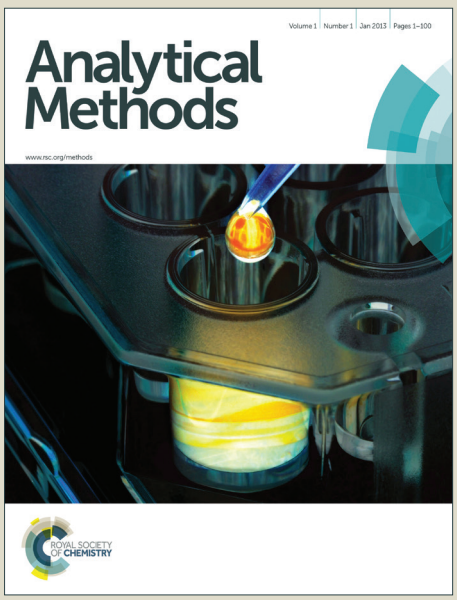
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

Metabolite fingerprint analysis of cervical cancer using LC-QTOF/MS and multivariate data analysis

Qun Liang*, Qian Yu, Haikun Wu, Yong-zhi Zhu, Ai-hua Zhang*

First Affiliated Hospital, School of Pharmacy, Heilongjiang University of Chinese Medicine, Heping Road 24, Xiangfang District, Harbin 150040, China

*Correspondence

Prof. Qun Liang and Ai-hua Zhang

First Affiliated Hospital, School of Pharmacy, Heilongjiang University of Chinese Medicine, Heping Road 24, Xiangfang District, Harbin 150040, China

Tel. & Fax +86-451-86053141

Email: qunliang1970@126.com

Abstract:

Cervical cancer (CC) is the second most common cancer in females worldwide, As yet, the metabolic alterations that are specific for the development of CC have not been fully determined, which also precludes the early diagnosis and prognosis of this pathology. In this pilot study, we determine the metabolic fingerprint of urine samples of CC women and normal people using LC (Agilent 1290 Infinity LC System) coupled with Q-TOF/MS (Agilent, 6550 iFunnel) and independent variable analysis. Urine fingerprints allowed for the discrimination of CC women from controls. In addition, we identified a set of metabolites with a strong discriminative power, such as 3-methylhistidine, citric acid, cytosine, indoleacetic acid, salicyluric acid, L-Methionine, aminomalonic acid, glutaric acid, ursodeoxycholic acid and N-Acetylornithine, involve the key metabolic pathways such as citrate cycle, lysine degradation, tryptophan metabolism, cysteine and methionine metabolism, etc. Finally, we provide the evidence for the implication of these compounds in metabolic routes that may be associated with the early genesis of CC, which highlights their potential use as prognostic markers for the identification of women at risk to develop CC. Urine fingerprints reveal disease-specific metabolic imbalances in CC women.

Keywords:

Cervical cancer; metabolomics; metabolites; biomarkers; identification; urine; LC/-Q/TOF-MS

INTRODUCTION

Cervical carcinoma (CC) is the second most common cancer among women worldwide [1]. CC is a highly malignant neoplasm that generally is diagnosed at advanced stage and is associated with a fatal outcome. Generally, imaging technique is of limited sensitivity for the detection of CC [2,3]. However, the sensitivity of diagnostic markers of CC is relatively low and difficult to get outcome immediately. In clinical practice, there is a need for more effective treatments in order to improve earlier diagnosis of CC patients. Major barriers for reducing morbidity, mortality, and costs of care include limited studies on noninvasive early diagnosis of CC and inadequate prognostic indicators as well as few studies on mechanisms of CC. Despite new diagnostic approaches, the definite diagnosis of this malignancy continues to be challenging. Fortunately, high-throughput metabolomics has been used to explore the particular metabolites, potentially diagnostic and prognostic biomarkers for deep understanding the essence of diseases [4-6].

Metabolomics, a science of systems biology, is the global assessment of endogenous metabolites within a biologic system and represents a "snapshot" reading of gene function, enzyme activity, and the physiological landscape [7]. Metabolite detection, either individual or grouped as a metabolomic profile, is usually performed in biofluids followed by sophisticated multivariate data analysis. Because loss of metabolic homeostasis is common in critical illness, the metabolome could have many applications, including biomarker and drug target identification. Metabolomics could also significantly advance our understanding of the complex pathophysiology of illnesses [8]. Meanwhile, it constructs a unique “metabolite fingerprint” through monitoring low-molecular-weight compounds in body fluids rather than focusing on individual metabolites, reflecting the terminal symptoms of metabolic network of globally biological systems, and aims at gathering the maximum amount of metabolic information for a total interpretation of biological systems [9]. The individual's metabolome is a potentially informative mirror of the impact of disease and its dynamics has led to promising developments in disease research, strongly geared toward the discovery of new biomarkers of understanding the disease mechanisms [10-13].

Metabolomics has been successfully applied to the study of Jaundice syndrome, hepatocellular carcinom, hepatitis b virus in human urine [14-16]. However, to date there are few works aimed at gaining deeper insights into CC through metabolomics. CC is characterized by the absence of reliable diagnostic biomarkers. Therefore, the aim of our study was to devise a multiplatform metabolic fingerprinting approach by use of the LC-QTOF-MS technology with multiple chemometric methods to obtain a comprehensive picture of the early metabolic alternations that define CC, and may reflect on the specific pathophysiological context of the disease. Future studies at later stages of CC will allow us to validate the identified metabolites as prognostic markers for the identification of women who are at risk to develop severe CC.

