Near-infrared spectroscopy of plasma amino acids with chemometrics towards breast cancer discrimination

Liying Zhu
Affiliated Hospital of Guizhou Medical University
Wen Qian
Affiliated Hospital of Guizhou Medical University
Haizhi Li
Affiliated Hospital of Guizhou Medical University
Chengcheng Li
Affiliated Hospital of Guizhou Medical University
Changyudong Huang
Affiliated Hospital of Guizhou Medical University
Yiqiong Zhang
Affiliated Hospital of Guizhou Medical University
Yunfeng Duan
Affiliated Hospital of Guizhou Medical University
Shuang Wang
Affiliated Hospital of Guizhou Medical University
YongJie Xu
Affiliated Hospital of Guizhou Medical University
Jingzhi Zhang
Affiliated Hospital of Guizhou Medical University
Mi Liu
Affiliated Hospital of Guizhou Medical University
Xing Li
Guizhou university of traditional Chinese medicine
Wei Pan (pw_6602@sina.com)
Affiliated Hospital of Guizhou Medical University

Article

Keywords: breast cancer, NIR spectroscopy, amino acids, chemometrics

Posted Date: April 25th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1560737/v1

Abstract

Breast cancer(BC) is the most prevalent cancer and the second-leading cause of cancer-related death for women in the worldwide. BC cells need more amino acids to meet the demand of rapid proliferation. In this paper, we focus on the role of near-infrared (NIR) spectroscopy to analyse blood plasma samples of breast cancer patients to differentiate healthy controls. The possibility of quantitative detection of 20 amino acids in breast cancer plasma by NIR spectroscopy was investigated for the first time in this study. 180 samples (80 BC patients, 30 benign breast disease patients and 70 healthy controls) were analysed. Canonical correlation analysis (CCA) was used to analyse the relationship between clinical biochemical parameters and amino acid metabolic profile for BC patients. In this study, plasma glutamine, histidine, threonine, proline, phenylalanine content of BC patients was higher than healthy control plasma ($p < 0.05$) by NIR spectroscopy. There was overall correlation (r = 0.935) between three clinical parameters(age, albumin, total triglyceride) and four amino acids (glutamic acid, tyrosine, valine and lysine) from BC patients. In this study we have built quantitative and qualitative model which is used to detect plasma amino acids of BC patients by NIR spectroscopy has good performance. It fully demonstrated the great potential and advantages of NIR spectroscopy combined with chemometrics in breast cancer research, which can provide a new research strategy and technical platform for the study of plasma amino acid metabolism in breast cancer patients, and provide basic experimental data reference for clinical diagnosis and treatment of breast cancer patients.

Intruduction

Breast cancer (BC) is the most common malignant tumor in women, accounting for 15% of all new cancer cases in China, and the mortality rate is increasing year by year, seriously endangering women's health[1, 2]. Currently, cancer antigen 15 − 3 (CA15-3) is mainly used for breast cancer screening, but it lacks specificity for early screening of breast cancer and is often highly expressed in patients with intermediate and advanced breast cancer[3]. Therefore, it is of great significance to find the relative specific markers of breast cancer screening.

Tumor cells have the characteristics of rapid proliferation, and need biomacromolecules to provide energy and complete biosynthesis to achieve continuous proliferation and escape biological characteristics such as immune monitoring of tumor cell invasion and drug resistance of tumor cells are related to matter energy metabolism. Pathological changes in cell function are often accompanied by metabolic recombination, including changes in amino acid metabolism[4]. Amino acid detection is very important in the screening of many inherited metabolic diseases. Cancer patients are in hypermetabolic state and hypercatabolism state, and protein synthesis and decomposition are improved compared with healthy people, resulting in changes in amino acid concentration and abnormal amino acid metabolism[5]. Studies have shown that amino acids are a very promising biomarker for screening tumors, such as Pancreatic cancer[6], Colorectal cancer[7], gastric cancer[8], Ovarian cancer[9], lung cancer[10]. However, the application of breast cancer screening in China is still lack of systematic research.

Nearinfrared (NIR) spectroscopy is a fast, simple, green and pollution-free spectral analysis method, which has been widely used in the rapid detection of liquid sample quality.NIR spectroscopy has been used in the fields of plasma components[11−13]. However, no reports were found in relation to the use of NIR spectroscopy for the

quantitative detection free amino acids in breast cancer plasma. To investigate the difference of serum amino acid levels between breast cancer patients and normal controls and its value in breast cancer screening, NIR spectroscopy was used to detect the levels of 20 kinds of amino acids in serum of breast cancer patients.The differences of amino acid levels were analyzed, and the efficacy of different serum amino acids in breast cancer screening was systematically evaluated, so as to find new biomarkers for breast cancer screening. Therefore, the aim of the present study was to investigate the possibility of NIR spectroscopy to quantitatively analyze twenty amino acids in breast cancer plasma without sophisticated methods.

Results

NIR spectroscopy analysis. NIR spectroscopy is a valuable tool capable of analysing different types of diseases by measuring biologically-derived samples. Herein, NIR spectroscopy was employed to detect blood plasma samples spectra of breast cancer patients. Figure 1 shows the spectra were cut in the region between 4000 to 10000nm, responsible for biomolecular-derived spectrochemical signatures. The basic statistics of the amino acid compositions for the calibration and validation sets are summarized in Table 1. The value for the correlation coefficient in calibration (r cal) was >0.97 for all of the amino acids. The root-mean-square error of cross-validation(RMSECV) was <0.05, except Gln. The value for the correlation coefficient in validation (r val) was > 0.82. The root-mean-square error of prediction (RMSEP) was >0.05, except Gln. The result indicated that the prediction of amino acids was also dependent on the specific of amino acids. The poor prediction of Gln might relate to the less accurate reference values due to the low response of Gln. Although the PLS models need to be improved and validated, the result obtained in this study suggested that NIR spectroscopy could be used as a tool to determine amino acids in breast cancer plasma rapidly and economically.

