Baseline Plasma Proteomic Analysis to Identify Biomarkers that Predict Radiation-Induced Lung Toxicity in Patients Receiving Radiation for Non-small Cell Lung Cancer

Xu-Wei Cai, MD,*†‡ Kerby A. Shedden, PhD,§ Shang-Hu Yuan, MD, PhD,† Mary A. Davis, PhD,‡ Lu-Ying Xu, MD,‡∥ Cong-Ying Xie, MD,‡∥ Xiao-Long Fu, MD, PhD,‡*† Theodore S. Lawrence, MD, PhD,‡ David M. Lubman, PhD,# and Feng-Ming (Spring) Kong, MD, PhD‡**

Purpose: To identify new plasma proteomic markers before radiotherapy start to predict later grade ≥2 radiation-induced lung toxicity (RILT2).

Methods: Fifty-seven patients with non-small cell lung cancer received radiotherapy (RT) were eligible. Forty-eight patients with minimum follow-up of 1 year, nine with RILT2 with tumor stage matched to 39 without RILT2, were enrolled for this analysis. Platelet-poor plasma was obtained within 2 weeks before radiotherapy. The plasma proteomes were compared using a multiplexed quantitative proteomics approach involving ExacTag labeling, reverse-phase high-performance liquid chromatography, and nano liquid chromatography electrospray ionization tandem mass spectrometry. Z scores and Bonferroni-adjusted p values for the two-sample mean comparison were used to identify the differential protein expression between patients with and without RILT2.

Results: More than 200 proteins were identified and quantified. After excluding proteins that were not detected in at least 40% of the 48 patient samples, C4b-binding protein alpha chain and vitronectin had significantly higher (p < 0.001 and p = 0.02) expression levels in patients with RILT2 compared with patients without RILT2. These two proteins were validated by Western blot. Ingenuity pathway analysis revealed that they both play important roles in the inflammatory response and are associated with the known pathways of radiation-induced lung damage.

Conclusions: This proteomic approach demonstrates new plasma protein biomarkers before treatment for future studies on RILT2 prediction.

Key Words: Non-small cell lung cancer, Proteomics, Biomarker, Radiation-induced lung toxicity.

(J Thorac Oncol. 2011;6: 1073–1078)

Approximately 60 to 70% of patients with lung cancer receive radiation therapy (RT) at least once during the course of their disease.1 Clinical trials showed that patients with non-small cell lung cancer (NSCLC) who received higher radiation doses may have better local control and survival.2–4 Nevertheless, the majority of patients do not receive an adequate radiation dose for tumor control because of concerns about risk estimates for the overall population to limit the incidence of radiation-induced lung toxicity (RILT) to within 5% in 5 years (tolerable dose [TD]5). Early prediction of RILT would help physicians to stratify patients according to their risk level for toxicity and determine an individualized treatment regimen rather than to use the average radiation tolerance of the whole population.

Unfortunately, there are no good predictive markers available to provide an individualized approach. A few small series of studies have shown that transforming growth factor beta 1 (TGF-β1) is elevated at the end or during the course of radiation, which might be correlated with radiation pneumonitis and, therefore, be used as a predictive marker.5–7 Nevertheless, the process of radiation damage is a complicated pathophysiological phenomena involving many cells and cytokines.7 We hypothesized that a test monitoring more proteins together in the blood would estimate the specific extent of the damage more accurately.

Pioneering work has already been performed using proteomics tools in the search for tumor markers, demonstrat-
ing the promise of such an effort. These studies were mostly focused on early detection, diagnosis, staging, treatment monitoring, and prognosis of lung cancer. A few studies have reported proteomics analysis of radiation-induced changes in lung cells or tissue. We have previously demonstrated in animal studies that radiation-induced changes in multiple proteins differ between lung fibrosis sensitive and resistant mouse strains by proteomic approaches. Using ExacTag labeling, reverse-phase high-performance liquid chromatography (RP-HPLC), and nano liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS), we recently assessed differential plasma proteomics in patients with and without grade ≥2 RILT (grade ≥2 radiation-induced lung toxicity [RILT2]) in 20 patients throughout the treatment course. We demonstrated that C4b-binding protein alpha chain (C4BPA), complement C3, and vitronectin (VTN) had significantly higher expression levels in patients with RILT2 compared with patients without RILT2, based on both the longitudinal data sets of RT start to 3 months post-RT (p < 0.01) and RT start to the end of RT (p < 0.01). The purpose of this study, using a larger number of patients, was to examine whether there is any potential proteomic biomarkers at baseline (before radiotherapy start) to predict the risk of RILT2.