2. MATERIALS AND METHODS

2.1 Ethical statement and subjects

All participants agreed to participate in this study were required to sign consent forms. The study protocol was approved by the Ethical Committee of Heilongjiang University of Chinese Medicine and was conducted according to the Declaration of Helsinki. Finally, twenty CC women and twenty healthy women (27-37 years) were matched and recruited in this study.

2.2 Collection and preparation of Urine Samples

The urine samples from all participants were centrifuged at 10,000 rpm for 10 min at 4 °C to remove any solid debris. Fractions (50 mL) of the urine supernatants were stored at - 80 °C until analysis. Thawed urine samples were collected via centrifugation at 13,000 rpm for 10 minutes at 4◦C, and filtered through a 0.22 um syringe filter (Millipore, Shanghai, China), 5 μL of the supernatant were injected into the LC/MS analysis..

2.3 Chemicals and reagents

HPLC grade acetonitrile was obtained from Merck (Darmstadt, Germany); methanol (HPLC grade) was purchased from Fisher Scientific Corporation (Loughborough, UK); formic acid was obtained from Honeywell Company (Morristown, New Jersey, USA); water was produced by a Milli-Q Ultra-pure water system (Millipore, Billerica, USA); leucine enkephalin was purchased from Sigma-Aldrich (St. Louis, MO, USA).

2.4 Fingerprinting of urine with LC/MS

A LC system (Agilent 1290 Infinity LC System), equipped with a degasser, two binary pumps, and a thermostated autosampler coupled with Q-TOF/MS (6550 iFunnel) system (Agilent), was used in the ESI^+ mode to increase the number of detected metabolite ions**.** Exact MS data from redundant m/z peaks were first used to help confirm the metabolite molecular mass. The column used was an HSS T3 column (50 mm×2.1mm i.d., 1.8μm, Waters Corporation, Milford, USA) at 35 ◦C for all analyses. The mobile phase consisted of a linear gradient system of (A) 0.1% formic acid and (B) 0.1% formic acid in water in acetonitrile. The gradient conditions of the mobile phase were as follows: 0-1 min, 2% A; 1-3 min, 2-15% A; 3-7 min, 15-35% A; 9-10.0 min, 35-90% A; 10.0-11 min, 90-99% A; 11-12 min, 99 A. The flow rate was 0.5 mL/min. Injection volume was 5μL. The source temperature of mass spectrometry was set at 100 °C, desolvation gas temperature was 400 °C, cone gas flow was 60h, desolvation gas flow was 400L/h. In positive ion mode, the capillary voltage was 2.5.0kV, the sampling and cone voltage was 35V. All analyses were acquired using the lock mass spray to ensure accuracy and reproducibility.

2.5. Data Processing.

To provide quality assurance of results, LC-MS data treatment was performed and normalized to the total ion intensity per chromatogram. Briefly, LC-MS raw data were cleaned of background noises and unrelated ions by the Molecular Feature Extraction tool (MassHunter Qualitative Analysis Software; Agilent). AMDIS software from NIST

(U.S. National Institute of Standards and Technology) was used for mass spectral deconvolution to identify coeluted compounds according to their retention indexes and retention times. The Independent variable analysis (ICA) and PLS were performed by the R software. The combined VIP- plot was carried out to select variables as potential markers for distinguishing CC patients from controls. Accurate mass information of metabolites were matched with those of metabolites obtained from HMDB (www.hmdb.ca), and METLIN (metlin.scripps.edu/) databases. Pathway analysis using the KEGG pathway database was carried out using Metaboanalyst. Statistically significant differences in mean values were tested by using 2-tailed, 2-sample Student's t-test, and $p<0.05$ was considered statistically significant.

3. RESULTS and DISCUSSION

3.1. Metabolic fingerprinting

Metabolic fingerprinting allowed for the simultaneous detection of potential compounds in urine. ICA analysis was used as an unsupervised method to get an overview and to detect trends within these data. For LC-MS data, a clear separation can be observed between CC and control groups, indicating metabolic changes inherent to CC (Fig. 1). As shown in Fig 1 in positive ion mode, the ICA plot showed that the different groups tended to cluster and the CC group was separated from controls. ICA is used in order to get a first glance of the properties of the MS data. As a result of this analysis applied to the totality of the samples, a clustering of some samples becomes apparent by eye (see Fig 1), which suggests that biochemical perturbation significantly happened in CC groups. ICA can really reflect a clear separation between the CC group and controls. In order to enhance the differences between the groups and isolate the masses that discriminate them, supervised statistical methods were used. From the supervised PLS loading plots, the ions furthest away from the origin may be therefore regarded as the differentiating metabolites (Fig. 2).

