

## Article

# Urinary $^1\text{H}$ -NMR Metabolic Signature in Subjects Undergoing Colonoscopy for Colon Cancer Diagnosis

Cristina Piras <sup>1</sup>, Monica Pibiri <sup>1</sup>, Vera Piera Leoni <sup>1</sup>, Francesco Cabras <sup>2</sup>, Angelo Restivo <sup>2</sup>, Julian Leether Griffin <sup>3</sup>, Vassilios Fanos <sup>4</sup>, Michele Mussap <sup>5</sup>, Luigi Zorcolo <sup>2</sup> and Luigi Atzori <sup>1,\*</sup>

<sup>1</sup> Department of Biomedical Science, University of Cagliari, 09124 Cagliari, Italy; cristina.piras@unica.it (C.P.); mpibiri@unica.it (M.P.); vleoni@unica.it (V.P.L.)

<sup>2</sup> Department of Surgery, Colorectal Cancer Surgery Center, University Hospital, University of Cagliari, 09124 Cagliari, Italy; francescocabras.1@gmail.com (F.C.); arestivo@unica.it (A.R.); lzorcolo@ao.ucagliari.it (L.Z.)

<sup>3</sup> Department of Metabolism, Digestion and Reproduction, Imperial College, London SW7 2AZ, UK; julian.griffin@imperial.ac.uk

<sup>4</sup> Department of Surgical Sciences, Neonatal Intensive Care Unit, Puericulture Institute and Neonatal Section, University Hospital, University of Cagliari, 09124 Cagliari, Italy; vafanos@tin.it

<sup>5</sup> Laboratory Unit, Department of Surgical Sciences, University of Cagliari, 09124 Cagliari, Italy; michele.mussap@unica.it

\* Correspondence: latzori@unica.it; Tel.: +39-0706758390

Received: 16 July 2020; Accepted: 1 August 2020; Published: 5 August 2020



**Featured Application:** Metabolomics can be a useful tool to support the diagnosis and the monitoring of colon cancer.

**Abstract:** Metabolomics represents a promising non-invasive approach that can be applied to identify biochemical changes in colorectal cancer patients (CRC) and is potentially useful for diagnosis and follow-up. Despite the literature regarding metabolomics CRC-specific profiles, discrimination between metabolic changes specifically related to CRC and intra-individual variability is still a problem to be solved. This was a preliminary case-control study, in which  $^1\text{H}$ -NMR spectroscopy combined with multivariate statistical analysis was used to profile urine metabolites in subjects undergoing colonoscopy for colon cancer diagnosis. To reduce intra-individual variability, metabolic profiles were evaluated in participants' urine samples, collected just before the colonoscopy and after a short-term dietary regimen required for the endoscopy procedure. Data obtained highlighted different urinary metabolic profiles between CRC and unaffected subjects (C). The metabolites altered in the CRC urine (acetoacetate, creatine, creatinine, histamine, phenylacetylglycine, and tryptophan) significantly correlated with colon cancer and discriminated with accuracy CRC patients from C patients (receiver operator characteristic (ROC) curve with an area under the curve (AUC) of 0.875; 95% CI: 0.667–1). These results confirm that urinary metabolomic analysis can be a valid tool to improve CRC diagnosis, prognosis, and response to therapy, representing a noninvasive approach that could precede more invasive tests.

**Keywords:** metabolomics; colorectal cancer; biomarkers;  $^1\text{H}$ -NMR spectroscopy

## 1. Introduction

Colorectal cancer (CRC) diagnosis is mainly based on invasive, costly, and time-consuming methods (e.g., endoscopic, histological, and radiographic techniques), as noninvasive methods, such as stool-based tests (e.g., fecal occult blood test, FOBT) or the carcinoembryonic antigen (CEA) test, lack sensitivity and specificity [1–3]. Indeed, the diagnostic power of the noninvasive techniques mentioned above, such as FOBT, is higher in the presence of advanced tumors, while survival and

prognosis for CRC are positively correlated with an early stage of neoplasia at the time of diagnosis [4]. Thus, new CRC diagnostic tools and a better understanding of the molecular events implicated in CRC onset and progression are needed for successful treatment [5]. Every cell function is based on metabolic pathways, and their alterations are reflected by changes in related metabolites. Systematic analysis of the full set of metabolites in a biological sample is the focus of metabolomics, a major constituent of a new type of integrated approach called “systems biology”, which might represent a promising noninvasive tool for the identification of potential markers involved in human tumorigenesis [6,7]. In the last decade, metabolomics analysis of samples from CRC patients has led to the identification of several molecules specifically expressed, and the monitoring of their fluctuations is emerging as an important way to detect the early stages of CRC [6,8–12]. On this basis, the present study aimed to analyze by  $^1\text{H}$ -NMR spectroscopy the urine metabolomics profile of CRC patients and cancer-free controls, to identify a possible specific set of metabolic biomarkers that could be proposed as a noninvasive diagnostic tool to identify CRC patients.

## 2. Materials and Methods

### 2.1. Patients and Sample Collection

The study was conducted on patients admitted to the Colorectal Surgical Unit, University Hospital of Cagliari, with a suspected diagnosis of CRC from fecal occult blood presence or for familiarity risk. Patients were subjected to colonoscopy. A few days before the colonoscopy procedure they started a low-fiber diet. The day before the colonoscopy they did not eat solid food, but only liquids. The night before the colonoscopy, patients took strong laxatives to clear the digestive tract. On the day of the colonoscopy, they could drink water up to 2 h before the procedure. Based on the outcome of this examination, the population was divided into two groups: CRC group, comprising 6 cases of rectal, 7 of colon, and 1 of caecum adenocarcinomas, and unaffected subjects (C) group, including 10 individuals without colon pathologies. Participant characteristics are indicated in Table 1. Patient status was confirmed by subsequent histological analysis. This study was approved by the Ethical Committee of the University Hospital of Cagliari, and all participants gave informed consent.