PATIENTS AND METHODS

Study Population

This work was part of prospective clinical trials, which were approved by the Institutional Review Board and performed at the University of Michigan and the Veterans Affairs Medical Center, Ann Arbor, MI. Written informed consent was obtained from each patient. Eligible subjects included patients with stages I to III NSCLC undergoing radiation alone or combined radiation and chemotherapy. Exclusion criteria included a life expectancy of less than 6 months, malignant pleural or pericardial effusion, or noncontiguous involvement of the parietal pleura. All the irradiation was given using the three-dimensional conformal technique, as described previously.

Plasma Sample Preparation

Platelet-poor plasma was used for proteomic analyses. Blood samples were collected within 2 weeks before RT start with Vacutainer plus collection tubes containing K2 ethylenediamine tetraacetic acid as the anticoagulant then placed in ice immediately after collection and centrifuged within 3 hours of collection at 3000 g for 30 minutes at 4°C. The upper one third of the supernatant was collected as platelet-poor plasma and stored at −80°C. Pooled normal human plasma was purchased from Innovative Research Inc. (Plymouth, MN, Cat #: IPLA-5) as a reference control.

Proteomic Analysis

Multiplex quantitative proteomics approaches, involving ExacTag labeling, RP-HPLC, and LC-ESI-MS/MS, were used in analyzing radiation-induced changes in plasma proteins between patients with and without RILT2. The detailed methods have been recently prescribed in our previous work. In brief, the proteomic analysis includes depletion of the 12 most abundant proteins, stable isotope sample labeling, protein separation by RP-HPLC, peptide separation by LC-ESI-MS/MS, protein identification by the SEQUEST algorithm and Trans-Proteomic Pipeline software, and protein quantification by ExacTag analysis software 3.0.

Western Blotting for Validation

Western blotting was used to further validate the measurement of proteomic approaches using anti-VTN antibody (ab11591), anti-affamin (AFM) antibody (ab49139; Abcam Inc., Cambridge, MA), and anti-C4BPA antibody (SAB2100305; Sigma-Aldrich, St Louis, MO). As a proof of the principle, we randomly selected eight samples, four with higher expression of VTN and four with lower expression of VTN. AFM was selected to be an internal reference because the expressions of AFM were almost the same in all samples, which were identified from our proteomic approaches.

Statistical Considerations

The primary end point in this study was grade ≥2 RILT (RILT2), as it is clinically significant. The details of the diagnosis and grading systems for RILT have been described previously. Z scores for the two-sample mean difference between log-transformed protein expression in RILT ≥2 and RILT less than 2 patients were used to identify univariate protein markers. The 80 proteins that were detected in at least 40% of the 48 patient samples were analyzed, based on the sample size and the number of events with weighted consideration on the number of proteins and cost of multiple testing. Bonferroni-adjusted p values for these 80 Z scores were used to identify statistically significant associations with RTIL2. Predictive analysis was performed using ridged logistic regression to adjust for potential errors from multilisting. Cross-validation was used to produce unbiased sensitivity and specificity estimates, which were displayed as a receiver operating characteristic curve. The logistic regression was based on selection of k proteins based on univariate Z scores, which were calculated internally to the cross-validation. Predictive models using log-scale protein expression and/or mean lung dose (MLD) were considered.

