research articles Journal of proteome research

Primary Tumor Xenografts of Human Lung Adeno and Squamous Cell Carcinoma Express Distinct Proteomic Signatures

Yuhong Wei,[†] Jiefei Tong,[†] Paul Taylor,[†] Dan Strumpf,[‡] Vladimir Ignatchenko,[‡] Nhu-An Pham,[‡] Naoki Yanagawa,[‡] Geoffrey Liu,[‡,§] Igor Jurisica,[‡,||,⊥] Frances A. Shepherd,[‡,§] Ming-Sound Tsao,[‡,⊥,#] Thomas Kislinger,[‡,⊥] and Michael F. Moran*[,†,¶,+]

Program in Molecular Structure and Function, Hospital For Sick Children, University Health Network, Ontario Cancer Institute/Princess Margaret Hospital and Campbell Family Institute for Cancer Research, and the Departments of Computer Science, Medical Biophysics, Medicine, Laboratory Medicine and Pathobiology, Molecular Genetics, and Banting and Best Department of Medical Research, University of Toronto, Toronto, Canada, McLaughlin Centre for Molecular Medicine, 101 College Street, MaRS Centre, University of Toronto

Received May 17, 2010

Nonsmall cell lung carcinoma (NSCLC) accounts for 80% of lung cancers. The most prevalent subtypes of NSCLC are adenocarcinoma (ADC) and squamous cell carcinoma (SCC), which combined account for approximately 90%. Ten resected NSCLC patient tumors (5 ADC and 5 SCC) were directly introduced into severely immune deficient (NOD-SCID) mice, and the resulting xenograft tumors were analyzed by standard histology and immunohistochemistry (IHC) and by proteomics profiling. Mass spectrometry (MS) methods involving 1- and 2-dimensional LC−MS/MS, and multiplexed selective reaction monitoring (SRM, or MRM), were applied to identify and quantify the xenograft proteomes. Hierarchical clustering of protein profiles distinguished between the ADC and SCC subtypes. The differential expression of 178 proteins, including a comprehensive panel of intermediate filament keratin proteins, was found to constitute a distinctive proteomic signature associated with the NSCLC subtypes. Epidermal growth factor receptor (EGFR) was expressed in ADC and SCC xenografts, and EGFR network activation was assessed by phosphotyrosine profiling by Western blot analysis and SRM measurement of EGFR levels, and mutation analysis. A multiplexed SRM/MRM method provided relative quantification of several keratin proteins, EGFR and plakophilin-1 in single LC−MS/MS runs. The protein quantifications by SRM and MS/MS spectral counting were associated with superior dynamic range and reproducibility but were otherwise consistent with orthogonal methods including IHC and Western immuno blotting. These findings illustrate the potential to develop a comprehensive MS-based platform in oncologic pathology for better classification and potentially treatment of NSCLC patients.

Keywords: lung cancer • xenograft • mudpit • SRM • MRM • biomarker • keratin • proteomic signature

Introduction

Lung cancer is the most common cause of death from cancer for both men and women, with a current worldwide mortality rate in excess of one million per year. Nonsmall cell lung carcinoma (NSCLC) is histologically heterogeneous, with adenocarcinoma (ADC), squamous cell carcinoma (SCC), and large cell carcinoma (LC) being the major subtypes.[1] Combined, these subtypes account for approximately 85% of lung cancers. In clinical practice, these subtypes have been treated similarly until recently, when new therapies (e.g., premetrexed) have shown differential responses in NSCLC subtypes.[2] However, despite improvements in surgical and chemotherapeutic treatments, and the development of drugs targeting the epidermal growth factor receptor (EGFR), which is a target in a subset of NSCLC, the 5-year survival rate associated with these cancers is poor, at approximately 15%. There is considerable variability in the molecular features between and within each of these NSCLC subtypes (e.g., EGFR expression level and mutational status), suggesting that additional stratification of tumors may facilitate more effective, tumor-specific treatments.[3]

The analysis of EGFR and various keratins by methods with limited dynamic range such as immunohistochemistry (IHC) are common practices in oncologic pathology. EGFR levels by IHC have not proven to be predictive of response to EGFR-directed drugs, despite initial studies suggesting that patients whose tumors demonstrate low expression have low response

* To whom correspondence should be addressed. E-mail: m.moran@utoronto.ca.
[†] Program in Molecular Structure and Function, Hospital For Sick Children.
[‡] University Health Network, Ontario Cancer Institute/Princess Margaret Hospital and Campbell Family Institute for Cancer Research.
[§] Department of Medicine, University of Toronto.
[||] Department of Computer Science, University of Toronto.
[⊥] Department of Medical Biophysics, University of Toronto.
[#] Department of Laboratory Medicine and Pathobiology, University of Toronto.
[¶] Department of Molecular Genetics, University of Toronto.
[+] Banting and Best Department of Medical Research, University of Toronto.

10.1021/pr100491e

Published on Web 09/03/2010

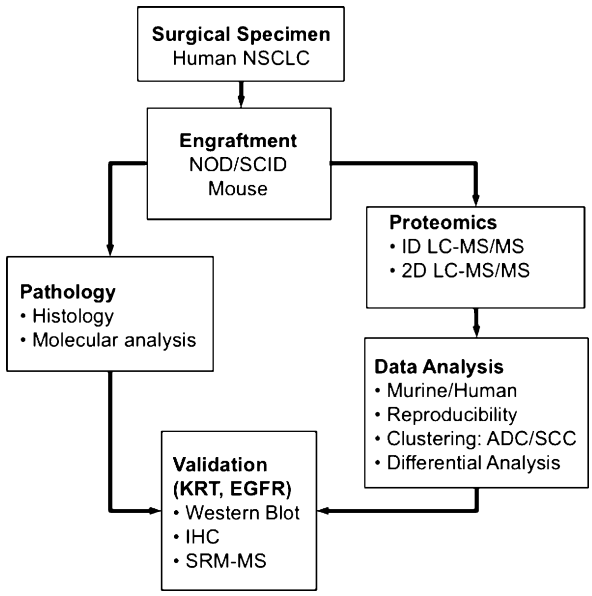

Figure 1. Experimental plan. Schematic diagram indicating the parallel collection of pathology and proteomics data sets, which were compared and subjected to further validation by immunohistochemistry (IHC) and Western blotting, and quantification by multiplexed SRM-MS (also known as MRM).

rates.[4] The keratins are relatively abundant proteins (i.e., expressed at high level) and are the major structural component of the intermediate filament-based epithelial barrier in tissue.[5] Keratin expression is stable during tumorigenesis, and the keratin expression pattern may signify tumor origin and type (reviewed in ref 5). However, while some keratins have been detected in blood and monitored as biomarkers (e.g., ref 6), to-date only a subset of the 54 human keratin proteins have been developed into clinically useful diagnostic biomarkers.[5] There remains an unmet need to develop sensitive and more quantitative methods to identify and quantify comprehensive sets of diagnostic biomarkers including drug targets such as the EGFR and their associated signaling network components, and protein classes such as the keratins whose function is involved in the epithelial tissue and tumor phenotypes, and which may inform of tumor subtypes.

Mass spectrometry (MS) has emerged as a powerful technology for proteomic analysis of tumors, and represents a promising approach to stratify tumors according to their protein profiles, and for drug target and biomarker discovery.[7] These methods have been extensively reviewed[7] and applied largely to study tumor-derived cell lines grown either in two-dimensional cultures or as xenograft tumors in immuno deficient mice. However, in either growth context, such established cell lines are mostly not representative of the more diversified or heterogeneous tumors in human cancers.[8] Another issue associated with MS analysis of human-murine xenograft systems is the recognition and assignment of human versus murine proteins, which share a large degree of sequence homology. Methods to recognize and quantify human tumor proteomes and to generate tissue models that faithfully retain or recapitulate their protein profiles are required.

