Changes in serum amino acid level in non-small cell lung cancer: a case-control study in Chinese population

Ke Liu [Equal first author, 1] , Jiaoyuan Li [Equal first author, 1] , Tingting Long [1] , Yi Wang [1] , Tongxin Yin [1] , Jieyi Long [1] , Ying Shen [Corresp., 1] , Liming Cheng [Corresp. 1]

[1] Department of Laboratory Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Corresponding Authors: Ying Shen, Liming Cheng
Email address: sying830@163.com, chengliming2015@163.com

Background: Previous studies have shown the alteration of amino acid (AA) profile in patients with non-small cell lung cancer (NSCLC). However, there is little data regarding AA profile in NSCLC in Chinese population. The aim of this study was to evaluate AA profile in Chinese NSCLC patients, explore its utility in sample classification and further discuss its related metabolic pathways. **Methods:** The concentrations of 22 AAs in serum samples from patients with 200 NSCLC and 202 healthy controls were determined by liquid chromatography-tandem mass spectrometer (LC-MS/MS). AA levels in different tumor stages and histological types were also discussed. The performance of AA panel in classifying the cases and controls was evaluated in the training data set and validation data set based on the receiver operating characteristic (ROC) curve, and the important metabolic pathways were identified. **Results:** The concentrations of tryptophan (Trp), phenylalanine (Phe), isoleucine (Ile), glycine (Gly), serine (Ser), aspartic acid (Asp), asparagine (Asn), cystein (Cys), glutamic acid (Glu), ornithine (Orn) and citrulline (Cit) were significantly altered in NSCLC patients compared with controls (all P-FDR<0.05). Among these, four AAs including Asp, Cys, Glu and Orn were substantially up-regulated in NSCLC patients (FC≥1.2). AA levels were significantly altered in patients with late-stage NSCLC, but not in those with early-stage when comparing with healthy controls. In terms of histological type, these AA profile was altered in both adenocarcinoma and squamous cell carcinoma. For discrimination of NSCLC from controls, the area under the ROC curve (AUC) was 0.80 (95%CI: 0.74-0.85) in the training data set and 0.79 (95%CI: 0.71-0.87) in the validation data set. The AUCs for early-stage and late-stage NSCLC were 0.75 (95%CI: 0.68-0.81) and 0.86 (95%CI: 0.82-0.91), respectively. Moreover, the model showed a better performance in the classification of squamous cell carcinoma (AUC=0.90, 95%CI: 0.85-0.95) than adenocarcinoma (AUC=0.77, 95%CI: 0.71-0.82) from controls. Three important metabolic pathways were involved in the alteration of AA profile, including Gly, Ser and Thr metabolism; Ala, Asp and Glu metabolism; and Arg biosynthesis.

Conclusions: The levels of several AAs in serum were altered in Chinese NSCLC patients. These altered AAs may be utilized to classify the cases from the controls. Gly, Ser and Thr metabolism; Ala, Asp and Glu metabolism and Arg biosynthesis pathways may play roles in metabolism of the NSCLC patient.

1 **Changes in serum amino acid level in non-small cell lung cancer: a case-control study in**
2 **Chinese population**
3

5 Ke Liu*[1], Jiaoyuan Li*[1], Tingting Long[1], Yi Wang[1], Tongxin Yin[1], Jieyi Long[1], Ying Shen[1],
6 Liming Cheng[1]
7
8 [1] Department of Laboratory Medicine, Tongji Hospital, Tongji Medical College, Huazhong
9 University of Science and Technology, Wuhan, 430030, China.
10
11 *****: These authors contributed equally to this work.
12
13 Corresponding Author:
14 Ying Shen1
15 Wuhan 430030, PR China
16 Email address: sying830@163.com
17 Liming Cheng1
18 sying830@163.com
19 Email address: chengliming2015@163.com

21 Abbreviations:
22 AA: amino acid; ENAAs: essential amino acids; NEAAs: non-essential amino acids; NSCLC:
23 non-small lung cancer; Lys: lysine; Met: methionine; Ile: isoleucine; Leu: leucine; Trp:
24 tryptophan; Phe: phenylalanine; Val: valine; Thr: threonine; His: histidine; Asp: aspartic acid;
25 Asn: asparagine; Ser: serine; Glu: glutamic acid; Gln: glutamine; Gly: glycine; Ala: alanine; Cys:
26 cystein; Tyr: tyrosine; Orn: ornithine; Cit: citrulline; Arg: arginine; Pro: proline; PA:
27 phosphatidic acid; Rheb: ras homologue enriched in brain; LEL: late lysosome/endosome
28 system; ROC: receiver operating characteristic; AUC: area under the ROC curve; SCC:
29 squamous cell carcinoma; TNM: tumor node metastasis; IQR: interquartile range; FDR: false
30 discovery rate; LOD: limit of detection; LOQ: limit of quantification; α-KG: α-ketoglutarate;
31 TCA: tricarboxylic acid cycle; ASNS: asparagine synthase; CH2-THF: 5, 10-methylene
32 tetrahydrofolate;

33 **Abstract**

34 **Background:** Previous studies have shown the alteration of amino acid (AA) profile in patients
35 with non-small cell lung cancer (NSCLC). However, there is little data regarding AA profile in
36 NSCLC in Chinese population. The aim of this study was to evaluate AA profile in Chinese
37 NSCLC patients, explore its utility in sample classification and further discuss its related
38 metabolic pathways.

39 **Methods:** The concentrations of 22 AAs in serum samples from patients with 200 NSCLC and
40 202 healthy controls were determined by liquid chromatography-tandem mass spectrometer (LC-
41 MS/MS). AA levels in different tumor stages and histological types were also discussed. The
42 performance of AA panel in classifying the cases and controls was evaluated in the training data
43 set and validation data set based on the receiver operating characteristic (ROC) curve, and the
44 important metabolic pathways were identified.

45 **Results:** The concentrations of tryptophan (Trp), phenylalanine (Phe), isoleucine (Ile), glycine
46 (Gly), serine (Ser), aspartic acid (Asp), asparagine (Asn), cystein (Cys), glutamic acid (Glu),
47 ornithine (Orn) and citrulline (Cit) were significantly altered in NSCLC patients compared with
48 controls (all *P*-FDR<0.05). Among these, four AAs including Asp, Cys, Glu and Orn were
49 substantially up-regulated in NSCLC patients (FC≥1.2). AA levels were significantly altered in
50 patients with late-stage NSCLC, but not in those with early-stage when comparing with healthy
51 controls. In terms of histological type, these AA profile was altered in both adenocarcinoma and
52 squamous cell carcinoma. For discrimination of NSCLC from controls, the area under the ROC
53 curve (AUC) was 0.80 (95%CI: 0.74-0.85) in the training data set and 0.79 (95%CI: 0.71-0.87)
54 in the validation data set. The AUCs for early-stage and late-stage NSCLC were 0.75 (95%CI:
55 0.68-0.81) and 0.86 (95%CI: 0.82-0.91), respectively. Moreover, the model showed a better
56 performance in the classification of squamous cell carcinoma (AUC=0.90, 95%CI: 0.85-0.95)
57 than adenocarcinoma (AUC=0.77, 95%CI: 0.71-0.82) from controls. Three important metabolic
58 pathways were involved in the alteration of AA profile, including Gly, Ser and Thr metabolism;
59 Ala, Asp and Glu metabolism; and Arg biosynthesis.

60 **Conclusions:** The levels of several AAs in serum were altered in Chinese NSCLC patients.
61 These altered AAs may be utilized to classify the cases from the controls. Gly, Ser and Thr
62 metabolism; Ala, Asp and Glu metabolism and Arg biosynthesis pathways may play roles in
63 metabolism of the NSCLC patient.

