Malaysian Journal of Chemistry, 2016, Vol. 18(1), 37–45

Colorectal Cancer Screening by Using Urine NMR Spectra

Vu Thi Kim Ngoc[1*], Le Thanh Tu[2], Phan Thanh Hai[3] and Pham van Huong[4]

[1] Center of Analytical Services and Experimentations (CASE) HCMC-Vietnam
[2] University of Natural Sciences HCMC-VN
[3] Medical Center MEDIC HCMC-VN
[4] University Bordeaux 1, Bordeaux / Talence, France

*Corresponding author (e-mail: ngocvtk@case.vn)

Colorectal cancer (CRC) is one of the most common types of cancer. Detecting CRC at an early stage improves survival rates. This study investigates the efficacy of metabolomic technology in screening colorectal cancer in human urine. Urine from both CRC patients and healthy persons was analyzed by NMR spectrometry in conjunction with a multivariate statistics technique. Specific differences have been observed, in particular in the spectral range corresponding to some metabolites. This analysis was to compare and verify important metabolic alterations between CRC patients and healthy persons. The NMR-based metabolomics approach could offer a non-invasive diagnosis tool and help to better understand the mechanism of carcinogenesis.

Key words: NMR metabolomics; urine metabolomics; colorectal cancer; carcinogenesis

Received: November 2014; Accepted: January 2016

Colorectal cancer (CRC) is the second leading cause of cancer mortality. CRC can be curable if identified early and prevented if found at the adenomatous stage [1]. In the present time, colonoscopy is the most effective screening tool for accurate diagnosis of precancerous lesions and cancer morbidity in the colon and rectum [2]. With this procedure, certain tumor biomarkers, such as carcinoembryonicantigen (CEA) and fecal occult blood testing (FOBT), have been clinically used, but with relatively poor sensitivity and specificity [3–5]. Metabolomics is an emerging field of research and is a quantitative collection of low molecular weight compounds, such as metabolic substrates and products, lipids, small peptides, vitamins and other protein cofactors, generated by metabolism [6]. It is currently being used as a mode of research in many disciplines of medicine, including oncology [7–10]. Metabolomics can be performed on urine, serum, tissue and less frequently, on fecal extracts, saliva and amniotic fluid [11]. To date, metabolomic methods include nuclear magnetic resonance spectroscopy (NMR), liquid chromatography-mass spectrometry (LCMS), gas chromatography-mass spectrometry (GC-MS) and capillary electrophoresis mass spectrometry (CE-MS) [12]. However, this technology is fairly new and few human studies have been carried out to validate the results of existing cellular and animal studies, especially in the field of colorectal oncology [13].

In the present study, urine samples from both CRC patients and healthy controls were analyzed by high-resolution 1H NMR. Their spectral profiles subjected to multivariate analysis by AMIX-VERSION 3.9.14 – Bruker to reduce the dimensionality for extracting useful information from complicated NMR data set. Specific differences have been observed, in particular in the spectral range corresponding to some metabolites. This analysis aimed to compare and

verify important metabolic alterations between CRC patients and healthy persons and would be extended to the screening of colorectal cancer.

EXPERIMENTS

Sample Collection and Handling

Urine samples were collected from 51 adults (33 males, 18 females, from 27 to 81 years old, including 23 CRC patients and 28 healthy persons, who had a physical examination at HoChiMinhCity-Vietnam during June 2013 and July 2014. Diagnoses of colorectal cancer were made on the basis of usual clinical and laboratory findings and were confirmed by colonoscopy. No patient had received chemotherapy, radiation therapy or surgery. Each individual (either patients or healthy persons) provided a middle stream sample of morning urine (after overnight lasting) in a sterile cup. Then, liquids of approximately 5 mL were transferred into sterile cryovials, frozen and stored at –80°C.

Sample Preparation

The urine samples were warmed up to room temperature (30 min), then centrifuged for 5 to 10 min at ~6000 c/min to separate any solid matter. For NMR analysis, mixing of 540 microliters of urine, 60 microliters of phosphate buffer pH 7.4 (1.5 M KH2P04 in D20 + 0.1% TSP and some NaN3. The resulting sample was transferred to a 5 mm NMR tube, which was capped and labeled.