In this pilot study of NSCLC proteomics, comprehensive protein identification, and in some instances quantification, were completed by application of LC−MS/MS methods to define the proteomes of NSCLC xenografts (Figure 1). The xenografts were generated by direct introduction of human NSCLC surgical specimens into nonobese diabetic, severe combined immune deficient (NOD/SCID) mice without intermediary growth in two-dimensional tissue cultures. Results presented herein include comprehensive strategies to distinguish human from murine peptides, which may represent tumor and stroma, respectively, and the documentation of ADC and SCC molecular signatures based on differential expression of keratins and various other proteins. The expression of certain keratins and EGFR were measured by application of various MS methods including selective reaction monitoring (SRM, also known as multiple reaction monitoring, or MRM)[9] and compared with one or more typical laboratory and clinical pathology approach including Western blotting, histological staining, and immunohistochemistry. These preliminary results provide a framework for the further development of a proteomics-based pathology platform to diagnose and treat NSCLC tumors.

Experimental Procedures

Xenograft Tumor Generation and Pathology. The protocols for the collection of patient tumor tissue for xenograft establishment and the animal study have been approved by the Research Ethics Board and Animal Care Committee of the University Health Network. Routinely harvested fresh human NSCLC were resected surgically at The University Health Network (Toronto) and directly implanted into nonobese diabetic and severe combined immune-deficient (NOD-SCID) mice to establish primary tumor xenograft models. Each tumor model was verified by at least three serial in vivo passages to demonstrate engraftment stability. OncoCarta MassARRAY Chip

Table 1. NSCLC Xenograft Histopathological and Molecular Features

				immunohistochemistry[b]						
xeno-graft ID	subtype and cellular differentiation	cellularity[a]	mutations	EGFR	EGFR pY1068	KRT5/6	KRT7	KRT14	KRT19	HMWK
ADC1	Adeno, moderate	70−75	−	>90	20	−	>90	−	>90	−
ADC2	Adeno, poor	>90	−	>90	>90	−	>90	>90	2−3	−
ADC3	Adeno, moderate	70−75	EGFR Δ746−750	>90	40	−	>90	−	70	−
ADC4	Adeno, poor	70	KRAS G12C	10	50	−	>90	−	>90	−
ADC5	Adeno, poor	>90	KRAS G12D	50	−	−	>90	−	40	−
SCC1	Squam, well	>90	−	60	30	>90	−	40	>90	>90
SCC2	Squam, well	80	PIK3CA E542K	−	−	>90	−	>90	>90	>90
SCC3	Squam, moderate	80−90	−	>90	40	>90	10	−	>90	>90
SCC4	Squam, poor	85−90	PIK3CA E545K	70	−	>90	−	>90	>90	>90
SCC5	Squam, moderate	>90	−	>90	−	>90	−	70	>90	>90

[a] Percent tumor cell abundance among cells present. [b] Percentage positively stained cells.

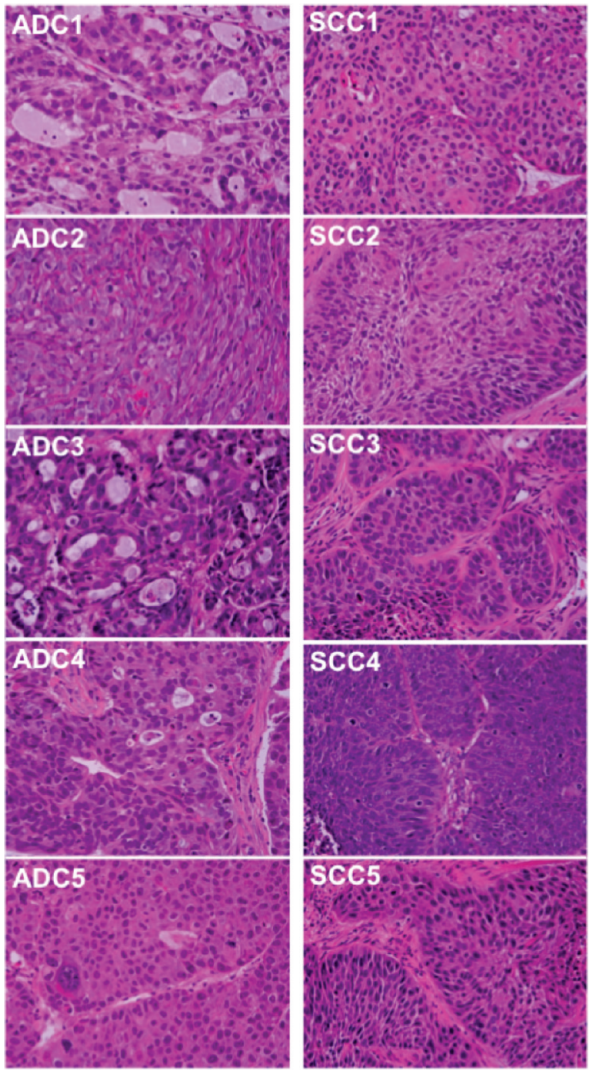

Figure 2. NSCLC histology. Hematoxylin/eosin stain of primary NSCLC xenografts including 5 adenocarcinoma (ADC) models and 5 squamous cell carcinoma (SCC) models.

(Sequenom) mutation screening was conducted to detect mutations in the EGFR, KRAS, and PIK3CA genes. These results were validated by direct DNA sequencing of tumor xenograft specimens.

Immunohistochemistry. Formalin-fixed paraffin-embedded tissue blocks were cut at 4 um thickness onto slides and dried in 60 °C oven overnight. Slides were further processed and stained in a fully automated process using the BenchMark XT (Ventana Medical Systems Inc.) with the following reagents and conditions listed in SI Table 7 (Supporting Information). Slides were scored by a pathologist (NY).

Sample Preparation and Western Immuno Blotting. Xenograft tissues were harvested from mice and immediately stored in liquid nitrogen. Aliquots of tissue (approximately 50 mg) were mixed with lysis buffer (1 mL buffer per 10 mg tissue; 20 mM HEPES pH 8.0, 9 M Urea, 1 mM sodium orthovanadate, 2.5 mM sodium pyrophosphate, 1 mM ss-glycerophosphate) and subjected to ultrasonication for 1 min, followed by centrifugation (20 000 × *g*) for 20 min. Aliquots of supernatants (clarified lysates; 50 μL) were set aside for Western blotting analysis. Routinely, the concentrations of clarified lysates were approximately 2 mg/mL (protein). An equal volume of 2 × SDS-PAGE sample loading buffer was added, and the samples were resolved by standard SDS-PAGE methods and then electrophoretically transferred to Immobilon-P membranes (Millipore) for Western blotting, essentially as described previously.[10]

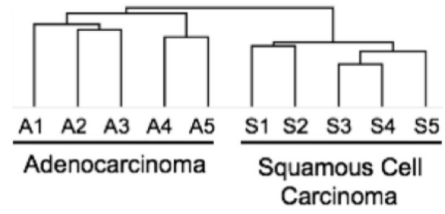

Figure 3. Recognition of ADC and SCC subtypes of NSCLC by 1D LC−MS/MS protein profiling. Dendrogram produced by hierarchical clustering of proteins measured by 1D LC−MS/MS and having ≥2 spectra and ≥2 sample incidence (541 proteins).

Table 2. Highly Differentially Expressed Proteins in ADC and SCC Xenografts Detected by 1D LC− MS/MS Protein Profiling

identified proteins (635)	ADC/SCC ratio	SCC/ADC ratio	*p*-value
CPS1	*24.3*	0.04	0.049
KRT7	*13.6*	0.07	0.001
AGR2	*10.1*	0.10	0.017
KRT14	0.08	11.9	0.022
KRT17	0.07	14.7	0.0001
KRT15	0.03	31.2	0.002
KRT16	0.03	31.6	0.011
KRT5	0.02	53.0	0.00003
KRT6A	0.02	59.1	0.00005
KRT6B	0.02	59.4	0.00006

For MS analysis, clarified lysates were reduced with 4.5 mM DTT, carboxamidomethylated by using 10 mM iodoacetamide, diluted 4-fold, and digested by incubation with Trypsin-TPCK for 12 h. Peptides were then desalted by using C18 resin as described previously.[11] The eluted peptides were aliquoted and lyophilized. Two micrograms or 15 μg dried, desalted peptides were dissolved in 0.1% formic acid and analyzed by 1D or 2D LC−MS/MS, respectively.