64

65 **Keywords:** non-small cell lung cancer; amino acids; metabolomics; liquid chromatography-
66 tandem mass spectrometer

Introduction

Lung cancer is one of the most common malignant tumors which threaten human health and life. According to global cancer statistics in 2020, lung cancer is the second most common cancer in terms of incidence, and the first leading cause of death globally, accounting for approximately 18% of global cancer deaths [1]. Non-small cell lung cancer (NSCLC) occupies over 85% of all the lung cancers, and the most common sub-types are adenocarcinoma and squamous cell carcinoma (SCC) among them [2]. The 5-year survival rate for localized lung cancer is 57%, while for metastatic lung cancer it is only 5% [3]. In China, approximately 733,300 lung cancer patients occurred in 2015 according to the latest data of the national cancer registries [4]. It is no doubt that lung cancer is still a serious public health problem in both China and the global.

As basic metabolites and regulatory factors, amino acids (AAs) play extensive effects on cancer cells [5]. For instance, AAs account for the majority of the carbon-based biomass production, and are the dominant nitrogen source for hexosamines, nucleotides, and other nitrogenous compounds in cancer cells [6-9]. AAs can also regulate gene expression and the protein phosphorylation cascade. For example, adequate essential amino acids (EAAs) stimulated MYC oncogene mRNA translation by inhibiting the GCN2-eIF2A-ATF4 stress pathway, leading to MYC-dependent transcriptional amplification [10]. Some AAs such as glutamine (Glu), arginine (Arg) and leucine (Leu) stimulated the cell-specific phosphorylation of mTOR1, thereby regulating the conversion of intracellular proteins [11]. Meanwhile, glutamine (Gln), citrulline (Cit) and methionine (Met) could also activate the mTOR1 signaling pathway through modulating the concentration of phosphatidic acids (PAs) and activity of Ras homologue enriched in brain (Rheb) in the late lysosome/endosome system (LEL) [12]. On the other hand, the metabolism of AAs is strictly regulated and closely related to other metabolic networks, such as lipid and glucose metabolism.

Although the concentrations of circulating AAs remain constant in a healthy state, numerous studies have found that various cancers, such as colorectal, lung and breast cancers, can influence the levels of AAs [13]. In recent years, the AminoIndex has aroused wide attention, which facilitates the evaluation of certain health conditions and the possibility of disease by establishing multivariate indices from blood free AA profile information [14]. Using these indices, researchers have discriminated various cancer types from healthy controls, including lung, colorectal, breast, gastric, prostate, gynecological and pancreatic cancers [13, 15]. For lung cancer, AminoIndex has also been utilized for monitoring cancer treatment [16]. These evidences suggested that AAs are suitable candidates for the study of NSCLC metabolomics, and it might be of potential value in clinical application.

At present, several studies have reported the alterations of blood free AA profile in lung cancer patients from Japan, Korea and Europe countries [17-20]. However, there is still little data regarding AA profile in NSCLC patients in China. Hence, a case-control study with 202 NSCLC patients and 200 healthy controls was conducted to characterize the feature of AA profile in Chinese NSCLC patients in this study. A total of 22 AAs, including 8 common EAAs and 14 common non-essential amino acids (NEAAs), were determined by a liquid chromatography-

110 tandem mass spectrometer (LC-MS/MS) method [21-23]. The alteration of AA profile in
111 NSCLC was identified and the related metabolic pathways were discussed. We also evaluated
112 the performance of the AA profile for the detection of NSCLC across different histological types
113 and stages in Chinese patients.
114

Materials & Methods

Study design and participants

117 The case group was comprised of 200 newly diagnosed NSCLC patients who were recruited
118 from Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology
119 from January 2017 to May 2018. All patients were histopathologically confirmed NSCLC
120 without radiotherapy, chemotherapy or surgery before admission. The histologic type of tumor
121 was identified based on results of histopathological examination, and the tumor stage was
122 confirmed according to the eighth edition of the tumor node metastasis (TNM) staging system
123 [24]. A total of 202 healthy subjects were randomly enrolled from healthy people who underwent
124 physical examinations in the same hospital during the same time as recruitment of patients. The
125 healthy controls were age and gender-matched with NSCLC cases, and none of them had a
126 history of lung cancer or other cancers. All serum samples of the participants were obtained
127 through intravenous collection after overnight fasting and stored at -80 °C until use. This case-
128 control study was approved by the Institutional Review Board of Tongji Medical College,
129 Huazhong University of Science and Technology, and it does not require the informed consent of
130 participants (IRB approval number: TJ-IRB20210806).
131

Determination of free AA profiles

133 The determination of AAs was performed on a LC-MS/MS (AB Sciex, USA) with an
134 electrospray ionization source in positive mode. The separation was achieved on a MSLAB-
135 45+AAC18 column (150*4.6 mm 5 μm, Beijing Mass Spectrometry Medical Research Ltd.,
136 China) with a flow rate of 1 mL/min. The column temperature was 50 ℃, and the injection
137 volume was 5 μL.
138

139 For AAs analysis, the AAs quantitation kits (Beijing Mass Spectrometry Medical Research Ltd,
140 China) were used. For sample preparation, 50 μL serum was firstly mixed with 50 μL protein
141 precipitation agent, and then vortexed for 30 s and centrifuged at 14,680 rpm for 10 mins. Then,
142 8 μL of the supernatant was mixed with 42 μL labeling buffer solution which containing
143 synthetic norvaline as internal standard. After that, and 20 μL of the derivatizing solution was
144 added for incubation at 55 ℃ for 15 min. After cooling down, the solution was mixed well and
145 centrifuged at 14,680 rpm for 10 min for direct LC-MS/MS analysis. The data acquisition and
146 quantitation were achieved by Analyst 1.6. (AB Sciex, USA).
147

148 We simultaneously detected 22 AAs: 8 EAAs including phenylalanine (Phe), tryptophan (Trp),
149 isoleucine (Ile), leucine (Leu), methionine (Met), lysine (Lys), valine (Val), and threonine (Thr),
150 and 14 NEAAs including glycine (Gly), serine (Ser), asparagine (Asn), aspartic acid (Asp),
151 cystein (Cys), glutamic acid (Glu), glutamine (Gln), histidine (His), proline (Pro), arginine (Arg),
152 tyrosine (Tyr), alanine (Ala), ornithine (Orn) and citrulline (Cit).