Table 1
Statistics of amino acids in breast cancer plasma (n = 80) in the calibration and validation set and the repeatability (Sr) of reference methods

analyte	spectrum range(cm^{-1})	Calibration set			validation set			RMSECV	no. of PLS factors
		n	r_{cal}	RMSEC	n	r_{val}	RMSEP		
Asp	7076.62–10001.03	80	0.9891	0.0004	30	0.9417	0.0010	0.0014	10
Glu	7400.85–9524.22	80	0.9999	0.0001	30	0.9608	0.0018	0.0037	6
Ser	5747.16–6504.46	80	1.0000	0.0000	30	0.9724	0.0068	0.0199	8
Asn	7432.64–7871.30	80	1.0000	0.0000	30	0.9684	0.0003	0.0008	8
Gly	5663.62–6896.78	80	0.9790	0.0144	30	0.9941	0.0044	0.0177	4
Gln	5687.15–6407.21	80	0.9742	0.3090	30	0.9396	0.4180	1.6192	7
His	7565.53–9177.72	80	0.9998	0.0003	30	0.9964	0.0015	0.0123	6
Thr	7591.57–9773.25	80	0.9990	0.0050	30	0.9986	0.0061	0.0846	8
Ala	5510.25 ~ 6461.71	80	0.9999	0.0001	30	0.9988	0.0004	0.0085	9
Arg	7189.00 ~ 7299.00	80	0.9999	0.0001	30	0.9961	0.0008	0.0043	9
Pro	7388.14 ~ 9880.24	80	1.0000	0.0005	30	0.8896	0.0063	0.0082	6
Tyr	7737.12 ~ 9908.87	80	0.9991	0.0001	30	0.9228	0.0015	0.0030	4
Val	5789.87–6275.63	80	0.9339	0.0032	30	0.9109	0.0037	0.0082	6
Met	7370.36 ~ 9657.73	80	0.9999	0.0001	30	0.9291	0.0042	0.0060	6
Cys	5395.56 ~ 6606.45	80	0.9733	0.0022	30	0.9919	0.0015	0.0066	4
Ile	5818.42 ~ 6199.65	80	0.9999	0.0001	30	0.9451	0.0024	0.0055	6
Leu	5582.63 ~ 5925.93	80	0.9846	0.0091	30	0.9917	0.0067	0.0188	8
Phe	7464.42 ~ 9829.38	80	0.9999	0.0013	30	0.9759	0.0235	0.0758	6
Trp	7901.91 ~ 9718.56	80	0.9951	0.0026	30	0.8237	0.0181	0.0206	4
Lys	7633.21 ~ 9552.23	80	0.9998	0.0002	30	0.9857	0.0016	0.0067	5

Average spectrum and spectral pretreatment. The average raw spectrum for each group of sample is depicted in Fig. 2. To reduce noise, the raw spectral data were pre-processed by the second derivative + Norris derivative filter (5, 5). There is a high degree of superposition between spectral features among categories; consequently, multivariate analysis tools are necessary to distinguish the categories (Fig. 3).

Discrimination results of plasma samples in three groups. In this study, the prediction accuracy of calibration set and validation set were used as indicators to select the best spectral region. The results showed that when 4119.20 ~ 9881.46cm^{-1} was selected as the spectral range, the accuracy of the results was the best, indicating that the three types of samples could be obviously distributed in three different regions (Fig. 4). According to the selected principal components, the 3D distribution diagram of principal component scores was obtained(Fig. 5) It was observed that the samples were separated clearly according to different types. It was possible that the aging of wines caused a variation in chemical components, which resulted in different spectral attributes.

Statistical analysis of clinical data. 80 blood plasma samples from breast cancer patients were analysed, with 30 samples originating from benign breast disease patients and 70 from healthy controls. Clinical parameter data of the participants can be observed in Table 2. Blood-glucose, urea nitrogen and creatinine were significant difference between breast cancer and the healthy controls. Furthermore, blood-glucose was significant difference between benign breast disease and the healthy controls.

Table 2
The clinical parameters of three different people

clinical parameters	Breast cancer	Benign breast disease	Healthy women
	(n = 80)	(n = 30)	(n = 30)
Age(year)	42.38 ± 8.45	41.00 ± 3.94	39.72 ± 6.34
BMI (kg/m^2)	22.23 ± 2.41	23.31 ± 1.52	22.88 ± 2.73
Total protein(g/L)	76.96 ± 5.73	74.60 ± 4.32	73.59 ± 5.85
Albumin (g/L)	44.85 ± 4.62	43.38 ± 3.78	42.89 ± 3.40
Glucose (mmol/L)	5.29 ± 0.41	5.12 ± 0.68	4.46 ± 0.46□△
Urea nitrogen (mmol/L)	4.93 ± 0.87	3.26 ± 0.63□	3.29 ± 0.67□
Creatinine (μmol/L)	51.4 ± 3.06	41.60 ± 5.20□	41.69 ± 5.00□
Total triglycerides (mmol/L)	4.16 ± 0.57	4.09 ± 0.68	4.06 ± 0.58
Total cholesterol (mmol/L)	2.66 ± 0.32	2.58 ± 0.12	2.67 ± 0.49

Determination of plasma amino acids in three groups. Compared with healthy women, the plasma levels of Gln, His, Thr, Pro, Phe in benign and breast cancer patients were statistically significant in Table 3. It was illustrated that the amino acids could be used as potential biomarkers and also could provide reference for clinical diagnosis and treatment.

Table 3
The 20 amino acids content in three different category women' plasma

analyte	Breast cancer (n = 80)	Benign breast disease (n = 30)	Healthy women (n = 30)
Asp	0.0102 ± 0.003	0.0107 ± 0.003	0.0101 ± 0.018
Glu	0.0205 ± 0.007	0.0200 ± 0.007	0.0210 ± 0.015
Ser	0.0311 ± 0.0310	0.0309 ± 0.008	0.0302 ± 0.015
Asn	0.0156 ± 0.008	0.0122 ± 0.004	0.0115 ± 0.003
Gly	0.0431 ± 0.036	0.0416 ± 0.013	0.0439 ± 0.016
Gln	2.2893 ± 1.775	1.2380 ± 0.421▯	1.1515 ± 0.253▯
His	0.0571 ± 0.024	0.0432 ± 0.007▯	0.0323 ± 0.023▯△
Thr	0.3618 ± 0.159	0.1104 ± 0.021▯	0.0974 ± 0.019▯
Ala	0.0364 ± 0.012	0.0465 ± 0.030	0.0551 ± 0.014▯
Arg	0.1230 ± 0.104	0.1156 ± 0.068	0.1512 ± 0.1073
Pro	0.0330 ± 0.016	0.1380 ± 0.017▯	0.1290 ± 0.015▯
Tyr	0.0147 ± 0.005	0.0187 ± 0.006	0.0205 ± 0.006
Val	0.0380 ± 0.012	0.0414 ± 0.011	0.0442 ± 0.013
Met	0.0072 ± 0.010	0.0068 ± 0.002	0.0081 ± 0.002
Cys	0.0383 ± 0.013	0.0396 ± 0.009	0.0409 ± 0.011
Ile	0.0258 ± 0.009	0.0272 ± 0.007	0.0298 ± 0.037
Leu	0.1348 ± 0.057	0.1151 ± 0.014	0.1127 ± 0.012
Phe	0.2097 ± 0.122	0.0302 ± 0.009▯	0.0331 ± 0.008▯
Trp	0.0898 ± 0.037	0.0939 ± 0.032	0.0763 ± 0.021
Lys	0.0301 ± 0.011	0.0369 ± 0.012	0.0328 ± 0.010