**Table 1.** Characteristics of the study population.

	CRC Patients	Controls
Number	14	10
Age (median, range)	67.4, 40–88	56.3, 47–67
Male/female ratio	11/3	7/3
CEA (>2.5 ng/mL)	5	0
<b>Cancer site</b>		
Colon	6	-
Rectum	7	-
Caecum	1	-
<b>FOBT</b>		
Positive	9	N/A
Negative	5	N/A
<b>Dukes' Stage</b>		
A	3	-
B	7	-
C	3	-
D	1	-

FOBT: Fecal Occult Blood Test; N/A: Not available.

### 2.2. Urine Samples Preparation

Urine samples were collected just before colonoscopies from participants in the study. An aliquot of 800  $\mu\text{L}$  of urine was transferred into a tube with 8  $\mu\text{L}$  of a 1% aqueous solution of  $\text{NaN}_3$ , to inhibit bacteria growth, and was stored at  $-80\text{ }^\circ\text{C}$ . Afterwards, to remove solid particles, the sample was

centrifuged at  $12,000\times g$  for 10 min at 4 °C. The supernatant (630  $\mu$ L) was mixed with 70  $\mu$ L of potassium phosphate buffer in D<sub>2</sub>O (1.5 M, pH 7.4) containing sodium 3-trimethylsilyl-propionate-2,2,3,3-d<sub>4</sub> (TSP) as an internal standard (98 atom% D, Sigma-Aldrich, Milan, Italy). An aliquot of 650  $\mu$ L was transferred to NMR glass tubes for <sup>1</sup>H-NMR analysis [13].

### 2.3. <sup>1</sup>H-NMR Spectroscopic Analysis

<sup>1</sup>H-NMR measurements of urine samples were carried out at 298 K using a Bruker DRX 500 spectrometer operating at 500 MHz (Bruker Biospin, Rheinstetten, Germany). <sup>1</sup>H-NMR spectra were obtained using a 1D Nuclear Overhauser Enhancement Spectroscopy (NOESY) pulse sequence to suppress water signals (relaxation delay of 3 s). For each sample, 128 free induction decays (FIDs) were collected into 64 K data points with a spectral width of 6000 Hz with a 90° pulse, an acquisition time of 2 s, and a mixing time of 150 ms. The FIDs were weighted by an exponential function with a 0.5 Hz line-broadening factor before Fourier transformation.

### 2.4. NMR Data Preprocessing and Multivariate Statistical Analysis

The phase and baseline of NMR spectra were corrected using ACDlab Processor Academic Edition (Advanced Chemistry Development, 12.01, 2010, Toronto, ON, Canada). The spectral region comprising the signal of residual water and urea (4.5–6.0 ppm) was removed. The final spectral regions considered were between 0.5 and 4.5 ppm as well as 6.0 and 9.5 ppm. The ACD Labs intelligent bucketing method was used for spectral integration [14]. A 0.01 ppm bucket width was defined with an allowed 50% looseness. The intelligent bucket method finds local minima in spectra and adjusts the buckets accordingly. In this way, a peak is integrated into one bucket. The area of bucketed regions was normalized using Median Fold Change Normalization [15], largely preferred to total sum normalization when studying urine samples, and a matrix was generated. The resultant data sets were then imported into SIMCA software (Version 15.0, Sartorius Stedim Biotech, Umea, Sweden) for multivariate statistical analysis. The data sets were then Pareto scaled. Pareto scaling, where each variable is divided by the square root of the standard deviation, gives greater weight to NMR variables with low intensity, but it is not as extreme as the UV (Unit Variance) scaling method. Principal component analysis (PCA) and Orthogonal Partial Least-Squares Discriminant Analysis (OPLS-DA) were used for multivariate statistical analyses of NMR data. PCA was performed to identify any possible relation (trends, outliers) between the samples. As far as the outliers are concerned, Hotelling's T<sup>2</sup> and DModX tests were applied. OPLS-DA analysis was used to reduce model complexity and to better highlight sample discrimination. The goodness of the model was evaluated using a 7-fold cross-validation and a "permutation test" (400 permutations). The permutation test was calculated by randomizing the Y-matrix (classification components) while the X-matrix (peak intensity in NMR spectra) was kept constant. The permutation plot shows the correlation coefficient between the original and the permuted y-variables on the x-axis, versus the cumulative R<sup>2</sup> and Q<sup>2</sup> on the y-axis, and draws the regression line. Q<sup>2</sup>Y intercept values are a measure of the overfit, and values <0.05 are indicative of a valid model. To highlight potential metabolites that mainly contributed to group separation, an S-plot for the OPLS-DA model was created. The S-plot reveals the contribution of each variable to the predictive component, matching the covariance  $p$  and the correlation  $p(\text{corr})$  obtained from the OPLS-DA model. The axes plotted in the S-plot from the predictive component are  $p_1$  versus  $p(\text{corr})_1$ , representing the magnitude (modelled covariance) and reliability (modelled correlation) respectively. The variables characterized by high magnitude and reliability values have an important role in the separation of different groups of samples. In the S-plot both magnitude (intensity) and reliability are plotted. The statistical significance of the difference in metabolite concentrations, quantified using Chenomx NMR suite 7.1 (Chenomx Inc., Edmonton, AB, Canada), was calculated using an unpaired Welch  $t$ -test with a 95% confidence interval. Chenomx NMR Suite software is useful for identifying and quantifying the metabolites in NMR spectra [16]. It is equipped with reference libraries containing numerous compound models that are identical to the spectra of pure compounds obtained under similar experimental conditions. Basically,

a Lorentzian peak shape model of each reference compound is created from database information and is overlapped with the actual spectrum. The linear combination of modeled metabolites gives rise to the total spectral fit, which can be assessed with a summation line [16]. The metabolites with both VIP > 1 (Variable Importance for the Projection) and a  $p$ -value  $\leq 0.05$  were considered statistically significant. The Metaboanalyst program (<https://www.metaboanalyst.ca/>) [17] was used to generate receiver operator characteristic curves (ROC), calculate sensitivity, specificity, and the area under the ROC curve (AUC), and Matlab (<http://it.mathworks.com/>) was used to generate the box-and-whisker plots. The Random Forest algorithm was used to construct the ROC. It identifies important features through repeated random sub-sampling cross-validation (CV). In each CV, two-thirds (2/3) of the samples are used to evaluate the importance of each feature based on decreases in accuracy. The top features are used to build classification/regression models that are validated on the one-third (1/3) of samples that were left out of the original model.