153
154 **Statistical analysis**
155 Continuous variables were presented as median and interquartile range (IQR), and categorical
156 variables were presented as counts and percentages. Due to the relatively large sample size of
157 this study, the distribution of the data was tested by Kolmogorov–Smirnov test. Chi-square test
158 was used to compare categorical variables, while T-test or Mann-Whitney U test was employed
159 to compare continuous variables in two groups. False discovery rate (FDR) was used to adjust *P*
160 values for multiple comparisons. An AA was identified as de-regulation if the fold change (FC)
161 of AA level in NSCLC and control groups exceeded the threshold (FC≥1.2 or≤0.8) with *P*-
162 FDR<0.05. However, considering the essential roles of AAs in maintaining physiological
163 homeostasis, to obtain more information and improve efficiency, all the significantly altered AAs
164 (*P*-FDR<0.05) in NSCLC patients were selected as candidates to perform pathway analysis and
165 construct the logistic regression model for NSCLC discrimination. To ensure the credibility of
166 the results in model construction, we randomly divided all the participants into two groups with a
167 ratio of 7:3: 70% of the samples assigned into training set and the other 30% into validation set.
168 Receiver operating characteristic (ROC) curve was assessed to illustrate and evaluate the
169 performance for case classification. The MetaboAnalyst 5.0 was used for metabolic pathway
170 analysis.
171
172 All statistical analyses were carried out using SPSS 22.0 (IBM Corporation, Armonk, NY, USA)
173 and GraphPad Prism 8.0 (GraphPad, San Diego, CA, USA). The visualization of the model for
174 NSCLC detection was achieved by the Deepwise & Beckman Coulter DxAI platform
175 (https://dxonline.deepwise.com). And two-sided *P* values of less than 0.05 were considered
176 statistically significant.
177
178 **Results**
179 **Method validation.**
180 This LC-MS/MS method was overall well validated with linearity, precision, accuracy, analytical
181 sensitivity and matrix effects as shown in Table S1-S5. The linear correlation coefficients were
182 all greater than 0.9855. The intra-day precision and inter-day precision at two levels were within
183 20% except that Ser, Met, Tyr and Trp were within 24% at the low level. The recovery rates
184 were in the range of 80-120% except that Pro, His and Trp were 121%, 128% and 66% at the
185 low level, respectively, and Cys was 128% at the high level. The limit of detection (LOD) and
186 the limit of quantification (LOQ) were also investigated to ensure the reliable results at low
187 concentration. A mixed experiment showed that the presence of the matrix effect didn't affect
188 the accurate quantification of the target analytes. All these ensured the reliability of the results in
189 this study.
190
191 **Demographics and clinical characteristics.**
192 The basic and clinical characteristics of the 200 NSCLC patients and 202 control subjects are
193 shown in Table 1. The results showed that 51.0% of the cases were smokers, while in controls,
194 the percentage of smokers was 40.1% (*P*=0.028). No significant difference was found between

195   the cases and the controls about drinking status. Besides, the average body mass index (BMI) of
196   the cases was significantly lower than that of the controls (24.0 kg/m² vs. 22.8 kg/m², P=0.002).
197   Among the patients, 66.5% were diagnosed with adenocarcinoma, while 28.5% were SSC and
198   the remaining 5% were other NSCLC. According to the eighth edition of the TNM classification
199   system, 2.0% of the NSCLC patients were categorized as stage 0, 33.5% as stage I, 7.0% as stage
200   II, 26.0% as stage III, 26.5% as stage IV, and the remaining 5% of patients as unclear.

202   In the subgroup of NSCLC, the percentage of smokers was 39.8% in the Adenocarcinoma group,
203   while 73.7% of the SCC were smokers (Table S6, P<0.001). There are more males in SCC
204   patients than adenocarcinoma patients (89.5% vs 57.9%, P<0.001). No significant difference was
205   found about age, drinking status and BMI between the adenocarcinoma and SCC group. There
206   are more male patients in late-stage than early-stage patients (77.1% vs 57.6%, P=0.004). The
207   percentage of smokers was 40.0% and 58.1% in early-stage and late-stage patients, respectively
208   (Table S7, P=0.013). While 17.6% and 36.2% of the early-stage and late-stage patients were
209   drinkers, respectively (P=0.005). No significant difference was found between the early-stage
210   and the late-stage patients about age and BMI.

212   **Alteration of serum AA concentrations in NSCLC patients.**
213   The comparison of serum AA concentrations between the case and control groups is presented in
214   Table 2. After multiple testing using FDR method, the concentration of Trp, Phe, Ile, Gly, Ser,
215   Asp, Asn, Cys, Glu, Orn, and Cit were significantly altered in serum of NSCLC patients when
216   comparing with healthy controls (all P-FDR<0.05). Among these, four AAs including Asp, Cys,
217   Glu and Orn were substantially up-regulated (FC≥1.2) in serum of NSCLC patients.

219   **AA concentrations in patients with different histological types and clinical stages.**
220   The results of subgroup analysis based on histological type are shown in Table 3. After multiple
221   testing using FDR method, the concentrations of Trp, Gly, Ser, Asp, Asn, Cys, Glu, Gln, Orn,
222   and Cit were significantly altered in serum of adenocarcinoma patients when comparing with
223   healthy controls (all P-FDR<0.05). Among these, three AAs including Asp, Cys and Orn were
224   substantially up-regulated (FC≥1.2) in serum of adenocarcinoma patients. And the
225   concentrations of Phe, Asp, Cys, and Glu were significantly altered in serum of SCC patients
226   when comparing with healthy controls (all P-FDR<0.05). Among these, three AAs including
227   Asp, Cys and Glu were substantially up-regulated (FC≥1.2) in serum of SCC patients. Besides,
228   the concentration of Asp was significantly altered and substantially down-regulated in serum of
229   adenocarcinoma patients when comparing with SCC patients (P-FDR<0.05, FC=0.69).

231   AA levels in patients with different clinical stages are shown in Figure1. After multiple testing
232   using FDR method, the concentration of Glu was significantly altered in serum of early-stage
233   NSCLC patients (P-FDR=0.044), whilst the concentrations of Phe, Trp, Ile, Asp, Cys, Glu, Orn,
234   Gly, Ser, Asn and Arg were significantly altered in serum of late-stage NSCLC patients when
235   comparing with healthy controls (all P-FDR<0.05). Four AAs including Asp, Cys, Glu and Orn
236   were substantially up-regulated (FC≥1.2) in serum of late-stage NSCLC patients. Although there

237   was no significant difference in AA levels between the early-stage and late-stage NSCLC,
238   several AA levels showed consecutive changes with the disease progression. For example, the
239   Trp concentration gradually decreased in the control group, early-stage and late-stage NSCLC
240   patients successively, while concentrations of AAs such as Ile, Gly, Ser, Asn, Asp, Cys, Glu and
241   Orn showed a gradually increased trend.
242

243   **Metabolic pathway analysis of AAs in NSCLC.**
244   Metabolic pathway analysis was performed based on the altered 11 AAs between the NSCLC
245   group and the control group. As illustrated in Figure 2 and Table S8, we identified three
246   significant pathways with important impact (impact factor >0.4, P-FDR<0.05), including Gly,
247   Ser and Thr metabolism which hits three AAs including Ser, Gly, Cys (P-FDR=0.016, impact
248   factor=0.463), Ala, Asp and Glu metabolism which hits three AAs including Glu, Asp, Asn (P-
249   FDR=0.017, impact factor=0.421), and Arg biosynthesis which hits four AAs including Glu,
250   Asp, Cit, Orn (P- FDR≤0.001, impact factor=0.406).
251

252   **The classification of NSCLC patients and control group.**
253   The AA panel for NSCLC classification was composed of the 11 AAs which of the levels were
254   altered in NSCLC patients compared with controls, including Phe, Trp, Ile, Gly, Ser, Asn, Asp,
255   Cys, Glu, Orn, and Cit, and the model was generated using logistic regression. As a result, the
256   area under the curve (AUC) of the ROC was 0.80 (95%CI: 0.74-0.85) in the training data set
257   (Figure 3A). In the validation data set, the AUC of the model also reached 0.79 (95%CI: 0.71-
258   0.87) (Figure 3A). These results demonstrated that the AA panel may be of potential value in
259   NSCLC classification and detection.
260

261   To further investigate the performance of the model in discriminating NSCLC patients across
262   different histological types and clinical stages, the AUCs for each histological type and stage
263   were also estimated. For early-stage NSCLC, the AUC was 0.75 (95%CI: 0.68-0.81), while for
264   late-stage NSCLC, the AUC reached 0.86 (95%CI: 0.82-0.91) (Figure 3B). Moreover, the model
265   showed large discrepancy in the detection of adenocarcinoma and SCC, with the AUC of 0.77
266   (95%CI: 0.71-0.82) for adenocarcinoma, while as high as 0.90 (95%CI: 0.85-0.95) for SCC
267   (Figure 3B).
268