1D LC−MS/MS. All samples were analyzed on a LTQ-Orbitrap XL. The instrument method consisted of one MS full scan (400−1800 *m*/*z*) in the Orbitrap mass analyzer, an AGC target of 500 000 with a maximum ion injection of 500 ms, 1 μscan and a resolution of 60 000 and using the preview scan option. Six data-dependent MS/MS scans were performed in the linear ion trap using the three most intense ions at 35% normalized collision energy. The MS and MS/MS scans were obtained in parallel. AGC targets were 10 000 with a maximum ion injection time of 100 ms. A minimum ion intensity of 1000 was required to trigger a MS/MS spectrum. The dynamic exclusion was applied using a maximum exclusion list of 500 with one repeat count with a repeat duration of 30 s and exclusion duration of 45 s.

2D LC−MS/MS Analyses. A fully automated 4-step two-dimensional chromatography sequence was set up as previously described.[12] Peptides were loaded on a 7 cm precolumn (150 μm i.d.) containing a Kasil frit packed with 3.5 cm 5 μ Magic C18 100 Å reversed phase material (Michrom Bioresources) followed by 3.5 cm Luna 5 μ SCX 100 Å strong cation exchange resin (Phenomenex, Torrance, CA). Samples were automatically loaded from a 96-well microplate autosampler by using an EASY-nLC system (Proxeon Biosystems, Odense, Denmark) at 3 μL/min. The precolumn was connected to an 8

Table 3. NSCLC Proteomics Profiling Summary

xenograft name	identified human proteins
ADC1_1	628
ADC1_2	650
ADC1_3	649
ADC1_Total	**671**
ADC2_1	890
ADC2_2	830
ADC2_3	811
ADC2_Total	**933**
ADC3_1	1140
ADC3_2	1185
ADC3_3	1022
ADC3_Total	**1271**
ADC4_1	634
ADC4_2	630
ADC4_3	545
ADC4_Total	**663**
ADC5_1	698
ADC5_2	686
ADC5_3	685
ADC5_Total	**734**
SCC1_1	701
SCC1_2	696
SCC1_3	708
SCC1_Total	**738**
SCC2_1	611
SCC2_2	620
SCC2_3	612
SCC2_Total	**645**
SCC3_1	695
SCC3_2	691
SCC3_3	690
SCC3_Total	**719**
SCC4_1	825
SCC4_2	797
SCC4_3	769
SCC4_Total	**854**
SCC5_1	665
SCC5_2	645
SCC5_3	665
SCC5_Total	**692**
Total:	**2015**

cm fused silica analytical column (75 μm i.d.) via a micro splitter tee (Proxeon) to which a distal 2.3 kV spray voltage was applied. The analytical column was pulled to a fine electrospray emitter by using a laser puller. For the peptide separation on the analytical column a water/acetonitrile gradient was applied at an effective flow rate of 400 nL/min, controlled by the EASY-nLC. Ammonium acetate salt bumps (8 μL) were applied at the following concentrations (0, 100, 300, and 500 mM), using the 96-well micro plate autosampler at a flow-rate of 3 mL/min in a vented-column setup.

SRM. SRM was carried out on duplicate 5 μg aliquots of each xenograft lysate. The peptides were captured on a 150 μm ID C18 precolumn and separated over a 75 μm ID analytical column constructed with an emitter tip. The separation was carried out with a gradient of 0−65% acetonitrile in 0.1% formic acid over 40 min using the EASY-nLC split-free HPLC system. The eluted peptides were monitored by using a TSQ Quantum Vantage triple quadrupole mass spectrometer (ThermoFisher, San Jose, CA). The dwell time was 20 ms and the scan width was 0.01 amu. The S-lens was varied with precursor m/z values and a 10 V declustering potential was used. Q1 and Q3 resolution were set to 0.2 and 0.7 amu, respectively. The transitions used are shown in Table 6. GAPDH is frequently used as a loading control in Western blot analyses, and the 2D LC−MS/MS experiments indicated that GAPDH was not statistically different in abundance between the ADC and SCC groups. Therefore, the SRM value for human GAPDH peptide LISWYDNEFGYSNR, based on 3 transitions, which averaged 5.7 $\pm$ 0.8 (SE) $\times$ 10^5 units across the 10 samples, was used to normalize the summed SRM transition measurements associated with individual peptides from corresponding xenograft samples. Collision energy was calculated by using the formula 3.41 + 0.034 $\times$ (m/z of parent peptide), with collision gas pressure at 1.5 mTorr as described by Prakash and colleagues.[13]

SRM transitions were developed manually by first examining all possible y ion transitions for ions of interest. From these results the best 3 for each peptide (5 for EGFR peptides) were included in the (multiplexed) SRM method files that were applied for analysis of replicate samples.

Clustering and Identification of Differentially Expressed Proteins. Spectral counts for peptides corresponding to a single protein were summed and protein level counts were normalized by obtaining the relative abundance ratio (dividing by the total sample spectral counts) then multiplying by the overall experimental spectral count. Spectra count values of zero were changed to 0.2 as part of the normalization routine similar, as recently reported.[14−17] For clustering analysis, the 3 replicates for each sample were averaged, and the normalized data was then filtered to include proteins that were detected in at least 2 samples in the entire data set of normalized protein data (i.e., sample incidence ≥2). Hierarchical clustering was applied to the normalized protein data in R (v2.10.0) via the "hclust" function using Spearman's rank correlation as distance metrics and the "average" agglomeration method, and was used to generate dendrograms. Heatmap plots were generated by using the heatmap.2 function in the R package "gplots" (v2.7.4), utilizing the log2 transformed normalized protein data.

The Wilcoxon Rank Sum Test ("wilcox.test" in R, v2.10.0) was used to identify differentially expressed proteins between ADC and SCC samples in the normalized protein data. Significance was assumed as p-value <0.05.

Protein Identification and Data Analysis. Raw data were converted to m/z XML using ReAdW and searched by X!Tandem against a locally installed version of a merged human and mouse IPI (http://www.ebi.ac.uk/IPI) protein sequence database (version 3.54; 75 427 human sequences and 55 985 mouse sequences). The searches were performed with a fragment ion mass tolerance of 0.4 Da, a parent ion mass tolerance of ±10 ppm. Complete tryptic digest was assumed. Carbamidomethylation of cysteine was specified as a fixed, and oxidation of methionine a variable, modification.

To estimate and minimize the false positive rate, the merged human and mouse protein sequence database also contained every IPI protein sequence in its reversed amino acid orientation (target-decoy strategy; total database size 262 824 sequences) as recently described.[18,19] For the presented study, we have set the value of total reverse spectra to total forward spectra to 0.5%, resulting in zero decoy sequences in the final output (0 reverse proteins for both the human and mouse assignments). Only fully tryptic peptides ≥7 amino acids, matching these criteria were accepted to generate the final list of identified proteins. Only proteins identified with at least two unique peptides per sample (across the triplicate analysis) were accepted. To minimize protein inference, we developed a database grouping scheme, and only reported proteins with