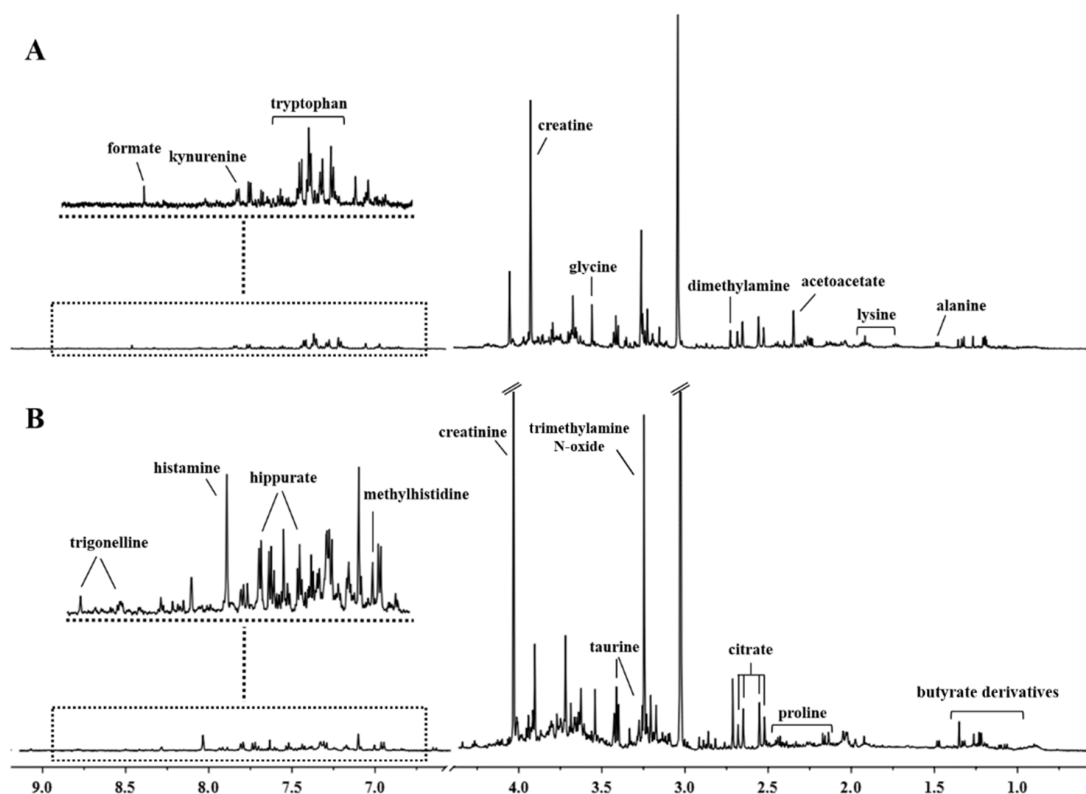
### 2.5. Serum Carcinoembryonic Antigen Level Determination

Blood samples for CEA analysis were collected just before colonoscopy, centrifuged, and stored at  $-80^{\circ}\text{C}$  before use. The cut-off for CEA 1 was  $>2.5\text{ ng/mL}$  [12].

## 3. Results

### 3.1. $^1\text{H}$ -NMR Spectra of Urine Samples

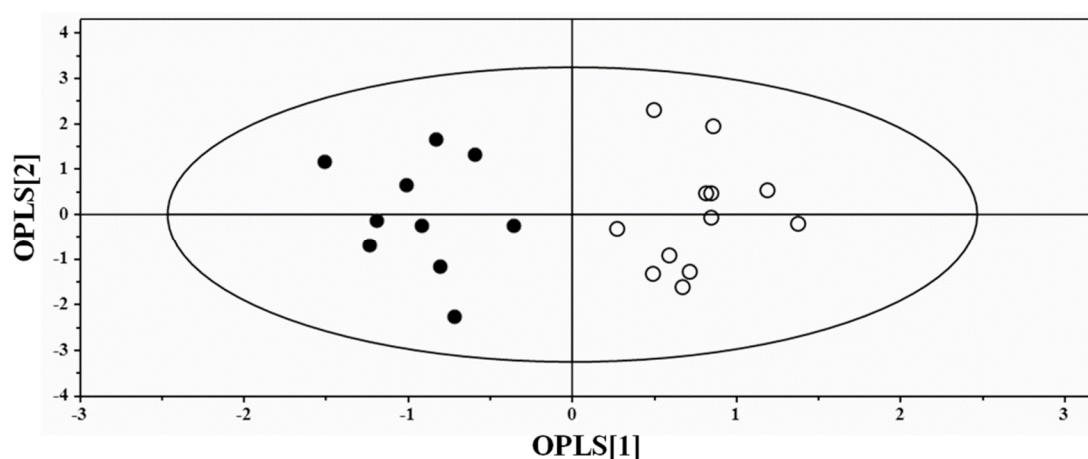
NMR spectral analysis of urine samples revealed a distinct metabolic signature between C and CRC subjects (Figure 1). The resonance of spectra was assigned to different metabolites based on data published in the literature [18] and by using the library from the Chenomx NMR suite. Representative spectra of C and CRC are shown in Figure 1A,B. Major peak assignments of urine samples are reported in Figure 1, while the chemical shifts of metabolites are summarized in Table S1.



**Figure 1.** Spectral area and major peak assignments of a representative  $^1\text{H}$ -NMR spectra of urine obtained from unaffected subjects (C) (A) or colorectal cancer patients (CRC) (B).