269   **Discussion**
270   Many previous reports have revealed that the metabolism was perturbed in cancer patients,
271   including the metabolism of AAs [25, 26]. As basic building blocks of proteins, AAs play crucial
272   roles in maintaining normal life activities and meeting the nutritional and energy requirements of
273   tumor cells as well [27, 28]. Previous studies showed that there were substantial changes in AA
274   profile in lung cancer, and an index derived from the altered AAs, the AminoIndex, has been
275   utilized for lung cancer screening and cancer treatment monitoring [13, 15-17, 19, 20, 29]. At
276   present, the common methods for determining the content of AAs in fluid samples include
277   capillary electrophoresis (CE), ion-exchange chromatography (IEX), high performance liquid
278   chromatography (HPLC), and LC-MS/MS, etc [30-33]. Compared with other methods, the LC-
279   MS/MS method which adopted in our study has the advantages of high throughput, good

280 selectivity, time-saving and cost-effective, thus has been widely used in clinical assessment of
281 AAs. Based on the application of amino derivative markers, this method has greatly improved the
282 accuracy, precision, and detection range of AAs determination [34]. In this study, we found that the
283 levels of 11 out of 22 AAs were significantly altered between the NSCLC and the healthy people
284 in Chinese population. Three important metabolic pathways were involved in the metabolism of
285 these AAs, including glycine, serine and threonine metabolism; alanine, aspartate and glutamate
286 metabolism; and arginine biosynthesis. In addition, this AA panel were of potential value in
287 sample classification of NSCLC.
288
289 The current study identified 11 altered AAs in serum of NSCLC, included Phe, Trp, Ile, Gly, Ser,
290 Asp, Asn, Cys, Glu, Orn, and Cit. In previous studies, circulating Phe, Ile, Ser, Asp and Orn
291 were found to be elevated while Trp showed a downward trend in lung cancer, which coincided
292 with our results [13, 17-20, 29, 35]. However, some inconsistent results were also found in
293 several studies. For instance, Zhao's team and Kim's group [19] reported down-regulated Gly
294 concentration, while Maeda's team, Miyagi's group [13] and our study observed that Gly was
295 up-regulated in NSCLC patients. Remarkably, abnormal levels of these AAs have been found in
296 different populations as well [17, 19, 20, 29]. The reasons for the inconsistent changes in AA
297 levels may be related to the diversity of genetic background, regional differences, nutritional
298 status, and lifestyle factors.
299
300 We found three important metabolic pathways which were involved in the metabolism of the
301 altered AAs. Among these, Glu, Asp and Asn were involved in alanine, aspartate and glutamate
302 metabolism. Glu could be converted to α-ketoglutarate (α-KG) and further entered the
303 tricarboxylic acid cycle (TCA), which is the main way for tumor cells to meet the demands of
304 bioenergy, biosynthesis and redox balance in tumor microenvironment [36]. Meanwhile, α-KG
305 can produce Asp through a series of TCA catalytic reactions in which Asn could also be
306 produced by the catalysis of Asn synthase (ASNS). Therefore, Asp and Asn are also involved in
307 the energy metabolism of the tumor [5]. Likewise, Gly and its precursor Ser participate in TCA
308 through serving as exchange factor of Asn. Notably, ASNS is frequently upregulated in tumors
309 and is associated with poor prognosis of several types of cancer, such as prostate cancer, glioma
310 and neuroblastoma [37, 38]. On the other hand, L-asparaginase, an enzyme catalyzing the
311 degradation of Asn, has been widely applied in clinical treatment against multiple types of
312 leukemia and lymphoma [39]. These evidences suggested a strong dependence of cancer cells for
313 Glu metabolism and energy from TCA.
314
315 We noticed that Asp, as well as Orn and Cit, also participated in the Arg biosynthetic pathway.
316 This pathway is a part of the urea cycle, in which intermediates including Orn, Cit, Asp, arginine
317 succinate and Arg are produced by the successive catalysis of carbamyl phosphate synthase 1,
318 arginine succinate synthase 1 and arginine succinate lyase [5]. Although the complete urea cycle
319 works primarily in the liver, some enzymes could relink the intermediates outside the liver.
320 These substances support the survival and proliferation of tumor cells through strong interactions
321 with other metabolic pathways [40, 41]. Arg is the most consumed AA in the inner necrotic core
322 of tumor mass, indicating its high demand for the survival of tumor cells. Preclinical studies

323 have reported essential roles of Arg in cancer metabolism. For example, Arg starvation caused
324 global transcriptional suppression of metabolic genes including those involved in oxidative
325 phosphorylation, mitochondrial functions, and DNA repair in breast and prostate cancer cells [42,
326 43]. Meanwhile, Arg can import into cells, directly activate mTOR pathway, a nutrient-sensing
327 kinase strongly implicated in carcinogenesis. By mTOR mediation, Arg thus has profound
328 impacts on protein synthesis, lipid synthesis and nucleotide synthesis of cancer cells [44, 45]. In
329 addition, Arg can also bind G-protein coupled receptor to activate down-stream RAS/ERK or
330 PI3K pathway, another essential pathway in cancer development and progression [46]. Therefore,
331 as the building blocks of Arg biosynthetic pathway, Asp, Orn and Cit might be signaling
332 metabolites and act as transcriptome reprogrammer in cancer metabolism. Besides, as mentioned
333 above, CIT could activate the mTOR1 signaling pathway through modulating the concentration
334 of PA and Rheb activity in the LEL, which also indicating the role of these AAs in cancer
335 signaling transduction [12].
336
337 In addition, numerous tumors rely on the availability of extra-cellular Ser to proliferate rapidly
338 [47]. In the conversion process of Ser to Gly, a single-carbon unit could be generated to assist the
339 production of 5, 10-methylene tetrahydrofolate (CH2-THF), while the Gly can also be cleaved
340 by the mitochondrial Gly cleavage system to produce CH2-THF. CH2-THF could enter the
341 folate cycle for the synthesis of other substances that are involved in cancer metabolism, such as
342 other AAs, nucleotides and lipids [48-50]. Moreover, Ser metabolic pathway probably contribute
343 to NSCLC by serving as a carbon source of nucleotide synthesis and DNA methylation [51].
344 Interestingly, in addition to the de novo-produced Ser, the exogenous donors perhaps affect one-
345 carbon and tumor behavior as well. For instance, a study has found that dietary Ser and Gly
346 deficiency can inhibit the growth of colon cancer cells in mice [52]. Besides, Cys might be
347 involved in cancer metabolism through regulating ferroptosis [53]. Thus, Ser metabolic pathway,
348 as well as Gly and Cys metabolic pathways, may play essential roles in NSCLC metabolism.
349
350 The established AA panel in our study was effective in discriminating NSCLC patients from
351 controls, as well as classification of NSCLC patients with different stages and histological types.
352 In the study of Kim et al. [19] and Shingyoji et al. [18], six AAs (Ser, Gln, Ala, His, Orn, and
353 Lys) were used to establish an index for distinguishing lung cancer cases from healthy controls.
354 Besides the Ser and Orn, we included some other AAs as well, such as Phe, Trp, Ile, Asp, Asn,
355 etc. Although we included more AAs, the AUCs were similar to the results of previous studies.
356 This may be related to sample size, sample composition of clinical stages and pathological types,
357 nutritional status, etc. Notably, a Polish study applied both the Korean and self-built models, and
358 found that the latter has better distinguishing ability (Aal, His, Phe, Cit, Asp, Asn) [20]. Taken
359 together, AAs may have potential value for the classification of NSCLC. However, it may be
360 necessary to establish population-specific AA panels in different countries and regions.
361
362 To our knowledge, this is the first case-control study in Chinese population to investigate the
363 alteration of AA levels as well as the involved AA pathways in NSCLC patients, and to evaluate
364 the potential value of AAs in the classification of NSCLC. The enrolled cases and controls were
365 comparable in the basic characteristics. And the approach for determination of AA levels were

