Serum Metabolic Disturbances in Lung Cancer Investigated through an Elaborative NMR-Based Serum Metabolomics Approach

Anjana Singh, Ved Prakash, Nikhil Gupta,* Ashish Kumar, Ravi Kant, and Dinesh Kumar*

ABSTRACT: Detection of metabolic disturbances in lung cancer (LC) has the potential to aid early diagnosis/prognosis and hence improve disease management strategies through reliable grading, staging, and determination of neoadjuvant status in LC. However, a majority of previous metabolomics studies compare the normalized spectral features which not only provide ambiguous information but further limit the clinical translation of this information. Various such issues can be resolved by performing the concentration profiling of various metabolites with respect to formate as an internal reference using commercial software Chenomx. Continuing our efforts in this direction, the serum metabolic profiles were measured on 39 LC patients and 42 normal controls (NCs, comparable in age/sex) using high-field 800 MHz NMR spectroscopy and compared using multivariate statistical analysis tools to identify metabolic disturbances and metabolites of diagnostic potential. Partial least-squares discriminant analysis (PLS-DA) model revealed a distinct separation between LC and NC groups and resulted in excellent discriminatory ability with the area under the receiver-operating characteristic (AUROC) = 0.97 [95% CI = 0.89–1.00]. The metabolic features contributing to the differentiation of LC from NC samples were identified first using variable importance in projection (VIP) score analysis and then checked for their statistical significance (with p-value < 0.05) and diagnostic potential using the ROC curve analysis. The analysis revealed relevant metabolic disturbances associated with LC. Among various circulatory metabolites, six metabolites, including histidine, glutamine, glycine, threonine, alanine, and valine, were found to be of apposite diagnostic potential for clinical implications. These metabolic alterations indicated altered glucose metabolism, aberrant fatty acid synthesis, and augmented utilization of various amino acids including active glutaminolysis in LC.

INTRODUCTION

Lung cancer (LC) is the foremost contributor to cancer-related deaths globally, and the number is increasing every year due to several factors including poor air-quality index. LC is also the most lethal cancer type, and for about 15% of LC diagnosed cases, the averaged survival time is no longer than five years.1,2 The primary contributing factor to the lethality of LC is lack of reliable noninvasive clinical markers, which can be used to distinguish LC from other lung diseases after a symptomatic stage and predict its prognosis.3 Therefore, a majority of patients are diagnosed at advanced stages following several clinical procedures or diagnostic tests including the painful procedure of bronchoscopy/lung biopsy.3 However, clinical intervention is unlikely to succeed if LC is diagnosed at an advanced stage, and therefore, there is paucity and need for other noninvasive diagnostic markers for LC to facilitate timely diagnosis and staging and for guiding treatment decisions.3,4

Starting our efforts in this direction, we explored here the metabolomics approach to identify reliable serum metabolic signatures that can help to improve the early diagnostic and prognostic screening of LC. Metabolomics studies on different cancer types revealed that tumor progression often led to aberrant metabolic changes in biofluids including blood serum.3,5,7−16 Analysis of the NMR spectra of biological fluids (such as urine, serum, bile, cerebrospinal fluid, etc.) using multivariate statistical analysis tools is an emerging metabolomics approach for the identification of the diagnostic panel of metabolic biomarkers for clinical surveillance.17 The serum-based metabolomics studies carried out in the recent past well demonstrate that the circulatory levels of serum metabolites can differentiate LC types and can predict the severity; thus, the approach has exquisite potential to aid early diagnosis of LC and its clinical management.3,5,15,16,18−27 Most of these metabolomics studies have been carried out in European, Chinese, Japanese, and Korean populations. Though the prevalence of LC is progressively increasing in India,28 however, there is scarcity of such metabolomics studies performed on Indian population. The present study, therefore,
aims to perform the serum metabolomics analysis in north Indian patients using NMR which offers several technical advantages like reproducibility and minimal sample handling. Clinically, LC is broadly classified into two subtypes: small-cell LC (SCLC) and non-small-cell LC (NSCLC); the ratio of NSCLC to SCLC is about 5:1. Among NSCLC, the most important histological subtypes include adenocarcinoma (ADC), squamous cell lung carcinoma (SqCC), and large-cell carcinoma, of which ADC and SqCC represent ∼90% of all cases. In the present study, we attempted to demonstrate the utility of NMR-based metabolomics analysis in LC detection and further to identify the distinctive metabolic patterns of clinical subtypes of LC.

#### RESULTS AND DISCUSSION

**Patient Characteristics.** The study involved 39 LC patients (36 male and 3 female) with median age at presentation 54 years, and the median disease duration was 3 years. The histological classification and other clinical and demographic characteristics of patients are tabulated in Table 1. As evident, there are 5 SCLC patients and 34 NSCLC patients, and among 34 NSCLC patients, there are 14 ADC patients and 20 SqCC patients. For a comparative analysis, the study involved about 42 normal healthy subjects comparable in age and sex.

**Serum Metabolic Disturbances Associated with LC.** Figure 1 compares the cumulative 1D 1H Carr–Purcell–Meiboom–Gill (CPMG) NMR spectra of serum samples obtained from LC patients and normal control (NC) subjects. The NMR peaks in the spectra are annotated as per the chemical shift assignments of metabolites tabulated in the Supporting Information (Table S1). Visual inspection revealed possible metabolic differences in the most abundant metabolites such as lipid/lipoproteins, glucose, and lactate. However, to confirm these changes and evaluate other subtle metabolic differences, we further performed the multivariate statistical analysis.