3.2 Discovery of metabolic markers

The detailed method for the compound identification was described in the previous work [18]. In brief, the molecular ion was found from the mass spectrometry and subsequently, accurate molecular weight was calculated. Databases of ChemSpider database, Mass Bank, PubChem, and MetFrag were used to identify the metabolite markers by comparing molecular weights. The molecular and structural formulas of the candidate compounds were retrieved by the comparison and then confirmed by MS/MS scans for the characteristic ions and fragmentation patterns of the metabolites. For further analysis of feature ions, the S-plot combined VIP- plot from the PLS were to select variables as potential markers for distinguishing CC patients from controls (Fig.3). Ten differentially expressed metabolites from CC patients were distinguished from those of the controls ($p < 0.05$, VIP>6). It was found that, among them, 10 compounds were 5 upregulated and 5 were downregulated (Supplementary table 1). The postulated elemental compositions for the ions of ursodeoxycholic acid are given in Fig.4. The metabolite, which gave an elemental composition of $C_{24}H_{41}O_4$ and molecular mass of [M+H]393.5717, were identified when subjecting to MS/MS

analysis after searching in the database. The following 10 metabolites were identified: 3-methylhistidine, citric acid, cytosine, indoleacetic acid, salicyluric acid, L-Methionine, aminomalonic acid, glutaric acid, ursodeoxycholic acid and N-Acetylornithine. Based on these findings, individual metabolites were compared, yielding statistical differences between control and CC women in metabolites in urine. Bar plots of Fig 4 show relative signal intensities for 10 metabolites in control and CC groups. Fig 5 shows metabolites that exhibited the highest significant differences between case and control groups.

3.3 Metabolic pathway and function analysis

The detailed analysis of the significantly changed pathways of CC was performed by MetaboAnalyst's tool. A total of 10 pathways which were identified together are important for the host response to CC (Fig 6 and supplementary Table S2). The predominant hits were citrate cycle, lysine degradation, tryptophan metabolism, cysteine and methionine metabolism, etc. These altered metabolism pathways with higher score had yield satisfactory results and clearly help us to better understand the underlying mechanisms of CC. Among 10 molecules identified, citric acid of interest in CC belongs to glyoxylate and dicarboxylate metabolism and TCA cycle, L-methionine, belong to cysteine and methionine metabolism.

The ability to detect metabolic changes in the body is essential for understanding the biological roles of low-molecular-weight compounds in body fluids [18-20]. In this study, high-throughput LC/MS and multiple data processing method and pathway analysis could provide a powerful approach to clearly differentiate patients with CC and identify the potential biomarkers. ICA revealed a significant separation between the CC and controls. Interestingly, 10 metabolites were identified in CC, and suggest a disrupted the citrate cycle, lysine degradation, tryptophan metabolism, cysteine and methionine metabolism, etc. Of note, 10 metabolism pathways were found that the altered pathway associated with CC. Furthermore, network analysis found that the abnormally expressed metabolites were tightly correlated with the metabolism pathway. These biochemical changes are helpful to understand the key features of CC. The relative intensities of the most significant metabolites in the analysis of urine were listed in Fig 5. The metabolites from citrate cycle, lysine degradation, tryptophan metabolism, cysteine and methionine metabolism, *etc.* were significantly altered. This study demonstrates that metabolomics could capture the subtle changes from CC, which may lead to an improved mechanistic understanding of CC. The metabolomic results not only supply a systematic view of the development and progression of CC but also provide a theoretical basis for the diagnosis, prevention or treatment of CC. These findings provided a non-invasive tool that can engender new insights into the improved mechanistic understanding of CC. Overall, this investigation demonstrates that metabolomics could capture the subtle metabolic changes resulting from CC.

4. CONCLUSIONS

Clinical metabolomics is the study of metabolic changes in the body and provides the small molecule metabolites

related to the disease. To our knowledge, this study represents the first non-targeted metabolome-wide analyses in urine of CC. We show that metabolite fingerprints in urine reveal metabolic imbalances that are specific for human CC. In this study, a LC/MS metabolomics has been successfully established for biomarker and mechanism exploration studies in CC. The selected variables were annotated and placed on pathway by using KEGG. Ten altered metabolic biomarkers indicated that CC could cause more severe disturbances in citrate cycle, lysine degradation, tryptophan metabolism, cysteine and methionine metabolism, etc. It showed that metabolomics has the potential as a promising screening tool for exploring essence of CC disease, and could provide the biological background to pharmacological interventions in the future. As this is a pilot study, future projects at later stages of CC will allow us to validate the identified biomarkers as prognostic tools to predict the onset of CC. Overall, it demonstrated that clinical metabolomics may highlight biomarkers and pathways and can capture subtle metabolic changes from CC, which may lead to an improved diagnosis and mechanistic understanding of CC.

ACKNOWLEDGMENTS

This work was supported by grants from Key Program of Natural Science Foundation of State (Grant No. 81302905). Foundation of Heilongjiang University of Chinese Medicine (Grant no. 201209).

CONFLICT OF INTEREST:

The authors declare no competing financial interests.

REFERENCES

1. McLaughlin-Drubin ME, Park D, Munger K. Tumor suppressor p16INK4A is necessary for survival of cervical carcinoma cell lines.Proc Natl Acad Sci U S A. 2013;110(40):16175-80.

2. Herfs M, Yamamoto Y, Laury A, Wang X, Nucci MR, McLaughlin-Drubin ME, Münger K, Feldman S, McKeon FD, Xian W, Crum CP. A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer.Proc Natl Acad Sci U S A. 2012;109(26):10516-21.