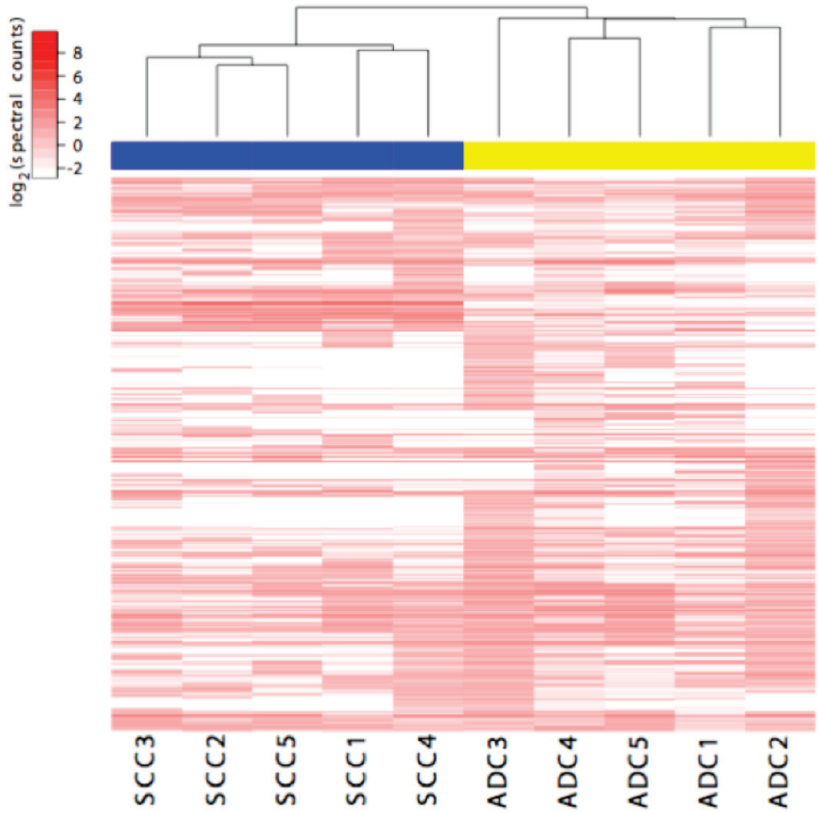

Figure 4. Cluster analysis of human proteins in NSCLC. Hierarchical clustering of 2D LC−MS/MS spectra of human proteins resolved ADC (yellow) and SCC (blue) subtypes. Included were the 1303 proteins with spectral counts and sample incidence ≥2.

substantial peptide information, as recently described.[15,18,20] For a protein to be assigned as "human" or "mouse", it required at least one unique peptide mapping uniquely to either a human or mouse entry in the mixed-species database.

Results

Experimental Plan. As part of a larger effort to establish and characterize >100 primary NSCLC xenografts, a pilot study was conducted with 10 NSCLC xenografts to test if a proteomics platform could be effectively applied to characterize these tumors (Figure 1). The analytical proteomics platform included tandem MS analysis of tryptic peptides resolved by one-dimensional (1D) or two-dimensional (2D), nano scale liquid chromatography. This provided peptide and protein identifications and relative semiquantification based on MS/MS spectral counting. In parallel, the xenografts were subjected to standard laboratory histopathology, which dictated their classification as ADC or SCC subtype. Analysis of MS data was completed with the objective to identify protein expression signatures characteristic of the ADC and SCC subtypes. Validation of protein expression involved limited applications of SRM-MS, IHC, and Western immuno blotting.

Xenograft Tumor Pathology. The establishment of the xenograft tumors and sample preparation are described in Experimental Procedures. Table 1 presents a summary of tumor information including limited histological and molecular features. The ten xenografts were classified as ADC or SCC based on the histologies of the primary tumors and corresponding xenografts; there was good concordance in differentiation grades (Figure 2). The tumors were screened for mutations by the OncoCarta v.1 MassARRAY system (Sequenom, San Diego, CA), and mutations in both primary and corresponding xenograft were confirmed by sequencing. Five tumors were found to have activating, coding mutations in the *EGFR*, *KRAS*, and *PIK3CA* genes (Table 1).

Analysis of NSCLC Xenograft Proteomes by 1D and 2D LC−MS/MS. 1D LC−MS/MS. A relatively rapid (i.e., approximately 2 h per sample analysis) 1D LC−MS/MS approach was used to complete an initial analysis of proteins expressed in the 10 xenograft samples. This was followed by a more comprehensive, but time-consuming (i.e., 8 h per sample analysis), 2D analysis (described below).[12] Results were tabulated as normalized MS/MS spectral counts per assigned gene product (see Experimental Procedures) and are summarized in the Supporting Information (SI Table 2). Variability in protein detection and expression were assessed by comparing results of technical and biological replicates at the protein and peptide levels (SI Figure 1, Supporting Information). When a sample (ADC1) was analyzed in triplicate, 493 proteins out of a total of 564 were identified in each replicate, indicating an overlap of 87%. To assess biological variation, equivalent portions of ADC1 were expanded in two different recipient mice and then analyzed. The overlap between the two samples was 88% at the protein level: 491 proteins out of a total of 559 were found in both samples. These numbers suggested a reasonable degree of reproducibility at the protein level between samples analyzed by 1D LC−MS/MS.

The 1D LC−MS/MS analysis of the 10 samples was considered a preliminary scan of the NSCLC xenograft proteomes, and allowed identification of 635 proteins. See Experimental Procedures for protein identification criteria and Supporting Information (SI) Table 1 and SI Table 2 for protein identification

and spectral counting data, respectively. As shown in the dendrogram in Figure 3, even by the relatively low resolution (in terms of proteome coverage) approach, hierarchical clustering of proteins based on normalized spectral counts separated the proteomes into two sets corresponding to the ADC and SCC subtypes (Figure 3). To semiquantify proteins and examine differential expression between the ADC and SCC subtypes, normalized spectral counts for proteins were summed across each subtype and compared. A subset of 57 proteins was significantly differentially expressed between the ADC and SCC subtypes (SI Table 3, Supporting Information). Ten proteins were deemed highly differentially expressed and displayed a >10-fold increase or decrease between ADC and SCC (Table 2). Prominent among the highly differentially expressed proteins were 8 keratin (KRT) gene products, with KRT7 highly expressed in ADC, and 7 others that were more highly expressed in SCC. The other two proteins found more highly expressed in ADC compared to SCC were the urea cycle enzyme Carbamoyl-Phosphate Synthase (CSP1) and Anterior Gradient homologue 2 (AGR2). These data indicate that the 1D platform was sufficient to resolve ADC and SCC subtypes based on their significantly different proteomes.

2D LC−MS/MS Analysis. To increase statistical significance and proteome coverage, a more rigorous protocol involving triplicate analysis by 2D LC−MS/MS, MudPIT (Multidimensional Protein Identification Technology), was applied.[12,21,22] Each sample was analyzed in triplicate and MS/MS spectral counts tabulated as described in Experimental Procedures.[14−17] As a product of the 30 individual 2D analyses, 2015 proteins were identified (Table 3). Details on the individual spectral counts and identified human and murine proteins are included in SI Tables 4 and 5 (Supporting Information), respectively. Expressed human proteins were subjected to hierarchical clustering analysis to determine if the ADC and SCC proteomes were distinct. Figure 4 shows the results of hierarchical clustering of identified human proteins. Similar to the 1D results described above, the 2D data set clustered into separate ADC and SCC sets.

By application of the Wilcoxon test, 178 human proteins were significantly different in their average expression between the ADC and SCC subtypes (SI Table 6, Supporting Information). Within this set, 50 proteins were increased or decreased >10-fold in ADC compared with SCC xenografts (Table 4). This highly differential subset included 8 keratins, including 6 that were identified as 10-fold differentially expressed in the 1D data set. The proteins CPS1 and AGR2 were again identified as more abundant in ADC: 5.0-fold ($p = 0.03$) for CPS1 and 17.4-fold ($p = 0.008$) for AGR2 (Table 4, SI Table 6, Supporting Information).

Of the 28 known human epithelial keratins (Moll et al. 2008), 22 were detected in the panel of xenografts, as summarized in Table 5. Ten KRT proteins were significantly differentially expressed between ADC and SCC subtypes (Table 5, boldface, KRTs 18, 7, 5, 14, 15, 6A, 16, 17, 4, 13). Table 5 is organized by grouping the KRT proteins according to their known expression in simple and stratified epithelia and, where known, in type I/II pairs that are known to assemble as obligate heterodimers for intermediate filament assembly.[5,23]

Validation of Expression of Keratins and EGFR in NSCLC. To validate and extend the information on these proteins that was generated by 1D and 2D tandem MS, additional analyses were performed. This included IHC on FFPE tissue sections, Western immuno blotting, and quantification by SRM-MS. Table 1 summarizes the IHC information related to the EGFR and certain keratins. Table 6 lists the peptides and corresponding transitions that were measured as part of a single, mulitplexed SRM (or MRM) method that was used to scan the xenografts.

