

RESEARCH ARTICLE

Quantitative Proteomic Analysis of Optimal Cutting Temperature (OCT) Embedded Core-Needle Biopsy of Lung Cancer

Xiaozheng Zhao,^{1,2,3} Kenneth E. Huffman,⁴ Junya Fujimoto,⁵ Jamie Rodriguez Canales,⁵ Luc Girard,⁴ Guangjun Nie,^{1,2} John V. Heymach,⁶ Igacio I. Wistuba,⁵ John D. Minna,⁴ Yonghao Yu³

¹CAS Center for Excellence in Nanoscience, National Center for Nanoscience and Technology, Beijing, 100190, People's Republic of China

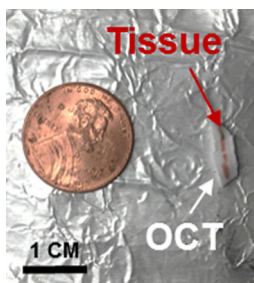
²University of Chinese Academy of Sciences, Beijing, 100049, People's Republic of China

³Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX, USA

⁴Hamon Center for Therapeutic Oncology Research, Simmons Comprehensive Cancer Center, Pharmacology and Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, USA

⁵Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, TX, USA

⁶Department of Head and Neck and Thoracic Oncology, University of Texas MD Anderson Cancer Center, Houston, TX, USA



Abstract. With recent advances in understanding the genomic underpinnings and oncogenic drivers of pathogenesis in different subtypes, it is increasingly clear that proper pretreatment diagnostics are essential for the choice of appropriate treatment options for non-small cell lung cancer (NSCLC). Tumor tissue preservation in optimal cutting temperature (OCT) compound is commonly used in the surgical suite. However, proteins recovered from OCT-embedded specimens pose a challenge for LC-MS/MS experiments, due to the large amounts of polymers present in OCT. Here we present a simple workflow for whole proteome analysis of OCT-embedded NSCLC tissue samples, which involves a simple trichloroacetic acid precipitation step. Comparisons of protein recovery between frozen versus OCT-embedded tissue showed

excellent consistency with more than 9200 proteins identified. Using an isobaric labeling strategy, we quantified more than 5400 proteins in tumor versus normal OCT-embedded core needle biopsy samples. Gene ontology analysis indicated that a number of proliferative as well as squamous cell carcinoma (SqCC) marker proteins were overexpressed in the tumor, consistent with the patient's pathology based diagnosis of "poorly differentiated SqCC". Among the most downregulated proteins in the tumor sample, we noted a number of proteins with potential immunomodulatory functions. Finally, interrogation of the aberrantly expressed proteins using a candidate approach and cross-referencing with publicly available databases led to the identification of potential druggable targets in DNA replication and DNA damage repair pathways. We conclude that our approach allows LC-MS/MS proteomic analyses on OCT-embedded lung cancer specimens, opening the way to bring powerful proteomics into the clinic.

Keywords: Proteomics, Lung cancer, Pathology, Biomarker, Drug target, Sample preparation

Received: 30 January 2017/Revised: 3 May 2017/Accepted: 4 May 2017/Published Online: 27 July 2017

Electronic supplementary material The online version of this article (doi:10.1007/s13361-017-1706-z) contains supplementary material, which is available to authorized users.

Correspondence to: Guangjun Nie; e-mail: niegj@nanocr.cn, Yonghao Yu; e-mail: Yonghao.yu@utsouthwestern.edu

Introduction

Lung cancer is the leading cause of cancer-related deaths in the United States, with an estimated 224,390 new cases and 158,080 deaths for 2016 [1]. Non-small cell lung cancer (NSCLC) is a highly heterogeneous disease accounting for 85% of all lung cancer diagnoses and is commonly divided

into three subtypes: adenocarcinoma (ADC), squamous cell carcinoma (SqCC), and large cell carcinoma (LCC). SqCC is an aggressive subtype accounting for 25%–30% of NSCLC cases, making it the second most common type of lung cancer [2]. SqCC is notoriously difficult to treat as patients tend to be older, present with advanced disease, and have centrally located tumors [3, 4]. Smoking is strongly associated with SqCC requiring clinical attention to other comorbidities when considering therapeutic options [5]. Five-year survival rates for SqCC patients are reasonable in stage I and II patients following surgical resection but drop precipitously among patients with stage III or disseminated disease [6, 7].

There has been significant improvement in NSCLC patient survival over the last 20 years due, in part, to clinical deployment of targeted therapies [8]. Over the past several years, there have been important strides made using “checkpoint inhibitors” for immunotherapy of NSCLC, particularly SqCC [9]. In determining which patient subsets benefited from checkpoint inhibitor therapy, measuring of tumor and microenvironmental biomarkers also with determining the number of neoantigens created were important [10]. SqCC is not strongly associated with established lung cancer driver mutations found in ADC such as EGFR, ALK, or KRAS [11, 12]. Recent work is beginning to highlight the complexity of SqCC with somatic mutation rates higher than that seen in glioblastoma, breast, or colorectal cancers [13]. Moreover, identification of potentially targetable amplifications in FGFR1, PDGFR, MET, and PIK3CA have bolstered the notion that targeted therapy for SqCC patients will be possible [13–16]. However, the need for novel drug and immunotherapy targets and biomarkers remains unmet.

A critical issue for the management of lung cancer patients is proper histologic subtyping, which ultimately informs therapeutic choice. Better understanding of NSCLC has led to discoveries that are differentiating standard of care between ADC and SqCC patients, making proper pretreatment diagnostics essential. Surveillance, Epidemiology and End Results (SEER) data from 2012 illustrate the complexities of lung cancer pathologic diagnosis with nearly 22,000 patients in the U.S.A. per year given “unclassified” or “not otherwise specified” (NOS) diagnoses [17]. While adequate tumor material is obtained at the time of surgical resection, in nearly 70% of lung cancer cases surgical resections is not possible leading to the need of obtaining tissue for diagnostic and biomarker studies to less invasive techniques such as core needle biopsies (CNB) [18]. Although standard histology is the current gold standard for lung cancer pathologic diagnosis, molecular and immunohistochemical (IHC) approaches can aid in subtype characterization [19, 20].

Recent advances in mass spectrometry (MS) have enabled quantitative proteome analysis in tissue with high reproducibility, sensitivity, and accuracy in a comparatively short time [21]. As important tools for biomarker discovery, MS technologies could be potentially deployed to (1) facilitate clinical diagnosis of a particular NSCLC subtype, (2) elucidate pathways that are aberrantly activated, (3) identify individuals with higher risk of

disease progression and recurrence, and (4) identify patients who will likely respond to specific therapies, including immunotherapy. Frozen tissue from resection or CNB would be ideal for proteomics analysis, but is not always available. Most clinical tumor samples, including CNBs, are preserved either in OCT for immediate pathology diagnostics use or are fixed in formalin and paraffin embedded (FFPE) for immunohistochemistry and long-term storage [22, 23]. However, formalin fixation produces inter- and intra-cross-linked proteins and peptide modifications that could compromise subsequent mass spectrometry-based proteomic analyses [24]. Conservative estimates suggest that at least 15%–20% of the proteome become inaccessible after FFPE fixation [25]. A number of other studies have also provided evidence that after deparaffinization, similar proteins could be identified between frozen and FFPE specimens of the same tissues [26]. Alternatively, OCT-embedded tissues are not exposed to fixatives and may better represent the original tissue proteome. OCT compound is a cryo-preserved consisting of nonreactive ingredients, including polyvinyl alcohol (PVA) and polyethylene glycol (PEG) [27]. Proteomic studies with OCT-embedded specimens, however, are challenged by the interference of the OCT polymer with MS analysis as they compete with peptides for LC column binding during sample preparation and separation, and suppress ionization leading to reduced peptide identification [28]. Toward this end, several methods have been developed to specifically remove OCT from tissue samples for downstream proteomic analysis [27, 28]. However, proteomic analysis of OCT-embedded samples still represent a difficult challenge when total amounts of protein are limited [29].

To begin to address this problem, we developed a simple protein extraction method from OCT samples to deliver to LC-MS/MS for analysis. As an example of this, we used resected tumor tissue from a lung cancer patient whose pathologic diagnosis was “poorly differentiated squamous cell carcinoma.” To demonstrate LC-MS/MS compatibility with other clinical pathology techniques, we optimized a workflow to purify whole proteome from SqCC chunk tissue embedded in OCT compound. Using a facile trichloroacetic acid (TCA)-based protein precipitation procedure, we were able to successfully remove interfering OCT polymers from the tumor protein preparation and rapidly identify more than 9200 proteins using shotgun proteomic analysis. Comparisons of OCT-embedded chunk tissue with frozen tissue derived from the same tumor demonstrated excellent correlation in both number and identity of the detected proteins, indicating excellent recovery of the proteome. Classification of the most highly expressed tumor proteins revealed a strong cell proliferation theme, characteristic of “poorly differentiated” tumors. We then applied our TCA protocol to more clinically challenging specimens of OCT-embedded CNBs. Utilizing a tandem mass tag (TMT)-based quantitative proteomic platform [30], we successfully identified and quantified more than 5400 proteins in OCT embedded CNBs from tumor and normal SqCC tissues. Replicate TMT analysis of CNB samples showed near perfect correlation. Quantitative comparison of differentially expressed proteins

between CNB-tumor and CNB-normal samples substantiated the poorly differentiated SqCC pathology diagnosis. We were able to quantify proteins relevant to squamous tumor biology and immune modulation, as well as identify potentially novel biomarkers and drug targets. Following the removal of the OCT compound using our protocol, no enrichment or systematic protein loss can be observed when comparing datasets for compartments/organelles.