NMR Experiments

Proton NMR spectra were obtained from each urine sample on a Bruker AVANCE III 500 Ultrashield Plus CryoNMR Spectrometer with BCU-05 pre-cooling device and a BBI-ATM probe. For each sample, a standard NOESY-1D spectrum was acquired, at 300 K, using a water suppression pulse sequence with water irradiation during relaxation delay and mixing time ('noesypr1d' in Bruker library, SW 10330.58Hz, TD 32 K data points, relaxation delay 4 s, mixing time 100 ms, 128 scans). All spectra were processed with a line broadening of 0.3 and a zero filling factor of 2, automatically phased, and baseline corrected. The chemical shifts were referenced internally to the TSP signal at δ 0.00.

Multivariate Statistical Analysis

The set of NMR data has been considered for multivariate analysis by Amix-version 3.9.14 – Bruker to reduce the dimensionality for extracting useful information from NMR data set. To build the data matrix, variable sized bucketing was performed in the δ 9.50–0.50 ppm region of the standard 1D spectra, excluding the subregion δ 6.00–4.55 ppm to remove variability arising from water suppression and from possible cross-relaxation effect on the urea signal via solvent exchanging protons. The procedure consisted of dividing the spectra into small regions (buckets) of 0.05 ppm and then merging some of these buckets to larger buckets in regions where chemical shift drifts (due to small pH differences or variations in sample ionic strength) were observed. After normalizing and scalling , Principal Component Analysis (PCA), Partial Least Squares Discriminant Analysis (PLS-DA) were performed. The set of marker metabolites was identified from NMR data by Chenomx NMR suite 7.0 software.

RESULTS AND DISCUSSION

Discrimination of CRC and Healthy Spectra

The spectra of CRC and healthy persons are showed on Figure 1.

Many differences in metabolite component of CRC spectra and healthy spectra could be observed in particular in the spectral range marked with 1* – 4* in Figure 1. In the zone marked 1* from 7.5–7.0 ppm and the zone marked 3* from 3.9–3.4 ppm, big increase could be seen at peak 23, which corresponded to tryptophan. While in the zones marked 2*, a significant change was found in CRC spectra at peak 19, which corresponded to phenylacetylglutamine. Otherwise, the presence of several special peaks was indicated in zone 4* on the CRC spectrum, the peaks 12 and 14 corresponded to glutamate and citrate, respectively.

Typically, CRC spectrum indicated the presence of specific biological component(s) which did not exist in healthy spectrum. Otherwise, the presence of several special peaks was indicated on the CRC spectrum. Their intensity varied depending on the relative concentration of the corresponding metabolites.