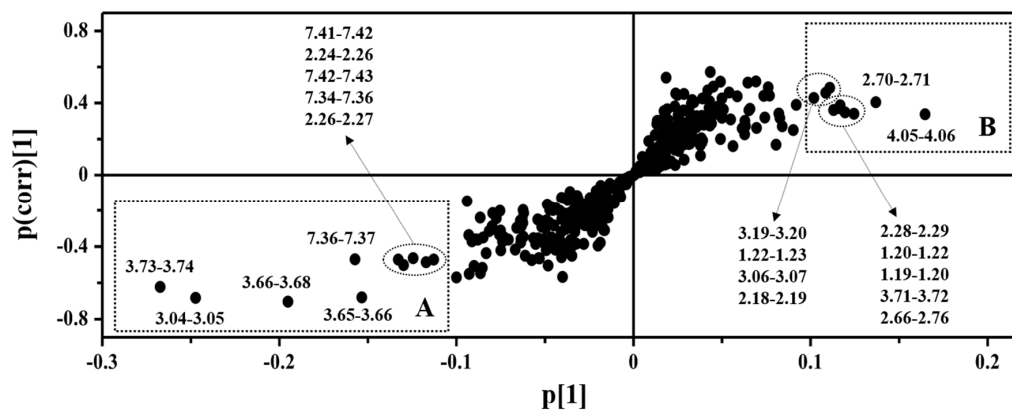
### 3.2. Multivariate Statistical Analysis of NMR Data

PCA was initially applied to the complete data set to highlight possible metabolic differences among the urine samples of control and CRC subjects and to identify potential outliers (Figure S1 in Supporting Information). Based on Hotelling's T2 and DModX tests, two samples were considered outliers and were removed from the analysis (Figure S2 in Supporting Information). To optimize the separation between the CRC and C urine samples, the supervised OPLS-DA model was applied. As displayed in Figure 2, the OPLS-DA scores plot showed clear separation between the two groups of samples. To test the validity of the model, a permutation test on the PLS-DA model was performed. Results showed that the model was statistically valid, with a Q2 intercept value of  $-0.192$ . To identify the metabolites that mainly contributed to group separation, an S-plot was constructed (Figure 3). As shown in Figure 3, the variables selected in the S-plot are indicated with a dotted rectangle and represent the metabolites responsible for differentiation in the OPLS-DA scores plot. Cutoff values for the covariance of  $|p| \geq 0.1$  and the correlation  $|p(\text{corr})| \geq 0.2$  were used. In the S-plot, the control samples were characterized, based on the discriminant regions, by high creatine, sn-glycero-3-phosphocholine, phenylacetylglutamine, and proline (Figure 3 square A), whereas an increase in citrate, creatinine, acetoacetate, 3-hydroxybutyrate, 3-aminoisobutyrate, tyrosine, tryptophan, histamine, methylhistidine, and fucose characterized the CRC samples (Figure 3, square B). The relative concentrations of metabolites highlighted in the S-plot were verified with Chenomx NMR Suite 7.1 and were subjected to the Welch *t*-test to identify significant variations in concentration in the two groups. Significantly discriminant metabolites were characterized by  $\text{VIP} > 1$  and  $p \leq 0.05$ . After this analysis, 14 metabolites exhibited  $\text{VIP} > 1$ , but only acetoacetate, creatine, creatinine, histamine, phenylacetylglutamine, and tryptophan showed significant variation (with  $p \leq 0.05$ ) (Table 2). The relative concentrations, calculated by normalization of the molar concentration of each metabolite to the total molar concentration of all 14 metabolites for each sample in the two groups, were compared using box-and-whisker plots. As shown in Figure 4, the data obtained demonstrated that the CRC group showed increased relative levels of acetoacetate, creatinine, and histamine, whereas creatine, phenylacetylglutamine, and tryptophan levels were lower compared to the C group. A ROC curve was constructed using only the metabolites with a significant statistical variation, and the area under the curve of the ROC analysis was found to be  $0.879$  (95% CI:  $0.667\text{--}1$ ), indicating high predictive accuracy of the model (Figure 5). Significantly, only 5/12 CRC patients had increased CEA levels ( $\text{CEA} > 2.5 \text{ ng/mL}$ ), showing 42% sensitivity.

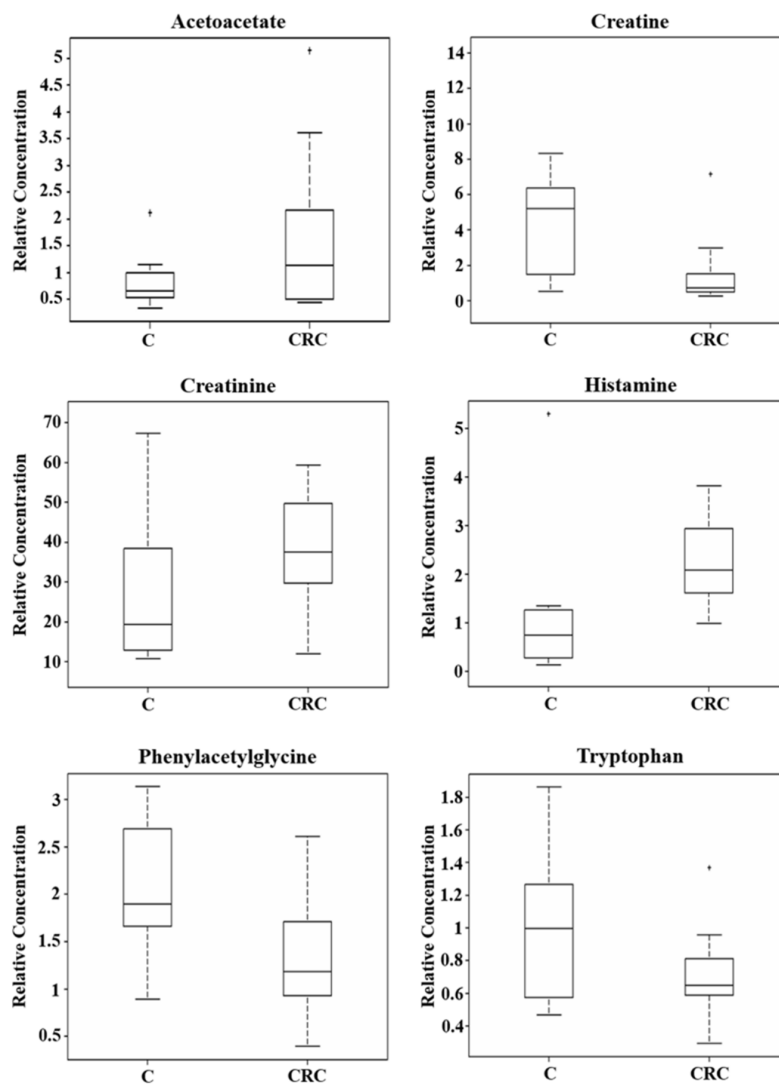


**Figure 2.** OPLS-DA scores plot of  $^1\text{H}$ -NMR spectra of urine samples: C (full circle) and CRC (open circle) subjects. The OPLS-DA model was performed using one predictive and one orthogonal component and is described by R2X, R2Y, and Q2 values of 0.328, 0.957, and 0.539, respectively.