First, the normalized spectral features were compared, and the score plot derived from the partial least-squares discriminant analysis (PLS-DA) analysis is shown in Figure 2A. An exquisite separation between LC and NC samples and clustering of samples in each study group were clearly evident from the score plot, suggesting that the two study groups are distinctively different in terms of their serum metabolic profiles. This was further corroborated by significantly higher ($R^2 > 0.9$; $Q^2 > 0.7$) validation parameters of the PLS-DA model (Figure 2B), suggesting the good predictive power of discriminatory models. Next, the serum metabolic features of the discriminatory potential were identified based on the indexing of variable importance in projection (VIP) scores derived from the PLS-DA analysis (Figure 2C). The VIP score value $> 1.0$ was considered as the criterion for significant contribution to the discriminatory model. Compared to NC, the sera of LC patients were characterized by (a) elevated levels of glucose and N-acetyl-glycoproteins and (b) decreased levels of alanine, lactate, myo-inositol, and diacyl glycoprotein; NAAL: N-acetyl-$\alpha$-amyl acetoyl glycoprotein; NAG: N-acetyl-glycoprotein; NAAL: N-alpha acetyl lysine.

#### Table 1. Clinical and Demographic Characteristics of Subjects Included in the Study

| parameter                  | LC     | NC     |
|----------------------------|--------|--------|
| total number (M/F)         | 39 (36:3) | 42 (34:8) |
| age (range)                | 54.13 ± 10.96 | 48.98 ± 7.27 |
| age of male subjects       | 53.42 ± 1.852 | 49.08 ± 1.303 |
| age of female subjects     | 62.67 ± 0.88 | 48.33 ± 0.88 |
| disease duration (months)  | 18 (12–24) | 19 (12–24) |
| smoking                    | 24 (61.5%) | 26 (61.9%) |

| clinical subtypes          | number (% age) | tumor location | smoking |
|----------------------------|----------------|----------------|---------|
| small cell lung carcinoma  | 5 (13%)        | peripheral = 5 | 3       |
| ADC                        | 14 (36%)       | central = 10   | 8       |
| SqCC                       | 20 (51%)       | central = 12   | 13      |

Figure 1. Stacking of cumulative 1D 1H NMR (CPMG) spectra recorded on serum samples of LC patients (red) and NC subjects (blue): the spectra in panel A represent chemical shift range $\delta$(0.55–4.65) ppm, whereas the spectra in panel B represent that from $\delta$(5.8–8.5) ppm with 8 times magnification compared to spectral region $\delta$(0.55–4.65) ppm for the purpose of clarity. Abbreviations used: His: histidine; 3HB: 3-hydroxybutyrate; LDL/VLDL: low-/very-low-density lipoprotein; TMAO: trimethyl-$\alpha$-oxide; NAG: N-acetyl glycoprotein; NAAL: N-alpha acetyl lysine.
in the sera of LC patients. However, the VIP score plot in our study showed a higher degree of metabolic redundancy, and it is clearly evident that the resonances of lipid/membrane metabolites, glucose, and lactate were dominating the VIP score plot (Figure 2C).

In order to avoid this redundancy and to screen other circulatory metabolites for their discriminatory relevance, the NMR spectra of LC and NC subjects were further analyzed, and the concentrations of 25 serum metabolites were estimated explicitly using Chenomx NMR suite program. These metabolites are 3-hydroxybutyrate, lactate, acetate, citrate, acetoacetate, pyruvate, succinate, alanine, betaine, creatine, creatinine, glutamine, glutamate, glucose, glycine, leucine, isoleucine, valine, proline, phenylalanine, threonine, tyrosine, histidine, trimethylamine-N-oxide (TMAO), and myo-inositol.

As described previously, these concentrations were further used to estimate other four relevant metabolic ratios such as glutamate to glutamine ratio (EQR), phenylalanine to tyrosine ratio (PTR), branched-chain amino acid to tyrosine ratio [BTR also referred as Fischer ratio; estimated as (leucine + isoleucine + valine)/tyrosine], and lactate-to-pyruvate ratio (LPR). The estimated mean and median values of these 29 serum metabolic features are tabulated in the Supporting Information (Table S2).

Next, these 29 metabolic concentration profiles were compared between the study groups using principal component analysis (PCA) and PLS-DA with orthogonal signal correction (OPLS-DA). Although PCA score plots revealed poor discrimination among the two groups (results not shown), the sample clustering and group separation among the two groups was clearly evident in the 2D score plots of OPLS-DA (with model validation parameters R²X = 0.729, R²Y = 0.648, and Q² = 0.2; Figure 3A). The results based on the OPLS-DA analysis were nearly similar to those based on the PLS-DA analysis (see Supporting Information, Figure S1). Further, the VIP score plot analysis was performed to identify the metabolites of discriminatory potential (Figure 3C, Supporting Information, Table S3). The VIP scores provide indexing of the metabolic features contributing to the discriminatory model. For normalized spectral features (more than 100 in number), the cutoff value is often selected is 1.0. However, for a limited number of variables (e.g., 29 in our present study), the VIP score cutoff value can be decided for legitimate selection of discriminatory features as described previously. For example, a cutoff value of 0.9 resulted in 15 metabolic features of discriminatory potential. All 29 circulatory profiles were further evaluated for their diagnostic potential and resulted in excellent discriminatory ability (Figure 3D,E). Among various multivariate receiver operating characteristic (ROC) curves generated, the curve based on top 10 discriminatory metabolites (selected based on the highest VIP scores in the PLS-DA model) showed exquisite diagnostic potential with the area under the ROC curve (AUROC) value = 0.97 [95% CI = 0.93–1.00] (Figure 3D) which was as good as the cumulative ROC curve generated based on all 29 circulatory profiles (AUROC value = 0.97) [95% CI = 0.89–1.00] (Figure 3E), suggesting that these serum-based metabolic profiles estimated in this study have good diagnostic potential as well.