Table 4. Proteins Highly Differentially Expressed in NSCLC

name	ADC/SCC	SCC/ADC	*p*-value
RPS3	*48.9*	0.02	0.027
GFPT1	*46.8*	0.02	0.012
KPNB1	*28.9*	0.03	0.027
KRT7	*28.4*	0.04	0.012
EIF3B	*25.5*	0.04	0.016
LPCAT1	*23.3*	0.04	0.016
C3	*22.5*	0.04	0.027
GALE	*21.2*	0.05	0.012
CRABP2	*21.1*	0.05	0.027
MGST1	*19.5*	0.05	0.012
AP2A2	*18.6*	0.05	0.021
AGR2	*17.4*	0.06	0.008
ICAM1	*17.1*	0.06	0.027
GORASP2	*16.1*	0.06	0.021
GLRX	*15.5*	0.06	0.012
IARS2	*15.1*	0.07	0.008
KIAA0368	*14.3*	0.07	0.012
FKBP10	*14.2*	0.07	0.027
S100A13	*13.7*	0.07	0.012
CRIP2	*12.3*	0.08	0.021
PGRMC1	*12.1*	0.08	0.027
AIFM1	*12.1*	0.08	0.027
SDF2L1	*12.0*	0.08	0.012
GSPT1	*11.2*	0.09	0.027
DNAJC10	*11.0*	0.09	0.021
VIM	*10.3*	0.10	0.036
RNF213	*10.3*	0.10	0.027
TFRC	0.10	10.2	0.016
NDUFS8	0.08	12.3	0.008
KRT14	0.07	13.7	0.016
DSP	0.07	14.4	0.008
TRIM29	0.07	15.1	0.008
GPC1	0.06	17.2	0.008
SPRR1B	0.06	17.6	0.008
GSTM4	0.05	21.6	0.028
DSC3	0.03	32.7	0.008
SPRR1A	0.03	33.3	0.028
CALML3	0.03	33.9	0.008
GBP6	0.02	43.8	0.028
DSG3	0.02	44.9	0.008
SPRR3	0.01	71.9	0.008
ADH7	0.01	74.3	0.008
PKP1	0.01	137	0.008
KRT4	0.00	213	0.028
CES1	0.00	284	0.008
KRT16	0.00	461	0.008
KRT15	0.00	483	0.008
KRT5	0.00	667	0.008
KRT6A	0.00	756	0.008
KRT13	0.00	1655	0.008

Keratins. In Figure 5A (and summarized in Table 1), IHC verified the expression of KRT7 in ADC samples, as well as low-level expression in SCC3. MS analysis by spectral counting and SRM provided quantitative results that were consistent with each other, and the IHC staining pattern (Figure 5B). The spectral counting analysis detected KRT7 in SCC3 to a greater extent than SRM. The SRM data were reproducible, and results from two distinct KRT7 peptides (detailed in Table 6) were very similar. All keratin peptides subjected to SRM analysis were

Table 5. Keratin Signatures in NSCLC[a]

Keratin type	Protein	Type	ADC1 AVG	ADC1 CV	ADC2 AVG	ADC2 CV	ADC3 AVG	ADC3 CV	ADC4 AVG	ADC4 CV	ADC5 AVG	ADC5 CV	SCC1 AVG	SCC1 CV	SCC2 AVG	SCC2 CV	SCC3 AVG	SCC3 CV	SCC4 AVG	SCC4 CV	SCC5 AVG	SCC5 CV	ADC /SCC	SCC /ADC	Wilcoxon p-value
Simple Epithelial	KRT8	II	93.2	18%	18.7	13%	64.8	43%	130.6	7%	54.6	11%	82.7	4%	22.9	2%	45.1	8%	31.1	35%	32.8	7%	1.7	0.6	0.310
	KRT18	**I**	**83.5**	**11%**	**20.4**	**20%**	**69.4**	**38%**	**108.9**	**3%**	**113.5**	**18%**	**22.5**	**6%**	**5.9**	**16%**	**33.7**	**6%**	**10.3**	**26%**	**18.6**	**13%**	**4.3**	**0.2**	**0.032**
	KRT20	I	4.9	45%	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	5.9	—	—
	KRT7	**II**	**87.9**	**12%**	**48.3**	**24%**	**29.9**	**46%**	**52.2**	**3%**	**80.6**	**12%**	—	—	—	—	**9.8**	**19%**	—	—	—	—	**28.4**	**0.04**	**0.008**
	KRT19	I	92.3	11%	4.7	16%	18.4	35%	115.6	26%	54.5	30%	295.0	8%	57.0	9%	61.3	2%	196.2	26%	94.9	9%	0.4	2.5	0.151
Stratified Epethelial	**KRT5**	**II**	—	—	—	—	—	—	—	—	—	—	**155.9**	**2%**	**179.0**	**7%**	**101.1**	**6%**	**190.7**	**23%**	**98.9**	**7%**	—	**667.3**	**0.008**
	KRT14	**I**	—	—	**73.5**	**48%**	—	—	—	—	—	—	**148.4**	**4%**	**457.4**	**7%**	**61.2**	**39%**	**110.0**	**18%**	**243.0**	**29%**	**0.1**	**13.7**	**0.016**
	KRT15	**I**	—	—	—	—	—	—	—	—	—	—	**183.1**	**3%**	**114.9**	**3%**	**41.1**	**6%**	**96.1**	**21%**	**90.4**	**34%**	—	**483.3**	**0.008**
	KRT6A	**II**	—	—	—	—	—	—	—	—	—	—	**160.1**	**4%**	**238.9**	**2%**	**88.6**	**4%**	**179.6**	**27%**	**154.7**	**0.3%**	—	**755.8**	**0.008**
	KRT6B	II	7.6	7%	—	—	—	—	—	—	6.5	13%	—	—	—	—	—	—	166.4	30%	—	—	0.1	11.3	1.000
	KRT6C	II	—	—	—	—	—	—	—	—	—	—	145.6	3%	215.2	3%	—	—	—	—	—	—	—	332.2	0.690
	KRT16	**I**	—	—	—	—	—	—	—	—	—	—	**118.0**	**3%**	**146.3**	**2%**	**13.5**	**12%**	**84.2**	**30%**	**139.6**	**46%**	—	**461.2**	**0.008**
	KRT17	**I**	**53.1**	**10%**	**26.9**	**51%**	**4.0**	**42%**	**0.3**	**0%**	**23.5**	**13%**	**55.6**	**2%**	**167.0**	**6%**	**121.7**	**7%**	**96.0**	**22%**	**134.3**	**57%**	**0.2**	**5.3**	**0.008**
	KRT4	**II**	—	—	—	—	—	—	—	—	—	—	**184.1**	**5%**	**18.3**	**27%**	**0.2**	**0%**	**20.9**	**18%**	**8.2**	**64%**	—	**213.0**	**0.032**
	KRT13	**I**	—	—	—	—	—	—	—	—	—	—	**505.4**	**5%**	**507.7**	**3%**	**139.8**	**1%**	**294.1**	**18%**	**352.5**	**37%**	—	**1654.8**	**0.008**
	KRT1	II	—	—	—	—	—	—	—	—	—	—	—	—	5.2	16%	—	—	9.8	48%	—	—	—	14.3	0.421
	KRT10	I	—	—	—	—	—	—	—	—	—	—	8.6	27%	—	—	—	—	13.5	28%	—	—	—	20.9	0.421
	KRT2	II	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	71.0	35%	—	—	—	66.1	—
	KRT3	II	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	52.2	27%	—	—	—	48.7	—
	KRT76	II	—	—	—	—	—	—	—	—	—	—	38.4	4%	58.0	6%	0.2	0%	54.0	27%	—	—	—	138.7	0.151
	KRT78	II	—	—	—	—	—	—	—	—	—	—	6.7	7%	—	—	—	—	—	—	—	—	—	6.8	N/A
	KRT80	II	—	—	—	—	—	—	—	—	—	—	38.1	2%	—	—	—	—	—	—	—	—	—	35.7	N/A

[a] Note that KRT10 is also found in simple epithelia (Moll et al., 2008).