Materials and Methods

Tissue Lysis

Chunk tissues (frozen or OCT-embedded tumor samples) were placed in a pre-chilled clean mortar filled halfway with liquid nitrogen, and were ground into powder with a pestle. The powder was then transferred to a tissue grinder (Fisher # K885452-0020), and was homogenized with an SDS lysis buffer (3% SDS, 10 mM HEPES, pH 7.0, 2 mM MgCl₂, 0.05% v/v universal nuclease 20 U/mL). Core-needled biopsy (CNB) samples (OCT-embedded normal or lung tumor samples) were directly placed into the tissue grinder, and homogenized with the abovementioned SDS lysis buffer. The samples were centrifuged at 18,000 × *g* and the supernatant was collected for subsequent proteomic analysis. The normal lung CNB sample was collected from the same lobe as the tumor within 30 min after the surgery procedure. The specimen was collected such that the normal sample is far from the tumor, and the normal sample was confirmed microscopically to be free of tumor contamination. All the surgically resected samples and CNBs were from the sample SqCC patient.

Trichloroacetic Acid (TCA) Precipitation

Ice-cold 100% (w/v) TCA was added to the lysates to a final concentration of 20% TCA. After incubation on ice for 20 min, precipitated proteins were collected by centrifugation for 20 min at 16,000 × *g* at 4 °C. The supernatant was discarded and the pellets were washed with 500 μL ice-cold 10% (w/v) TCA. This washing step was repeated four times. The pellets were then washed with 1 mL pre-chilled acetone (−20 °C) for another four times. After the final wash, the supernatant was carefully removed and the pellets were solubilized in freshly prepared 8 M urea buffer containing 50 mM Tris, 10 mM EDTA, pH 7.5.

Protein Digestion and Sample Preparation for MS Analysis

For chunk tissue samples, proteins were reduced by adding DTT to a final concentration of 5 mM and incubated at room temperature for 15 min. Proteins were then alkylated by adding iodoacetamide to a final concentration of 50 mM and were incubated in dark for 20 min. Proteins were digested with Lys-C (Wako, VA, USA) at 1:100 enzyme/protein ratio for 2 hours, then were diluted with freshly prepared 100 mM Ammonium Bicarbonate solution to 2 M urea. Proteins were then digested with

Trypsin (Promega, WI, USA) at 1:100 enzyme/protein ratio overnight at room temperature. Peptides were desalted using solid-phase C18 extraction cartridges (Sep-Pak, Waters, MA, USA) according to manufacturer's instructions and were lyophilized.

Desalted peptides then were resuspended in 200 mM HEPES (pH 8.5) to a final concentration of 1 μg/μL. Peptides were fractionated by basic pH reversed phase HPLC (bRPLC) [31] at a flow rate of 0.2 mL/min on a ZORBAX 300 Extend-C18 column (Narrow Bore RR 2.1 mm × 100 mm, 3.5 μm particle size, 300 Å pore size). Buffer A is (10 mM ammonium formate in H₂O, pH 10.0). Gradient was developed from 0% to 70% buffer B (1 mM ammonium formate, pH 10.0, 90% ACN). Seventeen fractions were collected, which were lyophilized, desalted, and analyzed by LC-MS/MS.

For core-needle biopsy samples, proteins were quantified by the BCA method [30]. Proteins were reduced, alkylated, and digested using the abovementioned protocol. Three replicate analyses were performed for each condition, i.e., 126, 127, and 128 for SqCC core-needle biopsy sample, whereas 129, 130, and 131 for the normal lung core-needle biopsy sample from the same patient. For each TMT channel, approximately 15 μg peptides were reacted with the corresponding amine-based TMT six-plex reagent (Thermo Fisher, CA, USA) for 1 h at room temperature. Hydroxylamine solution was added to quench the reaction and the labeled peptide samples were combined. Then the sample was desalted and fractionated by bRPLC.

LC-MS/MS Analysis

The LC-MS/MS experiments were performed as described previously [30]. Briefly, peptides were separated on a hand-pulled fused silica microcapillary (75 μm × 15 cm, packed with Magic C18AQ; Michrom Bioresources, Auburn, CA, USA) using a 75 min linear gradient ranging from 7% to 32% acetonitrile in 0.1% formic acid at 300 nL/min (Thermo EASY-nLC system). Samples were then analyzed on an LTQ Velos Pro Orbitrap mass spectrometer (Thermo, San Jose, Cal USA) using either a top 20 CID method (for protein identification) or a top 10 higher-energy collisional dissociation (HCD) method (for TMT protein quantification). Precursor ions were selected using an isolation window of 2 Th, and dynamic exclusion was enabled in which the same precursor ion was excluded from repeated MS₂ analyses for 60 s.

MS/MS spectra were searched against a composite database of the human IPI protein database (ver. 3.60, which contains a total of 80,412 entries) and its reversed complement using the Sequest (Rev28) algorithm. Search parameters allowed for a static modification of 57.02146, 229.16293, and 229.16293 Da for Cys (Carbamidomethyl), Lys (TMT), and peptide N-terminus (TMT), respectively. A dynamic modification of oxidation (15.99491 Da) was considered for Met. Search results were filtered to include <1% (at both peptide and protein levels) matches to the reverse database by the linear discriminator function [32]. Protein TMT ratios were determined using the procedure as described previously [30]. Specifically, a 0.03 Th window was scanned around the theoretical *m/z* of each

reporter ion (126:126.127725; 127:127.124760; 128:128.134433; 129:129.131468; 130:130.141141; 131:131.138176) to detect the presence of these ions, and the maximum intensity of each ion was extracted. For each reporter ion channel, the observed signal-to-noise ratio was summed across all quantified proteins and was normalized (equal amounts of proteins were digested and labeled for the six different channels). Gene ontology (GO) analyses were performed using David and Panther databases [30, 33, 34].

Meta-Analysis of Genes Expression in Lung ADC and SqCC

RNAseq data were downloaded from the TCGA portal. The archive filenames were “unc.edu_LUAD.IlluminaHiSeq_RNASeqV2.1.13.0” for ADC and “unc.edu_LUSC.IlluminaHiSeq_RNASeqV2.1.10.0” for SqCC. These files represent level 3 (processed) data. The unzipped files consisted of 490 ADC and 489 SqCC samples, one file per sample. The FPKM values were pooled into one spreadsheet, log₂-transformed, and median-normalized across all samples. Expression comparison between ADC and SqCC was performed and box plots generated with significance determined by unpaired *t*-test. To assess differences between overlapping RNA expression distributions for each gene in SqCC versus ADC comparisons, a Kolmogorov–Smirnov test was applied with D- and P-values calculated.

Results

Tissue with residual OCT compound contains polymers like polyvinyl alcohol and polyethylene glycol, which can interfere with LC-MS/MS performance [27, 28]. Supplementary Figure S1 shows an example of typical polymer signals generated from direct analysis of unprocessed OCT-embedded samples (MS1 spectrum extracted from retention time at approximately 59 min). In this case, the peptide signals were completely suppressed by ions derived from OCT polymers, which are in large fold-excess.

Because OCT contains polymers that are mostly water-soluble, we reason that they could potentially be removed by the TCA precipitation method [35]. We generated lysates from surgically resected SqCC samples that were embedded in OCT, and performed TCA precipitation. The extracted proteins were extensively washed to ensure complete removal of the OCT polymers.

Shotgun LC-MS/MS proteome analysis followed by database query resulted in the identification of 9228 proteins (1% FDR at the protein level, 7339 proteins identified with two or more unique peptides) in the OCT-embedded tissue sample. As a comparison, we also performed similar shotgun proteomic analysis on a frozen sample (both the OCT-embedded and frozen specimen were derived from the same tumor), from which we identified 9637 proteins (1% FDR at the protein level, 7710 proteins identified with two or more unique peptides). A total of 8583 proteins were commonly identified between the two samples (Figure 1a). Using gene ontology

(GO) analysis (www.pantherdb.org/), we observed no bias in protein recovery from any particular compartment or organelle and similar numbers of categorized proteins are recovered from either OCT embedded or frozen tissue (Figure 1b). Moreover, we observed no bias in identification of the whole proteome based on analysis of GO biological process with over 70% of proteins annotated in the top five processes, which include metabolic, cellular, biological regulation, localization, and development (Figure 1c). These results indicate that our TCA-based protocol resulted in excellent recovery of the proteome from OCT-embedded samples and did not have any appreciable effects on the number or identity of proteins detected during MS analysis.