Compared to NC, the LC patients showed significant alterations for 17 serum metabolites. The sera of LC patients were characterized by decreased serum levels of citrate, betaine, creatinine, and most of the amino acids (including valine, leucine, isoleucine, glycine, alanine, glutamine, proline, threonine, tyrosine, and histidine), whereas the circulatory levels of pyruvate, acetoacetate, and 3-hydroxybutyrate were decreased.
significantly increased. In this study, the metabolic changes observed were found well-consistent with previous serum-based studies, though some disagreements were observed as well, as evident from Table 2. The observed discrepancies may be attributed to the fact that the present study for the first time involved concentration profiling using formate as an internal reference. The example to be discussed here is that of circulatory glucose which, in principle, decreases in cancer patients due to Warburg effect (i.e., increased aerobic glycolysis in cancer). However, majority of metabolomics studies in the literature have failed to show significant difference in glucose levels. In our study, the comparison based on normalized spectral features clearly revealed that the circulatory glucose levels are elevated in LC patients, suggesting that analysis needs to be rectified and cross-checked. On the other hand, the circulatory glucose levels estimated with respect to formate (as an internal reference) clearly showed that the circulatory glucose levels are almost comparable between cancer patients and normal healthy controls. According to a recent metabolomics study, the circulating formate levels also decrease in LC patients relative to healthy control subjects; therefore possibly, the net decrease in glucose levels is counterbalanced by its normalization with respect to formate. Nevertheless, from the metabolite concentrations estimated for other circulatory metabolites including various amino acids and organic acids, a remarkable pattern of metabolic alterations has been found and demonstrated to be well consistent with various previous metabolomics studies. In simple words, the metabolite concentrations reported in this study can also be considered as ratiometric metabolic profiles.

The present study also aimed to evaluate the diagnostic utility of serum metabolic profiles estimated by NMR for differentiating LC patients from NC subjects. For this, the ROC curves were generated for serum metabolites of discriminatory relevance between LC and NC groups, and the AUROC analysis was performed to test their diagnostic ability. Setting the criteria for diagnostic potential as AUROC value more than 0.8 and p-value less than 0.001, six key metabolic entities (histidine, glutamine, glycine, threonine, alanine and valine) were selected as diagnostic markers of LC (Figure 4, Table 2). Among the selected metabolic ratios, the...
AUROCs for LPR were found to be greater than 0.8 (0.82, down) suggesting its diagnostic potential in LC as well, whereas the AUROCs for PTR (0.80, up) and EQR (0.7, up) were found to be in the moderate range and that for BTR (0.51, down) is of no diagnostic potential (see Supporting Information, Table S4).

Compared to NC, the decreased serum levels of various amino acids in LC might be related to their augmented utilization in LC to regulate various biological functions. Consistent with previous reports, the significantly decreased serum levels of glutamine might be related to activated glutaminolysis in LC patients to replenish the energy demand required for regulating complex immune-mediated inflammatory responses. Well-consistent with previous reports, the elevated serum levels of 3HB and acetoacetate (the end products of lipid-metabolism) in LC patients were indicative of aberrant lipid metabolism (or active fatty acid synthesis) in cancer including LC. Particularly, the important metabolite of histidine has its strong antioxidative and anti-inflammatory effects. Further, it is a precursor for histamine which serves as a key mediator for many pathological responses including immune-mediated chronic/acute inflammatory and hypersensitivity responses. A recent study from our lab has also demonstrated that the circulatory histidine levels significantly decrease in Takayasu arteritis patients with active disease (i.e., immune-mediated active inflammation). Therefore, the decreased serum levels of histidine might be related to its augmented utilization under conditions of elevated oxidative stress and inflammation (a common clinical manifestation of LC). The pathophysiological states, that is, oxidative stress and inflammation, were further indicated by the elevated blood levels of formate (ref 18 24, 38), tyrosine, proline, valine, histidine, and alanine (ref 18 23 24, 38, ref 3).

Table 2. Diagnostic Potential of Serum Metabolites Evaluated Based on AUROCs for Differentiating LC from HC

| metabolite | AUROC | p-value | fold change | relative change | consistent | not consistent |
|------------|-------|---------|-------------|----------------|------------|---------------|
| histidine  | 0.923 | 0.000   | -1.261      | †***           | refs 18 23 24, |
| glutamine  | 0.880 | 0.000   | -0.956      | †***           | refs 18 23 24, |
| glycine    | 0.869 | 0.000   | -1.062      | †***           | ref 18      |
| threonine  | 0.868 | 0.000   | -1.028      | †***           | refs 18 38,  |
| alanine    | 0.833 | 0.000   | -0.875      | †***           | refs 23 24,  |
| valine     | 0.826 | 0.000   | -0.812      | †***           | ref 24      |
| citrate    | 0.788 | 0.000   | -0.977      | †***           | ref 18      |
| tyrosine   | 0.774 | 0.000   | -0.714      | †***           | ref 24      |
| proline    | 0.771 | 0.000   | -0.767      | †***           | ref 23      |
| leucine    | 0.763 | 0.000   | -0.679      | †***           | ref 18      |
| isoleucine | 0.721 | 0.000   | -0.525      | †***           | ref 18      |
| pyruvate   | 0.695 | 0.002   | 0.710       | †***           | refs 18 24 38, |
| 3-hydroxybutyrate | 0.694 | 0.003 | 0.658       | †***           | ref 3       |
| betaine    | 0.694 | 0.004   | -0.507      | †**            |            |
| succinate  | 0.683 | 0.492   | -0.482      | †              |            |
| creatinine | 0.667 | 0.018   | -0.444      | †              | ref 18      |
| lactate    | 0.660 | 0.006   | -0.417      | †***           | refs 18 24, |
| creatine   | 0.645 | 0.186   | -0.340      | †              | ref 18      |
| acetooacetate | 0.628 | 0.011 | 0.440       | †              | ref 18      |
| myo-inositol | 0.616 | 0.529   | -0.219      | †              |            |
| phenylalanine | 0.606 | 0.109  | -0.163      | †              | ref 18 23,  |
| TMAO       | 0.586 | 0.318   | -0.234      | †              | ref 18      |
| acetate    | 0.579 | 0.952   | -0.193      | †              | ref 18      |
| glutamate  | 0.529 | 0.575   | -0.116      | †              | ref 3       |
| glucose    | 0.505 | 0.714   | -0.053      | †              | ref 18      |
| formate    | 0.551 | 0.714   | -0.053      | †              | ref 18 27,  |