Table 6. Transitions Measured by Multiplexed SRM/MRM

protein/peptide	parent ion	fragment	CE	ion
KRT7	721.90	657.4	28	y6
LPDIFEAQIAGLR		857.5	28	y8
		1004.6	28	y9
KRT7	636.86	729.5	25	y7
SLDLDGIIAEVK		844.5	25	y8
		1072.6	25	y10
KRT19	695.35	676.3	27	y6
AALEDTLAETEAR		890.5	27	y8
		1005.5	27	y9
KRT19	677.81	748.3	26	y6
SQYEVMAEQNR		847.4	26	y7
		1139.5	26	y9
KRT5	547.27	602.3	22	y5
AQYEEIANR		731.4	22	y6
		894.4	22	y7
KRT5	556.29	610.3	22	y6
ISISTSGGSFR		711.3	22	y7
		798.4	22	y8
KRT14	713.35	849.4	28	y9
APSTYGGGLSVSSSR		906.5	28	y10
		1069.5	28	y11
KRT15	911.45	1266.6	34	y16
GGSLLAGGGGFGGGSL		1323.6	34	y17
SGGGGSR		1394.6	34	y18
KRT15	688.31	811.4	27	y9
FVSSGSGGGYGGGMR		955.4	27	y11
		1129.5	27	y13
KRT13	624.85	715.4	25	y6
LKYENELALR		844.5	25	y7
		1007.5	25	y8
EGFR	625.35	402.2	25	y4
NLQEILHGAVR		539.3	25	y5
		765.5	25	y7
		894.5	25	y8
		1022.6	25	y9
EGFR	604.87	529.3	24	y4
IPLENLQIIR		548.3	24	y9 2+
		756.5	24	y6
		885.5	24	y7
		998.6	24	y8
GAPDH	882.40	743.3	33	y6
LISWYDNEFGYSNR		1101.5	33	y9
		1264.5	33	y10
PKP1	711.38	946.5	28	y9
GLMSSGMSQLIGLK		1033.6	28	y10
		1120.6	28	y11
PKP1	659.35	617.3	26	y6
NMLGTLAGANSLR		959.5	26	y10
		1072.6	26	y11

uniquely human and therefore not subject to interference from murine orthologs.

The IHC staining for keratins 5 and 6 (CK5/6) was negative for the ADC samples, and 100% positive for the SCC xenografts (Table 1, Figure 6A). This is consistent with the unique expression in SCC compared with ADC measured by 2D LC−MS/MS (Table 5). SRM was used to measure two distinct KRT5 peptides (Table 6). The results of 4 SRM measurements (2 technical replicates for 2 peptides), were almost superimposable (Figure 6B), indicating an accurate measure of KRT5 and showing a greater dynamic range than IHC.

Analysis of KRT19 by IHC and SRM was similar: Replicate measurement of two KRT19 peptides (Table 6) gave very similar results (Figure 6D), which were in general agreement with the spectral counting data (Table 5), and IHC, but with apparently greater dynamic range than IHC staining (Figure 6C, Table 1). A similar trend was observed for KRT14 wherein the SRM measurements (Figure 7B) were similar to spectral counting (Table 5), including the maximal signal seen in SCC2 and low-level signals in ADC2 and SCC3. This was generally consistent with the IHC results (Table 1, Figure 7A), except that SCC3 was scored negative by IHC (Table 1, Figure 7A).

Figure 8 provides additional SRM measurements that were simultaneously collected as part of the multiplexed method. KRT15 was measured by following the transitions of two distinct human peptides, and gave near identical results with minimal variance. This indicated SCC-specific expression of KRT15, consistent with spectral counting (Table 5). KRT13 was measured as a function of 3 transitions from a single peptide ion. Plakophilin-1 (PKP1) was also observed to show a distinctive SCC-positive, ADC-negative expression pattern by spectral counting, and this was confirmed by SRM measurement of two peptides that gave near identical results.

EGF Receptor. By 2D LC−MS/MS analysis, the EGFR was identified in 6 xenografts, 3-each ADC and SCC (SI Table 4, Supporting Information), but was not identified as differentially expressed between the two subtypes. To further examine EGFR expression and activation, additional data were generated by application of IHC (Table 1), Western immuno blotting, and SRM-MS. As shown in Figure 9, results obtained by MS/MS spectral counting (Figure 9A), Western blotting (Figure 9B, C) and SRM analysis of two different EGFR peptides (Figure 9D) were similar in their identification of SCC3 as having the relative highest EGFR expression level. Quantification of triplicate Western blot chemiluminescence was associated considerable variation, and apparently limited dynamic range compared to the MS methods, but was sensitive in its apparent detection of EGFR in samples SCC1 and SCC2, which was not detected by spectral counting. The SRM measurements were sufficiently sensitive to detect EGFR in all 10 samples. The SRM measurements were made in duplicate, and the results were associated with very minimal variation (Figure 9D). Moreover, a very similar pattern of expression was obtained by monitoring transitions associated with the two different EGFR peptides. The EGFR peptide m/z 604.87 is identical in murine and human species, whereas the peptide m/z 635.35 is distinctively human (Table 6).

Phosphotyrosine (pTyr or pY)-directed Western blotting was used to asses the activation of EGFR in the xenografts. Tyr^{1068} becomes phosphorylated upon EGFR activation, and through direct binding of the adaptor protein GRB2 is coupled to stimulation of the RAS-ERK signaling axis.[24] Probing with antibodies to pY^{1068} provided qualitative results indicating activation of EGFR to some extent in ADCs 1−3 and SCC3. Activation in ADC3 was expected since this xenograft harbors the EGFR kinase-domain-activating exon 19 deletion (Table 1). Staining of whole-tissue extracts with antibodies to pTyr revealed a prominent band at the expected size of the EGFR in ADC3 and SCC3, and to a lesser extent in ADCs 1 and 2. This is consistent the pY^{1068} staining, and suggests EGFR was activated in SCC3 in addition to its relatively high level of expression. The anti-pTyr blot also indicated distinct patterns of pY-proteins in the tumors. A strong signal migrating at M_r 55K−65K was evident in ADC1. Discernible bands at M_r 62K were present in ADCs 2 and 4, and SCCs 1−3, and bands at approximately M_r 38K in ADC3, ADC5, SCC4. Tissue micro array

