wjpmr, 2021,7(1), 188-195

SJIF Impact Factor: 5.922

WORLD JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.wjpmr.com

Research Article

ISSN 2455-3301

WJPMR

PLASMA METABOLOMIC TARGETED MASS SPECTROMETRY-BASED ANALYSIS FOR THE DISCOVERY OF POTENTIAL BIOMARKERS IN COLORECTAL CANCER PATIENTS

Carlos Augusto Rodrigues Veo[1,2*], Marcos Vinicius Denadai[1,2], Sandra Morini Silva[1,2], M. Gordian Adam[3], Udo Muller[3], Simon Schafferer[3], Ismael D. C. G. Da Silva[4], Isabella Alves Brunetti[5] and Delcio Matos[1]

[1]Health Evidence-Based Postgraduate Program– Escola Paulista de Medicina, Sao Paulo Federal University, São Paulo, Brazil.

[2]Barretos Cancer Hospital, Fundação Pio XII, Barretos, SP, Brazil.

[3]Biocrates Life Sciences AG, Innsbruck, Austria.

[4]Department of Gynecology and Obstetrics– Escola Paulista de Medicina, Sao Paulo Federal University, São Paulo, Brazil.

[5]LEEA- Laboratory of Ecotoxicology and Pesticide Efficacy.

***Corresponding Author: Carlos Augusto Rodrigues Veo**

Health Evidence-Based Postgraduate Program– Escola Paulista de Medicina, Sao Paulo Federal University, São Paulo, Brazil.

DOI: https://doi.org/10.17605/OSF.IO/URVJ3

Article Received on 06/11/2020 **Article Revised on 26/11/2020** **Article Accepted on 16/12/2020**

ABSTRACT

Metabolomics is an emerging analytical tool in oncology research which allows the identification of specific biomarkers and helps improve clinical results. The present study aimed to identify metabolic conditions in patients with colorectal cancer. This investigation used mass spectrometry to analyze the plasma of patients with colorectal adenocarcinoma. In the study, 85 plasma samples were analyzed and metabolite extraction was performed using kits. A targeted metabolomic approach with the Biocrates Absolute IDQ® p180 kit was used to quantify metabolites of various biochemical classes. The statistical analysis included data normalization and quality control, principal component analysis, hierarchical cluster analysis, univariate statistics with significance testing, and analysis of fold changes. The comparison between plasma and tissue samples showed differences in only a few analytes. The application of more restricted statistics with correction to control the false discovery rate (Benjamini-Hochberg adjustment) often did not show significant changes in the different groups of patients. Several sphingomyelins were significantly different in more than one group. The analysis without adjustment yielded significantly different metabolites, mainly sphingomyelins and phospholipids.

KEYWORDS: Colorectal cancer; biomarkers; spectrometry-based.

INTRODUCTION

Colorectal cancer (CRC) is one of the most highly prevalent types of cancer, causing high mortality rates in developed and developing countries. According to the American Cancer Society, about 1.7 million new cancer cases and around 600,000 deaths from the disease occurred in the United States in 2015, among which almost 100,000 new cases and 50,000 deaths were attributed to CRC.[1,2]

The need for non-invasive, specific and accurate screening methods for the early identification of CRC has led several researchers to use molecular techniques such as genomics, proteomics, and more recently, metabolomics to identify serum biomarkers.[3] Da Silva et al. in 2018 identified metabolic changes consistent with inborn-like errors that define a continuum from normal controls to elevated risk for invasive breast cancer and other types of adenocarcinoma.[4]

Metabolomics is the study of low-molecular weight metabolites in biological matrices. It is downstream of genomics, transcriptomics, and proteomics; thus, changes at the metabolomic level not only reflect genomic and proteomic alterations but also the effects of environmental factors. Differences in metabolite levels between a patient and a healthy individual are used to identify metabolic pathways for a particular disease. In recent years, metabolomic studies have been successfully used to identify biomarkers and altered metabolic pathways in various types of cancer, including gastric,[5] brain,[6] breast,[7] and lung cancer.[8]

Serum has been the specimen of choice for the identification of biomarkers as it reflects the metabolite profiles at the moment of sample collection. Changes observed in metabolite levels relative to normal profiles may serve as important indicators of disease states. Metabolomics has been used to identify early biomarkers

of CRC. Because altered metabolites can be influenced by biological and environmental factors, it is important to determine common differentiating metabolites, identified in previous studies, which act as potential biomarkers for CRC.[3]

The present study aimed to identify potential metabolomic changes using mass spectrometry to analyze plasma samples of patients with colorectal adenocarcinoma.