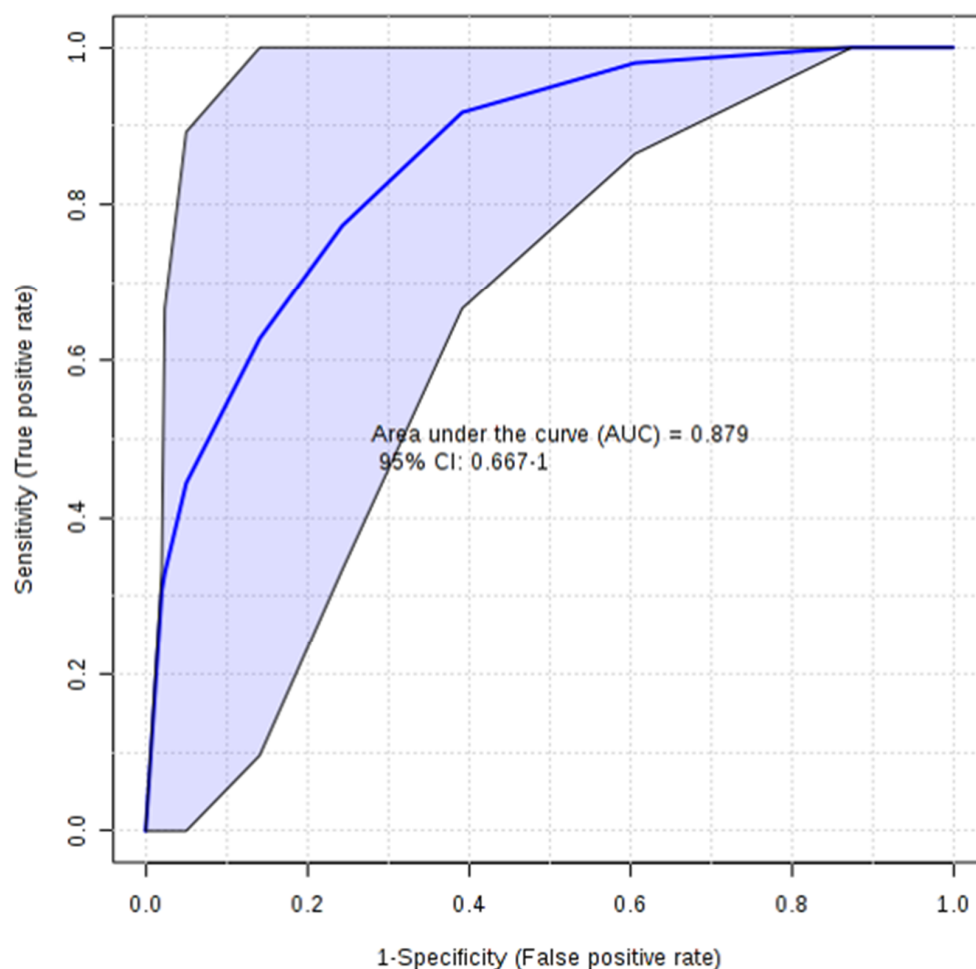




**Figure 3.** S-Plot corresponding to the OPLS-DA model used to characterize the most significant variables associated with group C (dotted rectangle A) and CRC (dotted rectangle B) subjects. Cutoff values for the covariance of  $|p| \geq 0.1$  and for the correlation  $|p(\text{corr})| \geq 0.2$  were used. Values of discriminant ppm regions are shown.



**Figure 4.** Box-and-whisker plots showing an altered expression of metabolites including acetoacetate, creatine, creatinine, histamine, phenylacetylglutamine, and tryptophan between C and CRC patients. Median values are highlighted by solid lines. Statistical significance of the metabolites ( $p \leq 0.05$ ) is presented in Table 2.



**Figure 5.** Receiver Operating Characteristic (ROC) analysis. Area under the curve (AUC) was found to be 0.879. CI: Confidence Interval.

**Table 2.** Statistical differences of metabolites characterized by Variable Importance for the Projection (VIP) > 1.

Metabolites	Mean (SD) of Group (mM) <sup>a</sup>		<i>p</i> -Value <sup>b</sup>	FC <sup>c</sup> (log2 CRC/C)
	C	CRC		
Phenylacetylglutamine	2.07 ± 0.65	1.34 ± 0.72	0.01	−0.623
Histamine	1.14 ± 0.49	2.26 ± 0.85	0.02	0.988
Creatine	6.64 ± 3.34	1.49 ± 0.98	0.03	−2.150
Tryptophan	1.01 ± 0.46	0.71 ± 0.27	0.03	−0.517
Acetoacetate	0.85 ± 0.28	1.64 ± 1.37	0.05	0.970
Creatinine	26.12 ± 17.96	37.83 ± 14.59	0.05	0.535
3-hydroxybutyrate	1.00 ± 0.54	1.64 ± 1.45	0.10	0.706
Tyrosine	0.82 ± 0.41	0.62 ± 0.37	0.12	−0.409
Proline	2.93 ± 1.12	3.38 ± 0.59	0.13	0.202
3-Aminoisobutyrate	1.09 ± 0.54	2.08 ± 2.87	0.15	0.933
Sn-glycero-3-phosphocholine	1.62 ± 0.67	1.88 ± 0.71	0.19	−0.215
Fucose	1.28 ± 0.54	1.16 ± 2.87	0.25	0.145
Methylhistidine	1.09 ± 0.67	1.23 ± 0.37	0.26	−0.180
Citrate	8.57 ± 4.12	7.53 ± 6.94	0.34	0.187

Metabolites were selected on VIP > 1 based on Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA).