CCA on the association between plasma amino acids and clinical parameters among breast cancer. CCA, first introduced by Hotelling, is a useful dimension reduction technique for exploring the relationship between two sets of variables. Figure 6A shows that the correlation coefficient of the first pair of typical correlation variables between clinical biochemical parameters (U1) and amino acid metabolic profile (V1) for breast cancer is 0.935, indicating that this statistical method has good application value. Figure 6B shows that the absolute values of the coefficients of age, albumin and total triglyceride in the clinical parameters of breast cancer patients are significantly higher than the others, indicating that these substances are closely related to breast cancer and should be paid more attention in clinical practice. Figure 6C shows that the absolute coefficients of Glu, Tyr,

Val and Lys were higher than others, indicating that these substances had a closer relationship with the occurrence of breast cancer.

Discussion

Amino acids, as substrates for protein synthesis, are the most important source of energy and nutrition in cells next only to glucose amino acids can also be used in energy production nucleoside synthesis and cell REDOX balance maintenance. Because cancer cells exist in a nutrient-poor microenvironment, they are particularly evident in cancers that are deficient in certain amino acids

There are few studies that relate amino acids of breast cancer and the search for new tools that will predict the possible diagnosis of breast cancer. To our knowledge, our group was the first to use the NIR spectroscopy to analyze the blood plasma amino acids collected from breast cancer patients.

In this study, 80 blood plasma samples taken from breast cancer,30 from benign breast disease patients and 70 from healthy women, were analysed using NIR spectroscopy to detect amino acids. The results of this study showed that compared with healthy women, the plasma levels of Gln, His, Thr, Pro, Phe in benign and breast cancer patients were statistically significant.

Glutamine is the most abundant amino acid in human plasma, which provides nitrogen source for the synthesis of purine pyrimidine nucleotides[14]. Glutamine is an essential raw material for the synthesis of glutathione, and is involved in maintaining the stability of reactive oxygen species and producing glutamine through the synthesis of glutamine synthase. However, the tumor cells can not meet the needs of rapid proliferation by their own synthesis of glutamine, and there is the phenomenon of glutamine dependence[15]. The incoming glutamine deamines in the mitochondria to form glutamate, which forms a-ketoglutarate and NADH/NADPH through glutamate dehydrogenase, and enters the TCA cycle and regulates intracellular redox homeostatic[16–19]. Overexpression of glutaminase (GLS1) in cancer cells and a subsequent increase of amino acids production have been reported in ER-negative breast cancer, with apparent association to poor prognosis[20]. Some subsets of breast cancer, especially MYC-overexpressing tumors most of which are TNBC, show Gln addiction, which could be a potential therapeutic target[21].

His can improve the function of human immune system, strengthen physiological metabolism, regulate the effective utilization rate of protein in the body, and promote the growth and development of the body [22–24].

Thr plays an important role in protecting cell membranes, promoting phospholipid synthesis and fatty acid oxidation in vivo, and promoting human development [25–27].

Pro can help the human body decompose protein and promote cell metabolism in the body, which is very important for maintaining the healthy growth of skin and tissues [28, 29].

Phe, an aromatic amino acid, is oxidized to tyrosine by phenylalanine hydroxylase in the body, and together with tyrosine, important neurotransmitters and hormones are synthesized to participate in sugar metabolism and fat metabolism in the body[30, 31] .

According to the variables selected by the CCA, we found that age,albumin,total triglycerides was closely related with breast cancer. Futhermore,we also found that plasma Glu, Tyr, Val, Lys of breast cancer was related with occurrence and development of breast cancer. It can provide basic experimental data reference for clinical analysis.

Conclusions

In this study, we have built quantitative and qualitative model which is used to detect plasma amino acids of BC patients by NIR spectroscopy has good performance. It fully demonstrated the great potential and advantages of NIR spectroscopy combined with chemometrics in breast cancer research, which can provide a new research strategy and technical platform for the study of plasma amino acid metabolism in breast cancer patients, and provide basic experimental data reference for clinical diagnosis and treatment of breast cancer patients. The number of cases in this study was relatively smaller, and it was only an exploratory study. Whether plasma amino acids can be used in the clinical practice of breast cancer screening needs to be verified by subsequent large samples.

Materials And Methods

Samples. 190 plasma samples were collected from breast cancer patients during 2016–2018. These samples were used for establishment of quantitative model of 20 plasma amino acids. The breast cancer patients were recruited from the breast surgeons of affiliated hospital of Guizhou medical university. 80 samples were assigned to the calibration set, 30 samples constituted to the validation set, while the remaining 80 samples were assigned to test. At the same time, 30 plasma samples were collected from benign breast disease and 70 plasma samples were collected from healthy females. All samples were stored at -80°C until analysis. All patients provided written informed consent, and ethical consent was granted from the Committees for Ethical Review of Research involving the Affiliated Hospital of Guizhou Medical University (Guizhou, China). All procedures were performed in compliance with the Declaration of Helsinki.

Reagents. All reagents used in this study were of analytical reagent grade. 20 amino acids were purchased from Sigma-Aldrich (America), and the concentration for each amino acid was 1000 μg/mL. The 20 amino acids used as standards were aspartic acid (Asp), glutamic acid (Glu), serine (Ser), Asparagine Asn (Asn), glycine(Gly), Glutamine Gln(Gln), Histidine(His), threonine (Thr), alanine (Ala), arginine (Arg), proline (Pro), tyrosine (Tyr), valine(Val), methionine (Met), cysteine(Cys), isoleucine (Ile), leucine (Leu), phenylalanine (Phe), tryptophan (Trp), lysine (Lys).