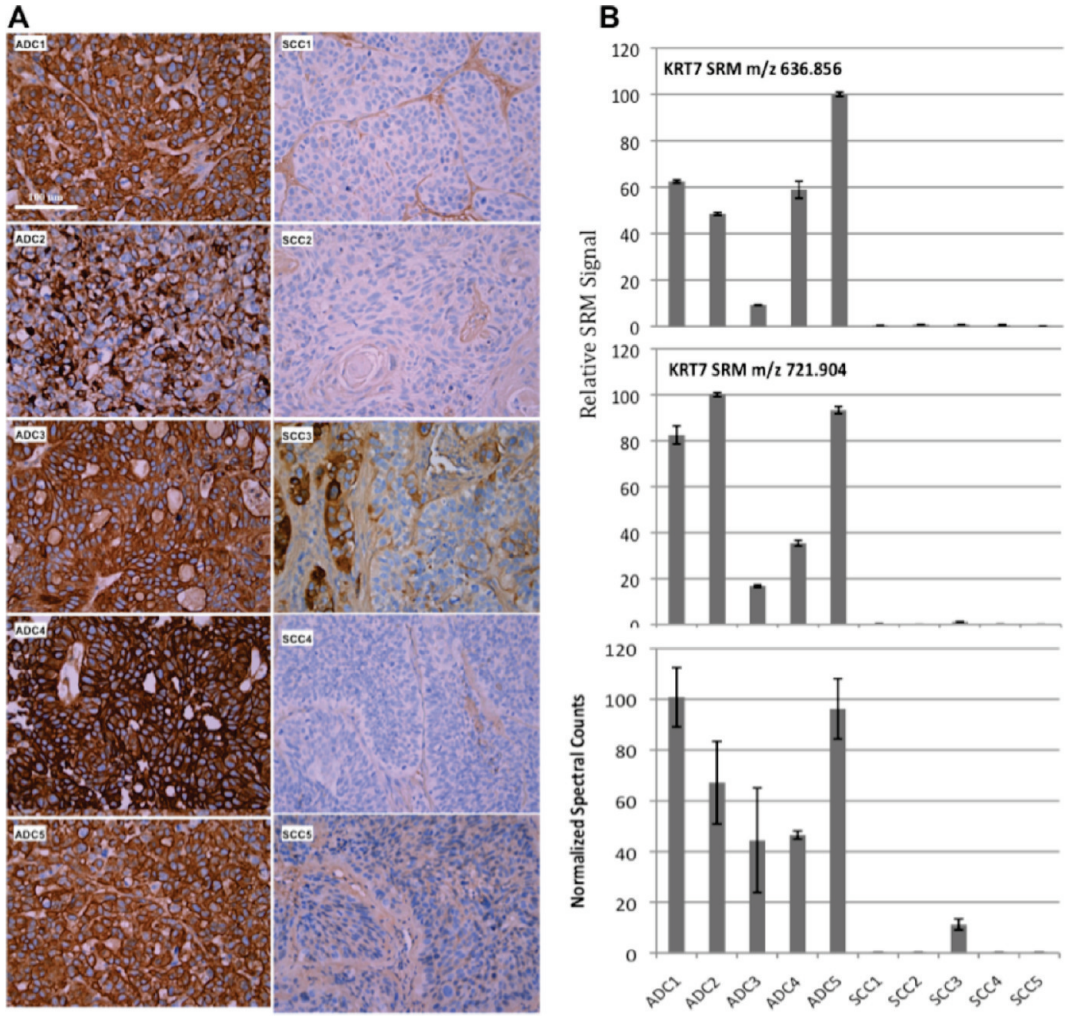

Figure 5. Comparison of KRT7 by immunohistochemistry and proteomics in NSCLC xenografts. A, KRT7 immunohistochemistry. B, Histograms presenting relative KRT7 expression measured by SRM (see Table 6 for peptide transitions), normalized to SRM-measured actin, in xenograft samples (upper two charts, $n = 2$, error bars shown range), and by spectral counting (lower chart, $n = 3 \pm$ SD).

(not shown) and subjective IHC scoring (Table 1) were in general agreement with the EGFR expression data presented in Figure 9 but did not highlight the relatively high level of EGFR in SCC3. The IHC fields shown in Figure 10 are consistent with highest EGFR expression in SCC3 and generally reflect well the profile of EGFR expression seen by SRM. Further examination of EGFR and protein-pY profiles will be required to identify and quantify the status of the EGFR and other signaling networks in these samples.

Discussion

Strategic Application of Proteomics for Tumor Profiling. The purpose of this pilot study was 2-fold. The first goal was to determine the feasibility of using a proteomics platform comprised of a high resolution LC−MS/MS instrument for 1D and 2D comprehensive protein signature discovery and combined with an LC-triple quadrupole instrument for multiplexed SRM-based relative quantification of signature proteins of interest. A similar approach of protein profiling leading to SRM/MRM-based quantification was effectively applied as part of a comprehensive platform to characterize a mouse model of breast cancer.[25] The second goal was to glean insights into the protein profiles expressed in a perceived information-rich resource represented by xenografts established from primary resected tumors. Proteomic profiles lacking detailed protein identifications have been shown to effectively stratify NSCLC tumors.[26] This report provides the most detailed analysis of protein expression in NSCLC to date and demonstrated effective recognition of ADC and SCC subtypes based on their unique proteomics signatures. The set of 10 tumors used in this pilot study did not include examples of the more rare, large cell carcinoma. The oncogenic mutations found conserved in both primary and corresponding xenograft tumors (Table 1) is consistent with these alleles functioning as "driver" mutations required for the transformed phenotype. We expect that as data accumulate by analysis of a greater diversity of xenografts, it is likely that the proteomics profiles will stratify into more groups than the traditionally recognized ADC, SCC, and large cell subtypes. The effective translation of proteomic signatures into multiplexed SRM (or MRM) assays, which may be applied to quantify proteins in minute surgical samples, represents a new strategy to stratify tumors. This will facilitate, in the first instance, case-controlled studies of outcome that may correlate with an expanded set of proteome-defined tumor subtypes.

The quantitative results from 2D LC−MS/MS spectral counting and SRM were remarkably consistent with each other, and complemented the IHC observations. Compared to the 2D method, the 1D LC−MS/MS approach was simpler and faster,