<sup>a</sup> Relative concentrations were calculated by normalization of the molar concentration of each metabolite to the total molar concentration of all 14 metabolites for each sample. <sup>b</sup> A Welch *t*-test was performed and the *p*-value reported. <sup>c</sup> FC indicates the fold change (FC) between CRC and C patients. FC with a positive value indicates a relatively higher concentration present in CRC patients, and a negative value shows a relatively lower concentration as compared to the C.

#### 4. Discussion

Previous studies [6,9–12,19] have suggested the use of a metabolomics approach to identify tumor-associated biological alterations in CRC patients.

In the present study,  $^1\text{H}$ -NMR spectroscopy coupled with pattern recognition was used to profile urine metabolites of subjects with (CRC) or without (C) CRC. Significantly, metabolic perturbations were evaluated between the urine samples of CRC patients and individuals undergoing colonoscopy for positive FOBT or familial risk of CRC, to whom a similar low-fiber, controlled diet had been administered for few days before they were subjected to the examination. Urine samples were collected just before the colonoscopy. Therefore, both CRC and C groups were evaluated in the same short-term dietary conditions, consisting of a low-fiber diet for a few days. Since diet represents an important factor that contributes to inter-individual variability, our sample collection approach reduced variability related to a different nutritional status. To date, our study, despite the small number of subjects analyzed, is the first to investigate the metabolomics profiles of potential CRC patients under a similarly controlled short-term diet. Based on identification of metabolites with significant statistical variations between CRC and C groups, a ROC curve was constructed to assess the sensitivity and specificity of the biomarker candidates for early detection of CRC. The AUC of the ROC analysis denoted high predictive accuracy of the model. Furthermore, CRC patients showed a urine metabolic profile characterized by a higher sensitivity compared to CEA serum levels. The sensitivity of CEA is well known to be low, but it is often used as a serum biomarker in CRC for integrating diagnosis and monitoring the follow up [20]. Our data confirmed CEA detection limitations, including its relatively low sensitivity and specificity [2], and suggested the possibility of integrating these data with the metabolomics approach. Our study indicates a better sensitivity and specificity of the urine profile test in CRC samples compared to those analyzed with  $^1\text{H}$ -NMR spectroscopy by Wang et al. [21] in adenomatous polyp samples. The improved sensitivity and specificity of urine profile tests in CRC samples may be due to a more advanced stage of colon disease in our samples in the CRC group. In our study, the metabolites identified were the end products of several metabolic pathways relating to lipid and amino acid metabolism, which are known to be perturbed during tumor cell proliferation [10]. We found that the levels of acetoacetate, creatinine, and histamine were increased, while those of creatine, phenylacetylglycine, and tryptophan decreased in the urine of the subjects affected by CRC. The increased concentrations of acetoacetate observed in CRC urine may be related to lipid metabolic changes associated with tumor development. The increased content of acetoacetate in our CRC samples could be related to increased  $\beta$ -oxidation due to the strong energy demands accompanying tumor growth. Indeed, in tumor cells, acetyl-CoA is not converted into citrate by the tricarboxylic acid cycle but is processed through an alternative pathway to form ketone bodies, which represent a very efficient fuel source, preferable even to glucose [11,22].

Cancer cells use some amino acids as an energy source [23]. In our study, we found that creatinine levels increased while those of creatine and tryptophan decreased in the urine of CRC patients. Urinary creatinine increase is normally connected to muscle catabolism, breakdown of proteins, or a combination of the two processes [24]. The increased creatinine levels observed in our tumor samples may be due to the non-enzymatic and irreversible conversion of creatine [25], resulting in significantly decreased levels in CRC urine samples. On the other hand, we found decreased tryptophan levels in CRC urine samples. Altered levels of tryptophan in urine have been observed in other types of tumors, such as breast and bladder cancer [26]. Tryptophan is an essential amino acid metabolized in tumor microenvironments, immune-privileged, or inflammation sites. Degradation of tryptophan by the enzyme indoleamine-2,3-dioxygenase (IDO) is considered to be an immune defense mechanism, which inhibits the growth of intracellular bacteria, viruses, parasites [27], and malignant tumor cells [28]. Based on different pieces of evidence [29,30], the catabolic products of tryptophan in cancer are considered important microenvironmental factors that suppress antitumor immune responses. Another metabolite that was increased in our CRC urine samples was histamine. It has been reported that histamine can regulate the proliferation and angiogenesis of cancer cells [31–33]. Finally, another metabolite decreased in our CRC urine samples was phenylacetylglycine, which is considered a minor



metabolite of fatty acids. Increased excretion of phenylacetylglutamine has been observed in gastric cancer patients, in which its levels correlated with the cancer T stage [6], and in rats treated with alkylating agents that produced precancerous colorectal lesions possibly related to gut microflora metabolism dysbiosis [34]. Significantly, even if the low number of subjects investigated does represent a limit of the work, a recent study by Wang et al. [11] analyzing the metabolomics profiles of CRC versus healthy patients reported similar results for acetacetate and phenylalanine, a precursor of phenylacetylglutamine. In addition, a recent study by Deng et al. [35] identified four metabolites as robust CRC urine biomarkers: proline, diacetylspermine, kynurenine, and glucose. Because kynurenine is a tryptophan metabolite, its increased levels correlate well with the decrease of tryptophan found in our CRC samples.