SqCC is usually identified by its classic pathologic features, which include keratinization, intercellular bridges, which distinguish it from other NSCLC subtypes [36]. One of the highly expressed proteins identified from our proteomic analysis is KRT17 (Figure 1d), but many of the other keratins were not detected. This was not a surprising result because the patient tumor used for this study was deemed “poorly differentiated” in standard pathology analysis. Recent work from Wilkerson et al. proposed a SqCC mRNA expression-based classification schema with four subgroups comprised of primitive, classic, basal, and secretory [37]. Amongst the four proposed classes, the “primitive” tumors exhibited the most poorly differentiated features, had the lowest overall survival (OS), worst recurrence free survival (RFS), and exhibited the strongest cell proliferation mRNA signature with MCM6 as the most predictive biomarker [37]. Interestingly, MCM6 was one of the highly expressed proteins detected in this SqCC tissue (Figure 1d).

Using our MS data, we chose to examine two groups of proteins; one group related to cell proliferation, which includes MCM family members, Ki67 and PCNA, and a second group commonly associated with overexpression in squamous tumors such as KRT17, SERPINB4, SERPINB3, SERPINB5, DCUN1D1, and S100A2 [36, 38–41]. The total peptide spectra count for all selected proteins from both OCT and frozen samples were normalized using GAPDH and ACTB. In this case, we observed no significant difference in the expression of these two groups of proteins between the two samples (Figure 1d). Moreover, GO analysis (David 6.7) revealed that the top 1000 highly expressed proteins (highest number of detected peptides) were enriched for translation and RNA processing, which are strongly related to cell proliferation (Supplementary Figures S2). These results indicate that the TCA precipitation procedure successfully removes OCT and does not significantly influence the number of proteins detected or the biological pathways they represent.

We then set out to demonstrate the applicability of our TCA-based protocol to more clinically challenging samples by using OCT-embedded CNB tumor and normal specimens. CNB samples preserved in OCT are difficult for quantitative proteomic analysis not only because of the purification required but also due to the minimal amount of tumor as demonstrated in Figure 2a. Specifically, after TCA precipitation, we performed BCA quantification, and found that a total of

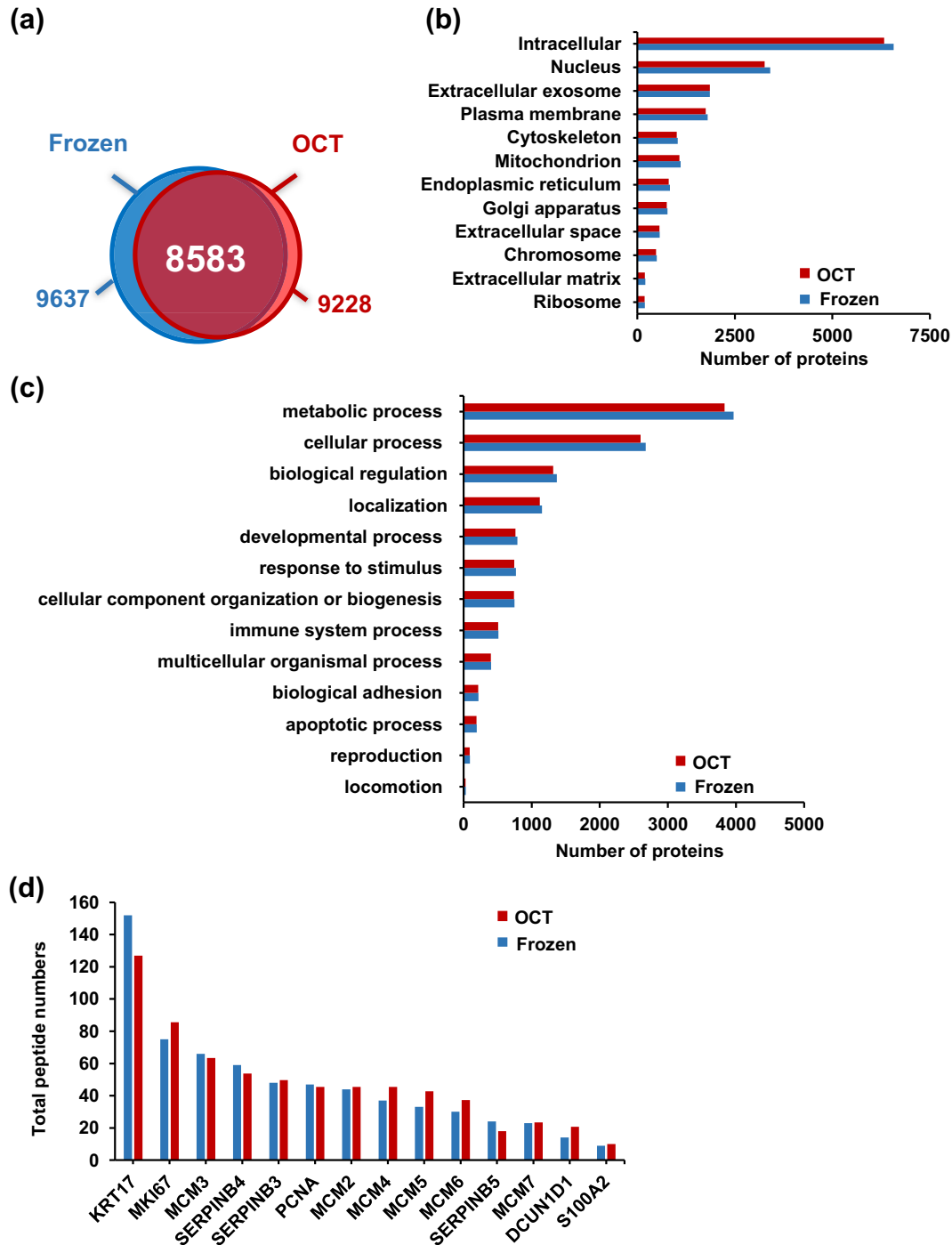


Figure 1. Qualitative evaluation of the TCA-based OCT-removal protocol for surgically resected tumor samples. **(a)** Proteins identified from a frozen surgically resected SqCC tumor, and the same tumor that was embedded in OCT. The Venn diagram shows the proteins that were commonly identified from the two types of samples. Gene ontology analysis shows the proteins identified from the frozen and OCT-embedded samples were equally represented in terms of cellular component **(b)** and biological process **(c)**. **(d)** The number of peptides identified from several representative proteins in the abovementioned two samples. Peptide numbers were normalized by GAPDH and actin

42 μ g and 166 μ g protein was recovered from the CNB-normal and CNB-tumor sample, respectively.

To assess the reproducibility of our method, we performed triple replicate analyses for each sample (channels 126,127, and 128 for CNB-tumor and channels 129,130, and 131 for

CNB-normal). After TMT labeling, peptides were combined, and were subject to bRPLC fractionation. A total of 17 fractions were collected, which were analyzed by LC-MS/MS experiments. From this set of samples, we were able to identify and quantify 5411 proteins (1% FDR at the protein level, 4505