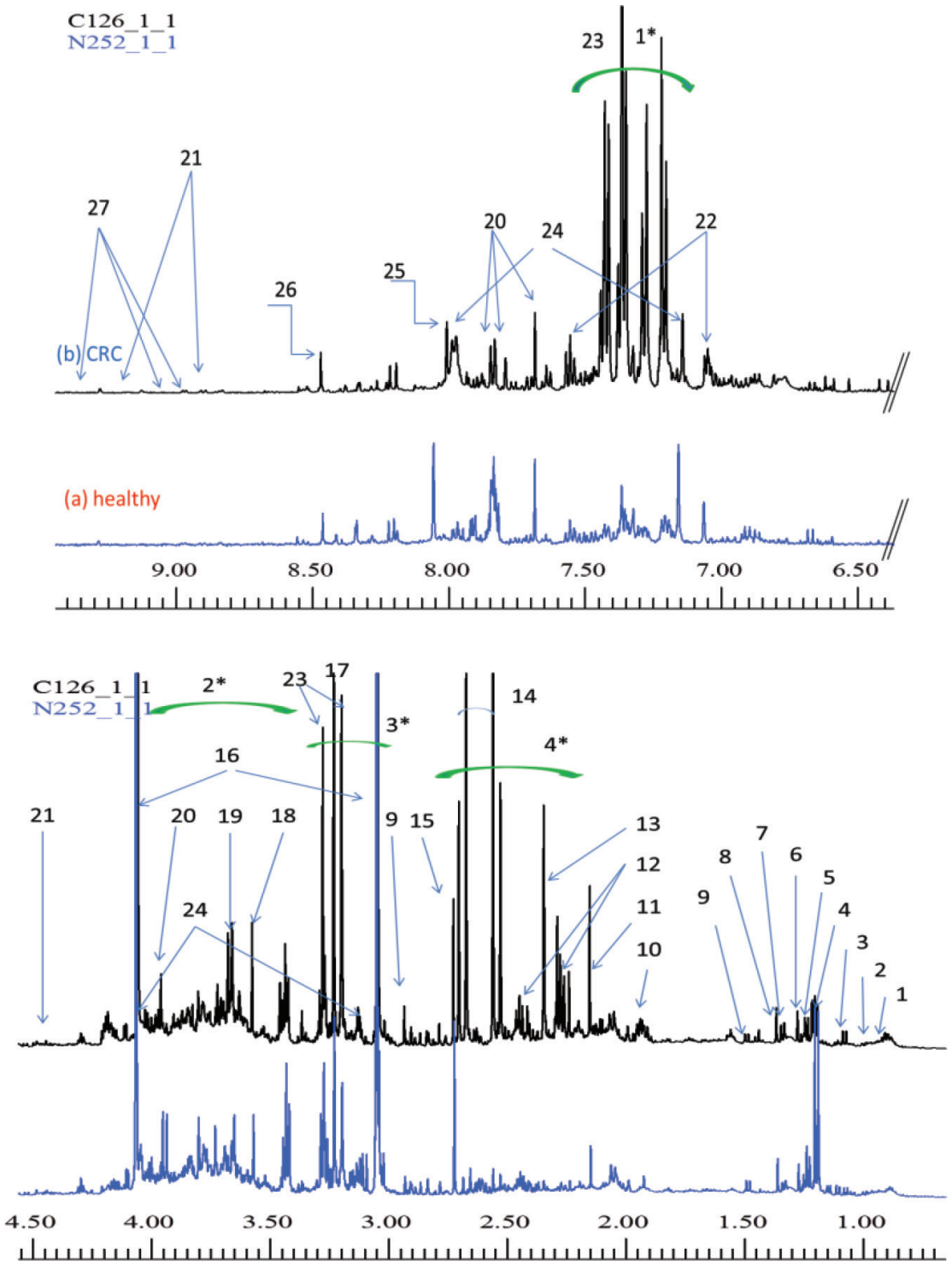

Figure 1. The 500 MHz 1H-NMR spectra of urine from (a) a healthy person, and (b) a colerectal cancer patient. Signal assignment:1,R-hydroxybutyrate; 2, valine; 3, isobutyrate; 4, -aminoisobutyrate; 5, methyl-hydroxybutyrate; 6, -hydroxyisovalerate; 7, lactic acid and threonine; 8,R-hydroxyisobutyrate; 9, alanine; 10, N-acetylglutamine; 11, p-cresol; 12,glutamate; 13, pyruvate/oxalacetate; 14, citrate; 15, dimethylamine; 16, creatinine; 17, trimethylamine-N-oxide and betaine; 18, glycine; 19,phenylacetylglutamine/glutamate; 20, hippurate; 21, trigonelline; 22,p-hydroxyphenylacetate; 23, trytophane; 24, histidine; 25, 3-methylhistidine; 26, formate and 27, trigonellinamide.

Statistical Models for CRC Screening

The high inter-individual variability in urinary profiles and their complexity made any attempt at visual comparison of these spectra an unproductive task. Instead, multivariate analysis allowed the finding of consistent variation patterns within the data set. To study the principal sources of variation among the results, detect intrinsic clustering and possible outliers, and distinguish cancer patients from healthy individuals, exploratory PCA was applied.

Identification of CRC fingerprint using unsupervised model: principal component analysis (PCA). In a first step, we investigated the CRC spectrum and healthy spectrum in an unsupervised approach by using principal component analysis (PCA). This approach was chosen because

