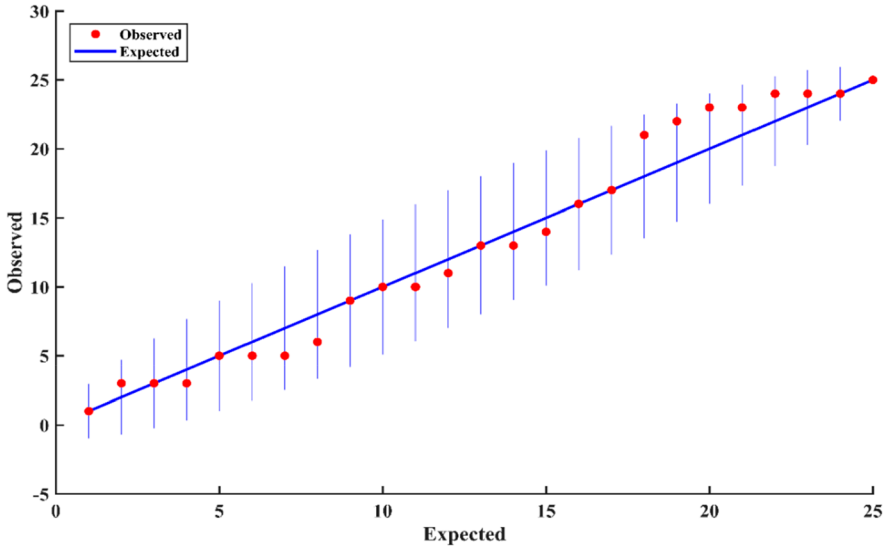


Figure 3. Outlier verification plot in samples with prostate cancer by the DD-SIMCA model in the training process.

(Figure 2). This data shows the PC sample classification obtained by the DD-SIMCA model, where it can be observed that no PC sample was found in the extreme's area (space between the green and red lines) or in the outlier's area (space above the red line). This data suggests that all PC samples were classified correctly because they are within the acceptance area, delimited by the green line with $\propto = 0.01$.

The verification of outlier samples in the training set was based on the analysis of the dependence on the observed number of the extreme samples versus the theoretically expected values, presented in the special extreme plot (Figure 3). In this data, it is observed that all samples fell within the tolerance area (shown in vertical lines), demonstrating that the training set does not contain outliers based on the DD-SIMCA model.

The evaluation of the model was performed with a test set composed of samples of target class (PC) and urine samples from 40 healthy individuals with no prior medical history of cancer (NPC). The results are presented in Table 2. The prediction specificity was 1, which suggests that all samples in the test set were correctly classified into their respective classes. In accordance with Table 1, further analysis using the acceptance plot showed correct classification of PC (Figure 4).

Table 2. Results of the Evaluation of the Test Set by the DD-SIMCA Model

alpha	principal components	samples	external objects	specificity
0.01	4	55	40	1

The acceptance plot shows the test set sample classification, where it can be observed that all PC samples are within the acceptance area, representing a correct classification. The urine samples of NPC were correctly classified as alien because they are outside the acceptance area (green line), indicating that there are differences between spectral profiles of these samples

https://doi.org/10.1021/acs.analchem.1c04004
Anal. Chem. XXXX, XXX, XXX−XXX

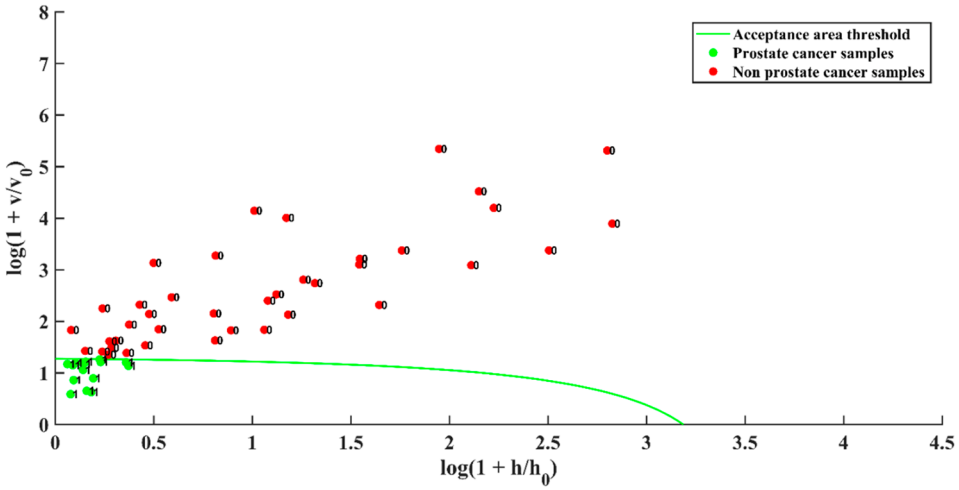


Figure 4. Acceptance plot of samples with and without prostate cancer by the DD-SIMCA model in the test process. The green line indicates the acceptance area threshold.

with the target class (PC) samples (Figure 4). Altogether, DD-SIMCA appears to be a robust model in PSI-MS-based metabolomics for accurate classification of a disease such as PC.