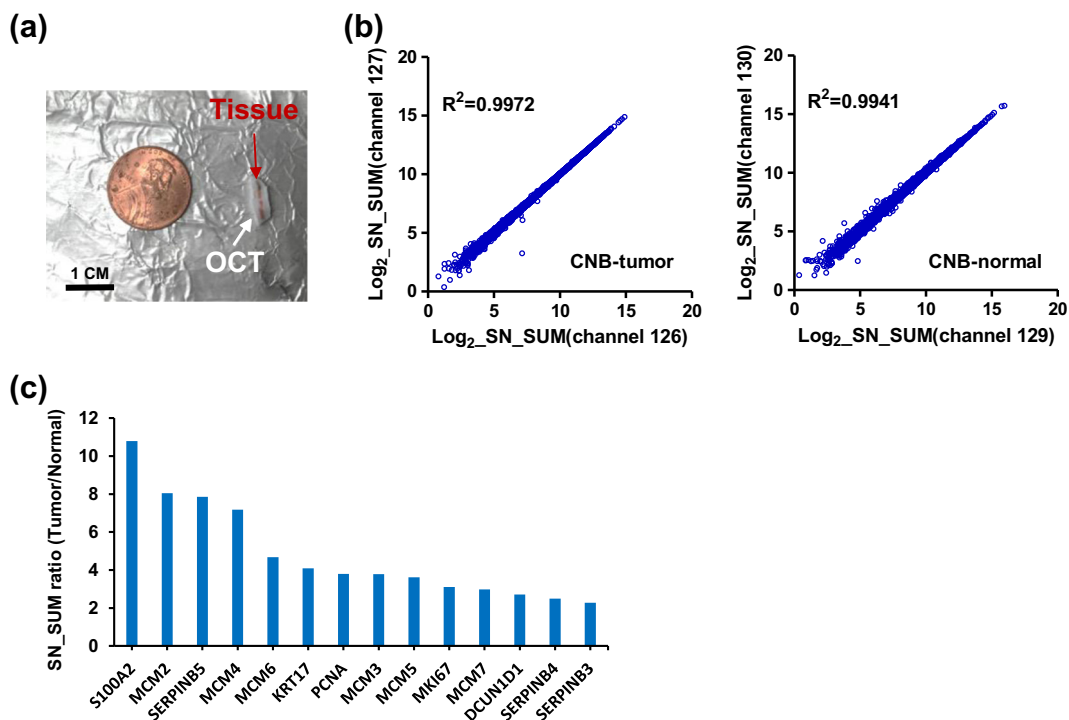


Figure 2. Quantitative proteomic analysis of OCT-embedded core needle biopsy (CNB) samples. **(a)** A photograph of the OCT-embedded CNB sample. Red and white arrows indicate the tissue and OCT compound, respectively. **(b)** Three technical replicate experiments were performed for each sample, i.e., 126,127, and 128 for CNB-tumor; 129, 130, and 131 for CNB-normal. The TMT reporter ion intensity (Signal-to-noise, SN) from all the peptides of a protein was summed, and was plotted, on a Log₂ scale, for the technical replicate experiments. **(c)** Representative proteins that are overexpressed in CNB-tumor compared with CNB-normal

proteins quantified based on two or more peptides). We achieved excellent correlation in signal intensities between TMT-replicates of tumor tissue (126 to 127, $R^2 = 0.9972$) and normal tissue (129 to 130, $R^2 = 0.9941$) (Figure 2b, Supplementary Figure S3A, B).

After proper quality control analysis, we interrogated the dataset for protein expression differences between CNB-tumor versus CNB-normal. To determine the expression level for each protein, we summed signal-to-noise (SN) intensities for channels 126,127, and 128 for CNB-tumor and SN intensities of channels 129,130, and 131 for CNB-normal. Each data point for the graph in Figure 3a represents the log₂-transformed value of the tumor/normal ratio for a single protein. We selected proteins that were up- or down-regulated by at least 3-fold in tumor compared with normal and performed GO analysis. Figure 3b demonstrates that proteins up-regulated in tumor tissue are enriched for biological processes such as translation, ribosome biogenesis, and DNA replication, which are supportive of the cell proliferation signature noted in poorly differentiated SqCC. Moreover, cellular component analysis for the tumor up-regulated protein indicates that these proteins represent a diverse array of cellular organelles, including membrane-enclosed lumen, nucleolus, and ribosome (Supplementary Figure S3C).

When we looked at the GO of the down-regulated proteins in the tumor samples (Figure 3c) we noted these proteins are strongly linked to acute inflammatory response, response to

wounding, cell adhesion, and coagulation with most of the proteins associated with the plasma membrane and the extracellular compartment (Supplementary Figure S3D). Down-regulation of cell adhesion related proteins (e.g., NCAM1, NCAM2, CADM1, and DPT) have been associated with more aggressive tumors and suggests that the poorly differentiated SqCC tumor may have increased mobility, invasive capacity, and poor prognosis [42–44]. Another set of tumor-associated down-regulated proteins that drew our attention was related to inflammatory response and wound healing (e.g., C1QA, C4BPA, C4A, C5, C7, C8A, C9, and C3). Many of these proteins are serum proteins and their downregulation in tumor specimens may reflect the high perfusion of normal lung tissue. These results nevertheless suggest that the microenvironment of the tumor is significantly different from that of the normal lung, with important immune components differentially expressed in these two types of samples.

We then chose a panel of the highly expressed proliferative as well as SqCC marker proteins identified from Figure 1d and investigated their expression in tumor versus normal comparisons (note that all the surgically resected samples and CNBs were from the sample patient). Indeed, we found that the MCM family (see below for further discussion) as well as MKI67, PCNA, KRT17, SERPINB4, SERPINB3, SERINB5, DCUN1D1, and S100A2 were expressed more than 2-fold higher in tumor compared with normal CNBs (Figure 2c and Figure 4). The protein showing the greatest bias in tumor

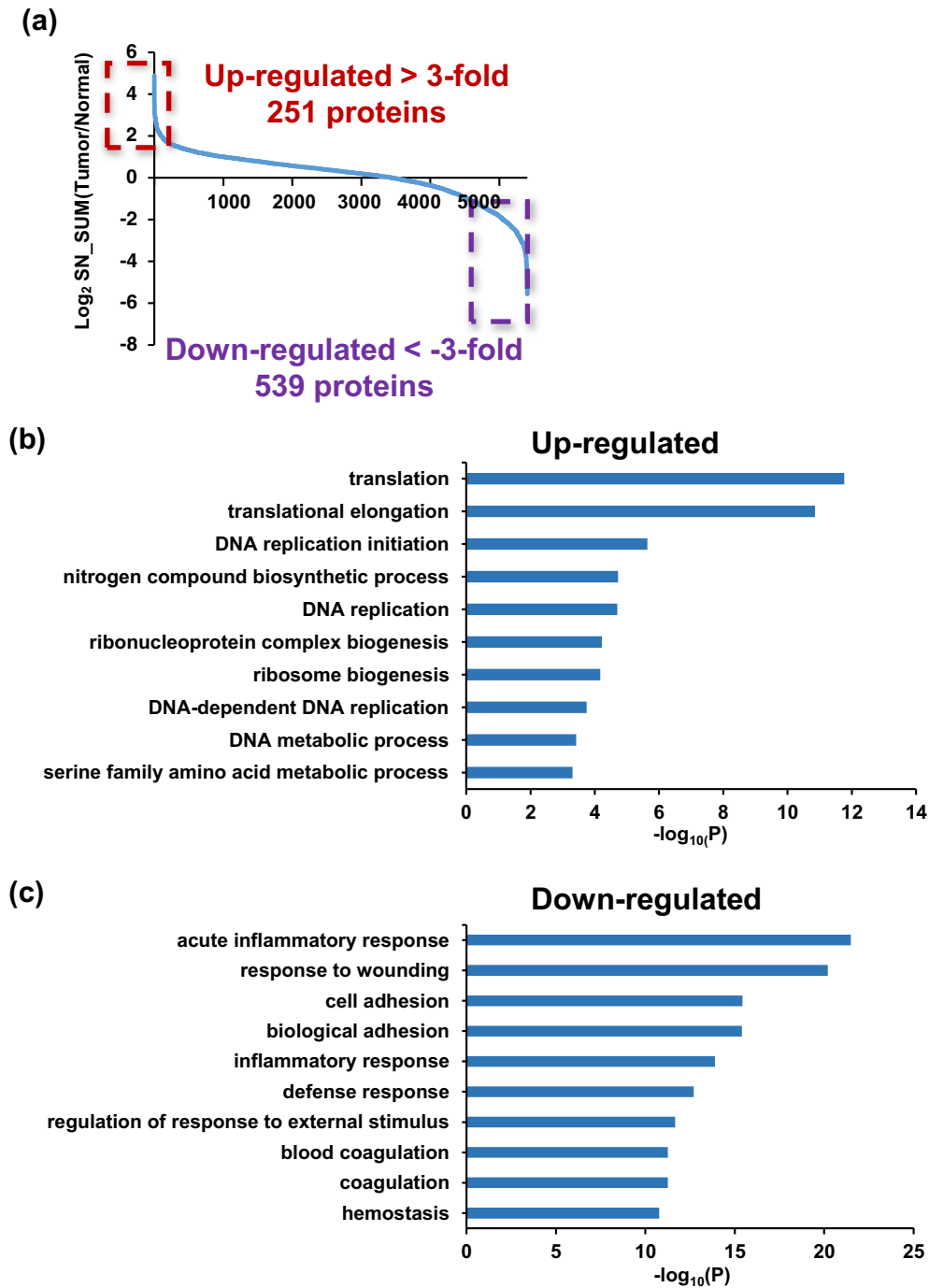


Figure 3. Gene ontology analysis of the SqCC proteome. (a) Comparison of protein expression between CNB-tumor and CNB-normal samples (ratio converted to a log₂ scale). The OCT-embedded CNB-tumor and CNB-normal samples were subject to the TCA-based OCT removal protocol. Proteins were extracted and digested. The resulting peptides were labeled with the TMT reagents, and were quantified using LC-MS/MS experiments. Proteins that are up-regulated or down-regulated by at least 3-fold are highlighted. Biological processes that are represented by these proteins are indicated in (b) (up-regulated) and (c) (down-regulated)

expression was S100A2, which has been identified previously as a potential squamous lung cancer marker [39]. TMT-MS analysis showed that the S100A2 protein level was up-regulated by more than 10-fold in the tumor. For example, we observed strong signal for an S100A2 peptide