NIRS analysis. Near-infrared(NIR)spectroscopy together with chemometrics has gained wide acceptance in the fields of food industry and agriculture in recent years, mainly because it is a low-cost, nondestructive method and generally requires minimal sample processing prior to analysis[32].NIR spectroscopy can record the response of chemical bonds in functional groups to the NIR spectrum, which is related to the primary structural components of organic molecules. Quan-titative NIR spectroscopy measurement is based on the correlation between sample composition, as determined by reference methods, and the absorption of NIR radiation by bonds between light atoms at different wavelengths in the NIR region[33].However, the absorption peaks of NIR spectra are broad and overlap, and it is impossible to make direct quantification analysis due to the high

dimension and complexity of NIR spectral data. Chemometrics methods, such as principal component analysis(PCA), principal component regression (PCR), and partial least-squares regression (PLSR), are often used to extract spectral features and investigate the correlation between the spectra and component concentrations.Before NIR spectra acquisition, all of the samples were stored in the laboratory at -80℃.All plasma samples were scanned using Antaris II Fourier transform near infrared spectrometer (Thermo Nicolet, USA). Spectra were collected using OMNIC software (Thermo Electron Corp.) and saved in absorbance format. Each sample was fitted in a 1 mm diameter cup that rotated during NIRS scanning. Spectra were the sum of 64 co-added scans across the spectral range of 4000-11,000 cm^{-1} with a spectral resolution of 8 cm^{-1}. Water background was taken each hour.

Calibration models were developed using PLSR, which is the most commonly used multivariate method for the evaluation of NIR spectra.

PLSR analysis was conducted using the spectra from the calibration dataset to develop an empirical equation for predicting the concentrations of total amino acids.

Here, Partial least squares (PLS) models with 1–15 factors were investigated, and the optimum number of factors used in PLSR was determined by the lowest value of the predicted residual error sum of squares (PRESS) to avoid overfitting. According to Williams, a calibration model should contain at least 100 samples, and so, due to our rather small sample set, leave-one-out cross-validation was employed to evaluate the established models[34]. Cross-validation has the advantage that all of the data available can be used to determine the calibration model,because no sample has to be held back in a separate validation set [35] Several studies, including that of Moron and Cozzolino [36], have shown that both procedures provided similar results. The performance of the calibration is assessed by the correlation coefficient in calibration (r_{cal}),root-mean-square error of calibration (RMSEC) and root-mean-square error of cross-validation(RMSECV) [37].In addition, to evaluate the prediction ability of the calibration model, the performance of the validation is assessed by the correlation coefficient (r_{val}) and root-mean-square error of prediction (RMSEP).To develop NIRS calibration models for prediction of amino acids, 80 samples was split up into a calibration set and 30 for a validation set. Three scans were conducted on each sample and the data were averaged before analysis.

Chemometrics and data analysis. In this study, spectra were exported from OMNIC software to TQ analyst software (version 9.7,Thermo Electron Corp.) for spectral pretreatment and chemometrics analysis.Two tailed Student's t-test was used for comparisons of two independent groups. A P value < 0.05 was considered to indicate statistical significance. Canonical Correlation Analysis (CCA) was used to find the correlation between plasma amino acid metabolism patterns and clinical parameters in breast cancer patients.

Declarations

Acknowledgements

The authors thank all of our colleagues working in the Department of

Medical Laboratory. This work was supported by grants from The Science and Technology Foundation of Guizhou province, Grant/Award

Number: [2021]483[2021] 366; Science and Technology Foundation of Guizhou Provincial Health Commission (GZWKJ2021-360, GZWJKJ2020-2-006);Doctor Fund from Affiliated Hospital of Guizhou Medical University(4 January 2019).

Conflict of interest

The authors have declared that no competing interests exist.

Availability of data and materials

All data and models generated or used during the study appear in the

submitted article.

Authors' contributions

All authors take public responsibility for the integrity of the data and the accuracy of analysis in the study. Xing Li and Wei Pan made substantial contributions to the concept and design of the present study. Liying Zhu and Haizhi Li wrote the main manuscript text and prepared figures. WenQian, ChangyudongHuang, Yiqiong Zhang ,Yunfeng Duan, Shuang Wang, Chengcheng Li ,YongJie Xu, Jingzhi Zhangand Mi Liu collected blood plasma samples. All authors reviewed the manuscript.

References

1. Xia, C., et al., *Cancer statistics in China and United States, 2022: profiles, trends, and determinants*. Chin Med J (Engl), 2022. **135**(5): p. 584–590.
2. Allemani, C., et al., *Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries*. Lancet, 2018. **391**(10125): p. 1023–1075.
3. Tang, S., et al., *CA153 in Breast Secretions as a Potential Molecular Marker for Diagnosing Breast Cancer: A Meta Analysis*. PLoS One, 2016. **11**(9): p. e0163030.
4. Manig, F., et al., *The why and how of amino acid analytics in cancer diagnostics and therapy*. J Biotechnol, 2017. **242**: p. 30–54.
5. Cheng, F., et al., *Investigation of salivary free amino acid profile for early diagnosis of breast cancer with ultra performance liquid chromatography-mass spectrometry*. Clin Chim Acta, 2015. **447**: p. 23–31.
6. Tumas, J., et al., *Towards a Personalized Approach in Pancreatic Cancer Diagnostics Through Plasma Amino Acid Analysis*. Anticancer Res, 2019. **39**(4): p. 2035–2042.
7. Kandasamy, P., et al., *Oncogenic KRAS mutations enhance amino acid uptake by colorectal cancer cells via the hippo signaling effector YAP1*. Mol Oncol, 2021. **15**(10): p. 2782–2800.
8. Jing, F., et al., *Discriminating gastric cancer and gastric ulcer using human plasma amino acid metabolic profile*. IUBMB Life, 2018. **70**(6): p. 553–562.
9. Plewa, S., et al., *Usefulness of Amino Acid Profiling in Ovarian Cancer Screening with Special Emphasis on Their Role in Cancerogenesis*. Int J Mol Sci, 2017. **18**(12).