In conclusion, metabolomics analysis showed a different urinary metabolic profile between CRC patients and healthy controls undergoing colonoscopy. Altered levels of some metabolites resulted in a significant correlation with colon cancer and were able to distinguish with accuracy CRC patients from C. Both our results and data from literature confirm that urinary metabolic profiling is an effective tool for identifying CRC, and it may be useful in improving diagnosis, prognosis, and response to therapy, representing a noninvasive diagnostic support to be used before other more invasive tests. Though a limit of the study is the number of subjects investigated, a strength of the study is the similar nutritional status of all the subjects undergoing colonoscopy. Our results suggest that for CRC identification, the metabolomics analysis does not need to be associated with colonoscopy considering that similar discriminant results can be observed without it. The next important step is the validation of these data in an independent larger group of patients to transfer this information into the clinic.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/16/5401/s1>, Figure S1: Principal Component Analyses (PCA) scores plot of 1H-NMR spectra of urine samples: C (full circle) and CRC (open circle) subjects. Two outliers are indicated with arrows. Figure S2: Hotelling's T2 versus DModX plot: C (full circle) and CRC (open circle) subjects. Table S1: Proton chemical shifts of the metabolites identified in the 500 MHz 1H-NMR spectra of urine samples.

**Author Contributions:** L.Z. and L.A. conceived the study and supervised the project. A.R., F.C. obtained the samples and clinical details. C.P. and V.P.L. performed metabolomics experiments and data analysis. C.P., M.P., and J.L.G. wrote the first draft of the manuscript, and C.P., M.P., L.A., V.F., and M.M., contributed to the final version. C.P., M.P., L.A., J.L.G., L.Z., V.F., and M.M. critically reviewed the data and the manuscript. All the authors have read and approved the whole content of this submitted manuscript.

**Funding:** This work was supported by R.A.S. LR7/2007. Grant Code CRP 18488 (L.A.).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Soerjomataram, I.; Lortet-Tieulent, J.; Parkin, D.M.; Ferlay, J.; Mathers, C.; Forman, D.; Bray, F. Global burden of cancer in 2008: A systematic analysis of disability-adjusted life-years in 12 world regions. *Lancet* **2012**, *380*, 1840–1850. [CrossRef]
2. El-Awady, S.; Lithy, R.; Morshed, M.; Khafagy, W.; Abd Monem, H.; Waleed, O.; Badr, S.; Fekry, A.; El Nakeeb, A.; Ghazy, H.; et al. Utility of serum preoperative carcinoembryonic antigen in colorectal cancer patients. *Hepato-Gastroenterology* **2009**, *56*, 361–366. [PubMed]
3. Hewitson, P.; Glasziou, P.; Watson, E.; Towler, B.; Irwig, L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): An update. *Am. J. Gastroenterol.* **2008**, *103*, 1541–1549. [CrossRef] [PubMed]
4. Ahlquist, D.A.; Hongzhi, Z.; Domanico, M.; Mahoney, D.W.; Yab, T.C.; Taylor, W.R.; Butz, M.L.; Thibodeau, S.N.; Rabeneck, L.; Paszat, L.F.; et al. Next-Generation Stool DNA Test Accurately Detects Colorectal Cancer and Large Adenomas. *Gastroenterology* **2012**, *142*, 248–256. [CrossRef]
5. Grady, W.M.; Markowitz, S.D. The molecular pathogenesis of colorectal cancer and its potential application to colorectal cancer screening. *Dig. Dis. Sci.* **2015**, *60*, 762–772. [CrossRef]
6. Jung, J.; Jung, Y.; Bang, E.J.; Cho, S.; Jang, Y.J.; Kwak, J.M.; Ryu, D.H.; Park, S.; Hwang, G.S. Noninvasive diagnosis and evaluation of curative surgery for gastric cancer by using NMR-based metabolomic profiling. *Ann. Surg. Oncol.* **2014**, *21*, 736–742. [CrossRef]