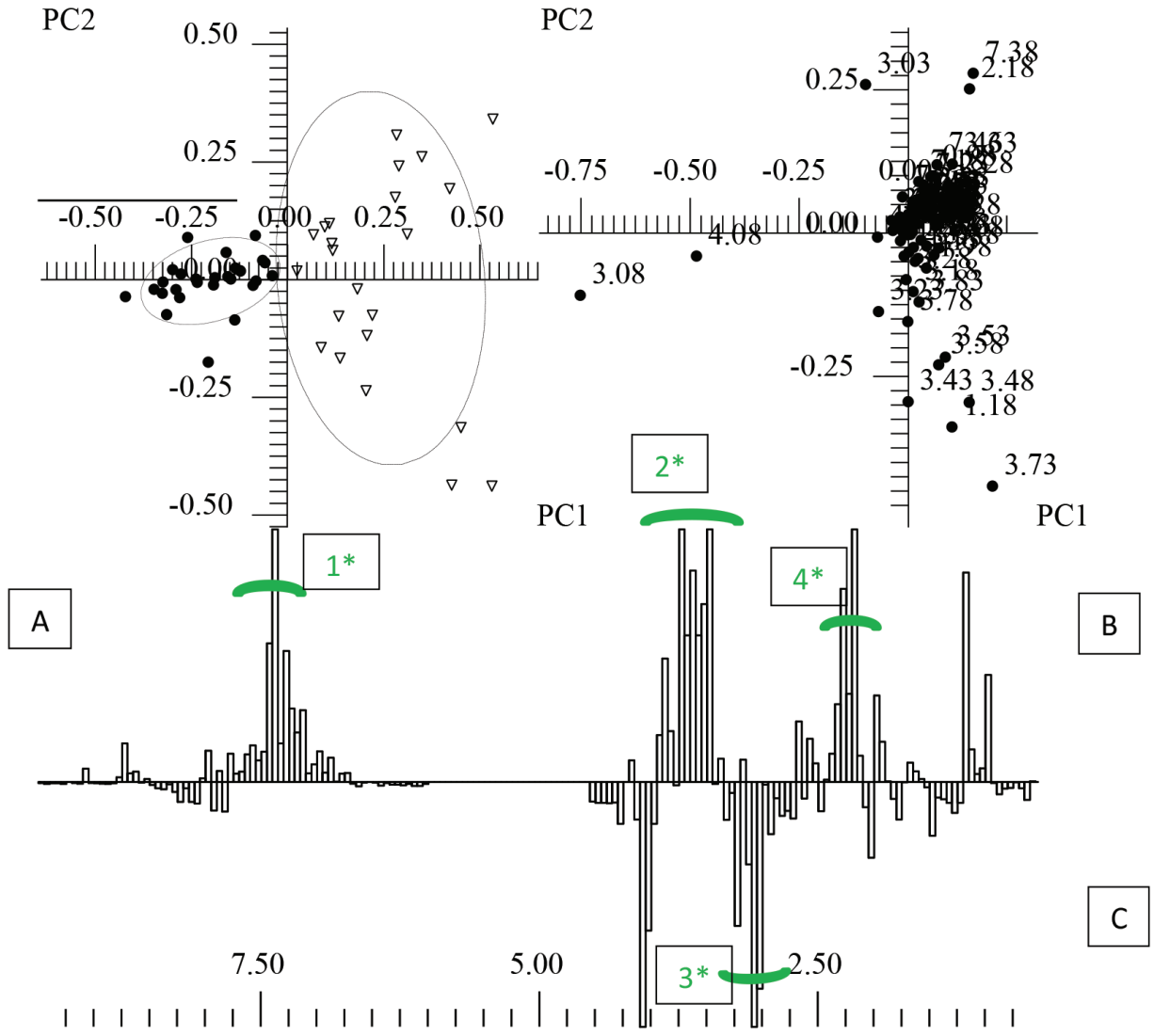

Figure 2. Principal component analysis model of CRC urine and healthy urine specimens. (A): PCA score plot displays the relationship between healthy samples (black dots) and CRC samples (triangles) with the proportion of variance explained by principal component PC1 and PC2. (B): 2D-Loading plot and (C) 1D-Loading plot show the resonances that contribute to the variance of the specific principal component.

unsupervised methods did not utilize clinical information about the samples. PCA methods reduce the original set of variables to a new set of principal components that retain the variance-covariance structure of the data, but use lesser dimensions of data space.

We analyzed the PCA results with six principal components (PC) to detect intrinsic clustering and possible outlier. Confidence level was 95% with 59.30% variances explained by PC1 and 9.94% variances explained by PC2. The PC1 vs PC2 scores scatter plot showed *a trend for separation of two groups, one corresponded to cancer patients and the other to healthy persons,* mostly along PC1. In Figure 2A, hostelling T2 ellipses assumed a Gaussian distribution of bucket variables in CRC region and healthy region with confidence interval of 75%.

After preliminary statistical analysis, Figure 2B depicted the loadings of variables on a plane defined by the first (PC1) and second (PC2) principal components. While Figure 2C looked like a 1H-spectra with four marked zones in Figure 1, the peaks indicated spectral regions which contributed significantly to PC1 in accordance. The set of marker metabolites corresponding to

Table 1. Representative differential metabolites contributed for the separation between the CRC patients and the healthy controls derived from NMR analysis.