“Abbreviations used: TMAO: tri-methylamine-N-oxide; note: the discriminatory analysis based on normalized spectral features revealed no change in formate. This was also well consistent with a previous NMR-based plasma metabolomics study performed at 900 MHz field strength and formed the basis for us to use it as an internal reference compound for performing concentration profiling in the software program Chenomx. The study showing decreased circulatory levels of formate in LC was based on gas chromatography–mass spectrometry.”

Differential Metabolic Signatures of ADC and SqCC. The sample size of 14 for ADC patients and 20 for SqCC allowed us to further compare the serum metabolic profiles between ADC and SqCC patients. The 3D score scatter plot derived from the PLS-DA analysis based on the concentration profiles of 29 metabolic entities listed in Table S2 is shown in Supporting Information (Figure S2A). A trend for clustering of ADC and SqCC samples and separation between these clustered samples was clearly evident from the score plot (shown in Supporting Information, Figure S2A), suggesting that serum metabolic profiles between two groups are distinctly different. However, the lower cross-validation (CV) parameters (>0.2) (see Supporting Information, Figure S2B) revealed that the generated PLS-DA model does not exhibit good discrimination and predictive ability. The poor performance of the discriminatory model partly may be attributed to the low sample size. The VIP score plot in
combination with the student t-test was used to identify the metabolic profiles of discriminatory and statistical significance (see Supporting Information, Figure S2C,D). Further, we performed the ROC curve analysis, and key metabolic entities (glycine, proline, creatinine, phenylalanine, myo-inositol, and glutamine) were selected as biomarkers of diagnostic potential for discrimination between ADC and SqCC groups (Figure 6).

We also observed that the circulatory metabolic ratios showed no statistical significant difference between ADC and SqCC groups, and the AUROC values were also less than 0.6 for each of these metabolic entities (see Supporting Information, Figure S3 and Table S5), suggesting their similar metabolic response in both the clinical subtypes of LC. To be mentioned here is that the results based on discriminatory analysis between ADC and SqCC are very preliminary, and future studies on large cohort of patients are required to validate these findings. As a reference for future studies, the results of PLS-DA-based discriminatory analysis performed between ADC and SqCC groups with respect to NC are summarized in Figure 7.

■ CONCLUDING REMARKS

In conclusion, the present study on Indian LC patients further supported that NMR-based serum metabolomics analysis has exquisite potential to provide metabolic markers for discriminating LC and NC subjects and has the ability to distinguish clinical subtypes of LC as well. The metabolic disturbances observed in LC patients were suggestive of augmented utilization of amino acids to support anabolic metabolism and cancer-induced inflammation and also the elevated oxidative stress, activated glutaminolysis, and altered energy metabolism (Figure 5). However, for translating these findings into clinical procedure, further studies are required on a large patient sample size, especially in each of the clinical subtype cohort. Additionally, procedural optimization will be required to improve the accuracy of the NMR-based tests. The limitations of the study are the following: (i) the sample size of clinical subtypes of LC is legitimately low and (ii) the

![Figure 4.](https://doi.org/10.1021/acsomega.1c06941) Top 9 putative metabolic biomarkers selected after the ROC curve analysis was performed with all 29 serum metabolic entities tabulated in Table 2. The ROC curve plots shown here are the diagnostic potential of these metabolic entities between LC and NC groups as evident from the AUROC value, and the computed 95% CI is in the faint blue background. The metabolic differences are further evident from the box-cum-whisker plots shown on the right side of each ROC curve plot.

![Figure 5.](https://doi.org/10.1021/acsomega.1c06941) Summary of key metabolic changes and their association with the underlying disease pathophysiology.
impact of staging on the serum metabolic profile is lacking. Nevertheless, the alternative strategy described in this study (i.e., concentration profiling of serum metabolites) and the consistency of the findings with the majority of previous reports suggest that the approach will open a new avenue for scientists involved in NMR-based serum metabolomics studies for encompassing the benefits of recent software tools like Chenomx used here for concentration profiling of circulatory metabolites in the NMR spectra of serum samples.

■ MATERIALS AND METHODS

Ethical Approval. The study was approved by the Human ethics Committee of King George’s Medical University (KGMU), Lucknow 226003, Uttar Pradesh, India (IEC no. 1758/Ethics/R.Cell-17; dated: 08/02/2017). The work was performed in strict accordance with the guidelines of the Institutional Ethical committee. Before withdrawing the blood from the subjects, the purpose of the study was explained to all participants, and a signed written informed consent was obtained.

Patient Selection and Sample Collection. A total of 39 consecutive newly diagnosed treatment naive LC patients were enrolled who consulted at the OPD of Department of Pulmonary Surgery, King George’s Medical University (KGMU), Lucknow 226003, Uttar Pradesh, India. Histopathological examinations of the biopsied or resected tissue samples were conducted to classify the LC patients into its clinical subtypes according to the 7th edition of the TNM staging system. For comparative analysis, 42 healthy subjects (comparable in age and sex) were recruited as NCs. From each participant, 2.0 mL of blood sample was drawn in plain vacutainer tubes (Becton Dickinson), and serum was extracted as described previously and stored at −80 °C.