MATERIAL AND METHODS

Study samples

In this study, 85 plasma samples from patients with CRC were analyzed. Patients with a scheduled colorectal resection after a CRC diagnosis were recruited. Sample collection was performed within 30 minutes after the beginning of the surgery. Patients undergoing chemotherapy and/or radiotherapy prior to surgery were excluded from the study.

All patients provided an informed written consent, and the research plan was approved by the Barretos Cancer Hospital – Ethics Committee of Pio XII Foundation under no. 1,570,292 and by the Ethics Committee of UNIFESP under no. 1,686,367.

Clinical and laboratory data were collected, including patient information (age, gender, relapse, survival) as well as tumor sample characteristics (tumor site, differentiation grade, venous invasion, lymph node invasion, perineural invasion, peritumoral infiltration, TNM stage, and AJCC staging). The samples were stored at -80°C immediately following collection until they were processed for plasma analysis. The collection was checked by the study technician in the pathology laboratory and in the pre-operative room before each resection procedure. All samples were shipped to Biocrates Life Sciences AG, Innsbruck, Austria, for plasma analysis and statistical data analyses.

Metabolite profiling

For plasma analysis, the samples were homogenized using Precellys® with ethanol phosphate buffer. For metabolite level measurements, the samples were centrifuged and the supernatants were used for analysis. The Biocrates Absolute IDQ® p180 kit was used to quantify metabolites of different biochemical classes, including amino acids, biogenic amines, acylcarnitines, glycerophospholipids, sphingolipids, and monosaccharides. The p180 kit is a widely used targeted metabolomics platform that produces highly reproducible results.[9]

The fully automated assay was based on phenylisothiocyanate (PITC) derivatization in the presence of internal standards followed by FIA-MS/MS (flow injection analysis tandem mass spectrometry) to detect acylcarnitines, (lyso-) phosphatidylcholines, sphingomyelins, and hexoses. LC-MS/MS (liquid chromatography tandem mass spectrometry) was performed to detect amino acids and biogenic amines using a SCIEX 4000 QTRAP® (SCIEX, Darmstadt, Germany) or a Waters XEVO™ TQMS (Waters, Vienna, Austria) instrument with electrospray ionization (ESI). The experimental metabolomic measurement technique has been described in detail in patent US 2007/0004044.[10]

For the LC−MS/MS quantification, the metabolite concentrations were calculated using stable isotope dilution and seven-point calibration curves. For the FIA-MS/MS analysis, the metabolites were quantified using a one-point internal standard calibration and were isotope-corrected.

Statistical Analysis

Different statistical methods were applied to identify the differences in metabolite levels. The analysis included data normalization and quality control, principal component analysis, hierarchical cluster analysis, univariate statistics with significance testing, multivariate analysis, and analysis of fold changes. In addition, the main biochemical pathways affected were analyzed. Figure 1 shows the analysis workflow.

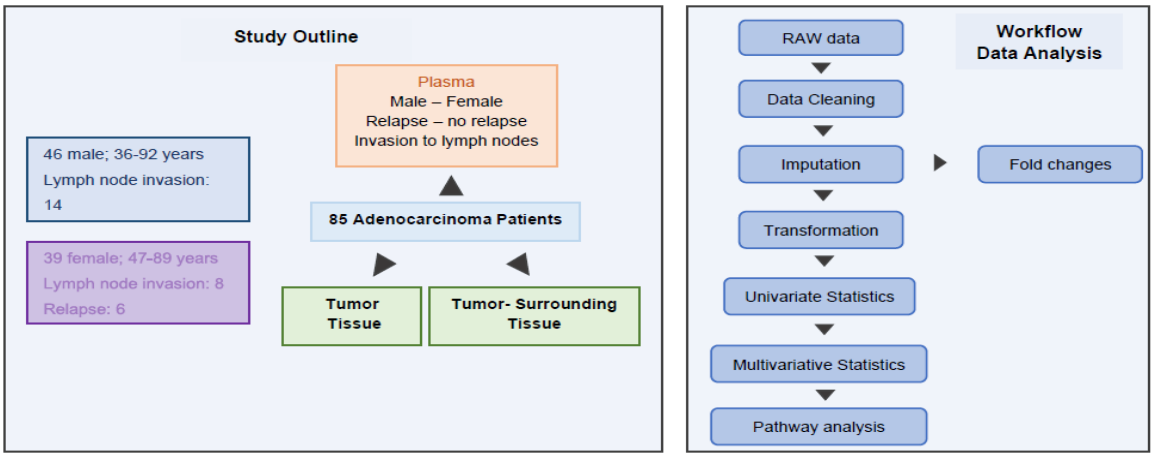

Figure 1: Schematic representation of the study outline and data analysis workflow.