366 validated with widely approved standards. However, several limitations should be addressed.
367 Firstly, there are many types of AAs and their derivatives, however, only 22 common ones were
368 selected in our study. More kinds of AAs thus need to be determined and compared in the future.
369 Secondly, AA levels in the human body are affected by a variety of factors including nutritional
370 status, lifestyle factors such as diet, physical activity and other underlying diseases. Although we
371 have matched the cases and controls with gender and age, our results might be still influenced by
372 other potential confounding bias. Therefore, our results still need to be verified in future studies.
373

374 **Conclusions**
375 In conclusion, the current study found that the level of Trp, Phe, Ile, Gly, Ser, Asp, Asn, Cys,
376 Glu, Orn, and Cit were significantly altered in serum of NSCLC patients in comparison with
377 controls. The altered AAs were involved in three important metabolic pathways, including Gly,
378 Ser and Thr metabolism, Ala, Asp and Glu metabolism, and Arg biosynthesis. In addition, this
379 AA panel might pose potential in NSCLC classification.
380

381 **Acknowledgements**
382 We wish to thank Xia Luo for assistance in the lab.

References

[1]   SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA Cancer J Clin, 2021,

[2]   GOLDSTRAW P, BALL D, JETT J R, et al. Non-small-cell lung cancer [J]. The Lancet, 2011, 378(9804): 1727-40.

[3]   SIEGEL R L, MILLER K D, JEMAL A. Cancer statistics, 2020 [J]. CA Cancer J Clin, 2020, 70(1): 7-30.

[4]   CHEN W, ZHENG R, BAADE P D, et al. Cancer statistics in China, 2015 [J]. CA Cancer J Clin, 2016, 66(2): 115-32.

[5]   WEI Z, LIU X, CHENG C, et al. Metabolism of Amino Acids in Cancer [J]. Front Cell Dev Biol, 2020, 8(603837.

[6]   HOSIOS A M, HECHT V C, DANAI L V, et al. Amino Acids Rather than Glucose Account for the Majority of Cell Mass in Proliferating Mammalian Cells [J]. Dev Cell, 2016, 36(5): 540-9.

[7]   MA Z, VOSSELLER K. O-GlcNAc in cancer biology [J]. Amino acids, 2013, 45(4): 719-33.

[8]   MENG M, CHEN S, LAO T, et al. Nitrogen anabolism underlies the importance of glutaminolysis in proliferating cells [J]. Cell Cycle, 2010, 9(19): 3921-32.

[9]   ZETTERBERG A, ENGSTROM W. Glutamine and the regulation of DNA replication and cell multiplication in fibroblasts [J]. Journal of cellular physiology, 1981, 108(3): 365-73.

[10]   YUE M, JIANG J, GAO P, et al. Oncogenic MYC Activates a Feedforward Regulatory Loop Promoting Essential Amino Acid Metabolism and Tumorigenesis [J]. Cell reports, 2017, 21(13): 3819-32.

[11]   WU G. Amino acids: metabolism, functions, and nutrition [J]. Amino acids, 2009, 37(1): 1-17.

[12]   BOURGOIN-VOILLARD S, GORON A, SEVE M, et al. Regulation of the proteome by amino acids [J]. Proteomics, 2016, 16(5): 831-46.

[13]   MIYAGI Y, HIGASHIYAMA M, GOCHI A, et al. Plasma free amino acid profiling of five types of cancer patients and its application for early detection [J]. PloS one, 2011, 6(9): e24143.

[14]   NOGUCHI Y, ZHANG Q W, SUGIMOTO T, et al. Network analysis of plasma and tissue amino acids and the generation of an amino index for potential diagnostic use [J]. American Journal of Clinical Nutrition, 2006, 83(2): 513S-9S.

[15]   MIKAMI H, KIMURA O, YAMAMOTO H, et al. A multicentre clinical validation of AminoIndex Cancer Screening (AICS) [J]. Scientific reports, 2019, 9(1): 13831.

[16]   HIGASHIYAMA M, MIYAZAKI R, YAMAMOTO H, et al. Preoperative AminoIndex Cancer Screening (AICS) abnormalities predict postoperative recurrence in patients undergoing curative resection for non-small cell lung cancer [J]. BMC cancer, 2020, 20(1): 1100.

[17]   MAEDA J, HIGASHIYAMA M, IMAIZUMI A, et al. Possibility of multivariate function composed of plasma amino acid profiles as a novel screening index for non-small cell lung cancer: a case control study [J]. BMC cancer, 2010, 10(690.

[18]   SHINGYOJI M, IIZASA T, HIGASHIYAMA M, et al. The significance and robustness

426   of a plasma free amino acid (PFAA) profile-based multiplex function for detecting lung cancer
427   [J]. BMC cancer, 2013, 13(

428   [19]   KIM H J, JANG S H, RYU J S, et al. The performance of a novel amino acid multivariate
429   index for detecting lung cancer: A case control study in Korea [J]. Lung Cancer, 2015, 90(3):
430   522-7.

431   [20]   KLUPCZYNSKA A, DEREZINSKI P, DYSZKIEWICZ W, et al. Evaluation of serum
432   amino acid profiles' utility in non-small cell lung cancer detection in Polish population [J]. Lung
433   Cancer, 2016, 100(71-6.

434   [21]   CHOI B H, COLOFF J L. The Diverse Functions of Non-Essential Amino Acids in
435   Cancer [J]. Cancers (Basel), 2019, 11(5):

436   [22]   SIVASHANMUGAM M, J J, V U, et al. Ornithine and its role in metabolic diseases: An
437   appraisal [J]. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 2017,
438   86(185-94.

439   [23]   PAPADIA C, OSOWSKA S, CYNOBER L, et al. Citrulline in health and disease.
440   Review on human studies [J]. Clin Nutr, 2018, 37(6 Pt A): 1823-8.

441   [24]   GOLDSTRAW P, CHANSKY K, CROWLEY J, et al. The IASLC Lung Cancer Staging
442   Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth)
443   Edition of the TNM Classification for Lung Cancer [J]. J Thorac Oncol, 2016, 11(1): 39-51.

444   [25]   LI Z, ZHANG H. Reprogramming of glucose, fatty acid and amino acid metabolism for
445   cancer progression [J]. Cellular and molecular life sciences : CMLS, 2016, 73(2): 377-92.

446   [26]   VANDER HEIDEN M G, CANTLEY L C, THOMPSON C B. Understanding the
447   Warburg effect: the metabolic requirements of cell proliferation [J]. Science (New York, NY),
448   2009, 324(5930): 1029-33.

449   [27]   VETTORE L, WESTBROOK R L, TENNANT D A. New aspects of amino acid
450   metabolism in cancer [J]. Br J Cancer, 2020, 122(2): 150-6.

451   [28]   TABE Y, LORENZI P L, KONOPLEVA M. Amino acid metabolism in hematologic
452   malignancies and the era of targeted therapy [J]. Blood, 2019, 134(13): 1014-23.

453   [29]   PROENZA A M, OLIVER J, PALOU A, et al. Breast and lung cancer are associated
454   with a decrease in blood cell amino acid content [J]. The Journal of Nutritional Biochemistry,
455   2003, 14(3): 133-8.

456   [30]   SMITH J T. Recent advancements in amino acid analysis using capillary electrophoresis
457   [J]. Electrophoresis, 1999, 20(15-16): 3078-83.

458   [31]   SMON A, CUK V, BRECELJ J, et al. Comparison of liquid chromatography with
459   tandem mass spectrometry and ion-exchange chromatography by post-column ninhydrin
460   derivatization for amino acid monitoring [J]. Clinica chimica acta; international journal of
461   clinical chemistry, 2019, 495(446-50.