Sample Preparation. Before carrying out the NMR experiments, the stored serum samples were thawed and homogenized using a vortex mixer for 5 min. The NMR samples were prepared following the procedure as described previously. Briefly, 250 μL of 0.9% saline sodium phosphate buffer of strength 50 mM (pH 7.4, prepared in 100% deuterium oxide i.e., D₂O) was added to 250 μL of serum to minimize the variation in pH. The resultant sample mixture was centrifuged at 16,278 g for 5 min, and then 450 μL of supernatant of this mixture was transferred to 5 mm NMR tubes (Wilmad Glass, USA). A sealed capillary tube holding 1.0 mM TSP (sodium salt of 3-trimethylsilyl (2,2,3,3,4,4)-propionic acid) dissolved in D₂O was inserted in the NMR tubes as an external reference (resulting final TSP concentration ∼ 0.1 mM). The NMR solvent (deuterium oxide with deuteration degree min. 99.95%) and the sodium salt of TSP were purchased from Merck Millipore and Sigma-Aldrich (St. Louis, MO, USA), respectively.

Data Collection and Preprocessing. The serum samples were prepared and examined on 800 MHz NMR following the procedure as mentioned previously. The recorded NMR spectra were analyzed using the PROCESSOR module of commercial software Chenomx (v8.2, Edmonton, Canada), and the CPMG data matrix containing 0.02 ppm spectral bins normalized with respect to the total spectral intensity was prepared for multivariate analysis as described previously. As glucose, lactate, and lipid/membrane metabolites (including low-/very-low-density lipoproteins) significantly and variably contribute to the total spectral intensity, the discriminatory analysis based on normalized spectral features may provide unreliable results as demonstrated previously. To avoid any such possibility, we additionally estimated the concentrations of 25 serum metabolites with respect to formate (as an internal reference and selecting pH of sample equal to 7.0 ± 0.5). The purpose of selecting formate is that its singlet NMR signal is present most downfield in the NMR spectrum of serum and does not have overlap with any other signal of serum metabolites. Further, the software Chenomx provides the option to use formate as a
calibration standard to overcome/minimize the analytical variations. A recent NMR-based serum metabolomics study showed that formate does not change significantly in the sera of LC patients compared to LC27 This further supported the use of formate as an internal reference. For spectral calibration, the concentration of formate was set to 10 μM, that is, nearly...
close to the detection limit of a 800 MHz NMR spectrometer and well within the reported circulatory range in the literature.\textsuperscript{38,34} The estimated concentrations of serum metabolites have been reported here in micromolar except for glucose and lactate for which the concentrations have been reported in millimolar (see Supporting Information, Table S2) and can simply be considered as ratiometric metabolic profiles estimated with respect to formate.

**Multivariate Statistical Analysis.** The multivariate data derived from NMR-based serum metabolic profiling was analyzed for comparison between LC and NC groups using multivariate statistical analysis tools such as unsupervised PCA and supervised PLS-DA performed using freely available web-based server named MetaboAnalyst.\textsuperscript{55,53,56} The PLS-DA analysis was performed following the details described previously.\textsuperscript{52} The performance of PLS-DA models was further improved by integrating it with orthogonal signal correction. The analysis is referred to as OPLS-DA and removes variability not relevant to class separation.\textsuperscript{57} The OPLS-DA analysis was performed using commercial software SIMCA (v14.0, Umetrics, Umeå, Sweden: https://umetrics.com/kb/simca-14).

The spectral features or metabolites of statistical significance and diagnostic potential were identified finally through performing the ROC curve analysis integrated with the the student t-test using Biomarker module of MetaboAnalyst. The level of statistical significance was set at \( p < 0.05 \). Continuous variables were expressed as the mean ± SD and categorical variables as the percentage.

**ASSOCIATED CONTENT**

- **Supporting Information**
  The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c06941.

Metabolite assignments of major resonances detected in the \(^1\)H NMR spectra of human serum samples; list of 29 serum metabolic features and their mean values estimated for LC and NC groups; VIP scores of metabolites derived from OPLS-DA-based discriminatory analysis between LC and NC groups; diagnostic potential of serum metabolites evaluated based on AUROC analysis for differentiating the LC from NC along with their respective adjusted \( p \)-values and fold change; diagnostic potential of serum metabolites evaluated based on AUROC for differentiating the major clinical subtypes of LC, that is, ADC and SqCC; synchronized 3D score plot (left) and loading plot (right) derived from the PLS-DA model-based discriminatory analysis performed online using MetaboAnalyst Software (v4.1) and the metabolites of discriminatory potential highlighted based on their VIP scores; 3D score plot derived from the PLS-DA model-based discriminatory analysis performed online using MetaboAnalyst software (v4.1) and showing exquisite separation between the clinical subtypes of LC, that is, ADC and SqCC; metabolites exhibiting statistically significant difference highlighted based on their \( p \)-values; and ROC curve analysis performed to evaluate the diagnostic potential of four circulatory ratios in discriminating the clinical subtypes of LC, that is, ADC and SqCC study groups (PDF).