Data cleaning, imputation, and transformation

A general cleaning of the dataset was performed to exclude analytes whose concentration values were missing or below the limit of detection (LOD). The cleaned dataset was then further used for scaling, transformation, and statistical analysis.[11] The concentration values for the whole dataset were cleaned and the 80-20 rule was applied. For the statistical analysis, the results of the samples had to contain at least 80% of the metabolite's valid values above LOD. If this criterion was fulfilled, the analyte was submitted for further statistical analysis. On the other hand, if 20% of the results of an analyte were below the LOD, it was excluded.

Missing value imputation is commonly used to replace the missing values with a non- zero value while maintaining the overall data structure. The remaining values below the LOD in the dataset were, therefore, imputed by applying a log-spline imputation method, which was developed for data being right censored, left censored, or interval censored.[12]

The study data were further processed by a log2 transformation. This technique is commonly applied to meet assumptions of statistical tests (e.g. data distribution, correction of heteroskedasticity, and skewness) and improve data interpretability and visualization.

Study analyses were performed using the freely available XCMS software (version 1.24.1), which runs on the R platform.

Univariate statistics

General measurements of central tendency and dispersion were made for the cleaned and imputed dataset to provide a quantitative description for the different groups. A paired t- test was performed with the transformed log concentration values to detect significantly altered analytes in the patients' plasma.

A significance level of $\alpha = 0.05$ was set, and p-values were calculated. To control the false discovery rate (FDR) during multiple comparisons, an adjusted p-value (Benjamini- Hochberg correction) was calculated.[13] Figure 4 shows all analytes with significantly altered concentrations ($p < 0.05$) and fold changes.

Multivariate statistics

Multivariate statistical methods were applied to detect changes in single metabolites between different groups and to show the dependency structures between individual analytes. In this case, principal component analysis (PCA), partial least squares discrimination analysis (PLS-DA), and hierarchical cluster analysis (HCA) were used as multivariate approaches. Multivariate analysis was performed with cleaned, imputed, and log-transformed data.

Hypothesis-free PCA was based on a linear mixture model to highlight the variance within the dataset while reducing the dimensionality and generating a smaller number of mutually decorrelated principal components (PCs). As a supervised linear mixture model, PLS regression was used to separate the predefined groups to the best possible extent based on metabolite concentrations.

An HCA was performed to visualize the samples according to intrinsic similarities in their measurements regardless of specific sample groupings. Here, the complete-linkage method was applied, which defines the cluster distance between two clusters as the maximum distance between their individual components.

RESULTS

The study comprised 39 female and 46 male patients; the two groups had a comparable mean age of 72.6 years (men) and 69.5 years (women). Figure 2 shows the metabolites analyzed for the gender group alone after the application of the 80-20 rule. Multivariate statistical analyses using PCA showed that the two groups were almost completely overlapped with regard to gender, while PSL-DA showed partial separation between the two groups (Fig. 3). Univariate statistical analyses using the t-test with Benjamini-Hochberg adjustment showed five metabolites that were significantly different between the male and female patients, while the non-adjusted test showed 15 significantly different metabolites (Fig. 4). Male patients had elevated creatinine and C3 acylcarnitine, and female patients had increased mean levels of three species of sphingomyelin (Fig. 5).

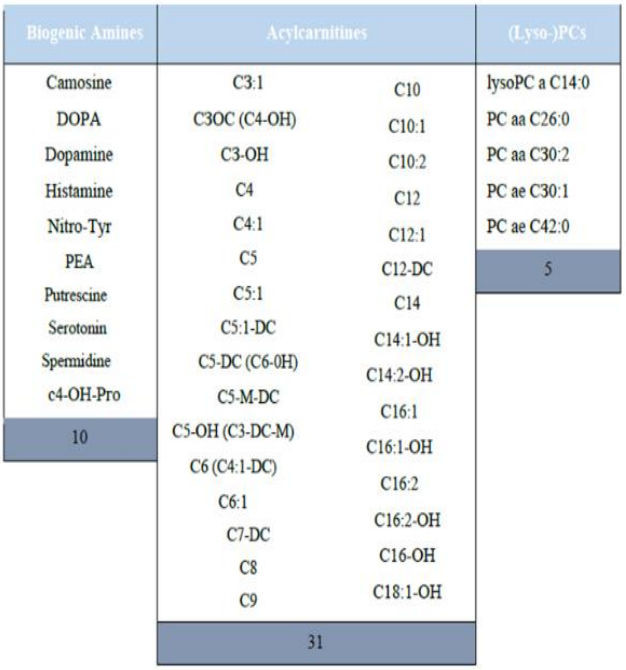

Figure 2: Metabolites analyzed after the application of the 80-20 rule.