CONCLUSION

The application of paper spray ionization mass spectrometry associated with the DD-SIMCA chemometric method allowed the development of a classification methodology for a noninvasive diagnosis of prostate cancer using only relevant information from the samples of interest (PC). The efficiency of this methodology was analyzed based on the sensitivity and specificity parameters that showed 100% correct classification of the training and test set samples. This analytical methodology is easy to interpret, efficient, and viable, because for its development, it does not require prior information from the urine samples of healthy individuals without a previous medical history of cancer.

AUTHOR INFORMATION

Corresponding Author

Timothy J. Garrett − *Southeast Center for Integrated Metabolomics, Clinical and Translational Science Institute and Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, Florida 32610, United States;* orcid.org/0000-0003-1623-009X; Phone: (352) 273-5050; Email: tgarrett@ufl.edu

Authors

Frederico G. Pinto − *Institute of Chemistry, Federal University of Viçosa, Rio Paranaíba, Minas Gerais 36570-900, Brazil;* orcid.org/0000-0003-3882-4657

Iqbal Mahmud − *Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, Florida 32610, United States;* orcid.org/0000-0001-7729-8122

Vanessa Y. Rubio − *Department of Chemistry, University of Florida, Gainesville, Florida 32603, United States*

Ademar Domingos Viagem Máquina − *Institute of Chemistry, Federal University of Uberlândia, Uberlândia, Minas Gerais 38400-902, Brazil;* orcid.org/0000-0001-6330-0455

Anízia Fausta Furtado Durans − *Institute of Chemistry, Federal University of Uberlândia, Uberlândia, Minas Gerais 38400-902, Brazil*

Waldomiro Borges Neto − *Institute of Chemistry, Federal University of Uberlândia, Uberlândia, Minas Gerais 38400-902, Brazil*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c04004

Author Contributions

F.G.P. and I.M. contributed equally. F.G.P. and I.M. analyzed and interpreted the data and wrote and revised the manuscript. V.Y.R. worked on the methodology and data acquisition. A.D.V.M., A.F.F.D., and W.V.N. worked on data analysis and revision. T.J.G. supervised the study, reviewed, and edited the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the National Institutes of Health Grant U24DK097209. We also acknowledge Prosolia for funding and donation of the paperspray source.

REFERENCES

(1) Siegel, R. L.; Miller, K. D.; Fuchs, H. E.; Jemal, A. *CA. Cancer J. Clin.* **2021**, *71* (1), 7−33.
(2) *Prostate-Specific Antigen (PSA) Test*; National Cancer Institute. https://www.cancer.gov/types/prostate/psa-fact-sheet (accessed 2020-04-01).
(3) Prostate Cancer Screening (PDQ®)−Patient Version; National Cancer Institute. https://www.cancer.gov/types/prostate/patient/prostate-screening-pdq (accessed 2020-04-01).
(4) Heitzer, E.; Haque, I. S.; Roberts, C. E. S.; Speicher, M. R. *Nat. Rev. Genet.* **2019**, *20* (2), 71−88.
(5) Mahmud, I.; Pinto, F. G.; Rubio, V. Y.; Lee, B.; Pavlovich, C. P.; Perera, R. J.; Garrett, T. J. *Anal. Chem.* **2021**, *93* (22), 7774−7780.
(6) Mahmud, I.; Kabir, M.; Haque, R.; Garrett, T. J. *Anal. Chem.* **2019**, *91* (23), 14784−14791.
(7) Mahmud, I.; Garrett, T. J. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 2013.
(8) Domenick, T. M.; Gill, E. L.; Vedam-Mai, V.; Yost, R. A. *Anal. Chem.* **2021**, *93* (1), 546−566.