10. Shingyoji, M., et al., *The significance and robustness of a plasma free amino acid (PFAA) profile-based multiplex function for detecting lung cancer*. BMC Cancer, 2013. **13**: p. 77.
11. Hirata, T. and Y. Kon, *Evaluation of the analytical capability of NIR femtosecond laser ablation-inductively coupled plasma mass spectrometry*. Anal Sci, 2008. **24**(3): p. 345–53.
12. Ajayakumar, P.V., et al., *FT-NIR spectroscopy for rapid and simple determination of nimesulide in rabbit plasma for pharmacokinetic analysis*. J Pharm Biomed Anal, 2012. **58**: p. 157–62.
13. Kapoor, S., *Re: Jehonathan H. Pinthus, Nir Kleinmann, Britton Tisdale, et al. Lower plasma adiponectin levels are associated with larger tumor size and metastasis in clear-cell carcinoma of the kidney. Eur Urol* 2008;*54:866 – 74.* Eur Urol, 2009. **55**(3): p. e52-3.
14. Cory, J.G. and A.H. Cory, *Critical roles of glutamine as nitrogen donors in purine and pyrimidine nucleotide synthesis: asparaginase treatment in childhood acute lymphoblastic leukemia*. In Vivo, 2006. **20**(5): p. 587–9.
15. Lukey, M.J., W.P. Katt, and R.A. Cerione, *Targeting amino acid metabolism for cancer therapy*. Drug Discov Today, 2017. **22**(5): p. 796–804.
16. Geck, R.C. and A. Toker, *Nonessential amino acid metabolism in breast cancer*. Adv Biol Regul, 2016. **62**: p. 11–17.
17. Yudkoff, M., et al., *Glutathione turnover in cultured astrocytes: studies with [15N]glutamate*. J Neurochem, 1990. **55**(1): p. 137–45.
18. DeBerardinis, R.J., et al., *Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis*. Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19345–50.
19. Metallo, C.M., et al., *Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia*. Nature, 2011. **481**(7381): p. 380–4.
20. Terunuma, A., et al., *MYC-driven accumulation of 2-hydroxyglutarate is associated with breast cancer prognosis*. J Clin Invest, 2014. **124**(1): p. 398–412.
21. Morotti, M., et al., *Increased expression of glutamine transporter SNAT2/SLC38A2 promotes glutamine dependence and oxidative stress resistance, and is associated with worse prognosis in triple-negative breast cancer*. Br J Cancer, 2021. **124**(2): p. 494–505.
22. Malevanets, A., et al., *Interplay of buried histidine protonation and protein stability in prion misfolding*. Sci Rep, 2017. **7**(1): p. 882.
23. Gracia-Marco, L., et al., *Amino acids intake and physical fitness among adolescents*. Amino Acids, 2017. **49**(6): p. 1041–1052.
24. McMahon, G.M., et al., *Urinary metabolites along with common and rare genetic variations are associated with incident chronic kidney disease*. Kidney Int, 2017. **91**(6): p. 1426–1435.
25. Ceulemans, T., et al., *Nutrient enrichment is associated with altered nectar and pollen chemical composition in Succisa pratensis Moench and increased larval mortality of its pollinator Bombus terrestris L*. PLoS One, 2017. **12**(4): p. e0175160.
26. Pinotsis, N. and G. Waksman, *Structure of the WipA protein reveals a novel tyrosine protein phosphatase effector from Legionella pneumophila*. J Biol Chem, 2017. **292**(22): p. 9240–9251.

27. Alvarez-Aznar, A., L. Muhl, and K. Gaengel, *VEGF Receptor Tyrosine Kinases: Key Regulators of Vascular Function*. Curr Top Dev Biol, 2017. **123**: p. 433–482.
28. Suzuki, Y., et al., *Clinical Implications of Plasma N-acetyl-seryl-aspartyl-lysyl-proline Level in Stable Kidney Transplant Recipients*. Clin Lab, 2016. **62**(7): p. 1323–1328.
29. Nair, A.G., et al., *Discovery of silyl proline containing HCV NS5A inhibitors with pan-genotype activity: SAR development*. Bioorg Med Chem Lett, 2016. **26**(5): p. 1475–9.
30. Mason, A., et al., *A four-compartment compartmental model to assess net whole body protein breakdown using a pulse of phenylalanine and tyrosine stable isotopes in humans*. Am J Physiol Endocrinol Metab, 2017. **313**(1): p. E63-E74.
31. Cheniany, M. and A. Ganjeali, *Developmental role of phenylalanine-ammonia-lyase (PAL) and cinnamate 4-hydroxylase (C4H) genes during adventitious rooting of Juglans regia L. microshoots*. Acta Biol Hung, 2016. **67**(4): p. 379–392.
32. Stubbs, T.L., A.C. Kennedy, and A.M. Fortuna, *Using NIRS to predict fiber and nutrient content of dryland cereal cultivars*. J Agric Food Chem, 2010. **58**(1): p. 398–403.
33. Llario, R., et al., *Determination of quality parameters of beers by the use of attenuated total reflectance-Fourier transform infrared spectroscopy*. Talanta, 2006. **69**(2): p. 469–80.
34. Kibugu, J., et al., *Improved Sample Selection and Preparation Methods for Sampling Plans Used to Facilitate Rapid and Reliable Estimation of Aflatoxin in Chicken Feed*. Toxins (Basel), 2021. **13**(3).
35. Mishra, P. and R. Nikzad-Langerodi, *A brief note on application of domain-invariant PLS for adapting near-infrared spectroscopy calibrations between different physical forms of samples*. Talanta, 2021. **232**: p. 122461.
36. Almendingen, K., et al., *Near infrared spectroscopy-a potentially useful method for rapid determination of fat and protein content in homogenized diets*. Eur J Clin Nutr, 2000. **54**(1): p. 20–3.
37. Saugo, M., et al., *Mineral equilibrium in commercial curd and predictive ability of near-infrared spectroscopy*. J Dairy Sci, 2021. **104**(4): p. 3947–3955.