RESULTS

Patient Characteristics

Of the 57 patients enrolled, 48 patients with longer than 2-year follow-up were included into this analysis. There were nine more than grade 2, 19 grade 1, and 20 grade 0 RILT. Our analysis focused on the difference between patients with RILT ≥2 (RILT2) and grouped RILT0 to 1 patients (without RILT2). The patients’ characteristics are listed in Table 1. Lung dosimetric factors such as MLD, V20, and normal tissue complication probability (NTCP) were not significantly different between patients with and without RILT2, largely due to the fact that RT dose prescription was individualized to certain fixed limits of MLD, V20, or NTCP in most patients. Of another note, MLD, V20, and NTCP were significantly correlated with each other (p < 0.001), we elected MLD as the representative dosimetric parameter of the lung.
Reproducibility of the Proteomics System

The reproducibility of the proteomics system, including isotope-labeling reagent, RP-HPLC, and nano LC-MS/MS system, has been confirmed and reported in our recent study.17

Plasma Proteomic Analysis between Patients with and without RILT2

More than 200 proteins were identified and quantified in plasma samples. A representative MS/MS spectrum of a unique peptide from the protein VTN is shown in Figure 1.

There were four proteins with significantly different expression between patients with and without RILT2 after Bonferroni correction for multiple testing. The expression of immunoglobulin (Ig) kappa chain V-III region Ti and region HAH in patients with RILT2 was significantly lower than in patients without RILT2, whereas the expression of C4BPA and VTN in patients with RILT2 was significantly higher than that in patients without RILT2 as shown in Figure 2 (adjusted p = 0.001 and p = 0.009, respectively).

Validation by Western Blotting

As shown in Figure 3, this monoclonal anti-VTN antibody (Abcam Inc., #ab11591) recognizes both the 75 kDa and 65 kDa bands of human purified and plasma VTN in an immunoblotting assay.20 VTN and C4BPA were detected at lower levels in samples 1 to 4 than in samples

---

**TABLE 1. Characteristics of Patients**

<table>
<thead>
<tr>
<th></th>
<th>Patients with RILT2</th>
<th>Patients without RILT2</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (male:female)</td>
<td>8:1</td>
<td>34:5</td>
<td>0.42</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>69 (53–81)†</td>
<td>68 (44–85)†</td>
<td>0.76</td>
</tr>
<tr>
<td>Radiation dose</td>
<td>66 (63–75.6) Gy</td>
<td>67.8 (45–81.4) Gy</td>
<td>0.78</td>
</tr>
<tr>
<td>MLD</td>
<td>16.5 (9.1–20.5) Gy</td>
<td>14.6 (3.6–23.4) Gy</td>
<td>0.37</td>
</tr>
<tr>
<td>V20</td>
<td>27 (17–34)%</td>
<td>24 (5–35)%</td>
<td>0.44</td>
</tr>
<tr>
<td>NTCP</td>
<td>9.8 (2.9–18.7)%</td>
<td>8.9 (0.8–26.2)%</td>
<td>0.58</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>77.8 %</td>
<td>76.9 %</td>
<td>0.96</td>
</tr>
</tbody>
</table>

*Data presented are median (range).

RILT2, grade ≥2 radiation-induced lung toxicity; MLD, mean lung dose; NTCP, normal tissue complication probability; V20, volume of both lungs (without including the gross tumor volumes) receiving ≥20 Gy.

---

**FIGURE 1.** MS/MS sequencing data of a unique peptide from vitronectin identified in fractions. MS/MS, tandem mass spectrometry.
5 to 8, whereas the reference protein AFM was detected at similar levels of expression, which are consistent with the results from proteomic approaches. This result suggests the reliability of the proteomic approaches used for protein quantification. VTN in the samples from patients with RILT2 presents smear bands at higher molecular weight and suggesting greater glycosylation.

Predictive Models for Risk of RILT2

Ridged logistic regression was used to perform predictive analysis of five models: C4BPA alone, VTN alone, MLD alone, C4BPA + VTN, and C4BPA + VTN + MLD. Cross-validation of unbiased sensitivity and specificity estimates is displayed as receiver operating characteristic curves with values of area under the curve (AUC), which are 0.40, 0.74, 0.49, 0.70, and 0.71 separately, as shown in Figure 4. Although all these models were not statistically significant ($p > 0.05$), this result showed that three models, VTN alone, C4BPA + VTN, and C4BPA + VTN + MLD, seemed to have better accuracy (AUC >0.7) than MLD and other models (AUC of 0.4 – 0.47) in predicting the occurrence of RILT2. Model parameters were limited by the number of events in this study.