(YSCQEGDK) in the tumor sample, whereas it was virtually undetectable in the normal CNB specimen (Figure 4c). We also show similar results for two other proteins, KRT17 (Figure 4a) and SERPINB5 (Figure 4b), both of which have been strongly associated with squamous lung cancer [41, 45]. One distinct

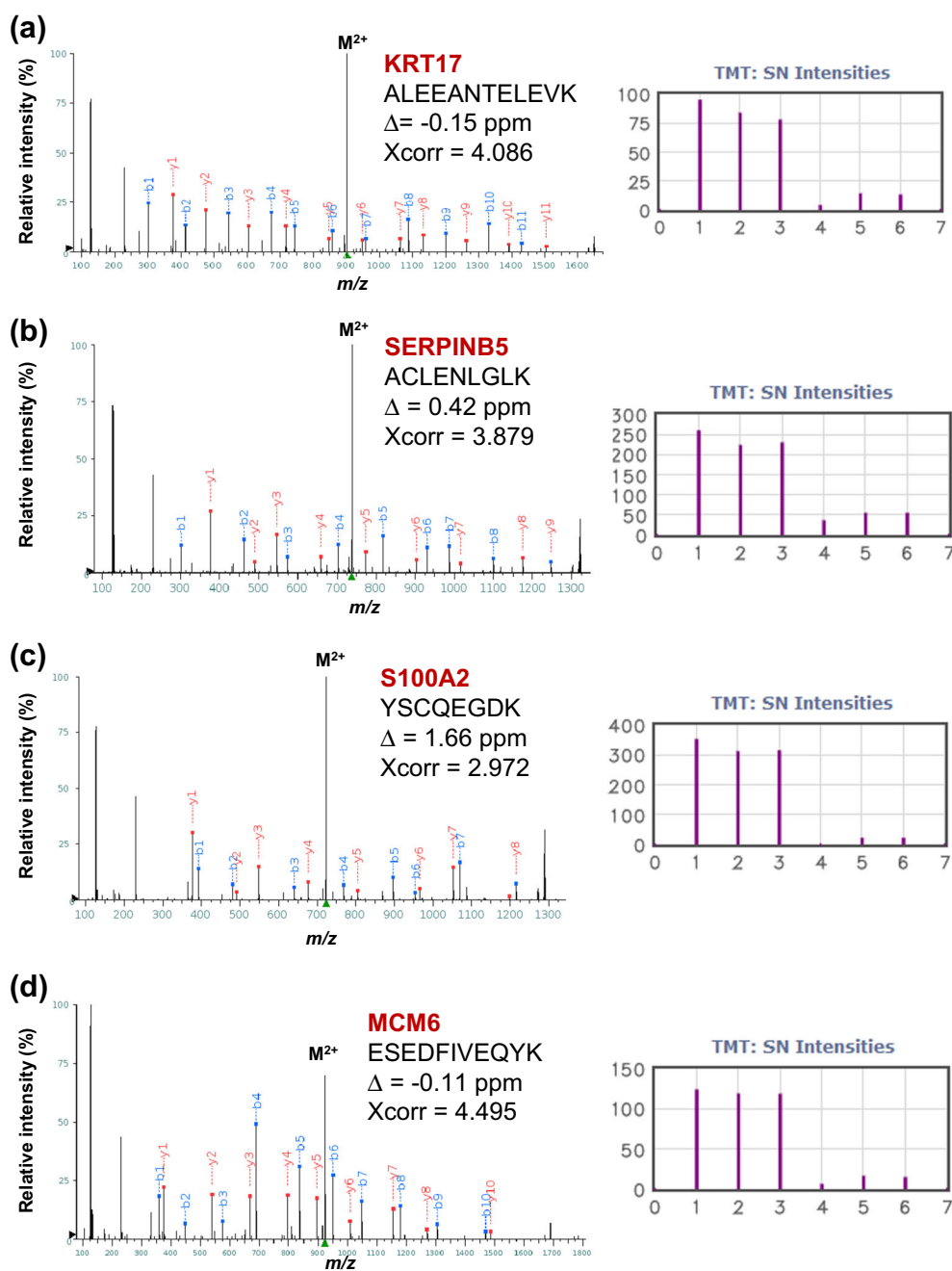


Figure 4. Representative overexpressed proteins identified from the OCT-embedded core needle biopsy samples. The MS2 spectra are shown for (a) KRT17, (b) SERPINB5, (c) S100A2, and (d) MCM6. Three replicate analyses were performed (126, 127, and 128 for CNB-tumor, and 129, 130, and 131 for CNB-normal). The TMT ion cluster is shown with the following channel information: channel 1, 2, 3, 4, 5, 6 for TMT 126, 127, 128, 129, 130, and 131, respectively

advantage of quantitative proteome analysis over standard pathology techniques is the ability to simultaneously monitor multiple closely related proteins. We successfully identified and quantified six related MCM polypeptides (MCM2 through MCM7), which are known to function as a DNA helicase complex required for DNA replication [46]. We highlight the data for MCM6 as it has been proposed to serve as a biomarker for primitive or poorly differentiated SqCC tumors (Figure 4d) [37]. Analysis of 489 SqCC and 490 ADC patient samples

from the TCGA mRNA expression dataset indicate that these markers (KRT17, SERPINB5, S100A2, and MCM6) are more highly expressed in SqCC than ADC (Figure 5). Further analysis using the Kolmogorov–Smirnov test (KS test) shows that although the expression distributions for each gene somewhat overlap, SqCC versus ADC comparisons are significantly different (Supplementary Figure S5). We believe that the identification and quantitation of these proteins highlights the potential applicability of TMT mass spectrometry as a diagnostic

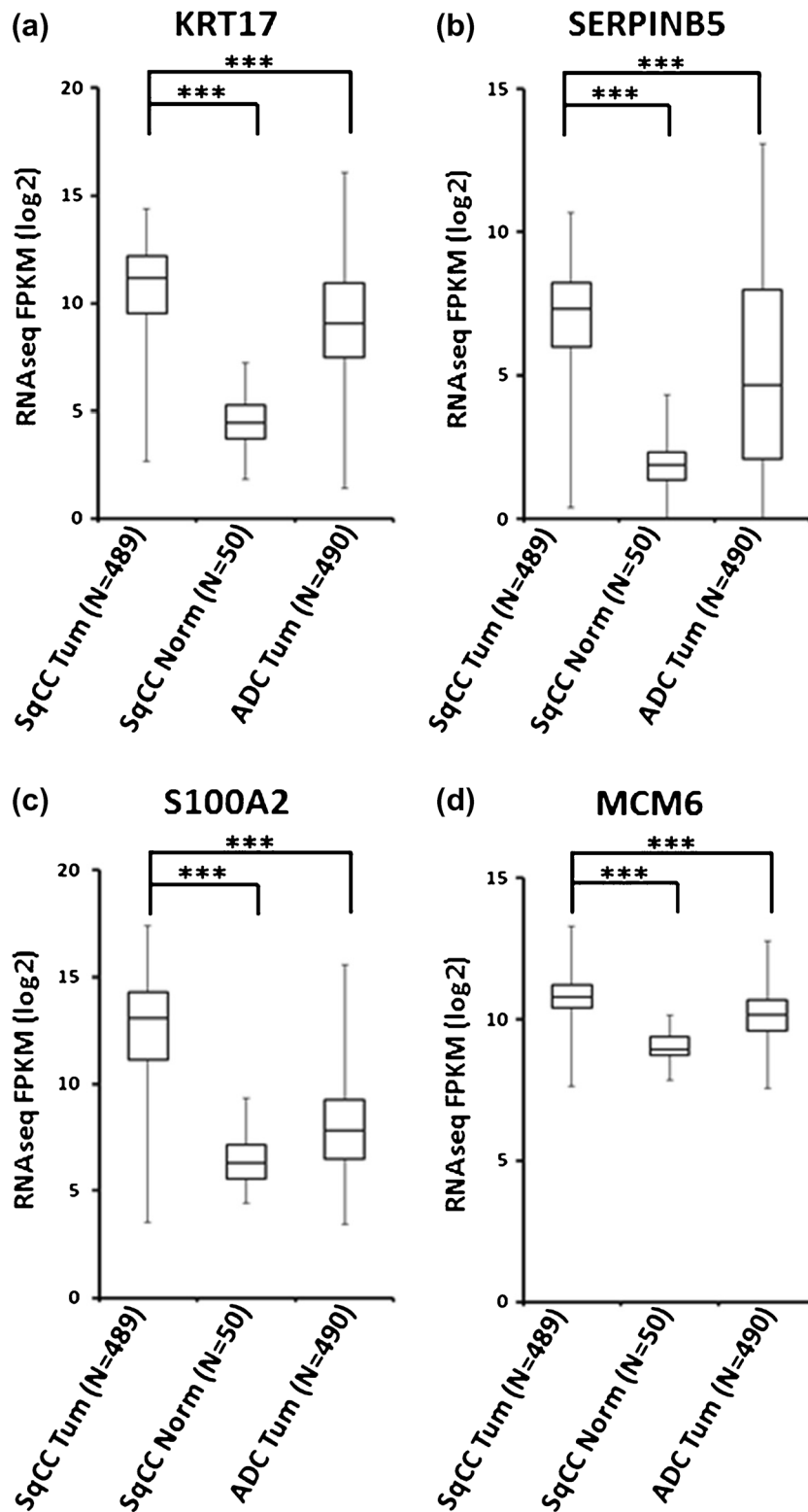


Figure 5. Meta-analysis of the expression of representative SqCC-specific proteins in different lung cancer subtypes. RNA expression levels for (a) KRT17, (b) SERPINB5, (c) S100A2, and (d) MCM6 were extracted from publicly available TCGA datasets on lung SqCC, ADC, and SqCC matched normal tissues, and are shown as box and whisker plots (***) = unpaired *t*-test, $P < 1E^{-15}$

tool. The successful identification of predictable squamous markers motivated us to look into the data for less well

characterized protein biomarkers and potentially interesting drug targets. For example, we identified DHFR (dihydrofolate

reductase, target for methotrexate) as a protein that is highly overexpressed in the SqCC tumor-CNB sample (~8-fold compared with normal) (Supplementary Figure S4A).