No.	Metabolite	Cluster (ppm)	Multiplicity[a]	Change of CRC from healthy[b]	Zone in Figure 2C
1	Trytophan	3.30	dd	++	
		3.48	dd		**3***
		4.05	dd		
		7.19	m		
		7.27	m		
		7.31	s		**1***
		7.53	d		
		7.72	d		
2	Histidine	3.14	m	-	
		3.24	m		**3***
		3.98	dd		
		7.09	s		
		7.88	s		**1***
3	4-hydroxyphenyl acetate	3.44	s	+	**3***
		6.85	d		
		7.15	d		**1***
4	Phenylacetylglutamine	3.65	s	+	**3***
5	Citrate	2.53	d	++	
		2.69	d		**4***
6	p-cresol	2.25	s		
		6.82	d	+	**4***
		7.13	d		
7	Glutamate	2.04	m	+	
		2.12	m		
		2.32	t		**4***
		2.36	t		
		3.75	dd		

a/d, m, q, s and t are doublet, multiplet, quartet, singlet and triplet, respectively.
b/ a positive sign indicates a relatively higher concentration present in CRC patients while a negative sign means a relatively lower concentration as compared to the healthy controls.

these variables was identified by Chenomx NMR Suite software and shown in Table 1with different intensities related to the health status.

Although this set of variables only explained 85.87% of the variability between the first two PCs, it was enough to deconvolute the set of cases in two subsets according to the healthy subjects and to cancer type.

Our results gave a complementary data to those obtained by Qiu *et al*. (2010) using mostly gas chromatography-mass spectroscopy [14] and by other workers in optical spectroscopies [15–29]. In particular we found a higher expression level of citrate in CRC urines, which might be produced from the nutritional supplementation.

Prediction of urine-type using supervised model partial least squares (PLS). In addition, we investigated a supervised method –PLS- in order to construct a predictive model that was capable of distinguishing between CRC and healthy urines. In the T/T scores of PLS model, a reasonably good separation was observed. The obtained class discrimination was explained X variances up to 76% and Y variances up to 90%.

An analysis of the PLS loadings was conducted to identify the metabolites contributing to the CRC fingerprint using the entire sets of patients with CRC and healthy subjects. The values of the relative concentrations of metabolites were estimated through the integration of the signals in the NMR spectra. By comparing the spectra of