**AUTHOR INFORMATION**

**Corresponding Authors**

Nikhil Gupta – Centre of Biomedical Research (CBMR), SGPGIMS, Lucknow, Uttar Pradesh 226014, India; Department of Chemistry, Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India; orcid.org/0000-0002-9818-2206; Phone: +91-8287048987; Email: nikhilgupta1532@gmail.com

Dinesh Kumar – Centre of Biomedical Research (CBMR), SGPGIMS, Lucknow, Uttar Pradesh 226014, India; Phone: +91-8953261506; Email: dinesh@cbmr.res.in, dineshbcmr@gmail.com

**Authors**

Anjana Singh – All India Institute of Medical Sciences (AIIMS), Rishikesh, Uttarakhand 249201, India; Pulmonary & Critical Care Medicine, King George’s Medical University, Lucknow, Uttar Pradesh 226003, India

Ved Prakash – Pulmonary & Critical Care Medicine, King George’s Medical University, Lucknow, Uttar Pradesh 226003, India

Ashish Kumar – Department of Chemistry, Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India

Ravi Kant – All India Institute of Medical Sciences (AIIMS), Rishikesh, Uttarakhand 249201, India

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c06941

**Author Contributions**

A.S. and V.P. contributed equally. A.S., V.P., and R.K. were involved in clinical evaluation, collecting patient samples, and compiling clinical details. N.G. and A.K. were involved in preparation of serum samples for NMR studies, NMR data collection, data processing, and manuscript drafting. D.K. was involved in metabolomics data analysis and manuscript preparation.

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

D.K. acknowledges the (a) Department of Science and Technology for financial assistance under SERB EMR Scheme (ref. no.: EMR/2016/001756) and (b) Department of Medical Education, Govt. of Uttar Pradesh for supporting the High Field NMR Facility at the Centre of Biomedical Research, Lucknow, India. N.G. acknowledges receipt of the SRF fellowship from CSIR, India [file no: 09/916(0096)/2018-EMR-1].

**ABBREVIATIONS**

- **2-COL**

| Abbreviation | Definition |
|--------------|------------|
| LC           | lung cancer|
| NC           | normal control|
| CT           | computerized tomography|
| ADC          | adenocarcinoma|
| SqCC         | squamous cell carcinoma|
| SCLC         | small cell lung carcinoma|
| NSCLC        | non-small-cell lung cancer|
| ROC          | receiver operating characteristic curve|
| AUROC        | area under ROC curve|
| CI           | confidence interval|
| VIP          | variable importance for the projection|
| TMAO         | tri-methylamine-N-oxide|
REFERENCE