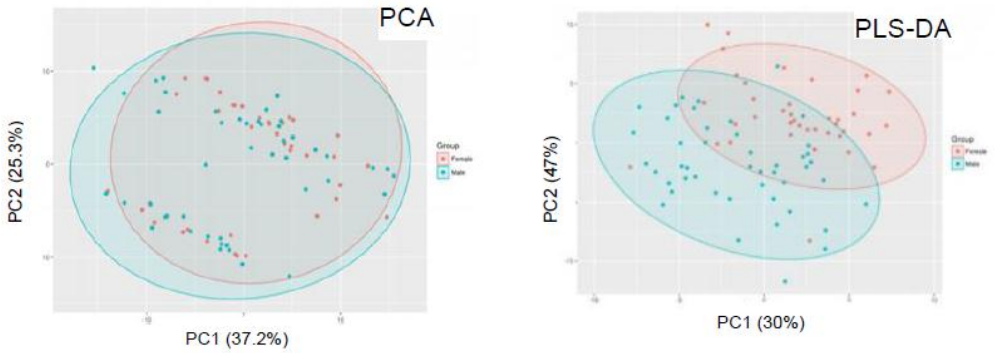

Figure 3: Principal component analysis and partial least squares discrimination analysis- multivariate statistical analyses by gender.

Figure 4: T-test with and without Benjamini-Hochberg adjustment, showing metabolites in female and male patients.

Metabolite	p-value	P-value adjusted	Mean Female	Mean Male	Meanlog2 Fold Change
SM.C22.3	<0. 000	0.001	- 2.53	- 5.09	- 1.01
SM.C18.1	<0. 000	0.008	4.05	3.62	0.16
Creatinine	<0. 000	0.016	5.99	6.39	- 0.09
SM.C20.2	0.001	0.023	- 0.49	- 0.87	- 0.83
C3	0.001	0.023	-2.3	- 1.8	0.35
PC.aa.C32.3	0.002	0.053	- 0.76	- 1.15	- 0.6
SM.C16.1	0.003	0.061	4.34	4.05	0.1
Gly	0.009	0.156	7.91	7.53	0.07
SM.C18.O	0.011	0.174	5.08	4.82	0.08
Kynurenine	0.012	0.177	0.99	1.38	- 0.48
alpha.AAA	0.025	0.297	- 0.45	- 0.17	1.42
Pro	0.026	0.297	7.04	7.27	- 0.04
SM.C26.1	0.029	0.297	- 0.83	-1.14	- 0.45
SM..OH..C22.2	0.029	0.297	3.65	3.39	0.11
PC.ae. C32.2	0.049	0.431	- 0.36	- 0.6	- 0.71

Figure 5: Significantly changed metabolites with adjustment.

Metabolite
SM.C22.3 SM.C18.1 Creatinine SM.C20.2 C3

Thirty patients had lymph node invasion, while 55 patients did not. Multivariate analysis using PCA showed almost complete overlapping between patients with and without lymph node invasion, and PLS-DA showed large overlapping between the two groups, with very few outliers (Fig. 6). Univariate analyses using the t-test with Benjamini-Hochberg adjustment did not demonstrate a significant difference in metabolite levels between patients with and without lymph node invasion. Without adjustment, 13 metabolites were significantly altered: SM.C18.0, SM.C16.0, SM.C24.1, SDMA, PC.aa.C32.0, SM.C22.3, SM.C20.2, PC.ae.C34.1, SM.OH.C16.1, SM.C26.1, PC.aa.C36.1, SM.C18.1, and taurine (Fig. 7). All sphingomyelins were significantly altered; phosphatidylcholines and SDMA were, on average, elevated in patients with lymph node invasion.