462   [32]   MO X M, LI Y, TANG A G, et al. Simultaneous determination of phenylalanine and
463   tyrosine in peripheral capillary blood by HPLC with ultraviolet detection [J]. Clinical
464   biochemistry, 2013, 46(12): 1074-8.

465   [33]   PHIPPS W S, CROSSLEY E, BORIACK R, et al. Quantitative amino acid analysis by
466   liquid chromatography-tandem mass spectrometry using low cost derivatization and an
467   automated liquid handler [J]. JIMD Rep, 2020, 51(1): 62-9.

468   [34]   LE A, NG A, KWAN T, et al. A rapid, sensitive method for quantitative analysis of

469 underivatized amino acids by liquid chromatography-tandem mass spectrometry (LC-MS/MS)
470 [J]. J Chromatogr B Analyt Technol Biomed Life Sci, 2014, 944(166-74.
471 [35]   ZHAO Q, CAO Y, WANG Y, et al. Plasma and tissue free amino acid profiles and their
472 concentration correlation in patients with lung cancer [J]. Asia Pac J Clin Nutr, 2014, 23(3): 429-
473 36.
474 [36]   TODISCO S, CONVERTINI P, IACOBAZZI V, et al. TCA Cycle Rewiring as Emerging
475 Metabolic Signature of Hepatocellular Carcinoma [J]. Cancers, 2020, 12(1):
476 [37]   SIRCAR K, HUANG H, HU L, et al. Integrative molecular profiling reveals asparagine
477 synthetase is a target in castration-resistant prostate cancer [J]. The American journal of
478 pathology, 2012, 180(3): 895-903.
479 [38]   ZHANG J, FAN J, VENNETI S, et al. Asparagine plays a critical role in regulating
480 cellular adaptation to glutamine depletion [J]. Molecular cell, 2014, 56(2): 205-18.
481 [39]   HASKELL C M, CANELLOS G P. l-Asparaginase resistance in human leukemia—
482 Asparagine synthetase [J]. Biochemical pharmacology, 1969, 18(10): 2578-80.
483 [40]   LEE J S, ADLER L, KARATHIA H, et al. Urea Cycle Dysregulation Generates
484 Clinically Relevant Genomic and Biochemical Signatures [J]. Cell, 2018, 174(6): 1559-+.
485 [41]   KESHET R, SZLOSAREK P, CARRACEDO A, et al. Rewiring urea cycle metabolism
486 in cancer to support anabolism [J]. Nature Reviews Cancer, 2018, 18(10): 634-45.
487 [42]   QIU F, CHEN Y R, LIU X, et al. Arginine starvation impairs mitochondrial respiratory
488 function in ASS1-deficient breast cancer cells [J]. Science signaling, 2014, 7(319): ra31.
489 [43]   CHANGOU C A, CHEN Y R, XING L, et al. Arginine starvation-associated atypical
490 cellular death involves mitochondrial dysfunction, nuclear DNA leakage, and chromatin
491 autophagy [J]. Proceedings of the National Academy of Sciences of the United States of America,
492 2014, 111(39): 14147-52.
493 [44]   NAOMOTO H B K S T Y M H T N M T T N M G N T Y. Arginine and Leucine regulate
494 p70 S6 kinase and 4E-BP1 in intestinal epithelial cells [J]. International journal of molecular
495 medicine, 2004, 13(4): 537-43.
496 [45]   SAXTON R A, SABATINI D M. mTOR Signaling in Growth, Metabolism, and Disease
497 [J]. Cell, 2017, 168(6): 960-76.
498 [46]   CHEN C L, HSU S C, ANN D K, et al. Arginine Signaling and Cancer Metabolism [J].
499 Cancers (Basel), 2021, 13(14):
500 [47]   JONES N P, SCHULZE A. Targeting cancer metabolism - aiming at a tumour's sweet-
501 spot [J]. Drug Discov Today, 2012, 17(5-6): 232-41.
502 [48]   TIBBETTS A S, APPLING D R. Compartmentalization of Mammalian folate-mediated
503 one-carbon metabolism [J]. Annu Rev Nutr, 2010, 30(57-81.
504 [49]   LOCASALE J W. Serine, glycine and one-carbon units: cancer metabolism in full circle
505 [J]. Nat Rev Cancer, 2013, 13(8): 572-83.
506 [50]   AMELIO I, CUTRUZZOLá F, ANTONOV A, et al. Serine and glycine metabolism in
507 cancer [J]. Trends in biochemical sciences, 2014, 39(4): 191-8.
508 [51]   PAVLOVA N N, THOMPSON C B. The Emerging Hallmarks of Cancer Metabolism [J].
509 Cell metabolism, 2016, 23(1): 27-47.
510 [52]   MADDOCKS O D, BERKERS C R, MASON S M, et al. Serine starvation induces stress
511 and p53-dependent metabolic remodelling in cancer cells [J]. Nature, 2013, 493(7433): 542-6.

512   [53]   POURSAITIDIS I, WANG X, CRIGHTON T, et al. Oncogene-Selective Sensitivity to
513   Synchronous Cell Death following Modulation of the Amino Acid Nutrient Cystine [J]. Cell
514   reports, 2017, 18(11): 2547-56.
515

Demographics and clinical characteristics

		Cases N=200	Controls N=202	*P*-value
Gender				0.568
	Males	138 (69.0%)	134 (66.3%)	
	Females	62 (31.0%)	68 (33.7%)	
Age		58 (34-81)	57 (45-80)	0.128
Smoking [a]				0.028
	Yes	102 (51.0%)	81 (40.1%)	
	No	98 (49.0%)	121 (59.9%)	
Drinking [a]				0.125
	Yes	57 (28.5%)	72 (35.6%)	
	No	143 (71.5%)	130 (64.4%)	
BMI [a]				0.002
	<18.5	22 (11.0%)	6 (3.0%)	
	18.5-24.0	106 (53.0%)	98 (48.5%)	
	24.0-28.0	63 (31.5%)	80 (39.6%)	
	≥28.0	9 (4.5%)	18 (8.9%)	
Histologic type of cancer				
	Adenocarcinoma	133 (66.5%)		
	SCC	57 (28.5%)		
	Others	10 (5.0%)		
Tumor stage				
	0	4 (2.0%)		
	I	67 (33.5%)		
	II	14 (7.0%)		
	III	52 (26.0%)		
	IV	53 (26.5%)		
	Unknown	10 (5.0%)		

1   [a] Data about 2 control subjects was absent.

2   Chi-square test was used to compare categorical variables. Chi-square test: (1) gender, $\chi^2=0.326$, df=1, (2) smoking,

3   $\chi^2=4.816$, df=1, (3) drinking, $\chi^2=2.353$, df=1, (4) BMI, $\chi^2=14.468$, df=3.

4   T-test (normal distribution) was employed to compare continuous variables. T-test: (1) age, F=33.772, df=400,

5   95%CI (-0.289-2.289).

Alteration in serum amino acid concentrations in NSCLC patients.