(1) Stewart, B. W.; Kleihues, P. World Cancer Report, 2003.
(2) Collins, L. G.; Haines, C.; Perkel, R.; Enck, R. E. Lung cancer: diagnosis and management. Am. Fam. Physician 2007, 75, 56–63.
(3) Hori, S.; Nishiumi, S.; Kobayashi, K.; Shinohara, M.; Hatateyama, Y.; Kotani, Y.; Hatano, N.; Maniwa, Y.; Nishio, W.; Bamba, T.; Fukusaki, E.; Azuma, T.; Takenawa, T.; Hatakeyama, Y.; Kotani, Y.; Hatano, N.; Maniwa, Y.; Nishio, W.; Bamba, T.; Fukusaki, E.; Azuma, T.; Takenawa, T.; Nishimura, Y.; Yoshida, M. A metabolomic approach to lung cancer. Lung Cancer 2011, 74, 284–292.
(4) Rivara, M. P.; Detterbeck, F.; Mehta, A. C. Diagnosis of lung cancer: the guidelines. Chest 2003, 123, 1295–1365.
(5) Puchades-Carrasco, L.; Jantus-Lewintre, E.; Pérez-Rambla, C.; Garcia-García, F.; Lucas, R.; Calabuig, S.; Blasco, A.; Dopazo, J.; Camps, C.; Pineda-Lucena, A. Serum metabolomic profiling facilitates the non-invasive identification of metabolic biomarkers associated with the onset and progression of non-small cell lung cancer. Oncotarget 2016, 7, 12904.
(6) Calejón-Leblic, B.; García-Barrera, T.; Grávalos-Guzmán, J.; Pereira-Vega, A.; Gómez-Arizá, J. L. Metabolic profiling of potential lung cancer biomarkers using bronchoalveolar lavage fluid and the integrated direct infusion/gas chromatography mass spectrometry platform. J. Proteomics 2016, 145, 197–206.
(7) Huang, Y.; Zhang, A.; Sun, H. Power of metabolomics in diagnosis and biomarker discovery of hepatocellular carcinoma. Hepatology 2013, 57, 2072–2077.
(8) Zhang, A.; Sun, H.; Yan, G.; Wang, P.; Wang, X. Metabolomics for biomarker discovery: moving to the clinic. BioMed Res. Int. 2015, 2015, 1.
(9) Claudino, W. M.; Goncalves, P. H.; di Leo, A.; Philip, P. A.; Sarkar, F. H. Metabolomics in cancer: a bench-to-bedside intersection. Crit. Rev. Oncol. Hematol. 2012, 84, 1–7.
(10) Griffin, J. L.; Shockcor, J. P. Metabolic profiles of cancer cells. Nat. Rev. Cancer 2004, 4, 551.
(11) Bu, Q.; Huang, Y.; Yan, G.; Cen, X.; Zhao, Y.-L. Metabolomics: A revolution for novel cancer marker identification. Comb. Chem. High Throughput Screen. 2012, 15, 266–275.
(12) Hanash, S. M.; Baik, C. S.; Kallioniemi, O. Emerging molecular biomarkers-blood-based strategies to detect and monitor cancer. Nat. Rev. Clin. Oncol. 2011, 8, 142.
(13) Zhang, A.-h.; Sun, H.; Qiu, S.; Wang, X.-j. Metabolomics in noninvasive breast cancer. Clin. Chim. Acta 2013, 424, 3–7.
(14) Wang, L.; Liu, X.; Yang, Q. Application of Metabolomics in Cancer Research: As a Powerful Tool to Screen Biomarker for Diagnosis, Monitoring and Prognosis of Cancer. Biomarkers J. 2018, 1, 12.
(15) Duarte, I. F.; Rocha, C. M.; Gil, A. M. Metabolic profiling of biofluids: potential in lung cancer screening and diagnosis. Expert Rev. Mol. Diagn. 2013, 13, 737–748.
(16) Mathé, E. A.; Patterson, A. D.; Haznadar, M.; Manna, S. K.; Krausz, K. W.; Bowman, E. D.; Shields, P. G.; Idle, J. R.; Smith, P. B.; Anami, K. Noninvasive urinary metabolic profiling identifies diagnostic and prognostic markers in lung cancer. Cancer Res. 2014, 74, 3259–3270.
(17) Guleria, A.; Kumar, A.; Kumar, U.; Raj, R.; Kumar, D. NMR Based Metabolomics: An Exquisite and Facile Method for Evaluating Therapeutic Efficacy and Screening Drug Toxicity. Curr. Top. Med. Chem. 2018, 18, 1827–1849.
(18) Hu, J.-M.; Sun, H.-T. Serum proton NMR metabolomics analysis of human lung cancer following microwave ablation. Radiat. Oncol. 2018, 13, 40.
(19) Zhang, X.; Zhu, X.; Wang, C.; Zhang, H.; Cai, Z. Non-targeted and targeted metabolomics approaches to diagnosing lung cancer and predicting patient prognosis. Oncotarget 2016, 7, 63437.
(20) Louis, E.; Adriaensens, P.; Guedens, W.; Bigirmurame, T.; Baeten, K.; Vanhove, K.; Vandeurzen, K.; Darquennes, K.; Vansteenkiste, J.; Dooms, C.; Shkedy, Z.; Mesotten, L.; Thomeer, M. Detection of lung cancer through metabolic changes measured in blood plasma. J. Thorac. Oncol. 2016, 11, S16–S23.
(21) Jelenek, K.; Wielak, P. Metabolome-based biomarkers: their potential role in the early detection of lung cancer. Contemp. Oncol. 2018, 22, 135.
(22) Mazzzone, P. J.; Wang, X.-F.; Beukenmann, M.; Zhang, Q.; Sceley, M.; Mohney, R.; Holt, T.; Pappan, K. L. Metabolite profiles of the serum of patients with non-small cell carcinoma. J. Thorac. Oncol. 2016, 11, 72–78.
(23) Klucznysza, A.; Derezinski, P.; Dyzskiewicz, W.; Pawlok, K.; Kasprzyk, M.; Kokot, Z. J. Evaluation of serum amino acid profiles’ utility in non-small cell lung cancer detection in Polish population. Lung Cancer 2016, 100, 71–76.
(24) Rocha, C. M.; Carrola, J.; Barros, A. S.; Gil, A. M.; Goodfellow, B. J.; Carreira, I. M.; Bernardo, J.; Gomes, A.; Sousa, V.; Carvalho, L.; Duarte, I. F. Metabolic signatures of lung cancer in biofluids: NMR-based metabolomics of blood plasma. J. Proteome Res. 2011, 10, 4314–4324.
(25) Derveaux, E.; Louis, E.; Vanhove, K.; Bertofts, L.; Mesotten, L.; Thomeer, M.; Adriaensens, P. Diagnosis of Lung Cancer: What Metabolomics Can Contribute. Lung Cancer-Strategies for Diagnosis and Treatment; IntechOpen, 2018.
(26) Banijj-Stocke, S.; van Berkel, V.; Miller, D. M.; Friebes, H. B. A review of metabolism-associated biomarkers in lung cancer diagnosis and treatment. Metabolomics 2018, 14, 81.
(27) Louis, E.; Cantrelle, F.-X.; Mesotten, L.; Reekmans, G.; Bertofts, L.; Vanhove, K.; Thomeer, M.; Lippens, G.; Adriaensens, P. Metabolic phenotyping of human plasma by 1H-NMR at high and medium magnetic field strengths: a case study for lung cancer. Magn. Reson. Chem. 2017, 55, 706–713.
(28) Mohan, A.; Garg, A.; Gupta, A.; Sahu, S.; Choudhari, C.; Vashistha, V.; Ansari, A.; Pandey, R.; Bhalla, A.; Madan, K.; Hadda, V.; Iyer, H.; Jain, D.; Kumar, R.; Mittal, S.; Tiwari, P.; Pandey, R.; Guleria, R. Clinical profile of lung cancer in India: A 10-year analysis of 1862 patients from a tertiary care center. Lung India 2020, 37, 190.
(29) Yu, L.; Li, K.; Zhang, X. Next-generation metabolomics in lung cancer diagnosis, treatment and precision medicine: mini review. Oncotarget 2017, 8, 115774.
(30) Moreno, P.; Jiménez-Jiménez, C.; Garrido-Rodriguez, M.; Calderón-Santiago, M.; Molina, S.; Lara-Chica, M.; Prieo-Capote, F.; Salviaterra, A.; Muñoz, E.; Calzado, M. A. Metabolic profiling of human lung tumor tissues-nucleotide metabolism as a candidate for therapeutic interventions and biomarkers. Mol. Oncol. 2018, 12, 1778–1796.
(31) Chang, S.-C.; Yang, W.-C. V. Hyperglycemia, tumorigenesis, and chronic inflammation. Crit. Rev. Oncol. Hematol. 2016, 108, 146–153.
(32) Kumar, U.; Kumar, A.; Singh, S.; Arya, P.; Singh, S. K.; Chaurasia, R. N.; Singh, A.; Kumar, D. An elaborate NMR based plasma metabolomics study revealed metabolic derangements in patients with mild cognitive impairment: a study on north Indian population. Metab. Brain Dis. 2021, 36, 957–968.
(33) Kumar, U.; Jain, A.; Guleria, A.; Misra, D. P.; Goel, R.; Danda, D.; Misra, R.; Kumar, D. Circulatory glutamine/glucose ratio for evaluating disease activity in Takayasu arteritis: a NMR based serum metabolomics study. J. Pharm. Biomed. Anal. 2020, 180, 113080.