DISCUSSION

In this study of 48 patients, we identified four plasma proteins that were significantly different before the start of treatment in patients with and without RILT2 (Figure 2): Ig kappa chain V-III region Ti and region HAH in patients with RILT2 were significantly lower in patients with RILT2, whereas the expression of C4BPA and VTN was higher in patients who later developed RILT2. This is the first human plasma proteomics study reporting potential plasma biomarkers for RILT2 in patients with NSCLC before RT start. These results represent a continuation of our recent study of 20 patients, which showed significant differences in angiotensinogen, complement C3, C4BPA, and VTN between patients with NSCLC with and without RILT2 during and after completion of RT course (from baseline to 3 months after the end of RT). This study, in a larger number of
patients, has further demonstrated a statistic significant difference in C4BPA and VTN in patients with RILT2 at baseline (before RT start), thus suggesting the possibility for predicting RILT2 before treatment start, which has more important clinical implications.

C4BPA\textsuperscript{21–23} and VTN\textsuperscript{24–27} are glycoproteins, which play important roles in radiation-induced lung damage. The detailed functions of these proteins are complex and have been summarized previously.\textsuperscript{17} C4BPA seems to be mainly associated with the inflammatory complement pathway, whereas VTN is involved in the fibrotic pathway. The high baseline level of these two proteins may suggest higher sensitivity of an individual to inflammatory or fibrotic pathogenesis, thus putting this patient at a higher risk for RILT.

Changes in abundance and alternations in glycan composition of plasma proteins and cell-surface proteins have been shown to correlate with cancer and other disease states. In fact, numerous clinical biomarkers and therapeutic targets are glycosylated proteins, such as CA125 for ovarian cancer and the prostate-specific antigen for prostate cancer. VTN is a glycoprotein with three N-glycosylation sites identified.\textsuperscript{28,29} Sano et al.\textsuperscript{20} reported that the changes in N-glycosylation of VTN modulate collagen binding during liver regeneration after partial hepatectomy. As shown in Figure 3, patients with RILT2 might have much higher levels of glycosylated VTN than patients without RILT2. Our future work will focus on glycoproteome analysis of human plasma to find more detailed information on the glycoproteins, which may be potential biomarkers for RILT2.

To reveal the association between our identified proteins and the known pathways for radiation-induced lung damage, we used ingenuity pathway analysis, a knowledge-based system derived from the literature containing information on interactions between genes, proteins, and other biological molecules, to reveal the associated biological pathways and generate global canonical pathways that are shown to be significantly associated with these candidates based on their interactions.\textsuperscript{30,31} Interestingly, these proteins were in the network of TGF-\(\beta\)1 and interleukin-8 (IL-8), which are known to play important roles in the molecular mechanism of radiation-induced lung damage.\textsuperscript{5–7,32} The network illustrated implies that C4BPA and VTN have interactions with TGF-\(\beta\)1.\textsuperscript{33,34} Human TGF-\(\beta\)1 increases production of hydrogen peroxide,\textsuperscript{35} and hydrogen peroxide increases expression of human C4BPA mRNA.\textsuperscript{36} This network is further supported by our previous findings that patients with RILT2 also had significantly higher plasma TGF-\(\beta\)1 during the course of radiation therapy than those without.\textsuperscript{32,37} Multicenter studies with a larger number of patients are needed to validate these interesting findings and to study whether combining all these markers together would improve the accuracy of RILT2 prediction.

It is interesting that increases in Ig kappa chain V-III region Ti\textsuperscript{19} or region HAH.\textsuperscript{39} These fragments may have some potential association with RILT protection and are worthy of our further study to identify their mechanism.

In summary, this study identified VTN and C4b-binding protein alpha chain as potential blood biomarkers for RILT2 in patients with NSCLC before RT start. A future independent study with a larger number of patients is needed to validate this result and confirm whether it can improve predictive accuracy from the models (such as MLD) in current use.

ACKNOWLEDGMENTS

Supported, in part, by the American Society of Clinical Oncology (ASCO) Career Developmental Award, a grant from Pardee foundation, and TRP Award #118 from Radiation Therapy Oncology Group (RTOG).

REFERENCES