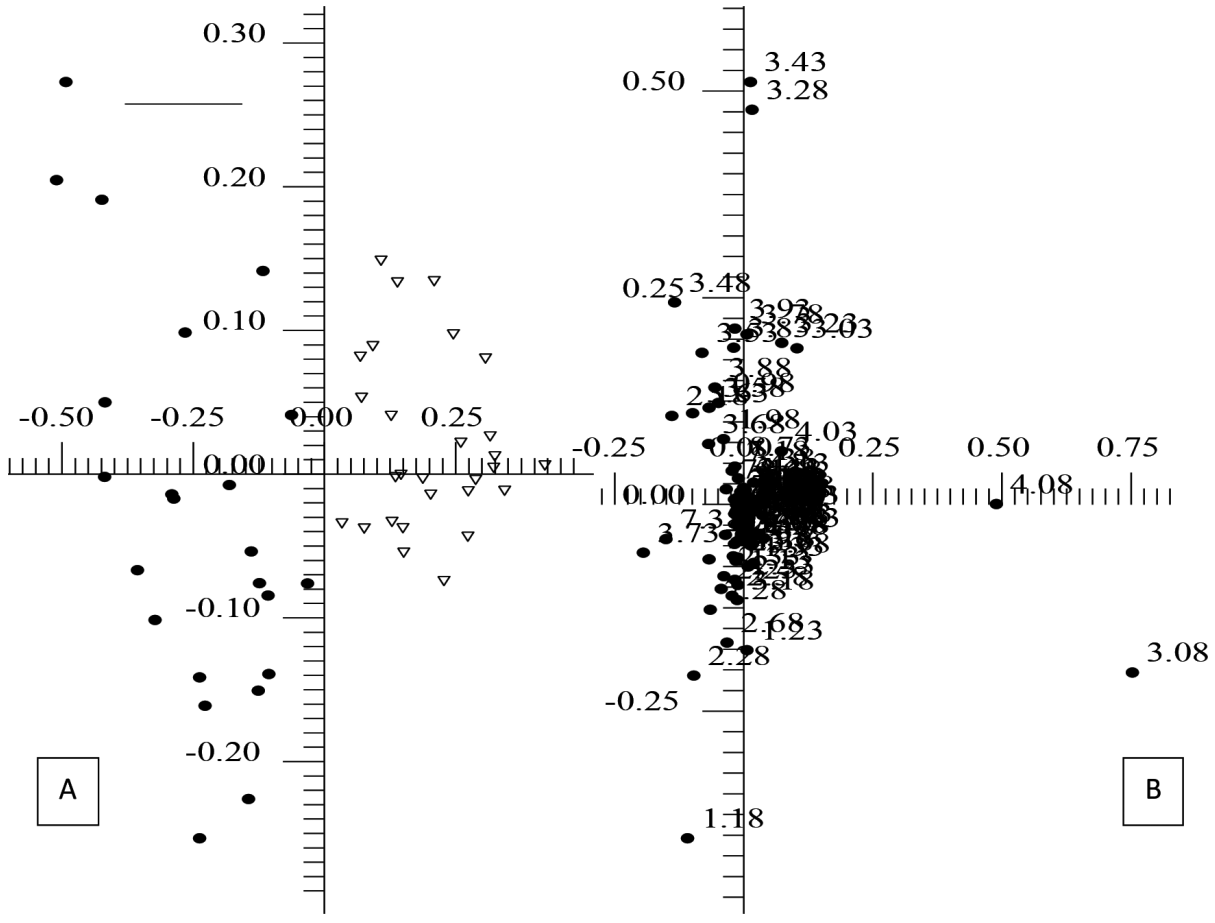

Figure 3. Partial Least Squares model of CRC urine and healthy urine specimens. (A): T/T score plot displays the discrimination of healthy samples (triangles) and CRC samples (black dots) with the proportion of variance explained by principal component PC1 and PC2. (B): 2D-Loading plot show the resonances that contribute to the variance of the specific principal component.

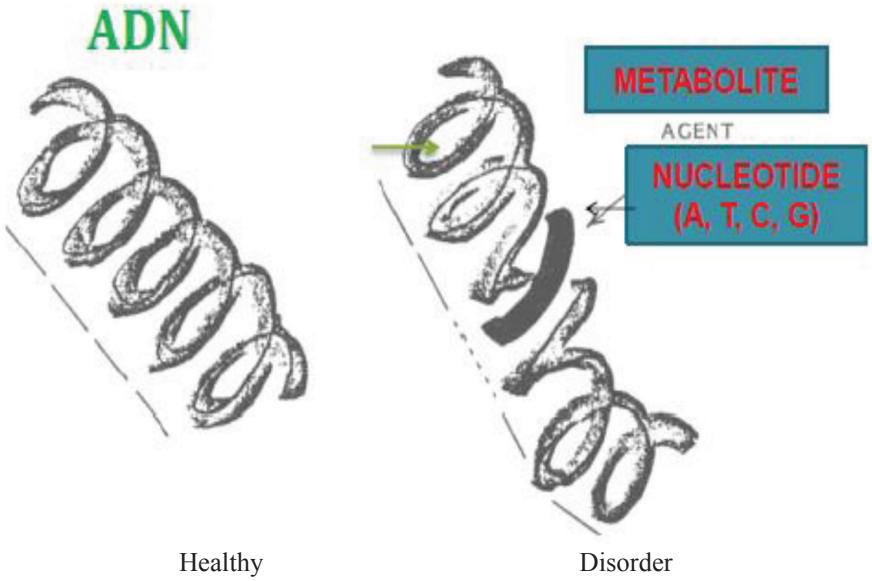

Figure 4. Formation of disordering complexes between metabolites and nucleotides inducing the lost of coding and messaging in the mechanism of carcinogenesis.

the urine samples of CRC patients with those of healthy subjects, it appears that the CRC patients are characterized by higherlevels of trytophan, 4-hydroxyphenyl acetate, phenylacetylglutamine, citrate, p-cresol, glutamate and lower level of histidine. The worsening of the clinical conditions thus corresponds to amore marked difference in the concentration of several of the metabolites that are responsible of the disease fingerprint.

Moreover, the presence of proton donors (-OH, -NH) and acceptors (>C=O, >N-) in the enhanced metabolite molecules suggests their possible bonding with ADN nucleotides creating disordering complexes and involves in the mechanism of carcinogenesis as suggested recently by PV Huong *et al*. [15].

The good discrimination of healthy and CRC urines on unbiased PCA and biased PLC indicated a high potential of this non-invasive urinary metabonomic strategy as a complementary screening tool for colorectal cancer. Of course, more samples would be examined to reinforce the above conclusion in the aim of a better understanding about the disease evolution. Once the metabolomic fingerprint of CRC is firmly established, the next step would be to test the accuracy of this fingerprint and these metabolites in a prospective blinded study against the reference standard by GC-MS or by LC/MS/MS. Success with metabolomics as a diagnostic and prognostic tool would likely help to fundamentally change the physicians' approach to human healthcare.