5519 https://doi.org/10.1021/acsomega.1c06941
(34) Kumar, U.; Sharma, S.; Durgappa, M.; Gupta, N.; Raj, R.; Kumar, A.; Sharma, P. N.; Krishna, V. P.; Kumar, R. V.; Guleria, A. Serum metabolic disturbances associated with acute-on-chronic liver failure in patients with underlying alcoholic liver diseases: An elaborative NMR-based metabolomics study. J. Pharm. BioAllied Sci. 2021, 13, 276.

(35) Muhammed, H.; Kumar, D.; Dubey, D.; Kumar, S.; Chaurasia, S.; Guleria, A.; Majumder, S.; Singh, R.; Agarwal, V.; Misra, R. Metabolomics analysis revealed significantly higher synovial Phe/Tyr ratio in reactive arthritis and undifferentiated spondyloarthropathy. Rheumatology 2020, 59, 1587–1590.

(36) Charlton, M. Branched-chain amino acid enriched supplements as therapy for liver disease. J. Nutr. 2006, 136, 2958–2988.

(37) Tajiri, K.; Shimizu, Y. Branched-chain amino acids in liver diseases. World J. Gastroenterol. 2013, 19, 7620.

(38) Pietzke, M.; Arroyo, S. F.; Sumpton, D.; Mackay, G. M.; Martin-Castillo, B. a.; Camps, J.; Joven, J.; Menendez, J. A.; Vazquez, A. Stratification of cancer and diabetes based on circulating levels of formate and glucose. Cancer Metab. 2019, 7, 3.

(39) Akarachantachote, N.; Chadcham, S.; Saithanu, K. Cutoff threshold of variable importance in projection for variable selection. Int. J. Pure Appl. Math. 2014, 94, 307–322.

(40) Chen, Z.; Lu, W.; Garcia-Prieto, C.; Huang, P. The Warburg effect and its cancer therapeutic implications. J. Bioenerg. Biomembr. 2007, 39, 267–274.

(41) Vanhove, K.; Derveaux, E.; Graulas, G.-J.; Mesotten, L.; Thomeer, M.; Noben, J.-P.; Guedens, W.; Adriaensens, P. Glutamine addiction and Therapeutic Strategies in Lung Cancer. Int. J. Mol. Sci. 2019, 20, 252.

(42) Beloribi-Djefallah, S.; Vasseur, S.; Guillaumond, F. Lipid metabolic reprogramming in cancer cells. Oncogene 2016, 35, No. e189.

(43) Luo, X.; Cheng, C.; Tan, Z.; Li, N.; Tang, M.; Yang, L.; Cao, Y. Emerging roles of lipid metabolism in cancer metastasis. Mol. Cancer 2017, 16, 76.

(44) Andou, A.; Hisamatsu, T.; Okamoto, S.; Chinen, H.; Kamada, N.; Kobayashi, T.; Hashimoto, M.; Okutsu, T.; Shimbo, K.; Takeda, T.; Matsumoto, H.; Sato, A.; Ohtsu, H.; Suzuki, M.; Hibi, T. Dietary histidine ameliorates murine colitis by inhibition of proinflammatory cytokine production from macrophages. Gastroenterology 2009, 136, 564–574.

(45) Wade, A. M.; Tucker, H. N. Antioxidant characteristics of L-histidine. J. Nutr. Biochem. 1998, 9, 308–315.

(46) Peterson, J. W.; Boldog, I.; Popov, V. L.; Saini, S. S.; Chopra, A. K. Anti-inflammatory and antisecreatory potential of histidine in Salmonella-challenged mouse small intestine. Lab. Invest. 1998, 78, 523–534.

(47) Nicoud, M. B.; Formoso, K.; Medina, V. A. Pathophysiological Role of Histamine H4 Receptor in Cancer: Therapeutic Implications. Front. Pharmacol. 2019, 10, 556.

(48) Shahid, M.; Tripathi, T.; Sobia, F.; Moin, S.; Siddiqui, M.; Khan, R. A. Histamine, histamine receptors, and their role in immunomodulation: an updated systematic review. Open Immunol. J. 2009, 2, 9.

(49) Kumar, U.; Mehta, P.; Kumar, S.; Jain, A.; Guleria, A.; Misra, R.; Kumar, D. Circulatory histidine levels as predictive indicators of disease activity in takayasu arteritis. Anal. Sci. Adv. 2021, 2, 527–535.

(50) Azad, N.; Rojanasakul, Y.; Vallyathan, V. Inflammation and lung cancer: roles of reactive oxygen/nitrogen species. J. Toxicol. Environ. Health, Part B 2008, 11, 1–15.

(51) Jain, A.; Kumar, D.; Guleria, A.; Misra, D. P.; Zanwar, A.; Chaurasia, S.; Kumar, S.; Kumar, U.; Mishra, S. K.; Goel, R.; Danda, D.; Misra, R. NMR-based serum metabolomics of patients with Takayasu arteritis: relationship with disease activity. J. Proteome Res. 2018, 17, 3317–3324.

(52) Goldstraw, P.; Crowley, J.; Chansky, K.; Giroux, D. J.; Groome, P. A.; Rami-Porta, R.; Postmus, P. E.; Rusch, V.; Sobin, L. The International Association for the Study of Lung Cancer International Staging Committee The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J. Thorac. Oncol. 2007, 2, 706–714.