3. Chen YF, Chiu WT, Chen YT, Lin PY, Huang HJ, Chou CY, Chang HC, Tang MJ, Shen MR. Calcium store sensor stromal-interaction molecule 1-dependent signaling plays an important role in cervical cancer growth, migration, and angiogenesis.Proc Natl Acad Sci U S A. 2011;108(37):15225-30.

4. Chung SH, Lambert PF. Prevention and treatment of cervical cancer in mice using estrogen receptor antagonists.Proc Natl Acad Sci U S A. 2009;106(46):19467-72.

5. Kendrick SF, Henderson E, Palmer J, Jones DE, Day CP. Theophylline improves steroid sensitivity in acute alcoholic hepatitis. Hepatology. 2010;52(1):126-31.

6. Wang X, Zhang A, Sun H. Power of metabolomics in diagnosis and biomarker discovery of hepatocellular carcinoma. Hepatology. 2013;57(5):2072-7.

7. Zhang AH, Sun H, Han Y, Yan GL, Yuan Y, Song GC, Yuan XX, Xie N, Wang XJ. Ultraperformance liquid chromatography-mass spectrometry based comprehensive metabolomics combined with pattern recognition and network analysis methods for characterization of metabolites and metabolic pathways from biological data sets. Anal Chem. 2013;85(15):7606-12.

8. Kim JW, Ryu SH, Kim S, Lee HW, Lim MS, Seong SJ, Kim S, Yoon YR, Kim KB. Pattern Recognition Analysis for Hepatotoxicity Induced by Acetaminophen Using Plasma and Urinary (1)H NMR-Based Metabolomics in Humans.Anal Chem. 2013;85(23):11326-34.

9. Knorr A, Monge A, Stueber M, Stratmann A, Arndt D, Martin E, Pospisil P. Computer-Assisted Structure Identification (CASI)-An Automated Platform for High-Throughput Identification of Small Molecules by Two-Dimensional Gas Chromatography Coupled to Mass Spectrometry. Anal Chem. 2013;85(23):11216-24.

10. Wang X, Zhang A, Yan G, Sun W, Han Y, Sun H. Metabolomics and proteomics annotate therapeutic properties of geniposide: targeting and regulating multiple perturbed pathways. PLoS One. 2013;8(8):e71403.

11. Naz S, García A, Barbas C. Multiplatform analytical methodology for metabolic fingerprinting of lung tissue. Anal Chem. 2013;85(22):10941-8.

12. Sun H, Zhang A, Yan G, Piao C, Li W, Sun C, Wu X, Li X, Chen Y, Wang X. Metabolomic analysis of key regulatory metabolites in hepatitis C virus-infected tree shrews. Mol Cell Proteomics. 2013;12(3):710-9.

13. Manwaring V, Boutin M, Auray-Blais C. A metabolomic study to identify new globotriaosylceramide-related biomarkers in the plasma of Fabry disease patients. Anal Chem. 2013;85(19):9039-48.

14. Wang X, Zhang A, Han Y, Wang P, Sun H, Song G, Dong T, Yuan Y, Yuan X, Zhang M, Xie N, Zhang H, Dong H, Dong W. Urine metabolomics analysis for biomarker discovery and detection of jaundice syndrome in patients with liver disease. Mol Cell Proteomics. 2012;11(8):370-80.

15. Zhang A, Sun H, Han Y, Yan G, Wang X. Urinary metabolic biomarker and pathway study of hepatitis B virus infected patients based on UPLC-MS system. PLoS One. 2013;8(5):e64381

16. Zhang AH, Wang P, Sun H, Yan GL, Han Y, Wang XJ. High-throughput ultra-performance liquid chromatography-mass spectrometry characterization of metabolites guided by a bioinformatics program. Mol Biosyst. 2013;9(9):2259-65.

17. Zhang A, Sun H, Han Y, Yan G, Wang X. Urinary metabolic biomarker and pathway study of hepatitis B virus infected patients based on UPLC-MS system. PLoS One. 2013;8(5):e64381.

18. Chen S, Kong H, Lu X, Li Y, Yin P, Zeng Z, Xu G. Pseudotargeted metabolomics method and its application in serum biomarker discovery for hepatocellular carcinoma based on ultra high-performance liquid chromatography/triple quadrupole mass spectrometry.Anal Chem. 2013;85(17):8326-33.

19. Kasukawa T, Sugimoto M, Hida A, Minami Y, Mori M, Honma S, Honma K, Mishima K, Soga T, Ueda HR. Human blood metabolite timetable indicates internal body time. Proc Natl Acad Sci U S A. 2012;109(37):15036-41.

20. Arakaki AK, Skolnick J, McDonald JF. Marker metabolites can be therapeutic targets as well. Nature. 2008;456(7221):443.