(9) Wishart, D. S. *Nat. Rev. Drug Discovery* **2016**, *15* (7), 473−484.
(10) Lee, B.; Mahmud, I.; Marchica, J.; Dereziński, P.; Qi, F.; Wang, F.; Joshi, P.; Valerio, F.; Rivera, I.; Patel, V.; Pavlovich, C. P.; Garrett, T. J.; Schroth, G. P.; Sun, Y.; Perera, R. J. *Sci. Rep.* **2020**, *10* (1), 1−17.
(11) Lynch Kelly, D.; Farharfar, N.; Starkweather, A.; Garrett, T. J.; Yao, Y.; Wingard, J. R.; Mahmud, I.; Menzies, V.; Patel, P.; Alabasi, K. M.; Lyon, D. *Biol. Blood Marrow Transplant.* **2020**, *26*, 1803.
(12) Mahmud, I.; Tian, G.; Wang, J.; Lewis, J.; Waddell, A.; Lydon, M. L.; Zhao, L. Y.; Li, J.-L.; Purayil, H. T.; Huo, Z.; Daaka, Y.; Garrett, T. J.; Liao, D. DAXX-SREBP Axis Promotes Oncogenic Lipogenesis and Tumorigenesis. 2021, 2020.12.31.424997. https://www.biorxiv.org/content/10.1101/2020.12.31.424997v1 (accessed 2022-01-14), DOI: 10.1101/2020.12.31.424997.
(13) Tan, S. K.; Mahmud, I.; Fontanesi, F.; Puchowicz, M.; Neumann, C. K. A.; Griswold, A. J.; Patel, R.; Dispagna, M.; Ahmed, H. H.; Gonzalgo, M. L.; Brown, J. M.; Garrett, T. J.; Welford, S. M. *Cancer Discovery* **2021**, *11* (8), 2072−2093.
(14) Garrett, T. J.; Coatsworth, H.; Mahmud, I.; Hamerly, T.; Stephenson, C. J.; Yazd, H. S.; Ayers, J.; Miller, M. R.; Lednicky, J. A.; Dinglasan, R. R. Niclosamide Reverses SARS-CoV-2 Control of Lipophagy. 2021, 2021.07.11.451951. https://www.biorxiv.org/content/10.1101/2021.07.11.451951v1 (accessed 2022-01-14), DOI: 10.1101/2021.07.11.451951.
(15) Sindelar, M.; Stancliffe, E.; Schwaiger-Haber, M.; Anbukumar, D. S.; Albrecht, R. A.; Liu, W.-C.; Travis, K. A.; García-Sastre, A.; Shriver, L. P.; Patti, G. J. Longitudinal Metabolomics of Human Plasma Reveals Robust Prognostic Markers of COVID-19 Disease Severity. *medRxiv*. 2021, 2021.02.05.21251173. https://www.medrxiv.org/content/10.1101/2021.02.05.21251173v1 (accessed 2022-01-14), DOI: 10.1101/2021.02.05.21251173.
(16) Wilkins, J. M.; Trushina, E. *Front. Neurol.* **2018**, *8*, 719.
(17) Gill, E. L.; Koelmel, J. P.; Yost, R. A.; Okun, M. S.; Vedam-Mai, V.; Garrett, T. J. *Anal. Chem.* **2018**, *90* (5), 2979−2986.
(18) Gill, E. L.; Koelmel, J. P.; Meke, L.; Yost, R. A.; Garrett, T. J.; Okun, M. S.; Flores, C.; Vedam-Mai, V. *J. Proteome Res.* **2020**, *19* (1), 424−431.
(19) O'Kell, A. L.; Garrett, T. J.; Wasserfall, C.; Atkinson, M. A. *Sci. Rep.* **2017**, *7* (1), 9467.
(20) Menzies, V.; Starkweather, A.; Yao, Y.; Kelly, D. L.; Garrett, T. J.; Yang, G.; Booker, S.; Swift-Scanlan, T.; Mahmud, I.; Lyon, D. E. *Biol. Res. Nurs.* **2021**, *23*, 119.
(21) Pinto, F. G.; Mahmud, I.; Harmon, T. A.; Rubio, V. Y.; Garrett, T. J. *J. Proteome Res.* **2020**, *19*, 2080.
(22) Wang, H.; Bennett, P. *Bioanalysis* **2013**, *5* (10), 1249−1267.
(23) Wang, H.; Manicke, N. E.; Yang, Q.; Zheng, L.; Shi, R.; Cooks, R. G.; Ouyang, Z. *Anal. Chem.* **2011**, *83* (4), 1197−1201.
(24) Manicke, N. E.; Bills, B. J.; Zhang, C. *Bioanalysis* **2016**, *8* (6), 589−606.
(25) Frey, B. S.; Damon, D. E.; Badu-Tawiah, A. K. *Mass Spectrom. Rev.* **2020**, *39* (4), 336−370.
(26) Raja, G.; Jung, Y.; Jung, S. H.; Kim, T. J. *Process Biochem.* **2020**, *99*, 112−122.