Considering the strong correlation of smoking with SqCC along with the proliferative signature noted in poorly differentiated (primitive) tumors, we examined proteins involved in DNA damage repair [47]. We focused on proteins up-regulated by more than 3-fold and discovered four proteins (FEN1, POLA1, SOD2, and TRIP13) involved with double-strand break repair (P -value = 0.066, data not shown). From this group, we provide more data on TRIP13 (Supplementary Figure S4B), which is a protein that has recently been shown to bind to DNA-PKcs and mediate non-homologous end joining (NHEJ) repair of DNA-strand breaks. Overexpression of TRIP13 is associated with aggressive tumors, and may contribute to platinum-resistance in squamous head and neck cancers [48].

Finally, to extend these findings, we further examined the top 200 most overexpressed proteins in tumor versus normal comparison and entered them into the Drug Gene Interaction Database (DGIdb, <http://dgidb.genome.wustl.edu/>). After curating the list, we chose six examples of genes and their associated drugs, some of which are already used in the cancer clinic and some that are in various stages of clinical development (Table 1). As an example of a novel squamous drug target identified in this screen, we focused on KIF11, which is involved in spindle dynamics during mitosis [49]. KIF11 mRNA is also overexpressed in TCGA comparison of SqCC to ADC (data not shown). The effectiveness of drugs targeting kinesin motors like ispinesib and filanesib are being explored in clinical trials currently [50], but we are not aware of any publications discussing these drugs in lung SqCC.

Discussion

In this study, we provide a simple method for purifying whole proteome from OCT embedded lung tumor tissue. The use of OCT embedding in the cancer clinic and surgical suite is common because of its simplicity and the ability to aid in the delivery of quick diagnostic information about unknown tissue masses. To our knowledge, the long-term tissue and molecular stability of OCT tumor-embedded specimens has not been well studied but it is believed that they are stable for reasonable

durations if care is taken to keep them frozen [23]. OCT-tumor embedded specimens are not widely used by pathologists who need to perform general H and E staining and downstream IHC subtyping analysis as FFPE produces superior results for sectioning and staining. Tumors fixed and embedded in FFPE have proven to be very stable over long durations (decades) and have been successfully used in many archival genomics studies. However, for characterization of whole proteome, the inherent nature of fixation becomes problematic in terms of protein recovery and extensive chemical modification of native proteins. A number of attempts have been made to characterize the proteome from FFPE using LC-MS/MS, with varying success [51–53]. Conservative estimates suggest 15%–20% of the proteome may be inaccessible after formalin fixation [25]. With regards to OCT-embedded specimens, which do not involve any chemical cross-linking reagents, whole proteome analysis could be complicated by the large amounts of soluble polymers present in OCT [27, 28]. We have addressed this issue by developing a pipeline that allows for quantitative proteome analysis using a simple TCA-based OCT removal and protein purification scheme.

Implementation of our protocol resulted in identification of more than 9200 and 5400 proteins from OCT-embedded surgically resected and CNB samples, respectively. We demonstrated excellent concordance with respect to protein identity and intracellular localization between frozen SqCC tumor tissue and OCT-embedded tissue from the same patient. Pathology (H and E and IHC) subtyped this patient's SqCC tumor as "poorly differentiated," which has been noted for strong cell proliferation signatures in other studies. MS examination of the most highly expressed proteins in tumor tissue was in agreement, as many of the proteins are suggestive of high proliferative capacity.

Because resected tumor is not always available, we challenged the detection limits of our platform using OCT-embedded core needle biopsies of both tumor and normal tissue. Using a quantitative TMT platform, we were able to identify and quantify more than 5400 proteins from OCT-embedded tumor and normal tissue. Moreover, we were able to demonstrate that results from our TCA-based protein purification protocol led to the identification of predictably up-regulated squamous specific proteins as well as potentially novel squamous related proteins. It is important to note that although the magnitude of change as measured by the TMT assay could be suppressed due to co-isolated ions, the proteins defined as regulated using the TMT approach are usually undistorted [54].

The most recent report from the World Health Organization (2015) has begun to recognize the inherent heterogeneity in lung cancer and the need for further classification of lung cancer based on expanded IHC for small biopsies, cytology, and resected specimens [55]. Publication of these new guidelines (the first since 2004) is tacit recognition of the rapid discoveries regarding the heterogeneity and targeted therapeutic options being discovered in lung cancer [56]. This means that more high quality, in-depth pathologic diagnostics will be

Table 1. Potential Druggable Proteins that are Overexpressed in the SqCC Tumor Sample (OCT-embedded core needle biopsy samples)

Gene	T/N ratio	Drugs
POLA1	7.27	Clofarabine, fludarabine, nelarabine, cladribine
PDE4D	4.46	Dyphylline, etazolol, otryxyphylline, piclamilast, ibudilast, rofulimast, iloprost
ESRRA	4.29	Genistein, bichanin A, XCT790, daidzein, troglitazone
PDK1	3.97	CPI-613
KIF11	3.67	Filanesib, ispinesib, monastrol, SB-743921
RRM1	3.45	Gemcitabine, fludarabine, clofarabine, cladribine, imexon

T = tumor; N = normal.

required for lung cancer since that is the clinical basis for determining therapeutic regimen. Although IHC is and remains the gold standard for clinical diagnostics, the results of our analysis are encouraging in terms of moving forward with the use of LC-MS/MS as a future tool that could aid in pathology and diagnostic analysis. In this regard, the ability to detect a larger panel of proteins in a high throughput manner, while difficult for standard clinical pathology, is a distinct advantage of the proteomic technologies, which could potentially provide more information with regards to tumor biomarkers and diagnostics.

Acknowledgments

This work was supported in part by grants from the Welch Foundation (I-1800 to Y.Y.), National Institutes of Health (NIH) (GM114160 to Y.Y.), American Cancer Society (RSG-15-062-01-TBE to Y.Y.), University of Texas Specialized Program of Research Excellent in Lung Cancer (NIH CA70907 to J.D.M.) and National Natural Science Foundation of China (31571021 to G.N.). Y.Y. is a Virginia Murchison Linthicum Scholar in Medical Research and a CPRIT Scholar in Cancer Research. G.N. is a scholar of National Distinguished Young Scientist of China.

All the raw files have been deposited into the Chorus database, and can be downloaded using the following links: <https://chorusproject.org/anonymous/download/experiment/70bfa6688de0436d867d6c52654dcac0>
<https://chorusproject.org/anonymous/download/experiment/c7f7408fbafc44daa2dd4b47d8ba4edc>
<https://chorusproject.org/anonymous/download/experiment/0ed6fd79d2af4b6eab30e7d05bdcd20c>