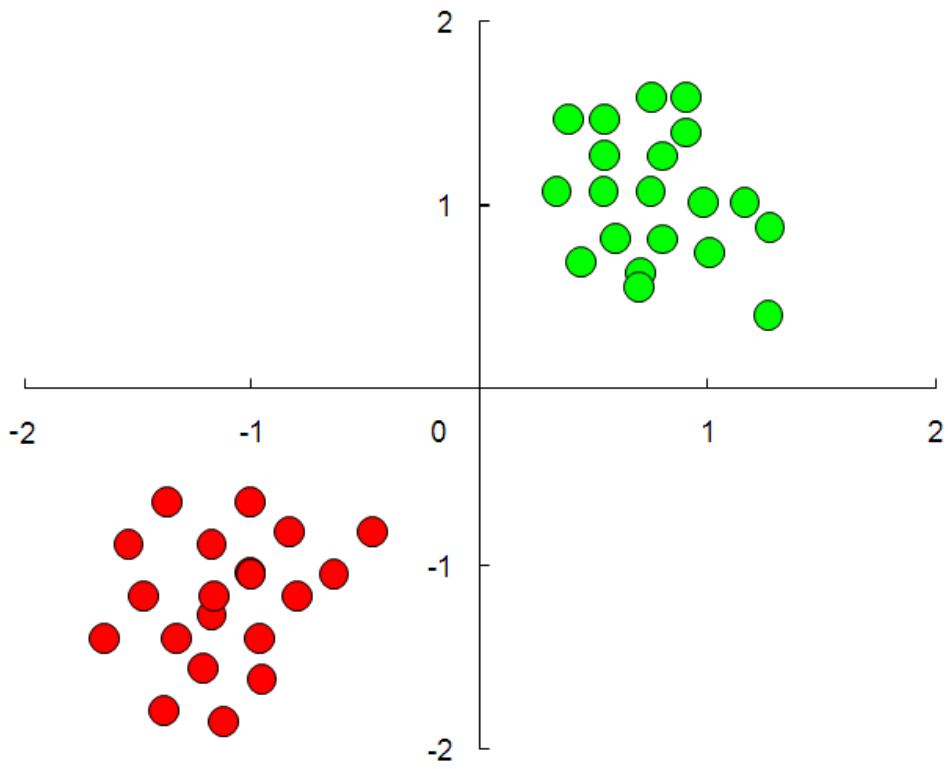


Fig. 1. Score plots of urine metabolic profiles obtained for control (●) and cervical cancer women (●).