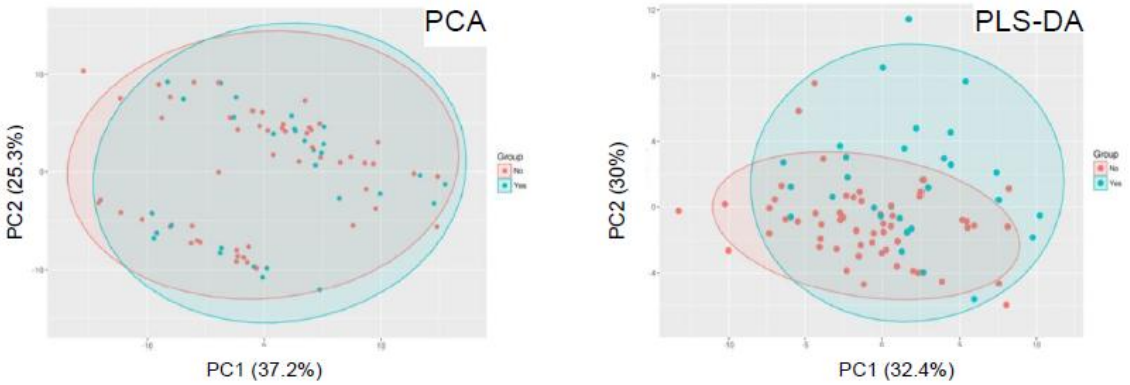

Figure 6: Principal component analysis and partial least squares discrimination - multivariate statistical analyses and lymph node invasion.

Figure 7: Metabolites significantly altered without adjustment and lymph node invasion.

Metabolite	p-value	P-value adjusted	Mean No Invasion	Mean Invasion	Meanlog2 Fold Change
SM.C18.8	0.007	0.414	4.84	5.13	- 0.08
SM.C16.0	0.010	0.414	6.98	7.2	- 0.05
SM.C24.1	0.015	0.414	5.73	5.94	- 0.05
SDMA	0.019	0.414	- 0.59	- 0.2	1.58
PC.aa.C32.0	0.019	0.414	3.43	3.67	- 0.1
SM.C22.3	0.021	0.414	- 4.43	- 2.97	0.58
SM.C20.2	0.028	0.414	- 0.79	- 0.52	0.59
PC.ae.C34.1	0.032	0.414	2.65	2.89	- 0.12
SM..0H..C16.1	0.034	0.414	2.28	2.5	- 0.14
SM.C26.1	0.035	0.414	- 1.11	- 0.8	0.47
PC.aa.C36.1	0.037	0.414	5.08	5.32	- 0.07
SM.C18.1	0.037	0.414	3.73	3.98	- 0.09
Taurine	0.038	0.414	5.69	5.97	- 0.07

With regard to the relapse variable, 10 patients relapsed and 75 did not. Owing to the differences in the sizes of the groups, a robust statistical analysis could not be performed. In the multivariate analysis, PCA showed almost complete overlapping, PSL-DA showed large overlapping (70%) between the two groups, and the t-test with and without Benjamini- Hochberg adjustment did not show significant differences in metabolites between the patients with and without relapse (Fig. 8).

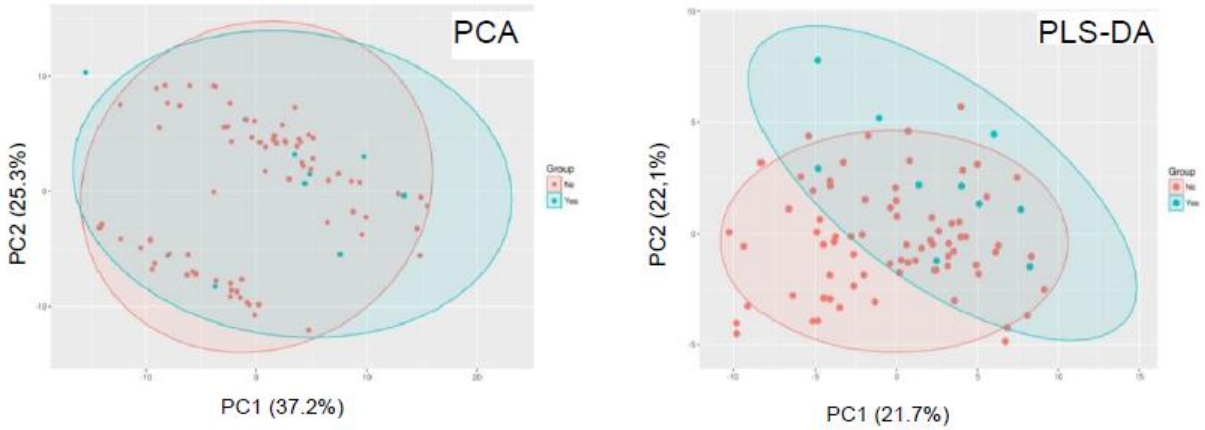

Figure 8: Schematic representation of relapse between the groups of male and female patients.