Figures

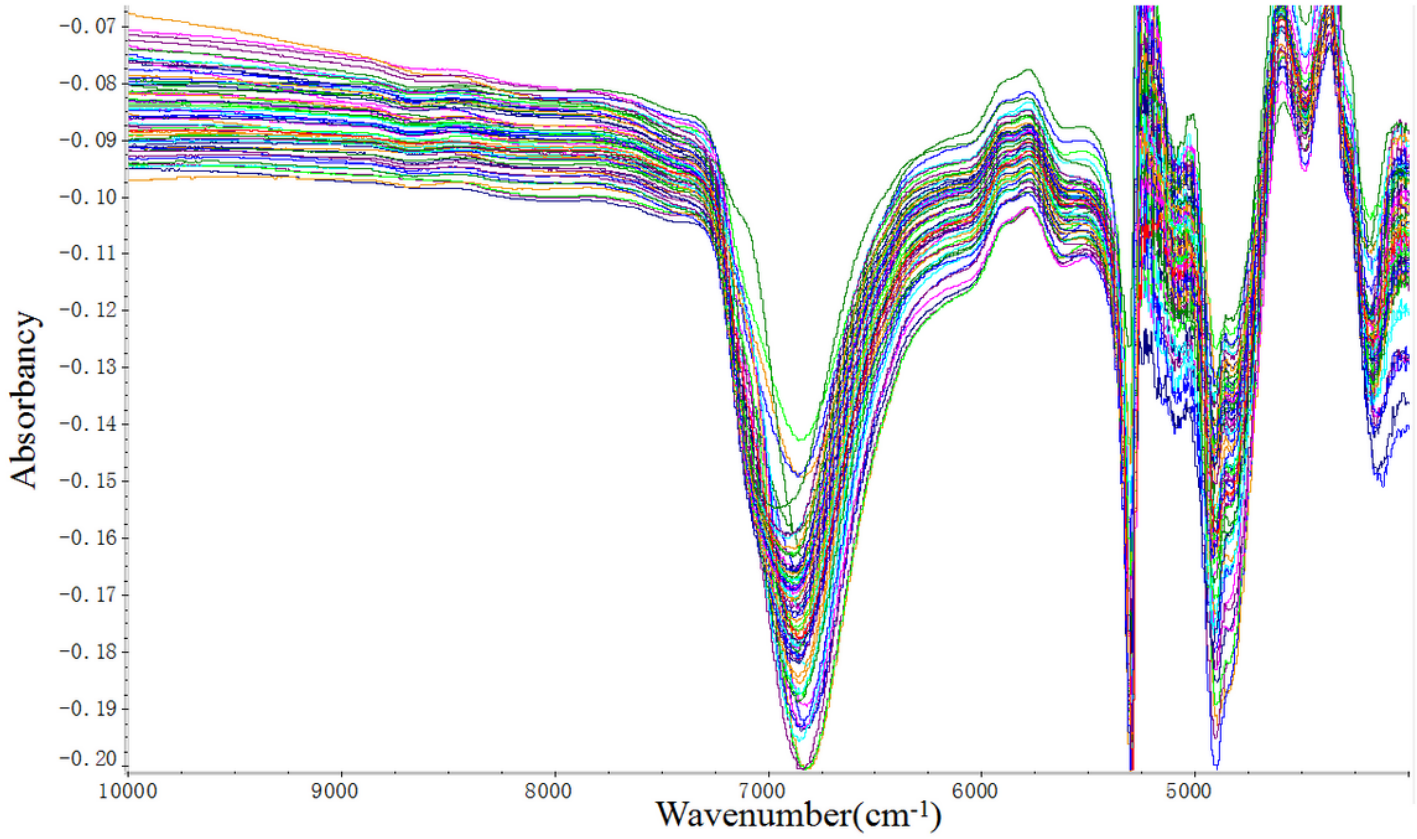

Figure 1

NIRS of breast cancer plasma samples (wave number, (nm) in abscissa and absorbance in ordinate).