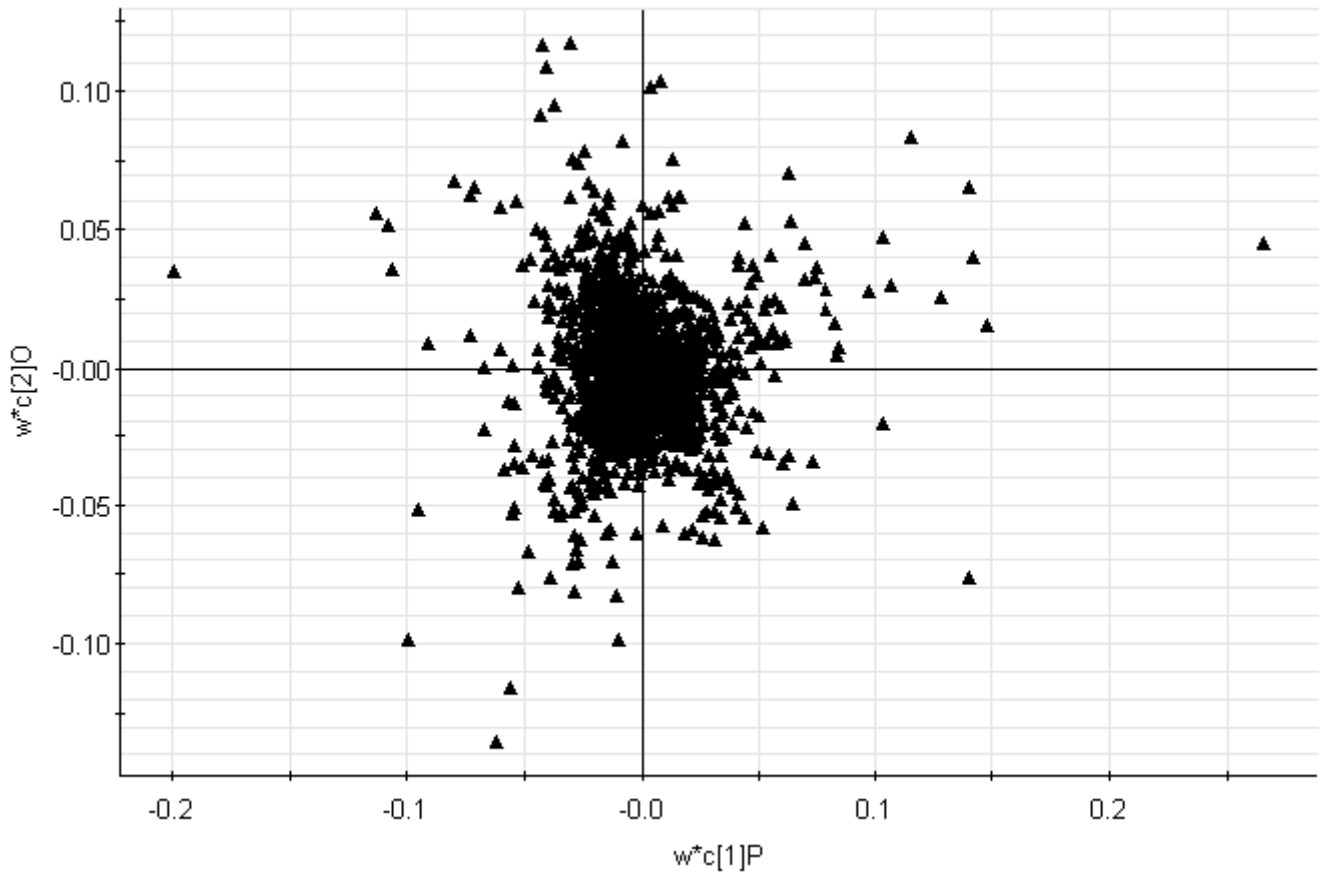


Fig. 2 Loading plot of supervised PLS analysis built for the two groups (CC women versus healthy controls).

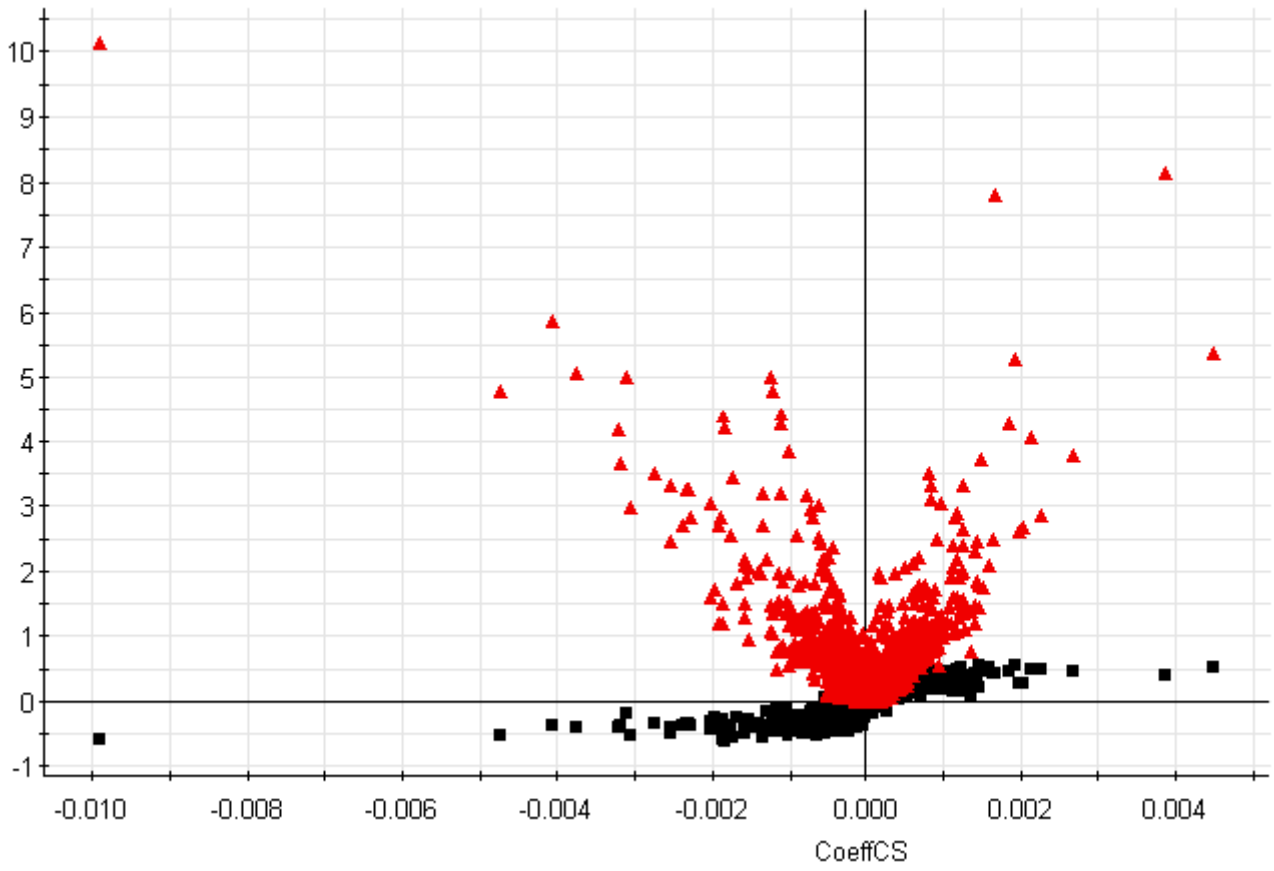


Fig. 3 VIP-score plots constructed from the supervised PLS analysis of urine metabolic profiles obtained for control and cervical cancer women. Ions with the highest abundance and correlation in the CC group with respect to the controls are present on the upper far right hand quadrant, whereas ions with the lowest abundance and correlation in the CC group with respect to the control group are residing in the lower far left hand quadrant.