References

- Siegel, R.L., Miller, K.D., Jemal, A.: Cancer statistics, 2016. *CA Cancer J. Clin.* **66**, 7–30 (2016)
- Travis, W.D.: Pathology of lung cancer. *Clin. Chest Med.* **32**, 669–692 (2011)
- Johnson, D.H., Fehrenbacher, L., Novotny, W.F., Herbst, R.S., Nemunaitis, J.J., Jablons, D.M., Langer, C.J., DeVore, R.F.I.I.I., Gaudreault, J., Damico, L.A., Holmgren, E., Kabbinnavar, F.: Randomized phase II trial comparing bevacizumab plus carboplatin and paclitaxel with carboplatin and paclitaxel alone in previously untreated locally advanced or metastatic non-small cell lung cancer. *J. Clin. Oncol.* **22**, 2184–2191 (2004)
- Quoix, E.: Therapeutic options in older patients with metastatic non-small cell lung cancer. *Ther. Adv. Med. Oncol.* **4**, 247–254 (2012)
- Pesch, B., Kendzia, B., Gustavsson, P., Jockel, K.H., Johnen, G., Pohlbeln, H., Olsson, A., Ahrens, W., Gross, I.M., Bruske, I., Wichmann, H.E., Merletti, F., Richiardi, L., Simonato, L., Fortes, C., Siemiatycki, J., Parent, M.E., Consonni, D., Landi, M.T., Caporaso, N., Zaridze, D., Cassidy, A., Szeszenia-Dabrowska, N., Rudnai, P., Lissowska, J., Stucker, I., Fabianova, E., Dumitru, R.S., Bencko, V., Foretova, L., Janout, V., Rudin, C.M., Brennan, P., Boffetta, P., Straif, K., Bruning, T.: Cigarette smoking and lung cancer-relative risk estimates for the major histological types from a pooled analysis of case-control studies. *Int. J. Cancer* **131**, 1210–1219 (2012)
- Deneffe, G., Lacquet, L.M., Verbeken, E., Vermaut, G.: Surgical treatment of bronchogenic carcinoma: a retrospective study of 720 thoracotomies. *Ann. Thorac. Surg.* **45**, 380–383 (1988)
- Mountain, C.F.: Assessment of the role of surgery for control of lung cancer. *Ann. Thorac. Surg.* **24**, 365–373 (1977)
- Molina, J.R., Yang, P., Cassivi, S.D., Schild, S.E., Adjei, A.A.: Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin. Proc.* **83**, 584–594 (2008)
- Brahmer, J., Reckamp, K.L., Baas, P., Crino, L., Eberhardt, W.E., Poddubskaya, E., Antonia, S., Pluzanski, A., Vokes, E.E., Holgado, E., Waterhouse, D., Ready, N., Gainor, J., Aren Frontera, O., Havel, L., Steins, M., Garassino, M.C., Aerts, J.G., Domine, M., Paz-Ares, L., Reck, M., Baudelet, C., Harbison, C.T., Lestini, B., Spigel, D.R.: Nivolumab versus docetaxel in advanced squamous-cell non-small cell lung cancer. *N. Engl. J. Med.* **373**, 123–135 (2015)
- McGranahan, N., Furness, A.J., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S.K., Jamal-Hanjani, M., Wilson, G.A., Birkbak, N.J., Hiley, C.T., Watkins, T.B., Shafi, S., Murugaesu, N., Mitter, R., Akarca, A.U., Linares, J., Marafioti, T., Henry, J.Y., Van Allen, E.M., Miao, D., Schilling, B., Schadendorf, D., Garraway, L.A., Makarov, V., Rizvi, N.A., Snyder, A., Hellmann, M.D., Merghoub, T., Wolchok, J.D., Shukla, S.A., Wu, C.J., Peggs, K.S., Chan, T.A., Hadrup, S.R., Quezada, S.A., Swanton, C.: Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016)
- Pikor, L.A., Ramnarine, V.R., Lam, S., Lam, W.L.: Genetic alterations defining NSCLC subtypes and their therapeutic implications. *Lung Cancer* **82**, 179–189 (2013)
- Sun, S., Schiller, J.H., Gazdar, A.F.: Lung cancer in never smokers—a different disease. *Nat. Rev. Cancer* **7**, 778–790 (2007)
- Cancer Genome Atlas Research Network: Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012)
- Weiss, J., Sos, M.L., Seidel, D., Peifer, M., Zander, T., Heuckmann, J.M., Ullrich, R.T., Menon, R., Maier, S., Soltermann, A., Moch, H., Wagener, P., Fischer, F., Heynck, S., Koker, M., Schottle, J., Leenders, F., Gabler, F., Dabow, I., Querings, S., Heukamp, L.C., Balke-Want, H., Ansen, S., Rauh, D., Baessmann, I., Altmüller, J., Wainer, Z., Conron, M., Wright, G., Russell, P., Solomon, B., Brambilla, E., Brambilla, C., Lorimier, P., Sollberg, S., Brustugun, O.T., Engel-Riedel, W., Ludwig, C., Petersen, I., Sanger, J., Clement, J., Groen, H., Timens, W., Sietsma, H., Thunnissen, E., Smit, E., Heideman, D., Cappuzzo, F., Ligorio, C., Damiani, S., Hallek, M., Beroukhi, R., Pao, W., Klebl, B., Baumann, M., Buettner, R., Ernestus, K., Stöbel, E., Wolf, J., Numberg, P., Perner, S., Thomas, R.K.: Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Sci. Transl. Med.* **2**, 62–93 (2010)
- Ramos, A.H., Dutt, A., Mermel, C., Perner, S., Cho, J., Lafargue, C.J., Johnson, L.A., Stiedl, A.C., Tanaka, K.E., Bass, A.J., Barretina, J., Weir, B.A., Beroukhi, R., Thomas, R.K., Minna, J.D., Chiriac, L.R., Lindeman, N.I., Giordano, T., Beer, D.G., Wagner, P., Wistuba, I.I., Rubin, M.A., Meyerson, M.: Amplification of chromosomal segment 4q12 in non-small cell lung cancer. *Cancer Biol. Ther.* **8**, 2042–2050 (2009)
- Takahashi, Y., Ishii, G., Taira, T., Fujii, S., Yanagi, S., Hishida, T., Yoshida, J., Nishimura, M., Nomori, H., Nagai, K., Ochiai, A.: Fibrous stroma is associated with poorer prognosis in lung squamous cell carcinoma patients. *J. Thorac. Oncol.* **6**, 1460–1467 (2011)
- Girard, L., Rodriguez-Canales, J., Behrens, C., Thompson, D.M., Botros, I.W., Tang, H., Xie, Y., Rekhman, N., Travis, W.D., Wistuba, I.I., Minna, J.D., Gazdar, A.F.: An expression signature as an aid to the histologic classification of non-small cell lung cancer. *Clin. Cancer Res.* **22**, 4880–4889 (2016)
- Travis, W.D., Brambilla, E., Noguchi, M., Nicholson, A.G., Geisinger, K., Yatabe, Y., Ishikawa, Y., Wistuba, I., Flieder, D.B., Franklin, W., Gazdar, A., Hasleton, P.S., Henderson, D.W., Kerr, K.M., Petersen, I., Roggli, V., Thunnissen, E., Tsao, M.: Diagnosis of lung cancer in small biopsies and cytology: implications of the 2011 International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification. *Arch. Pathol. Lab. Med.* **137**, 668–684 (2013)
- Rekhman, N., Ang, D.C., Sima, C.S., Travis, W.D., Moreira, A.L.: Immunohistochemical algorithm for differentiation of lung adenocarcinoma and squamous cell carcinoma based on large series of whole-tissue sections with validation in small specimens. *Mod. Pathol.* **24**, 1348–1359 (2011)
- Cregger, M., Berger, A.J., Rimm, D.L.: Immunohistochemistry and quantitative analysis of protein expression. *Arch. Pathol. Lab. Med.* **130**, 1026–1030 (2006)