Amino acid	Case N=200		Control N=202		Fold Change	*P*-value	*P*-FDR
	Median	IQR	Median	IQR			
Phenylalanine [a]	104.00	45.10	96.90	35.70	1.07	0.002	**0.006**
Tryptophan [a]	51.15	25.10	57.70	27.50	0.89	0.005	**0.012**
Isoleucine [a]	87.30	39.20	81.85	32.60	1.07	0.016	**0.032**
Leucine [a]	171.50	62.00	161.00	62.30	1.07	0.033	0.061
Methionine [a]	28.30	11.80	29.65	12.20	0.95	0.755	0.755
Lysine [a]	227.50	117.30	222.00	113.80	1.02	0.390	0.477
Valine [a]	292.50	112.30	275.00	94.30	1.06	0.131	0.206
Threonine [a]	152.00	74.80	148.00	63.00	1.03	0.268	0.369
Glycine [b]	366.00	175.30	326.00	138.30	1.12	0.001	**0.004**
Serine [b]	191.00	80.80	171.00	67.80	1.12	0.001	**0.004**
Asparagine [b]	63.05	25.30	57.70	24.70	1.09	0.002	**0.006**
Aspartic acid [b]	46.75	31.80	32.55	16.90	**1.44**	<0.001	**<0.001**
Cysteine [b]	68.05	49.10	53.75	44.00	**1.27**	<0.001	**<0.001**
Glutamic acid [b]	126.00	82.80	104.50	63.40	**1.21**	<0.001	**<0.001**
Glutamine [b]	524.00	221.80	504.50	238.50	1.04	0.076	0.129
Histidine [b]	62.45	33.50	60.05	36.70	1.04	0.559	0.647
Proline [b]	218.50	101.50	213.50	82.50	1.02	0.286	0.37
Arginine [b]	112.00	71.40	107.50	63.80	1.04	0.150	0.22
Tyrosine [b]	75.90	51.10	72.65	50.00	1.04	0.733	0.755
Alanine [b]	471.50	227.30	474.00	169.80	0.99	0.604	0.664
Ornithine [b]	113.00	79.30	92.50	59.10	**1.22**	<0.001	**<0.001**
Citrulline [b]	30.55	18.90	26.75	15.90	1.14	0.016	**0.032**

1   [a] essential amino acids.

2   [b] not essential amino acids.

3   Mann-Whitney U test was used for continuous variables that did not normally distributed.

4   Benjamini-Hochberg procedure was used for multiple testing correction.

Subgroup analysis of amino acid concentrations in patients based on histological types.

Amino acid	Adenocarcinoma N=133					SCC N=57					Fold Change [e]	P-value [e]	P-FDR [e]
	Median	IQR	Fold Change [c]	P-value [c]	P-FDR [c]	Median	IQR	Fold Change [d]	P-value [d]	P-FDR [d]			
Phenylalanine [a]	99.60	34.80	1.03	0.212	0.301	114.00	59.70	1.18	<0.001	**<0.001**	0.87	0.008	0.088
Tryptophan [a]	51.70	25.00	0.90	0.006	**0.019**	48.10	25.80	0.83	0.087	0.216	1.07	0.925	0.969
Isoleucine [a]	86.50	35.00	1.06	0.049	0.098	87.00	52.20	1.06	0.113	0.216	0.99	0.789	0.882
Leucine [a]	170.00	50.00	1.06	0.180	0.292	175.00	108.00	1.09	0.038	0.120	0.97	0.245	0.797
Methionine [a]	28.90	11.00	0.97	0.856	0.856	26.70	13.10	0.90	0.460	0.633	1.08	0.316	0.797
Lysine [a]	234.00	106.00	1.05	0.219	0.301	210.00	158.50	0.95	0.879	0.938	1.11	0.519	0.882
Valine [a]	281.00	95.50	1.02	0.285	0.369	295.00	149.00	1.07	0.250	0.394	0.95	0.759	0.882
Threonine [a]	154.00	68.00	1.04	0.065	0.119	143.00	86.50	0.97	0.553	0.715	1.08	0.103	0.567
Glycine [b]	378.00	162.50	1.16	0.001	**0.006**	349.00	224.50	1.07	0.113	0.216	1.08	0.582	0.882
Serine [b]	191.00	68.00	1.12	0.007	**0.019**	190.00	88.50	1.11	0.023	0.100	1.01	0.719	0.882
Asparagine [b]	63.90	24.40	1.11	0.005	**0.018**	58.80	35.00	1.02	0.115	0.216	1.09	0.802	0.882
Aspartic acid [b]	39.90	26.70	**1.23**	<0.001	**<0.001**	57.50	32.90	**1.77**	<0.001	**<0.001**	**0.69**	<0.001	**<0.001**
Cysteine [b]	67.40	47.80	**1.25**	<0.001	**0.009**	76.10	47.80	**1.42**	0.001	**0.006**	0.89	0.280	0.797
Glutamic acid [b]	122.00	77.50	1.17	<0.001	**<0.001**	135.00	93.90	**1.29**	<0.001	**<0.001**	0.90	0.689	0.882
Glutamine [b]	528.00	213.50	1.05	0.011	**0.027**	488.00	287.50	0.97	0.733	0.896	1.08	0.042	0.308
Histidine [b]	65.90	31.60	1.10	0.186	0.292	57.80	45.60	0.96	0.881	0.938	1.14	0.311	0.797
Proline [b]	217.00	97.00	1.02	0.493	0.542	223.00	124.00	1.04	0.330	0.483	0.97	0.675	0.882
Arginine [b]	111.00	72.80	1.03	0.492	0.542	114.00	85.70	1.06	0.118	0.216	0.97	0.326	0.797
Tyrosine [b]	75.80	49.70	1.04	0.612	0.641	76.00	57.90	1.05	0.896	0.938	1.00	0.635	0.882

Alanine [b]	477.00	201.50	1.01	0.411	0.502	433.00	299.00	0.91	0.954	0.954	1.10	0.593	0.882
Ornithine [b]	116.00	76.80	**1.25**	<0.001	**<0.001**	110.00	90.20	1.19	0.031	0.112	1.05	0.611	0.882
Citrulline [b]	30.20	15.60	1.13	0.017	**0.037**	32.00	27.30	1.20	0.186	0.315	0.94	0.991	0.991

1 [a] essential amino acids.

2 [b] not essential amino acids.

3 [c] compared between adenocarcinoma and control.

4 [d] compared between squamous cell carcinoma and control.

5 [e] compared between adenocarcinoma and squamous cell carcinoma.

6 Mann-Whitney U test was used for continuous variables that did not normally distributed.

7 Benjamini-Hochberg procedure was used for multiple testing correction.

Figure 1

Subgroup analysis of amino acid concentrations in patients based on clinical stages.