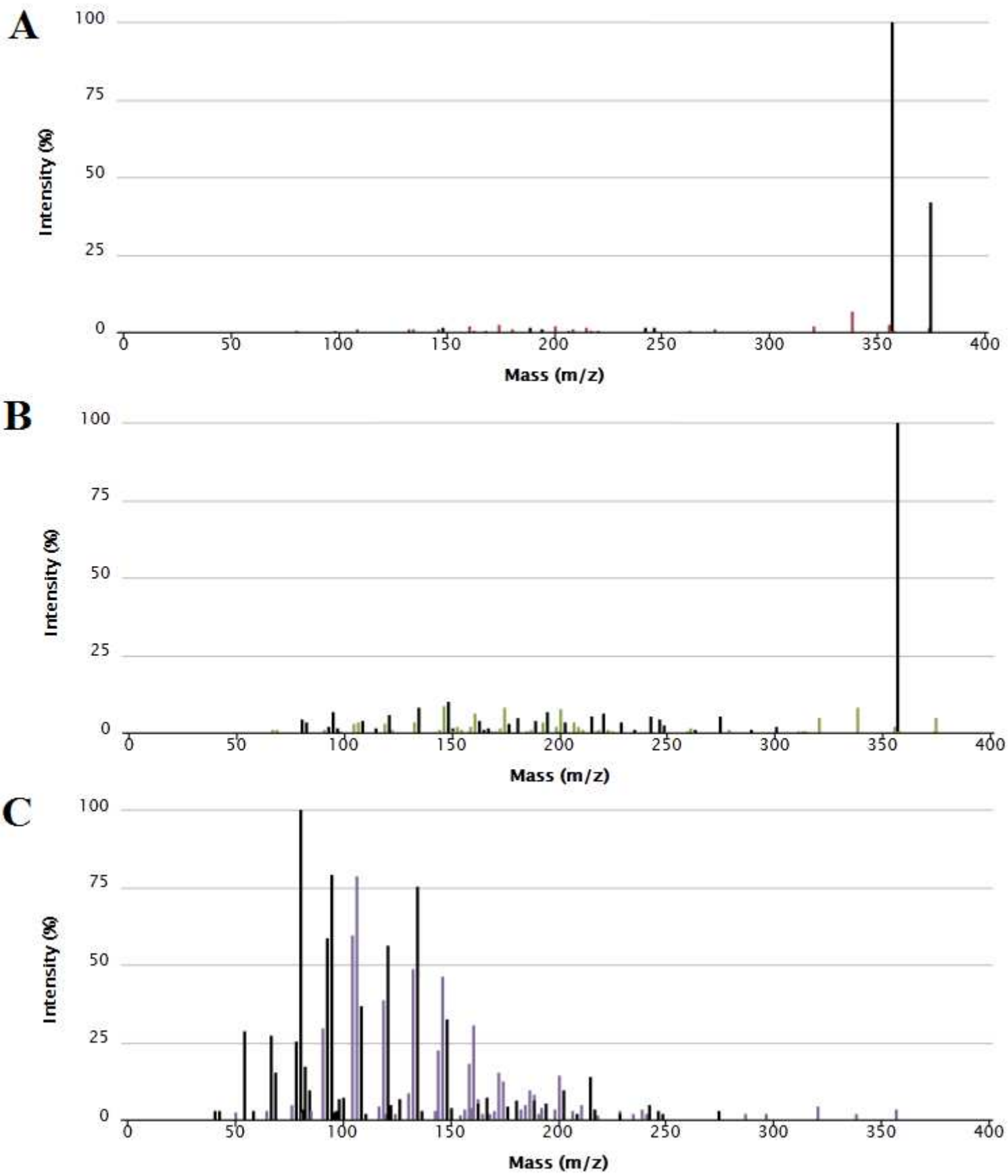


Fig.4 Mass spectrum of full scan and product ion scan of ursodeoxycholic acid. The basic information: Rt=5.44 and $[M+H]^+$=393, A: 10V, B:20 V, C:40 V.