(27) Deev, V.; Solovieva, S.; Andreev, E.; Protoshchak, V.; Karpushchenko, E.; Sleptsov, A.; Kartsova, L.; Bessonova, E.; Legin, A.; Kirsanov, D. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **2020**, *1155*, 122298.
(28) Drivelos, S. A.; Danezis, G. P.; Halagarda, M.; Popek, S.; Georgiou, C. A. *Food Chem.* **2021**, *338*, 127936.
(29) Rodionova, O. Y.; Titova, A. V.; Pomerantsev, A. L. *TrAC - Trends in Analytical Chemistry* **2016**, *78*, 17−22.
(30) Rodionova, O. Ye; Titova, A. V.; Pomerantsev, A. L. *TrAC Trends Anal. Chem.* **2016**, *78*, 17−22.
(31) Zontov, Y. V.; Rodionova, O. Ye; Kucheryavskiy, S. V.; Pomerantsev, A. L. *Chemom. Intell. Lab. Syst.* **2017**, *167*, 23−28.
(32) de Santana, F. B.; Borges Neto, W.; Poppi, R. J. *Food Chem.* **2019**, *293*, 323−332.
(33) Gupta, K.; Bhavsar, A.; Sao, A. K. *Comput. Biol. Med.* **2019**, *111*, 103328.
(34) Pantazi, X. E.; Moshou, D.; Tamouridou, A. A. *Comput. Electron. Agric.* **2019**, *156*, 96−104.
(35) Máquina, A. D. V.; Sitoe, B. V.; Santana, F. B. d.; Santos, D. Q.; Neto, W. B. *Anal. Lett.* **2021**, *54*, 790.
(36) Mazivila, S. J.; Páscoa, R. N. M. J.; Castro, R. C.; Ribeiro, D. S. M.; Santos, J. L. M. *Talanta* **2020**, *216* (March), 120937.
(37) Neves, M. D. G.; Poppi, R. J. *Talanta* **2020**, *219*, 121338.
(38) Adenan, M. N. H.; Moosa, S.; Muhammad, S. A.; Abrahim, A.; Jandric, Z.; Islam, M.; Rodionova, O.; Pomerantsev, A.; Perston, B.; Cannavan, A.; Kelly, S. D.; Othman, Z.; Abdullah Salim, N. A.; Sharif, Z.; Ismail, F. *Forensic Chem.* **2020**, *17*, 100197.
(39) Chamberlain, C. A.; Hatch, M.; Garrett, T. J. *Metabolites* **2021**, *11* (5), 308.
(40) Chamberlain, C. A.; Rubio, V. Y.; Garrett, T. J. *Anal. Chem.* **2019**, *91* (8), 4964−4968.
(41) He, L.; Diedrich, J.; Chu, Y.-Y.; Yates, J. R. *Anal. Chem.* **2015**, *87* (22), 11361−11367.
(42) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *BMC Bioinformatics* **2010**, *11* (1), 395.
(43) Ferreira, M. M. C. *Quimiometria: Conceitos, Métodos e Aplicações*; Ed.a da Unicamp: Campinas-SP, 2015; DOI: 10.7476/9788526814714.
(44) Rodionova, O. Ye; Oliveri, P.; Pomerantsev, A. L. *Chemom. Intell. Lab. Syst.* **2016**, *159* (5), 89−96.
(45) Pomerantsev, A. L. *J. Chemom.* **2008**, *22* (11−12), 601−609.
(46) Chen, T.; Morris, J.; Martin, E. *Chemom. Intell. Lab. Syst.* **2007**, *87* (1), 69−71.
(47) Pomerantsev, A. L.; Rodionova, O. Y. *J. Chemom.* **2014**, *28* (6), 518−522.
(48) Oliveri, P. *Anal. Chim. Acta* **2017**, *982* (5), 9−19.
(49) Moya, M. M.; Koch, M. W.; Hostetler, L. D. One-Class Classifier Networks for Target Recognition Applications. In *International Neural Network Society; Proc. Word Congr.*; Neural Networks: Portland, OR, 1993; pp 797−801.
(50) Pomerantsev, A. L.; Rodionova, O. Y. *J. Chemom.* **2014**, *28* (5), 429−438.
(51) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M.-A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau; David, D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; MacInnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. *Nucleic Acids Res.* **2007**, *35* (Database issue), D521−D526.
(52) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. *Metabolomics Off. J. Metabolomic Soc.* **2007**, *3* (3), 211−221.

https://doi.org/10.1021/acs.analchem.1c04004
Anal. Chem. XXXX, XXX, XXX−XXX