Patients lost to follow-up or deceased due to other causes were removed from the dataset; hence, the remaining group had 18 deceased patients and 61 living patients. Multivariate analysis by PCA showed almost complete

overlapping between the two groups, and PLS-DA showed a group of living patients within the group of deceased patients. Owing to the different sizes of the groups, a robust statistical analysis could not be performed (Fig. 9). In univariate analyses, the t-test with Benjamini-Hochberg adjustment did not show a significant difference in metabolites between the living and deceased patients. Without adjustment, 11 metabolites were significantly altered: SM.C24.1, PC.aa.C32.0, PC.aa.C36.2, SM.C16.0, PC.ae.c44.6, PC.ae.C44.5, SM.C18.0, PC.ae.C42.5, SM.C20.2, PC.ae.C34.1, and SM.C22.3 (Fig. 10).

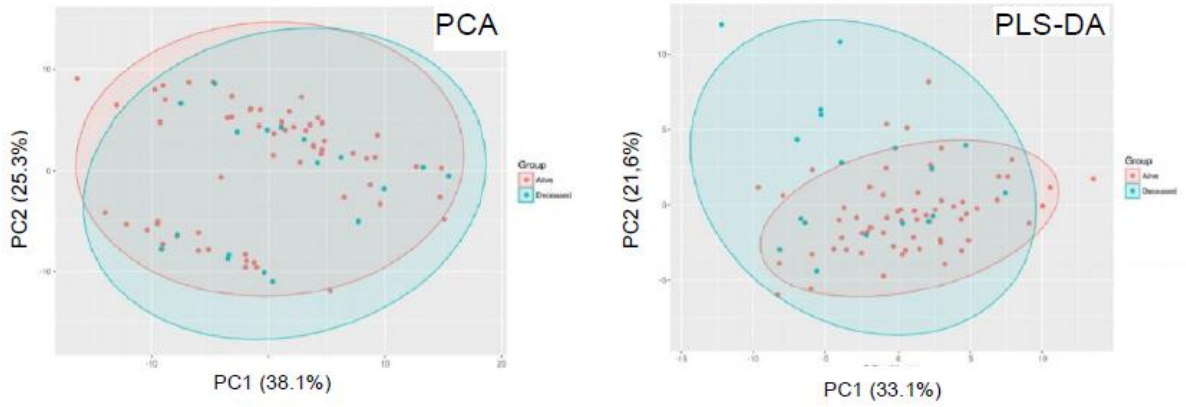

Figure 9: Schematic representation of survival between male and female patients within the group of deceased patients.

Figure 10 T-test without adjustment, showing 11 significantly different metabolites between living and deceased patients.

Metabolite	p-value	P-value adjusted	Mean Alive	Mean Deceased	Meanlog2 Fold Change
SM.C24.1	0.010	0.595	6	- 1.05	- 0.07
PC.aa.C32.0	0.016	0.595	3.74	- 1.09	- 0.12
PC.ae.C44.6	0.023	0.595	0.48	- 4.71	- 2.24
PC.aa.C36.2	0.028	0.595	7.39	- 1.04	- 0.06
SM.C16.0	0.029	0.595	7.23	- 1.03	- 0.05
PC.ae.C44.5	0.030	0.595	0.47	- 5.06	- 2.34
SM.C18.0	0.037	0.595	5.15	- 1.06	- 0.08
PC.ae.C42.5	0.042	0.595	1.23	- 1.23	- 0.3
SM.C20.2	0.043	0.595	- 0.46	1.65	0.72
PC.ae.C34.1	0.046	0.595	2.91	- 1.1	- 0.14
SM.C22.3	0.048	0.595	- 2.7	1.54	0.63

With regard to TNM staging, the group of 85 patients was distributed as follows: 9 patients in T1, 15 patients in T2, 50 patients in T3, and 11 patients in T4. Multivariate analysis by PCA and PLS-DA showed almost complete overlapping between the two groups. Owing to the differences in the sizes of the groups, a robust statistical analysis could not be performed (Fig. 11). Univariate analysis using analysis of variance (ANOVA) with Benjamini-Hochberg adjustment did not show a significant difference in metabolites between the different stages. ANOVA without adjustment showed significant differences in two metabolites: taurine and SM22: 3 (Figure 12). There was high interindividual variability in sphingomyelin SM22: 3 among patients in the T1, T2, and T4 stages as well as significant differences in the gender, relapse, and lymph node invasion groups.