21. Nilsson, T., Mann, M., Aebersold, R., Yates III, J.R., Bairoch, A., Bergeron, J.J.: Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat. Methods* **7**, 681–685 (2010)
22. Lipson, D., Capelletti, M., Yelensky, R., Otto, G., Parker, A., Jarosz, M., Curran, J.A., Balasubramanian, S., Bloom, T., Brennan, K.W., Donahue, A., Downing, S.R., Frampton, G.M., Garcia, L., Juhn, F., Mitchell, K.C., White, E., White, J., Zwirko, Z., Peretz, T., Nechushtan, H., Soussan-Gutman, L., Kim, J., Sasaki, H., Kim, H.R., Park, S.I., Ercan, D., Sheehan, C.E., Ross, J.S., Cronin, M.T., Janne, P.A., Stephens, P.J.: Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. *Nat. Med.* **18**, 382–384 (2012)
23. Steu, S., Baucamp, M., von Dach, G., Bawohl, M., Dettwiler, S., Storz, M., Moch, H., Schraml, P.: A procedure for tissue freezing and processing applicable to both intra-operative frozen section diagnosis and tissue banking in surgical pathology. *Virchows Arch.* **452**, 305–312 (2008)
24. Maes, E., Broeckx, V., Mertens, I., Sagaert, X., Prenen, H., Landuyt, B., Schoofs, L.: Analysis of the formalin-fixed paraffin-embedded tissue proteome: pitfalls, challenges, and future perspectives. *Amino Acids* **45**, 205–218 (2013)
25. Crockett, D.K., Lin, Z., Vaughn, C.P., Lim, M.S., Elenitoba-Johnson, K.S.: Identification of proteins from formalin-fixed paraffin-embedded cells by LC-MS/MS. *Lab. Investig.* **85**, 1405–1415 (2005)
26. Sprung Jr., R.W., Brock, J.W., Tanksley, J.P., Li, M., Washington, M.K., Slebos, R.J., Liebler, D.C.: Equivalence of protein inventories obtained from formalin-fixed paraffin-embedded and frozen tissue in multidimensional liquid chromatography-tandem mass spectrometry shotgun proteomic analysis. *Mol. Cell. Proteomics* **8**, 1988–1998 (2009)
27. Zhang, W., Sakashita, S., Taylor, P., Tsao, M.S., Moran, M.F.: Comprehensive proteome analysis of fresh frozen and optimal cutting temperature (OCT) embedded primary non-small cell lung carcinoma by LC-MS/MS. *Methods* **81**, 50–55 (2015)
28. Shah, P., Zhang, B., Choi, C., Yang, S., Zhou, J., Harlan, R., Tian, Y., Zhang, Z., Chan, D.W., Zhang, H.: Tissue proteomics using chemical immobilization and mass spectrometry. *Anal. Biochem.* **469**, 27–33 (2015)
29. Tian, Y., Bova, G.S., Zhang, H.: Quantitative glycoproteomic analysis of optimal cutting temperature-embedded frozen tissues identifying glycoproteins associated with aggressive prostate cancer. *Anal. Chem.* **83**, 7013–7019 (2011)
30. Hu, R., Huffman, K.E., Chu, M., Zhang, Y., Minna, J.D., Yu, Y.: Quantitative secretomic analysis identifies extracellular protein factors that modulate the metastatic phenotype of non-small cell lung cancer. *J. Proteome Res.* **15**, 477–486 (2016)
31. Paulo, J.A., Gygi, S.P.: Nicotine-induced protein expression profiling reveals mutually altered proteins across four human cell lines. *Proteomics* **17**(1–2), 1600319 (2017)
32. Huttlin, E.L., Jedrychowski, M.P., Elias, J.E., Goswami, T., Rad, R., Beausoleil, S.A., Villen, J., Haas, W., Sowa, M.E., Gygi, S.P.: A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143**, 1174–1189 (2010)
33. Mi, H., Muruganujan, A., Casagrande, J.T., Thomas, P.D.: Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**, 1551–1566 (2013)
34. da Huang, W., Sherman, B.T., Lempicki, R.A.: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009)
35. Rajalingam, D., Loftis, C., Xu, J.J., Kumar, T.K.: Trichloroacetic acid-induced protein precipitation involves the reversible association of a stable partially structured intermediate. *Protein Sci.* **18**, 980–993 (2009)
36. Drilon, A., Rekhtman, N., Ladanyi, M., Paik, P.: Squamous-cell carcinomas of the lung: emerging biology, controversies, and the promise of targeted therapy. *Lancet Oncol.* **13**, e418–e426 (2012)
37. Wilkerson, M.D., Yin, X., Hoadley, K.A., Liu, Y., Hayward, M.C., Cabanski, C.R., Muldrew, K., Miller, C.R., Randell, S.H., Socinski, M.A., Parsons, A.M., Funkhouser, W.K., Lee, C.B., Roberts, P.J., Thorne, L., Bernard, P.S., Perou, C.M., Hayes, D.N.: Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. *Clin. Cancer Res.* **16**, 4864–4875 (2010)
38. Calabrese, F., Lunardi, F., Balestro, E., Marulli, G., Perissinotto, E., Loy, M., Nannini, N., Valente, M., Saetta, M., Agostini, C., Rea, F.: Serpin B4 isoform overexpression is associated with aberrant epithelial proliferation and lung cancer in idiopathic pulmonary fibrosis. *Pathology* **44**, 192–198 (2012)
39. Tsuta, K., Tanabe, Y., Yoshida, A., Takahashi, F., Maeshima, A.M., Asamura, H., Tsuda, H.: Utility of 10 immunohistochemical markers including novel markers (desmocollin-3, glypican 3, S100A2, S100A7, and Sox-2) for differential diagnosis of squamous cell carcinoma from adenocarcinoma of the lung. *J. Thorac. Oncol.* **6**, 1190–1199 (2011)
40. Wang, J., Qian, J., Hoeksema, M.D., Zou, Y., Espinosa, A.V., Rahman, S.M., Zhang, B., Massion, P.P.: Integrative genomics analysis identifies candidate drivers at 3q26-29 amplicon in squamous cell carcinoma of the lung. *Clin. Cancer Res.* **19**, 5580–5590 (2013)
41. Shi, I., Hashemi Sadraei, N., Duan, Z.H., Shi, T.: Aberrant signaling pathways in squamous cell lung carcinoma. *Cancer Informat.* **10**, 273–285 (2011)
42. Cmic, I., Strittmatter, K., Cavallaro, U., Kopfstein, L., Jussila, L., Alitalo, K., Christofori, G.: Loss of neural cell adhesion molecule induces tumor metastasis by up-regulating lymphangiogenesis. *Cancer Res.* **64**, 8630–8638 (2004)
43. Faraji, F., Pang, Y., Walker, R.C., Nieves Borges, R., Yang, L., Hunter, K.W.: Cadm1 is a metastasis susceptibility gene that suppresses metastasis by modifying tumor interaction with the cell-mediated immunity. *PLoS Genet.* **8**, e1002926 (2012)
44. Yamatoji, M., Kasamatsu, A., Kouzu, Y., Koike, H., Sakamoto, Y., Ogawara, K., Shiiba, M., Tanazawa, H., Uzawa, K.: Dermatopontin: a potential predictor for metastasis of human oral cancer. *Int. J. Cancer* **130**, 2903–2911 (2012)
45. Borczuk, A.C., Shah, L., Pearson, G.D., Walter, K.L., Wang, L., Austin, J.H., Friedman, R.A., Powell, C.A.: Molecular signatures in biopsy specimens of lung cancer. *Am. J. Respir. Crit. Care Med.* **170**, 167–174 (2004)
46. Tye, B.K.: MCM proteins in DNA replication. *Annu. Rev. Biochem.* **68**, 649–686 (1999)
47. Heist, R.S., Sequist, L.V., Engelman, J.A.: Genetic changes in squamous cell lung cancer: a review. *J. Thorac. Oncol.* **7**, 924–933 (2012)
48. Banerjee, R., Russo, N., Liu, M., Basur, V., Bellile, E., Palanisamy, N., Scanlon, C.S., van Tubergen, E., Inglehart, R.C., Metwally, T., Mani, R.S., Yocum, A., Nyati, M.K., Castilho, R.M., Varambally, S., Chinnaiyan, A.M., D'Silva, N.J.: TRIP13 promotes error-prone nonhomologous end joining and induces chemoresistance in head and neck cancer. *Nat. Commun.* **5**, 4527 (2014)
49. Rapley, J., Nicolas, M., Groen, A., Regue, L., Bertran, M.T., Caelles, C., Avruch, J., Roig, J.: The NIMA-family kinase Nek6 phosphorylates the kinesin Eg5 at a novel site necessary for mitotic spindle formation. *J. Cell Sci.* **121**, 3912–3921 (2008)
50. Chandrasekaran, G., Tatrai, P., Gergely, F.: Hitting the brakes: targeting microtubule motors in cancer. *Br. J. Cancer* **113**, 693–698 (2015)
51. Prieto, D.A., Hood, B.L., Darfler, M.M., Guiel, T.G., Lucas, D.A., Conrads, T.P., Veenstra, T.D., Krizman, D.B.: Liquid tissue: proteomic profiling of formalin-fixed tissues. *Biotechniques* **38**(Suppl.), 32–35 (2005)
52. Bronsert, P., Weisser, J., Biniotsek, M.L., Kuehs, M., Mayer, B., Drendel, V., Timme, S., Shahinian, H., Kusters, S., Wellner, U.F., Lassmann, S., Werner, M., Schilling, O.: Impact of routinely employed procedures for tissue processing on the proteomic analysis of formalin-fixed paraffin-embedded tissue. *Proteomics Clin. Appl.* **8**, 796–804 (2014)
53. Craven, R.A., Cairns, D.A., Zougman, A., Harnden, P., Selby, P.J., Banks, R.E.: Proteomic analysis of formalin-fixed paraffin-embedded renal tissue samples by label-free MS: assessment of overall technical variability and the impact of block age. *Proteomics Clin. Appl.* **7**, 273–282 (2013)
54. Rauniyar, N., Gao, B., McClatchy, D.B., Yates III, J.R.: Comparison of protein expression ratios observed by sixplex and duplex TMT labeling method. *J. Proteome Res.* **12**, 1031–1039 (2013)
55. Travis, W.D., Brambilla, E., Nicholson, A.G., Yatabe, Y., Austin, J.H., Beasley, M.B., Chirieac, L.R., Dacic, S., Duhig, E., Flieder, D.B., Geisinger, K., Hirsch, F.R., Ishikawa, Y., Kerr, K.M., Noguchi, M., Pelosi, G., Powell, C.A., Tsao, M.S., Wistuba, I.: Panel, WHO: the 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J. Thorac. Oncol.* **10**, 1243–1260 (2015)
56. Lin, J.J., Shaw, A.T.: Resisting resistance: targeted therapies in lung cancer. *Trends Cancer* **2**, 350–364 (2016)