7. Nicholson, J.K.; Holmes, E.; Kinross, J.M.; Darzi, A.W.; Takats, Z.; Lindon, J.C. Metabolic phenotyping in clinical and surgical environments. *Nature* **2012**, *491*, 384–392. [\[CrossRef\]](#)
8. Chan, E.C.Y.; Koh, P.K.; Mal, M.; Cheah, P.Y.; Eu, K.W.; Backshall, A.; Cavill, R.; Nicholson, J.K.; Keun, H.C. Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS). *J. Proteome Res.* **2009**, *8*, 352–361. [\[CrossRef\]](#)
9. Barberini, L.; Restivo, A.; Noto, A.; Deidda, S.; Fattuoni, C.; Fanos, V.; Saba, L.; Zorcolo, L.; Mussap, M. A gas chromatography-mass spectrometry (GC-MS) metabolomics approach in human colorectal cancer (CRC): The emerging role of monosaccharides and amino acids. *Ann. Transl. Med.* **2019**, *7*, 727. [\[CrossRef\]](#)
10. Cheng, Y.; Xie, G.; Chen, T.; Qiu, Y.; Zou, X.; Zheng, M.; Tan, B.; Feng, B.; Don, T.; He, P.; et al. Distinct urinary metabolic profile of human colorectal cancer. *J. Proteome Res.* **2012**, *11*, 1354–1363. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Wang, Z.; Lin, Y.; Liang, J.; Huang, Y.; Ma, C.; Liu, X.; Yang, J. NMR-based metabolomic techniques identify potential urinary biomarkers for early colorectal cancer detection. *Oncotarget* **2017**, *8*, 105819–105831. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Nicholson, B.D.; Shinkins, B.; Pathiraja, I.; Roberts, N.W.; James, T.J.; Mallett, S.; Perera, R.; Primrose, J.N.; Mant, D. Blood CEA levels for detecting recurrent colorectal cancer. *Cochrane Database Syst. Rev.* **2015**, *12*, CD011134. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Piras, C.; Pintus, R.; Pruna, D.; Dessì, A.; Atzori, L.; Fanos, V. Pediatric Acute-onset Neuropsychiatric Syndrome and Mycoplasma Pneumoniae Infection: A Case Report Analysis with a Metabolomics Approach. *Curr. Pediatr. Rev.* **2019**, *15*. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Sousa, S.A.A.; Magalhaes, A.; Ferreira, M.M.C. Optimized bucketing for NMR spectra: Three case studies. *Chemometr. Intell. Lab.* **2013**, *122*, 93–102. [\[CrossRef\]](#)
15. Wu, Y.; Li, L. Sample normalization methods in quantitative metabolomics. *J. Chromatogr. A* **2016**, *1430*, 80–95. [\[CrossRef\]](#)
16. Weljie, A.M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C.M. Targeted Profiling: Quantitative Analysis of <sup>1</sup>H-NMR Metabolomics Data. *Anal. Chem.* **2006**, *2*, 4430–4442. [\[CrossRef\]](#)
17. Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D.S.; Xia, J. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucl. Acids Res.* **2018**, *46*, W486–W494. [\[CrossRef\]](#)
18. Schicho, R.; Shaykhtudinov, R.; Ngo, J.; Nazyrova, A.; Schneider, C.; Panaccione, R.; Kaplan, G.G.; Vogel, H.J.; Storr, M. Quantitative metabolomic profiling of serum, plasma, and urine by <sup>1</sup>H-NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. *J. Proteome Res.* **2012**, *11*, 3344–3357. [\[CrossRef\]](#)
19. Bathe, O.F.; Farshidfar, F. From Genotype to Functional Phenotype: Unraveling the Metabolomic Features of Colorectal Cancer. *Genes* **2014**, *5*, 536–560. [\[CrossRef\]](#)
20. Lin, J.K.; Lin, C.C.; Yang, S.H.; Wang, H.S.; Jiang, J.K.; Lan, Y.T.; Lin, T.C.; Li, A.F.Y.; Chen, W.S.; Chang, S.C. Early postoperative CEA level is a better prognostic indicator than is preoperative CEA level in predicting prognosis of patients with curable colorectal cancer. *Int. J. Colorectal Dis.* **2011**, *26*, 1135–1141. [\[CrossRef\]](#)
21. Wang, H.; Tso, V.; Wong, C.; Sadowski, D.; Fedorak, R.N. Development and Validation of a Highly Sensitive Urine-Based Test to Identify Patients with Colonic Adenomatous Polyps. *Clin. Trans. Gastroenterol.* **2014**, *5*, e54. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Jobard, E.; Pontoizeau, C.; Blaise, B.J.; Bachelot, T.; Elena-Herrmann, B.; Trédan, O. A serum nuclear magnetic resonance-based metabolomic signature of advanced metastatic human breast cancer. *Cancer Lett.* **2014**, *343*, 33–41. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Argiles, J.; Azco Bieto, J. The metabolic environment of cancer. *Mol. Cell. Biochem.* **1988**, *81*, 3–17. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Eisner, R.; Stretch, C.; Eastman, T.; Xia, J.; Hau, D.; Damaraju, S.; Greiner, R.; Wishart, D.S.; Baracos, V.E. Learning to predict cancer-associated skeletal muscle wasting from <sup>1</sup>H-NMR profiles of urinary metabolites. *Metabolomics* **2011**, *7*, 25–34. [\[CrossRef\]](#)
25. Wyss, M.; Kaddurah-Daouk, R. Creatine and creatinine metabolism. *Physiol. Rev.* **2000**, *80*, 1107–1213. [\[CrossRef\]](#)
26. Bryan, G.T. The role of urinary tryptophan metabolites in the etiology of bladder cancer. *Am. J. Clin. Nutr.* **1971**, *24*, 841–847. [\[CrossRef\]](#)

27. Denz, H.; Orth, B.; Weiss, G.; Herrmann, R.; Huber, P.; Wachter, H.; Fuchs, D. Weight loss in patients with hematological neoplasias is associated with immune system stimulation. *Clin. Investig.* **1993**, *71*, 37–41. [[CrossRef](#)]
28. Carlin, J.M.; Ozaki, Y.; Byrne, G.I.; Brown, R.R.; Borden, E.C. Interferons and indoleamine 2,3-dioxygenase: Role in antimicrobial and antitumor effects. *Experientia* **1989**, *45*, 535–541. [[CrossRef](#)]
29. Murr, C.; Bergant, A.; Widschwendter, M.; Heim, K.; Schrocksnadel, H.; Fuchs, D. Neopterin is an independent prognostic variable in females with breast cancer. *Clin. Chem.* **1999**, *45*, 1998–2004. [[CrossRef](#)]
30. Platten, M.; Wick, W.; Van den Eynde, B.J. Tryptophan catabolism in cancer: Beyond IDO and tryptophan depletion. *Cancer Res.* **2012**, *72*, 5435–5440. [[CrossRef](#)]
31. Medina, V.A.; Rivera, E.S. Histamine receptors and cancer pharmacology. *Br. J. Pharmacol.* **2010**, *161*, 755–767. [[CrossRef](#)] [[PubMed](#)]
32. Cricco, G.P.; Mohamad, N.A.; Sambuco, L.A.; Genre, F.; Croci, M.; Gutiérrez, A.S.; Medina, V.A.; Bergoc, R.M.; Rivera, E.S.; Martín, G.A. Histamine regulates pancreatic carcinoma cell growth through H3 and H4 receptors. *Inflamm. Res.* **2008**, *57*, 23–24. [[CrossRef](#)] [[PubMed](#)]
33. Medina, V.; Croci, M.; Crescenti, E.; Mohamad, N.; Sanchez-Jiménez, F.; Massari, N.; Nuñez, M.; Cricco, G.; Martín, G.; Bergoc, R.; et al. The role of histamine in human mammary carcinogenesis: H3 and H4 receptors as potential therapeutic targets for breast cancer treatment. *Cancer Biol. Ther.* **2008**, *7*, 28–35. [[CrossRef](#)] [[PubMed](#)]
34. Qiu, Y.; Cai, G.; Su, M.; Chen, T.; Liu, Y.; Xu, Y.; Ni, Y.; Zhao, A.; Cai, S.; Xu, L.X. Urinary metabonomic study on colorectal cancer. *J. Proteome Res.* **2010**, *9*, 1627–1634. [[CrossRef](#)]
35. Deng, L.; Ismond, K.; Liu, Z.; Constable, J.; Wang, H.; Alatise, O.I.; Weiser, M.R.; Kingham, T.P.; Chang, D. Urinary Metabolomics to Identify a Unique Biomarker Panel for Detecting Colorectal Cancer: A Multicentre Study. *Cancer Epidemiol. Prev. Biomark.* **2019**, *28*, 1283–1291. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).