CONCLUSION

The burden of colorectal cancer is growing worldwide and with it a more desperate need for better tools to detect, diagnose and monitor the disease. The interesting information obtained from this study proved that the diseased state of CRC conducted characteristic changes in the small molecule metabolites secreted in urine. To fully develop this technology, further study is necessary on a much larger data set, in the order of 500 samples, to yield a better resolution. A larger sample group would result in a better predictive model which could then be used to diagnose and potentially. Such a model could be used in human clinical medicine as a non-invasive diagnostic tool, detecting more cases of CRC at an early stage, and increasing survival rates. Effective determination of the significance of each bin, in conjunction with other techniques which have been efficient in the identification of nanomaterials and bio-objects

[16–26], could also lead to potential biomarkers for CRC. Once cancerous biomarkers are identified, they might be explored for potential drug targets which could lead to more effective, safer treatments for a serious human health threat. Furthermore, some metabolites with strong enhancement variation would intervene in the interaction with ADN nucleotides creating disorder complexes in the mechanism of carcinogenesis.

ACKNOWLEDGMENTS

This work was supported by funding from the Science Research Program of Science and Technology Department –HCMC-Vietnam and by grants from Bruker-BioSpin-Asia Pacific.

REFERENCES

1. Sung, J. J., Lau, J. Y., Goh, K. L. and Leung, W. K. (2005) Increasing incidence of colorectal cancer in Asia: implications for screening, *Lancet Oncology*, **6(11)**, 871–876.
2. Winawer, S. J., Zauber, A. G., May Nah, H., O'Brien, M. J., Gottlieb, L. S., Sternberg, S. S., Waye, J. D., Schapiro, M., Bond, J. H., Panish, J. F., Ackroyd, F., Shike, M., Kurtz, R. C., Hornsby-Lewis, L., Gerdes, H., Stewart, E. T., Lightdale, C. J., Edelman, M. and Fleisher, M. (1993) Prevention of colorectal cancer by colonoscopic polypectomy, *N. Engl. J. Med.*, **329(27)**, 1977–1981.
3. Fletcher, R. H. (1986) Carcinoembryonic antigen, *Ann. Intern. Med.*, **104(1)**, 66–73.
4. Kronborg, O., Fenger, C., Olsen, J., Jo¨rgensen, O. D. and So¨ndergaard, O. (1996) Randomised study of screening for colorectal cancer with faecal-occult-blood test, *Lancet*, **348(9040)**, 1467–1471.
5. Mandel, J. S., Bond, J. H., Church, T. R., Snover, D. C., Bradley, G. M., Schuman, L. M. and Ederer, F. (1993) Reducing mortality from colorectal cancer by screening for fecal occult blood, *N. Engl. J. Med.*, **328(19)**, 1365–1371.
6. Claudino, W. M., Quattrone, A., Biganzoli, L., Pestrin, M., Bertini, I. and Di Leo A. (2007) Metabolomics: Available results, current research projects in breast cancer, and future applications, *J. Clin. Oncol.*, **25(19)**, 2840–2846.
7. Kind, T., Tolstikov, V., Fiehn, O. and Weiss, R. H. (2007) A comprehensive urinary metabolomic approach for identifying kidney cancer, *Analytical Biochem.*, **363(2)**, 185–195.
8. Kaddurah-Daouk, R., McEvoy, J., Baillie, R. A. *et al.* (2007) Metabolomic mapping of atypical antipsychotic effects in schizophrenia, *Mol. Psychiatry*, **12(10)**, 934–945.
9. Broadhurst, D. I., Brown, M., Dunn, W. B. *et al.* (2008) The application of UPLC/LTQ-Orbitrap mass spectrometry to human metabolite biomarker discovery. A case study: pre-eclampsia, *Reprod. Sci.*, **15(2)**, 208A–208A.
10. Wang, H., Tso, V. K., Slupsky, C. M. and Fedorak, R. N. (2010) Metabolomics and detection of colorectal cancer in humans: a systematic review, *Future Oncol.*, **6(9)**.
11. Siffedeen, J. S., Rankin, K. N., Fu, H., Dieleman, L. A., Slupsky, C. and Fedorak, R. (2008) Metabolomics in inflammatory bowel disease: differentiating patients with or without IBD, *Gastroenterology*, **134(4)**, A203–A203.
12. Lenz, E. M. and Wilson, I. D. (2007) Analytical strategies in metabonomics, *Journal of Proteome Research*, **6(2)**, 443–458.
13. Issaq, H. J., Van, Q. N., Waybright, T. J., Muschik, G. M. and Veenstra, T. D. (2009) Analytical and statistical approaches to metabolomics research, *J. Sep. Sci.*, **32**, 2183–2199.
14. Qiu, Y., Cai, G., Su, M., Chen, T., Liu, Y., Xu, Y., Ni, Y., Zhao, A., Cai, S., Xu, L. X. and Jia, W. (2010) Urinary metabonomic study on colorectal cancer, *J. Proteome. Res.*, **9(3)**, 1627–1634.
15. Pham, V. H., Vu, K. N., Thai, V. H., Tu, L. T. and Hai, P. T. (2015) Studies of colorectal cancer by nuclear magnetic resonance spectroscopy, *Int J. Magn. Nucl. Sci.*, **1(1)**,1–5.
16. Huong, P. V. and Verma, A. L. (1993) Presence of tri-iodide ions in iodine-intercalated IBi2Sr2CaCu2O8 superconductors, *Phys. Rev B.*, **48(13)**, 9869–9872.
17. Huong, P. V., Ajayan, P. M., Cavagnat, R. and Stephan, O (1995) Temperature-dependent vibrational spectra of carbon nanotubes, *Phys. Rev. B.*, **51(15)**, 10048–10051.
18. Huong, P. V., Jerome, D., Auban-Senzier, P. and Bernier, P. (1995) Sciences and technology of fullerene materials, in *MRS Proceedings*, Boston, Vol. 359.
19. Huong, P. V., Khoi, P. H., Tam, N. T. T., Hoa, P. L. P. and Tuong, L. T. C. (1999) A Raman spectroscopic study of photoluminescent porous silicon fibres, *Int. J. Inorg. Mater.*, **1(3)**, 209–212.