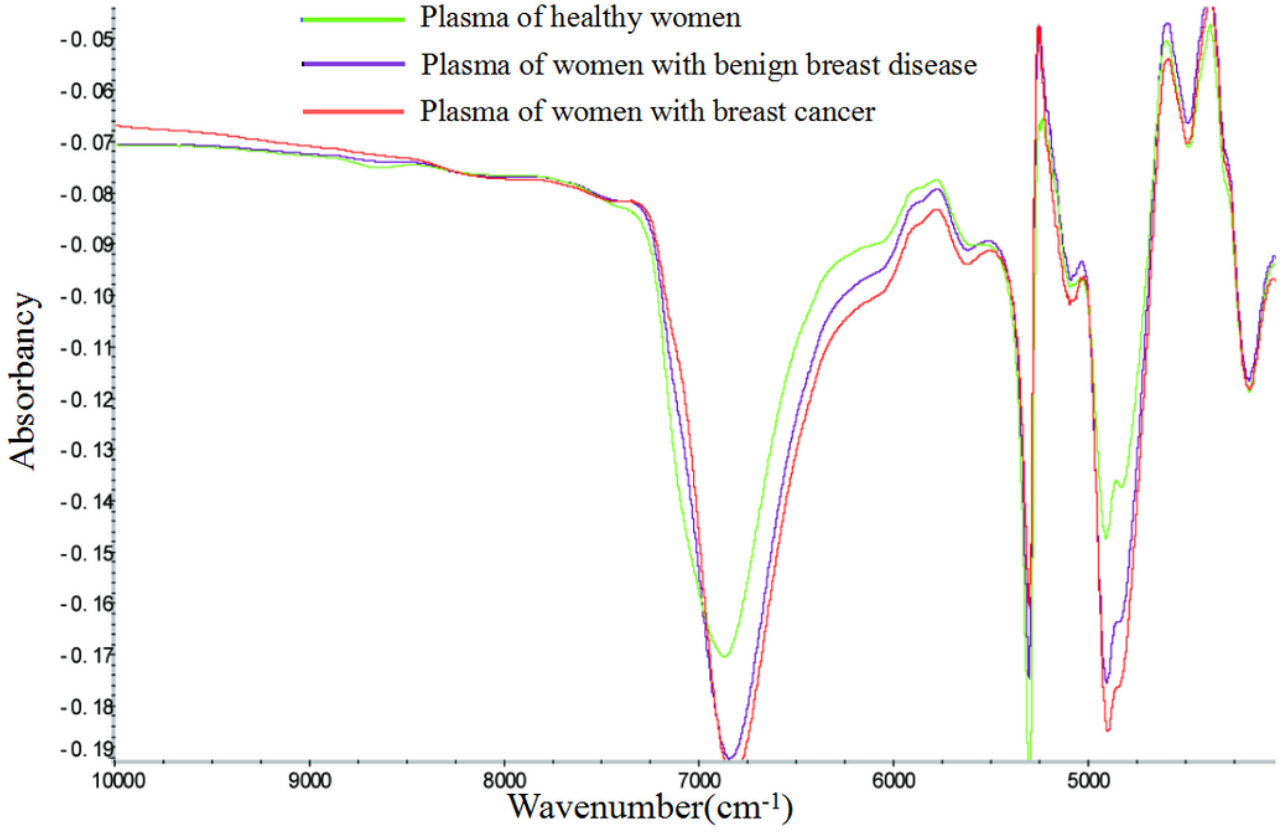

Figure 2

The average raw spectrum for each group of sample

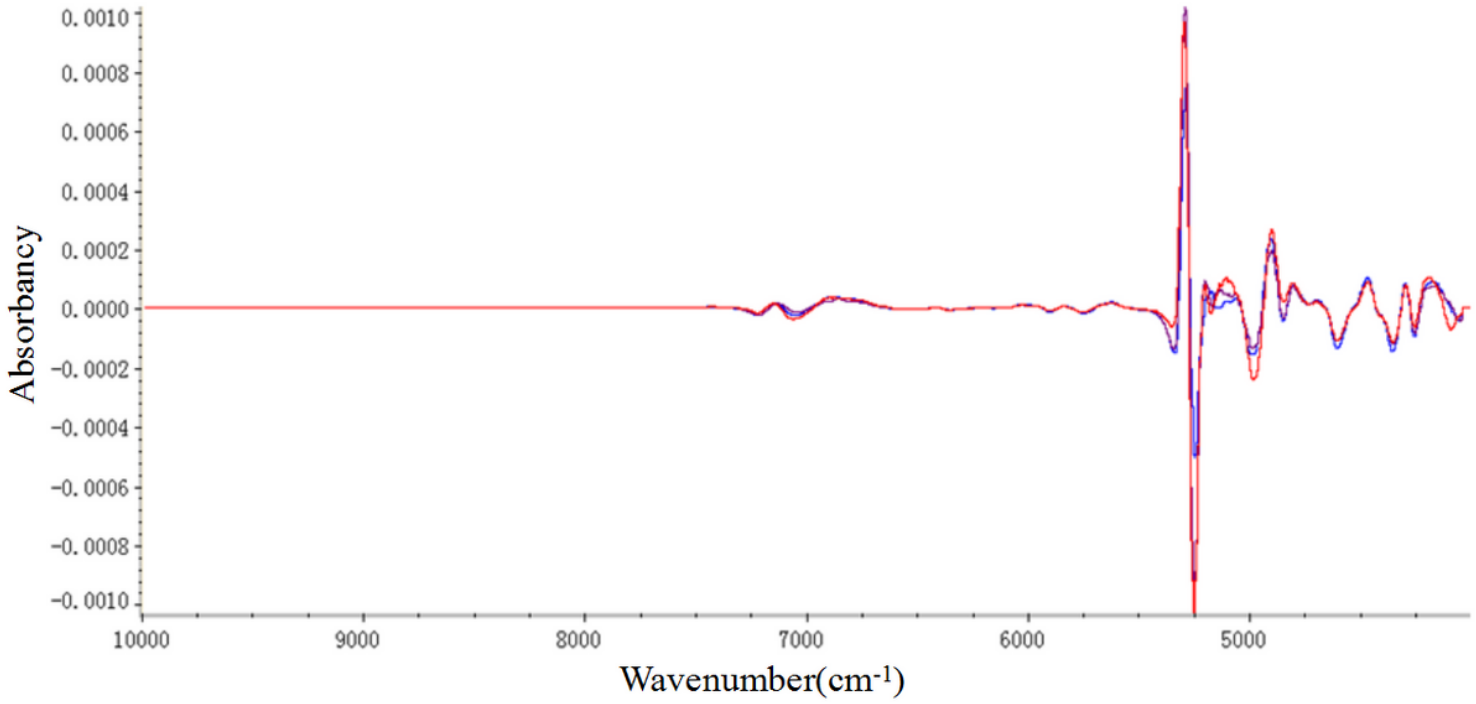

Figure 3

The average spectrum after pre-processed by the second derivative +Norris derivative filter (5, 5)

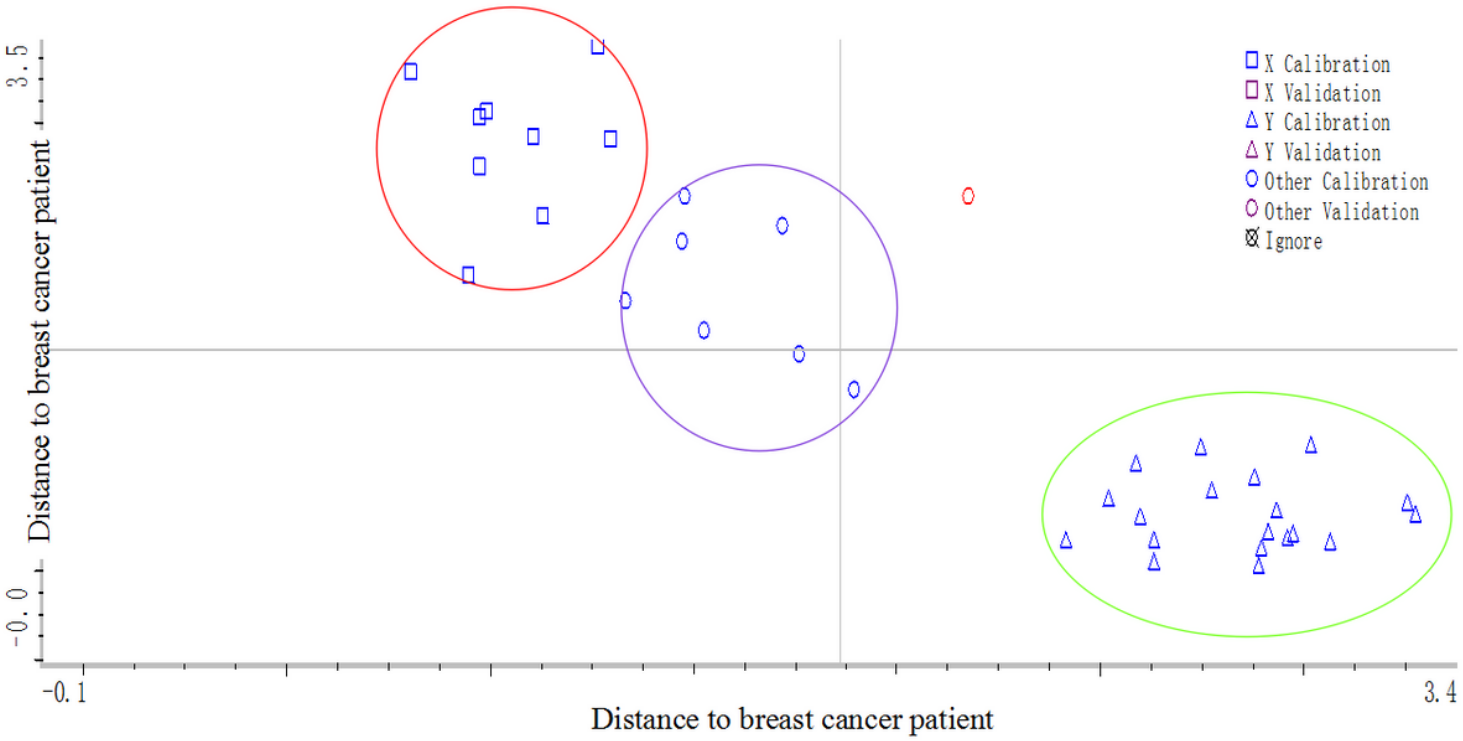

Figure 4

Discriminant analysis of different varieties of plasma

Note: The square icon is breast cancer plasma; The circular icon is the plasma of patients with benign breast diseases; The triangle shows healthy female plasma