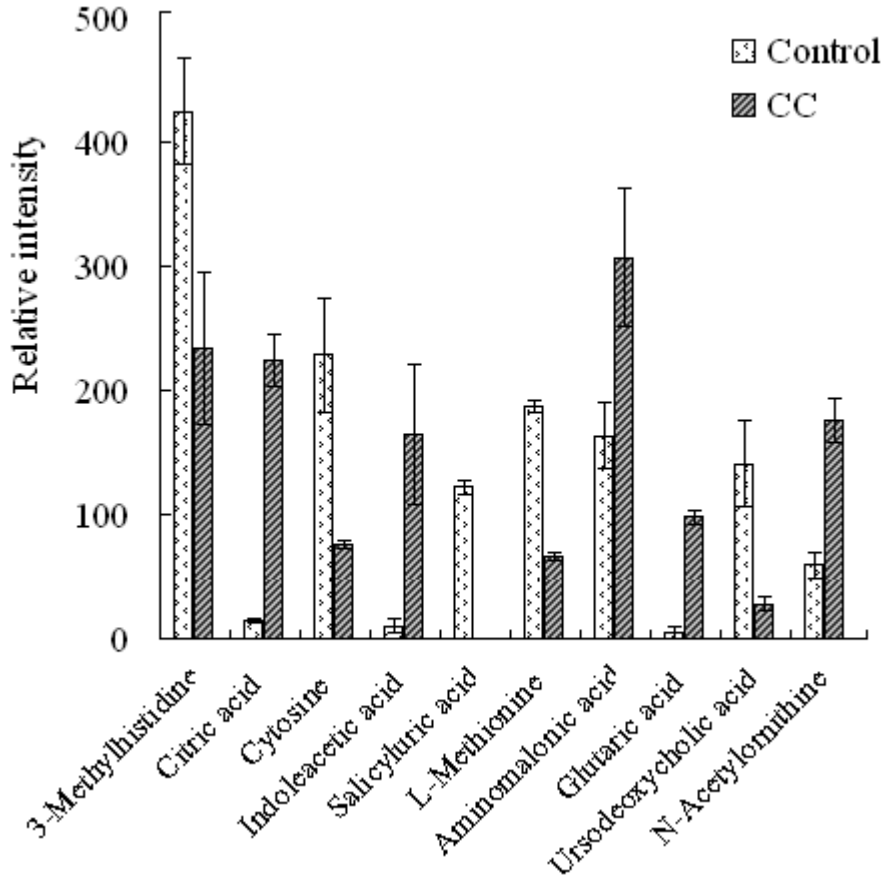


Fig. 5 Histogram plots show the relative signal intensities for metabolites in control and CC groups.

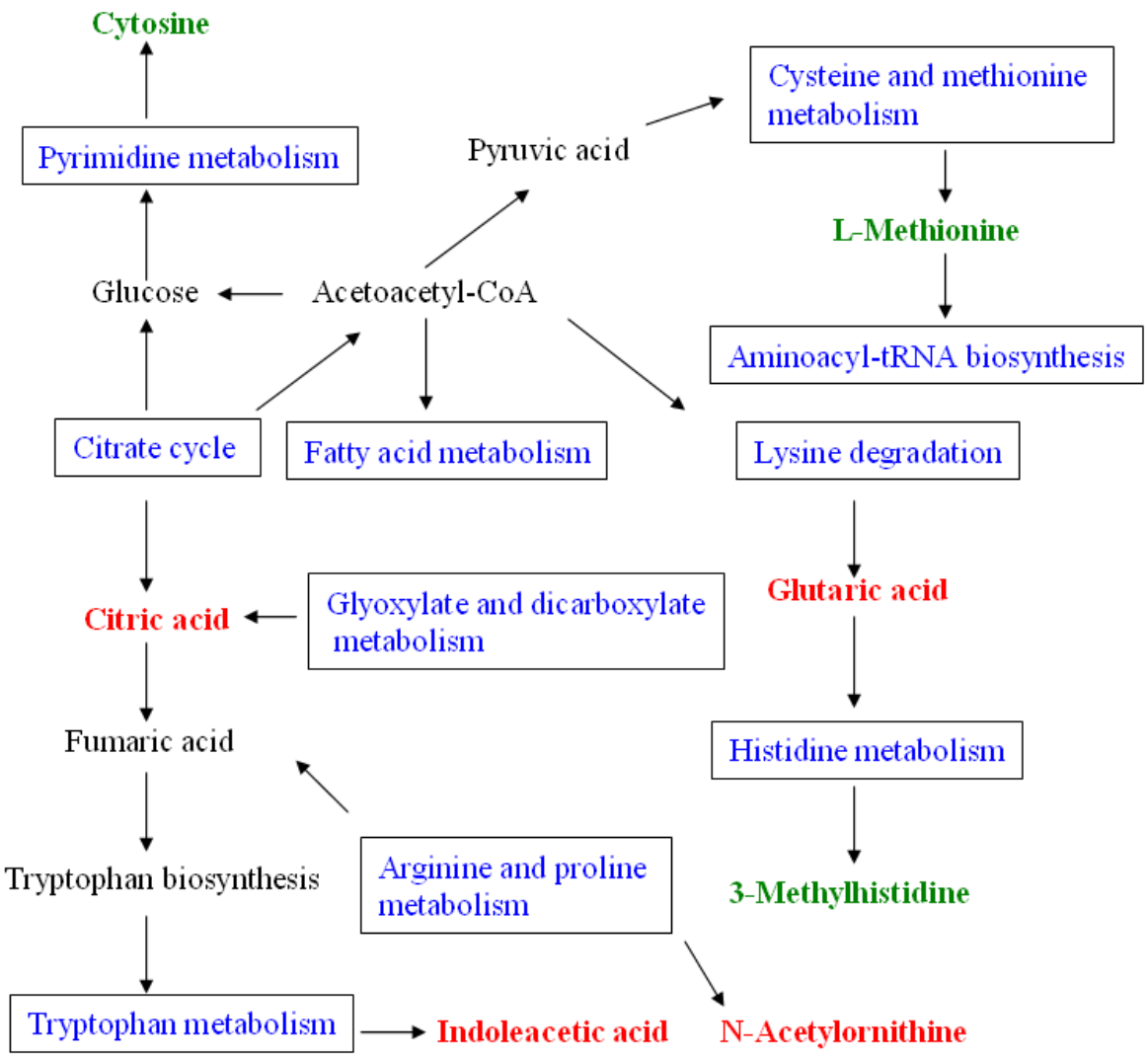


Fig. 6 Schematic overview of the metabolites and relevant pathways changing for CC modulation according to the KEGG PATHWAY database. Metabolites in red and green represent elevation (up regulated) and inhibition (down regulated), respectively. Metabolites in black means they were not detected in our experiment. The blue words are pathway's names.