Almost no amino acid or biogenic amine was significantly altered in the plasma samples.

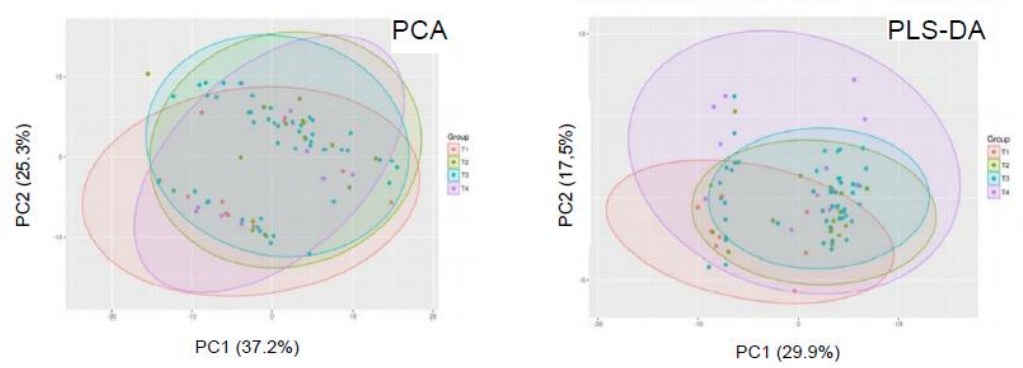

Figure 11: Schematic representation of TNM staging in the groups of male and female patients.

Figure 12: Metabolites SM.C22:3 and taurine with significant changes. SM.C22:3 exhibited a high interindividual variability in T1, T2, and T4 patients and was also significantly different in the groups of gender, relapse, and lymph node invasion.

Metabolite	p-value	P-value adjusted	Mean T1	Mean T2	Mean T3	Mean t4
SM.C22.3	0.016	0.955	- 5.89	- 4.63	- 3.73	- 2.18
Taurine	0.030	0.955	5.97	5.57	5.73	6.2

DISCUSSION

Tissue markers, including genetic markers, allow the determination of a different prognosis and a distinct therapeutic approach for each case. However, they lack the ability of the plasma markers to determine a diagnosis or indicate tumor relapse. They are limited to determining staging, prognosis, and probable therapeutic response to a given lesion.[14]

Our findings regarding plasma metabolites are in line with those reported by other authors who also showed differences in the serum and plasma levels of sphingolipids according to gender. Higher levels of C 18 -SM, C 18: 1-SM, C 18-Cer1P, and total dhCers, specifically C 18-, C 22-, and C 24-dhCer, were more commonly observed among women than men. The higher levels of high density lipoprotein particles observed in women may be at least partially responsible for the higher levels of sphingolipids in women. However, other hormonal and metabolic factors may contribute to the differences in sphingolipids between male and female patients.[15,16]

In several studies, phosphorylation of the myosin light chain by serine/threonine protein kinases led to increased cellular contractility and consequent blebbing of the membrane. This process is apparently associated with the regulation of cell morphology, cell migration, and neoplastic and lymph node invasion in patients with CRC.[17-19]

An increase in the mean concentration of phosphatidylcholine (PC) was observed in deceased patients, and PC.aa.C32:0 was also significantly increased in patients with lymph node invasion. All sphingomyelins that were significantly increased in deceased patients were also elevated in patients with lymph node invasion. These data suggest that poor survival may be related to the occurrence of lymph node invasion.

The results of this investigation and the comparison with our previously published findings,[20] in normal and tumor tissue samples from the same patients included in the present study show that very few metabolites found in the plasma were different from those found in tissues. The application of more restricted statistics with correction to control the false discovery rate (Benjamini-Hochberg adjustment) most often did not result in significant changes in the different groups of patients. The analysis of the results without adjustment showed significant differences in metabolite levels, especially sphingomyelins and phospholipids.

The role of sphingomyelins and lysoPCs as potential plasma biomarkers for CRC should be further investigated. Although potential interindividual differences may occur, additional studies are required to confirm our findings.