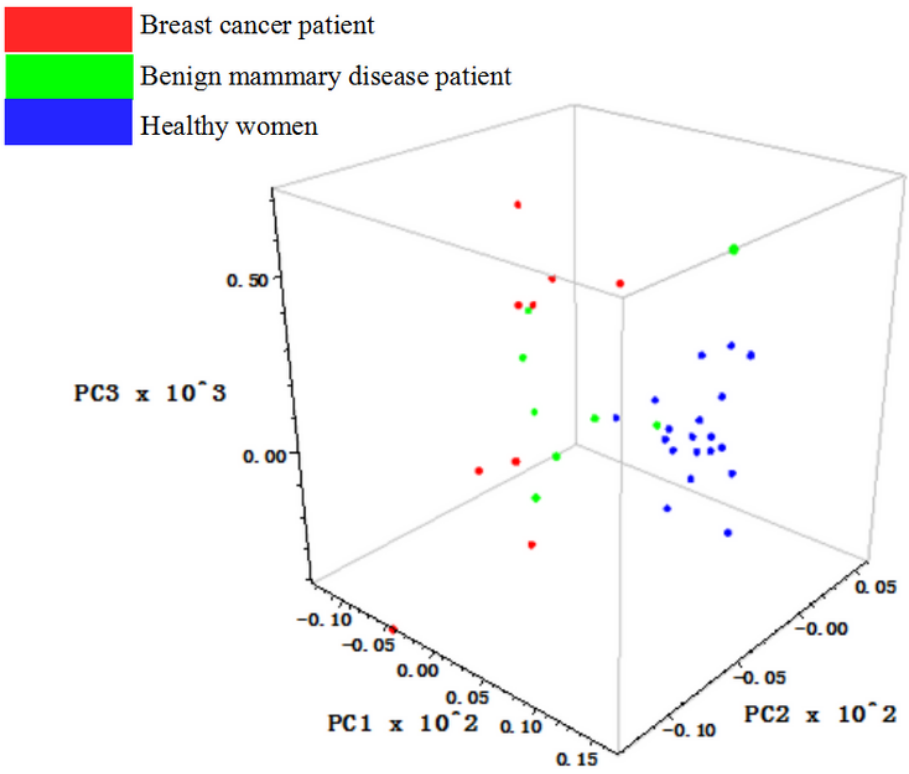

Figure 5

Three-dimensional principal component score plot derived from

raw spectra.

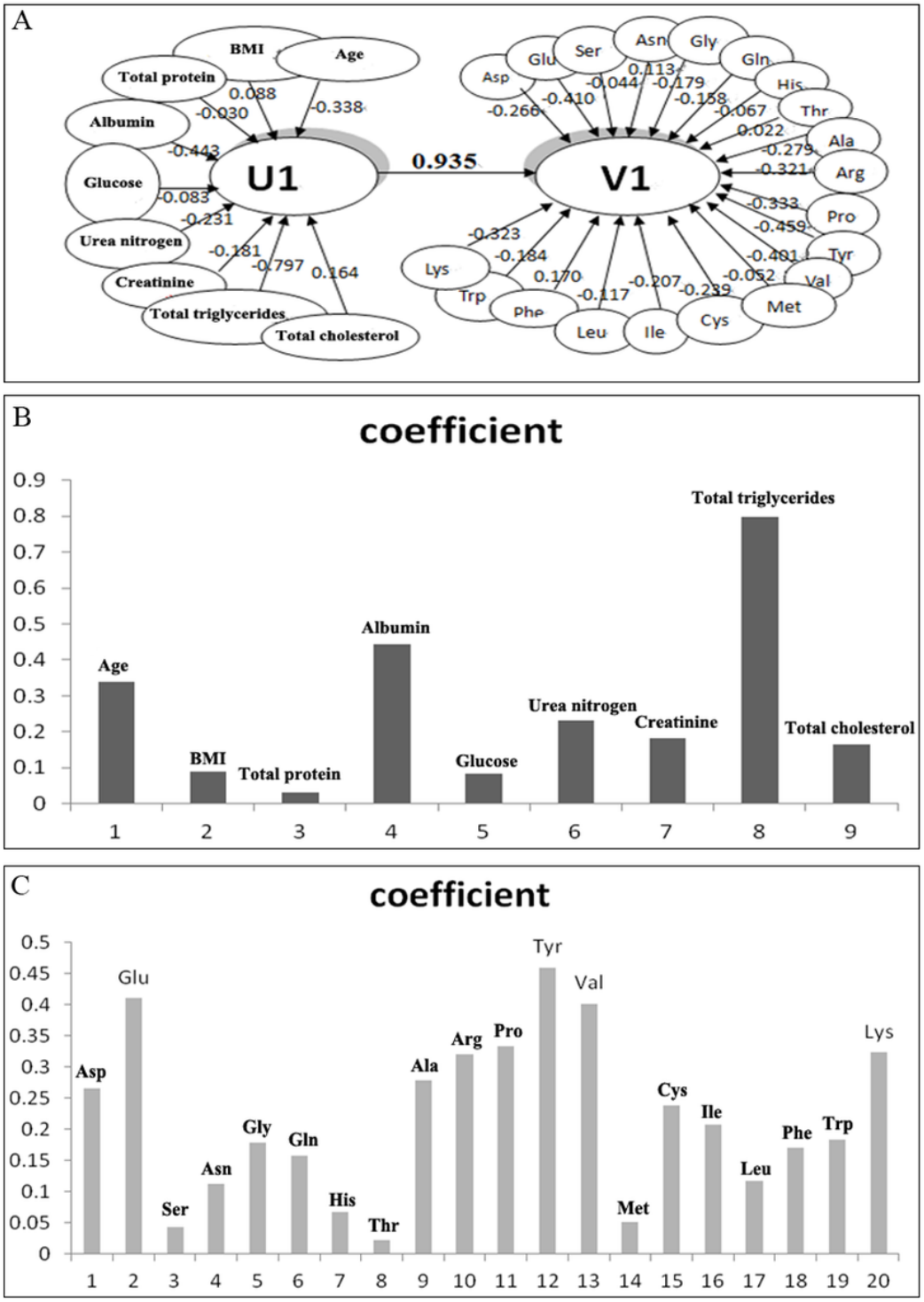

Figure 6

Canonical correlation analysis on the association between plasma amino acids and clinical parameters among breast cancer.（A）.The typical structure diagram of the first couple typical variant. （B）The absolute clinical parameter coefficient value of the first couple typical variant. （C）The absolute amino acid parameter coefficient value of the first couple typical variant