(A-K) The concentrations of Phe (**A**), Trp (**B**), Ile (**C**), Gly (**D**), Ser (**E**), Asn (**F**), Asp (**G**), Cys (**H**), Glu (**I**), Arg (**J**), Orn (**K**) in patients with different stages. Mann-Whitney U test was used for continuous variables that did not normally distributed. Benjamini-Hochberg procedure was used for multiple testing correction. See the supplementary document 5 for the statistical result report.
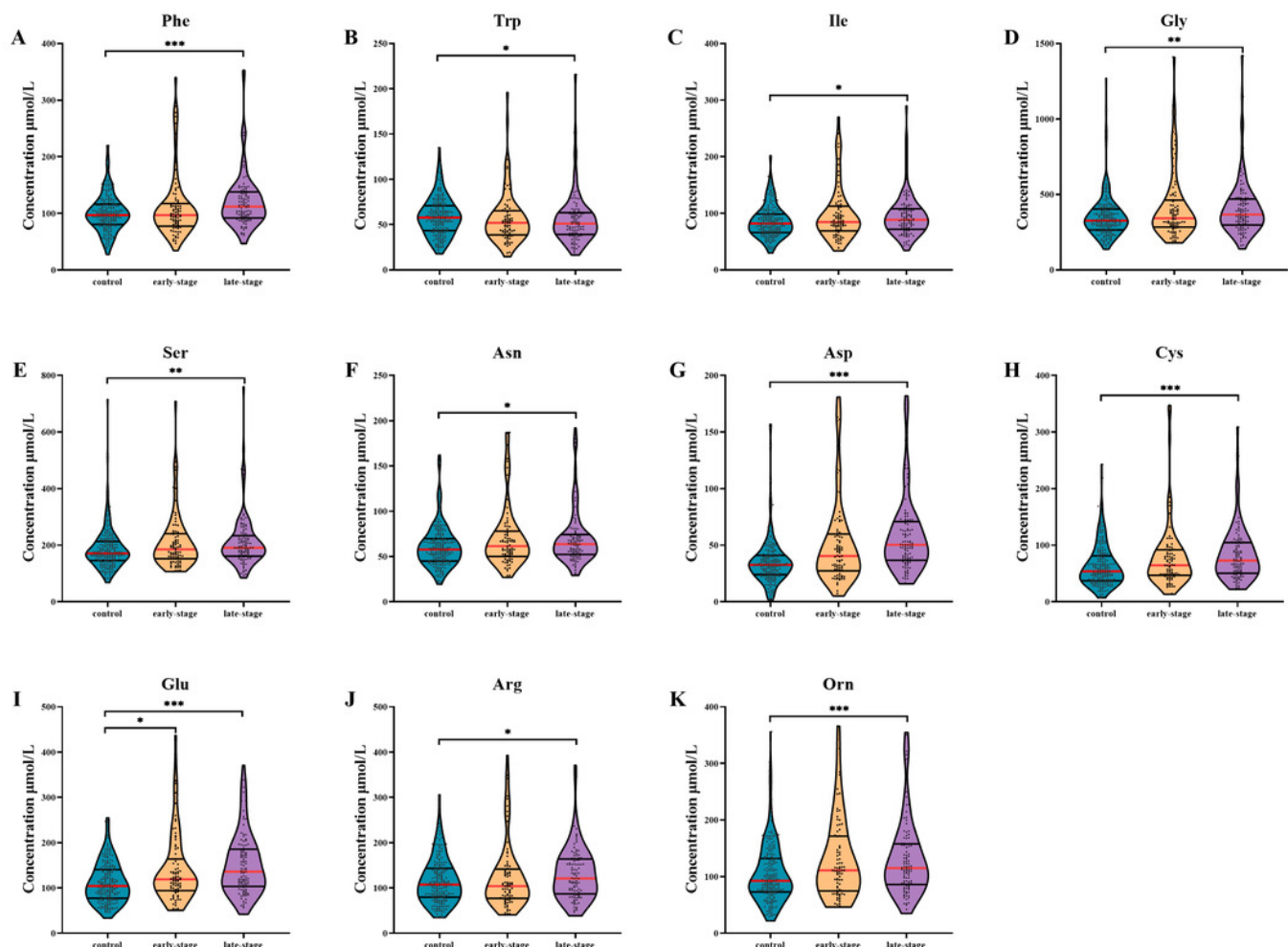
Figure 2

Analysis of amino acid metabolic pathways in NSCLC patients.

The MetaboAnalyst 5.0 was used for metabolic pathway analysis. See the supplementary document 4 for the statistical result report.
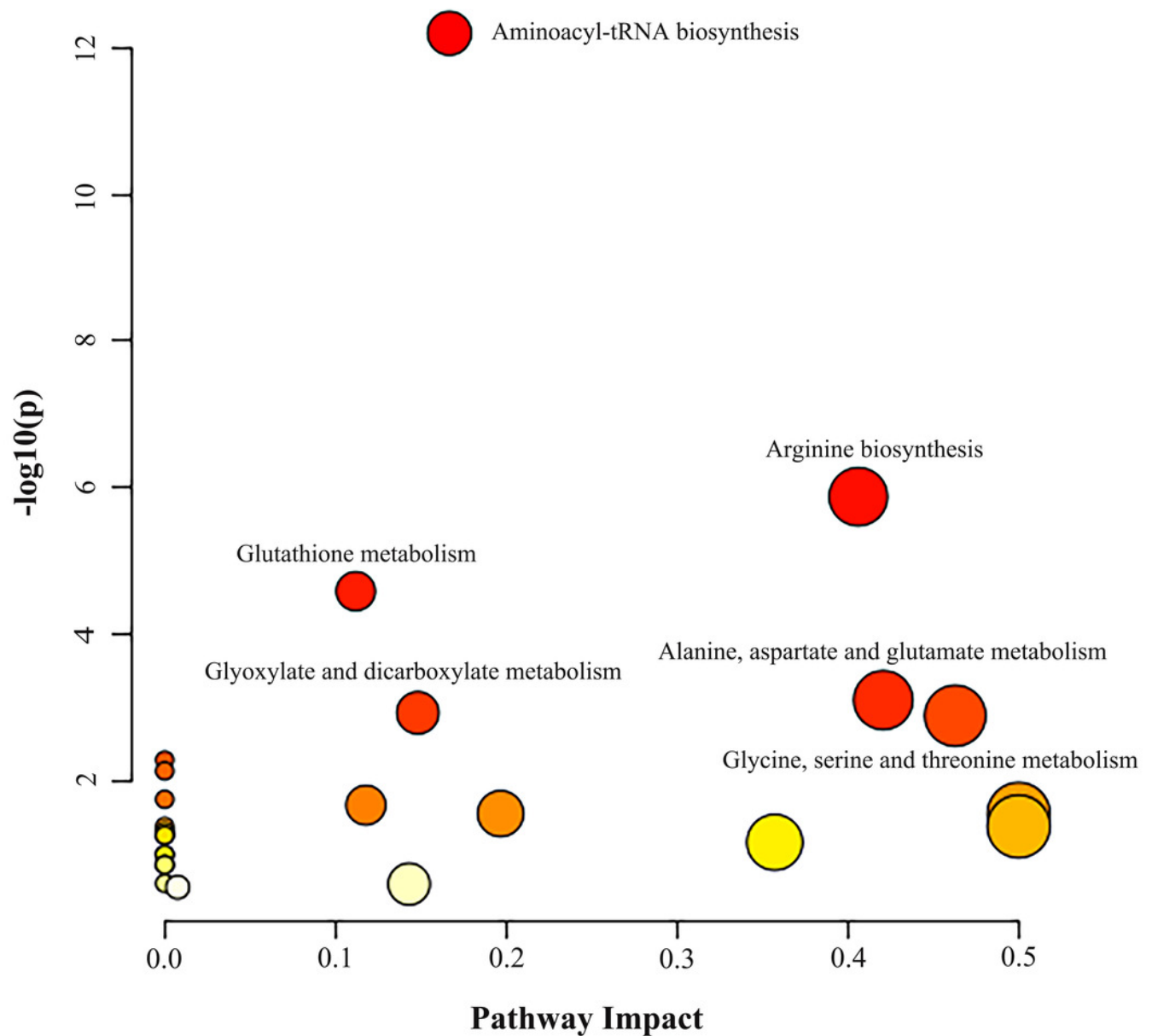
Figure 3

The classification of NSCLC patients and control groups.

The amino acid panel including 11 differential amino acids (Phe, Trp, Ile, Gly, Ser, Asn, Asp, Cys, Glu, Orn, Cit). **(A)** The ROC curve for the training data set and the validation data set. The training data set: AUC=0.80, 95%CI (0.74-0.85), SE=0.03, *P*<0.001 . The validation data set: AUC=0.79, 95%CI=0.71-0.87, SE=0.04, *P*<0.001. **(B)** The ROC curve for NSCLC patients with different stages and histological types. Early-stage NSCLC: AUC=0.75, 95%CI (0.68-0.81), SE=0.03, *P*<0.001. Late-stage NSCLC: AUC=0.86, 95%CI (0.82-0.91), SE=0.02, *P*<0.001. Adenocarcinoma: AUC=0.77, 95%CI (0.71-0.82), SE=0.03, *P*<0.001. SCC: AUC=0.90, 95%CI (0.85-0.95), SE=0.03, *P*<0.001.