REFERENCES

1. Weitz J, Koch M, Debus J, Höhler T, Galle PR, Büchler MW. Colorectal cancer. Lancet, 2005; 365: 153–65.
2. Siegel RL, Miller, KD, Jemal A. Cancer statistics. Ca-Cancer J Clin, 2015; 65: 5–29.
3. Hashim NAA, Ab-Rahim S, Suddin LS, Saman MSA, e Mazlan M. Global serum metabolomics profiling of colorectal cancer (Review). Mol Clin Oncol, 2019; 11: 3-14.
4. da Silva I, da Costa Vieira R, Stella C, Loturco E, Al, C Veo C, et al. Inborn-like errors of metabolism are determinants of breast cancer risk, clinical response and survival: a study of human biochemical

individuality. Oncotarget, 2018; 9(60): 31664-81.
5. Wang D, Li W, Zou Q, Yin L, Du Y, Gu J, Suo J. Serum metabolomic profiling of human gastric cancer and its relationship with the prognosis. Oncotarget, 2017; 8: 110000-15.
6. Chinnaiyan P, Kensicki E, Bloom G, Prabhu A, Sarcar B, Kahali S, Eschrich S, et al.. The metabolomic signature of malignant glioma reflects accelerated anabolic metabolism. Cancer Res, 2012; 72: 5878-88.
7. Hadi NI, Jamal Q, Iqbal A, Shaikh F, Somroo S, Musharraf SG. Serum metabolomic profiles for breast cancer diagnosis, grading and staging by gas chromatography-mass spectrometry. Sci Rep, 2017; 7: 1715.
8. Kumar N, Shahjaman M, Mollah MNH, Islam SMS, Hoque MA. Serum and plasma metabolomic biomarkers for lung cancer. Bioinformation, 2017; 13: 202-8.
9. Siskos AP, Jain P, Römisch-Margl W, Bennett M, Achaintre D, Asad Y, et al. Interlaboratory Reproducibility of a Targeted Metabolomics Platform for Analysis of Human Serum and Plasma. Anal Chem, 2017; 89: 656-65.
10. Ramsay SL, Stoegg WM, Weinberger KM, Graber A, Guggenbichler W. Apparatus and method for analyzing a metabolite profile. EP 1875401 A211-Jan-2007.
11. Di Guida R, Engel J, Allwood JW, Weber RJM, Jones MR, Sommer U, et al. Non- targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. Metabolomics, 2016; 12: 93.
12. Kooperberg C, Stone CJ. Logspline density estimation for censored data. J Comput Graph Stat, 1992; 1: 301–28.
13. Benjamini Y, Hochberg Y. Benjamini-1995.pdf. J R Stat Soc B, 1995; 57: 289–300.
14. Fernandes, L C, Matos, D. Tumor markers in colorectal cancer. Marcadores Tumorais. 2002; 29(2): 106-11.
15. Hammad SM, Pierce JS, Soodavar F, Smith K J, Al Gadban MM, Rembiesa B, et al. Blood sphingolipidomics in healthy humans: impact of sample collection methodology. J Lipid Res, 2010; 51(10): 3074–87.
16. Mielke MM, Bandaru VVR, Han D, An Y, Resnick SM, Ferrucci L, Haughey NJ. Factors affecting longitudinal trajectories of plasma sphingomyelins: the Baltimore Longitudinal Study of Aging. Aging Cell, 2014; 14(1): 112–21.
17. Short SP, Thompson JJ, Bilotta AJ, Chen X, Revetta FL, Washington MK, et al. Serine threonine kinase 17A maintains the epithelial state in colorectal cancer cells. Mol Cancer Res, 2019; 4: 882–94.
18. Sanjo H, Kawai T, Akira S. DRAKs, novel serine/threonine kinases related to death associated protein kinase that trigger apoptosis. J Biol Chem, 1998; 273(44): 29066–71.
19. Jin X, Liu K, Jiao B, Wang X, Huang S, Ren W, et al. Vincristine promotes migration and invasion of colorectal cancer HCT116 cells through RhoA/ROCK/Myosin light chain pathway. Cell Mol Biol (Noisy-le-grand), 2016; 62(12): 91–6.
20. Denadai, MV, Veo, CAR, Silva, SM, Adan, MG, Muller, U, Schafferer, S, Silva, IDCG, Brunetti, IA, Matos, D. Significant metabolomic differences between the tumor and non-tumor adjacent mucosa in colorectal cancer – a targeted mass-spectrometry-based study to identify potential biomarkers. World J Pharm Med Res, 2020; 6: 7.