20. Choy, J. H., Kim, Y. I., Hwang, S. J. and Huong, P. V. (2000) Trigonal planar (D-3h) AuI3 complex stabilized in a solid lattice, *J. Phys. Chem. B.*, **104(31)**, 7273–7277.

21. Choy, J. H., Kim, Y. I., Hwang, S. J., Muraoka, Y., Ohnishi, N. *et al*. (2000) HRTEM and micro-Raman studies on superconducting-superionic conducting nanohybrid, Ag1.17I1.54Bi2Sr2CaCu2Oy, *J. Phys. Chem. B.*, **104(39)**, 9086–9090.

22. Kwon, S. J., Choy, J. H., Jung, D. W. and Huong, P. V. (2002) Heterostructured high-Tc superconducting nanohybrid: (Me3S)2HgI4- Bi2Sr2CaCu2Oy, *Phys. Rev. B.*, **66(22)**, 224510–224522.

23. Hwang, J., Kwon, C. W., Portier, J., Campet, G. and Park, H. S. *et al*. (2002) Local crystal structure around manganese in new potassium-based nanocrystalline manganese oxyiodide, *J. Phys. Chem. B.*, **106(16)**, 4053–4060.

24. Pham, V. H. and Giege, R. (1989) Molecular dynamics in biomolecules, in *Spectroscopy of biological molecules*, Bertolluza, C. and Fagano, M. eds., Bologne (Italy), 163–166.

25. Plouvier, S. R., Huong, P. V. and Lambert, P. (1985) Bacterial chromophores and metals in cancer, leukemias and chronic diseases. A study by scanning electron microscopy, laser Raman and electron microspectroscopies, in *4th Eur. Cce on Clinical Haemorheology*, Siena, Italy.

26. Huong, P. V. (1986) Organometallic Interactions in Biological Systems, *Journal of Molecular Structure*, **141**, 203–209.

27. Plouvier, S. R. and Lambert, P. (1988) Organometallic complexes of microbial origin, identified in malignant tumors and leukemia by Raman and X-ray microspectroscopy, in *XVII International Congress of Pathology*, Dublin.

28. Huong, P. V. (1989) *Raman Spectroscopy for Biological Applications in Analytical Raman Spectroscopy*, Wiley, New York., **114**, 397–423.

29. Huong, P. V. (1989) *Metallo-organic complexes and Carcinogenesis in Molecules in Physics, Chemistry and Biology*. Springer, Netherlands, 87–109.