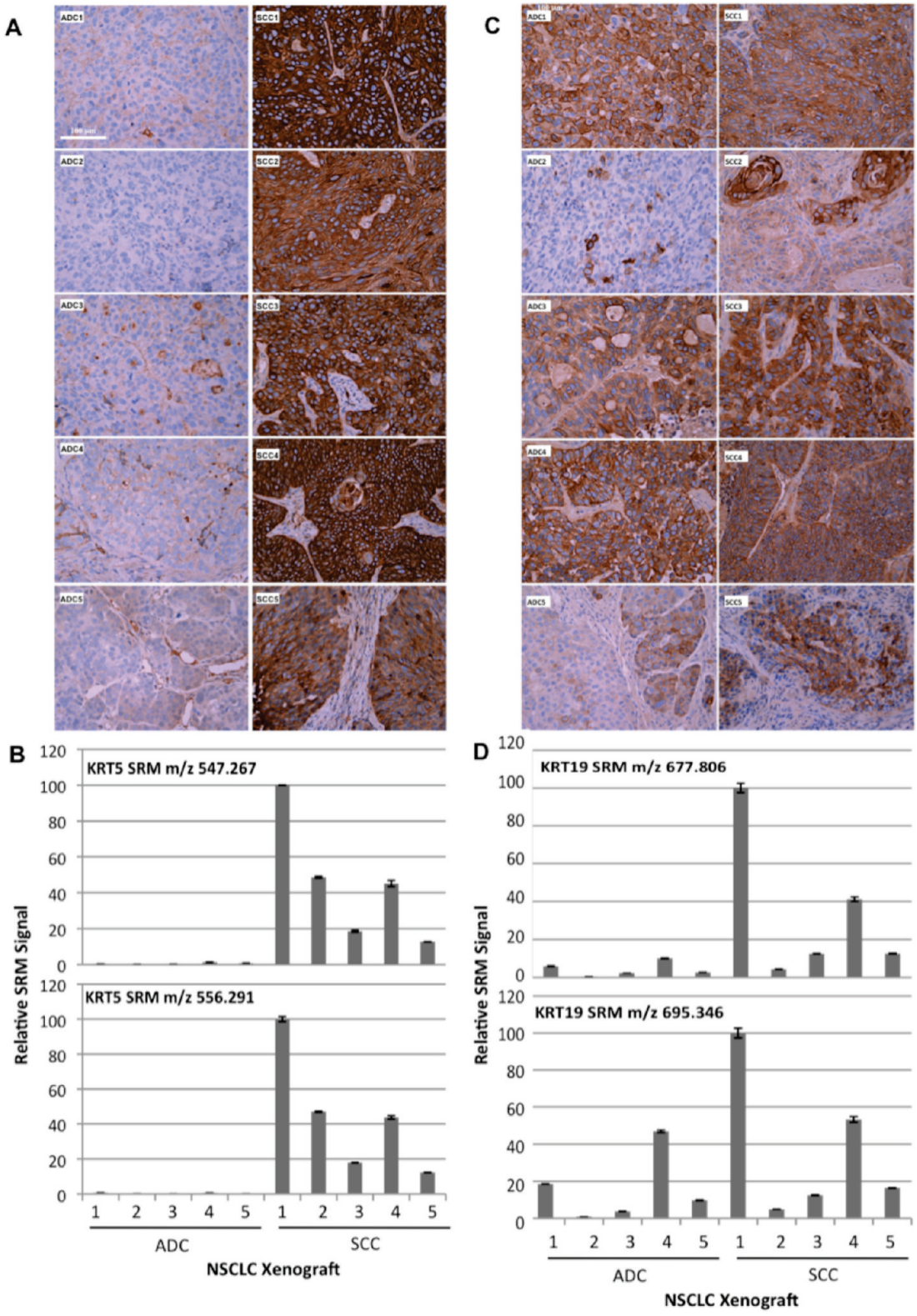

Figure 6. Immunohistochemistry and SRM analysis of KRT5 and KRT19 in NSCLC xenografts. (A) KRT5 immunohistochemistry. (B) SRM analysis of KRT5 peptides, as listed in Table 6. (C) KRT19 immunohistochemistry. (D) SRM analysis of KRT19 peptides, as listed in Table 6.

both technically and with respect to data analysis. It revealed the highly differential expression of several keratins, the urea cycle enzyme CPS1, and AGR2. The 2D findings verified the results of the 1D analysis and provided a more statistically robust data set, and with greater proteome coverage. Both approaches resulted in the identification of a set of highly differentially expressed proteins that were detected at levels differing by at least 10-fold between the 5 ADC and 5 SCC tumor-derived xenografts. More than 200 proteins (216) were identified as statistically different between ADC and SCC

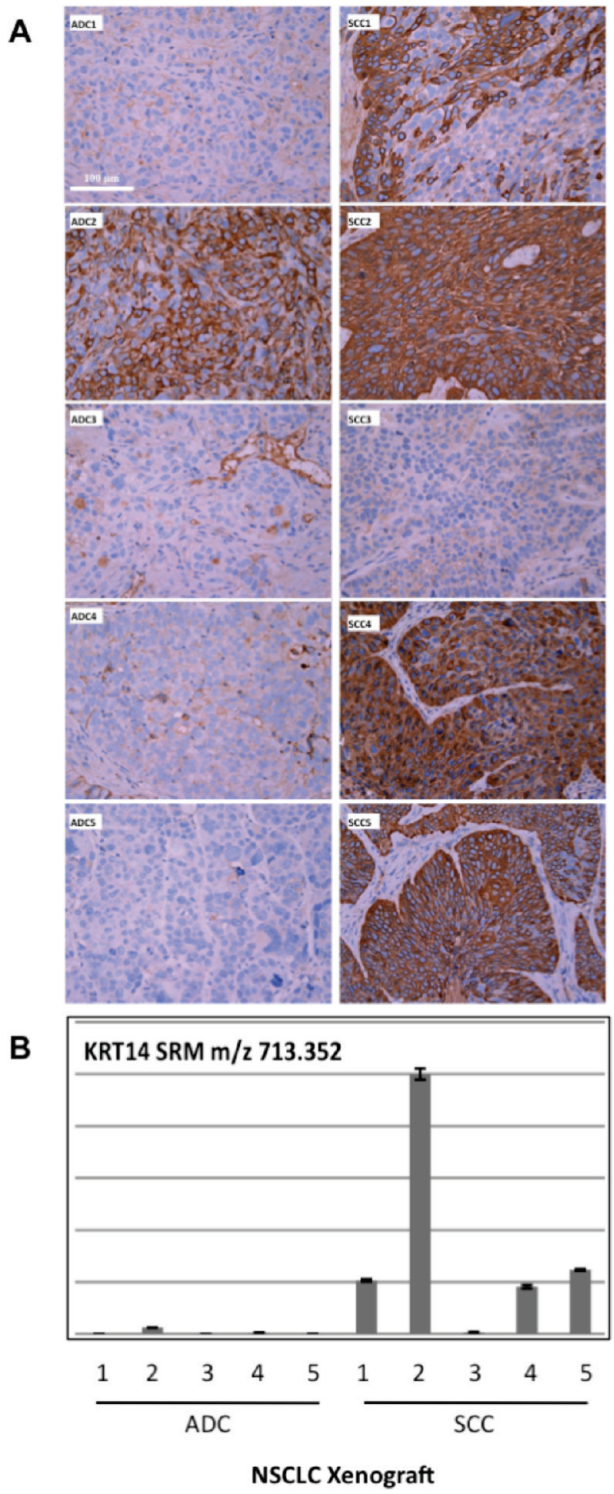

Figure 7. Immunohistochemistry and SRM analysis of KRT14 in NSCLC xenografts. (A) KRT14 immunohistochemistry. (B) SRM analysis of a KRT14 peptide, as listed in Table 6.

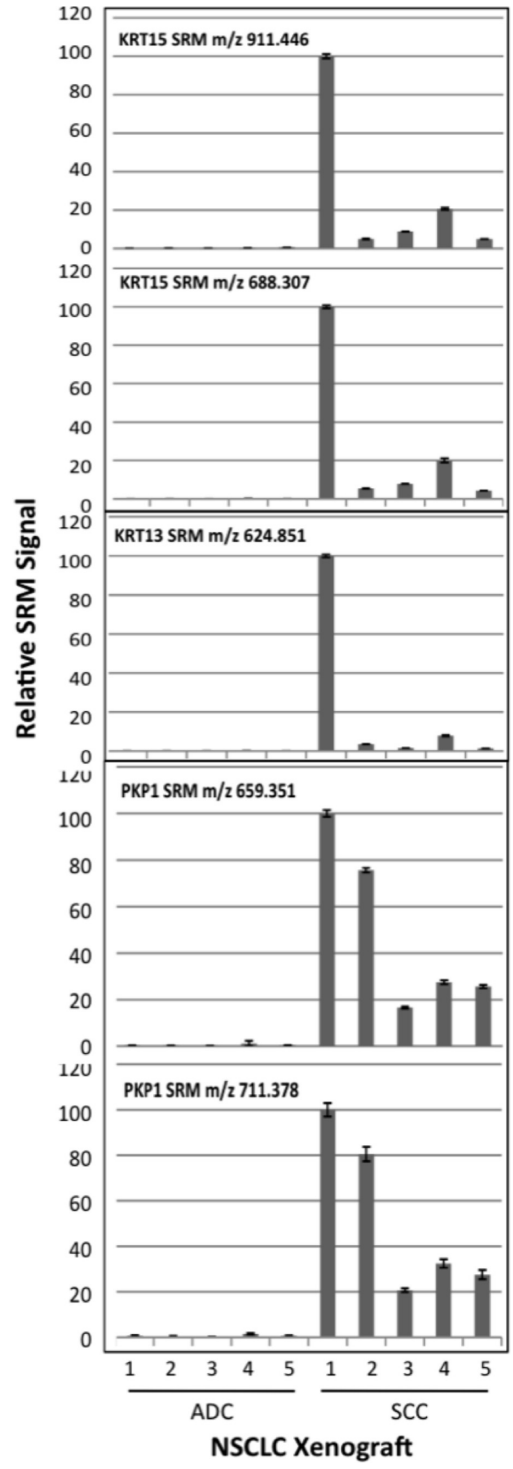

Figure 8. SRM measurements of KRT15, KRT13, and plakophilin-1 (PKP1) in NSCLC xenografts. See Table 6 for SRM transitions and associated peptides.

tumors, which includes 178 identified the 2D data set and another nonoverlapping 38 from the 1D analysis. A key element of the experimental plan was the comparison of traditional pathology laboratory methods such as IHC with quantitative proteomics. The MS-based proteomics results were verified by the IHC and Western data. In terms of tumor characterization, the SRM results complemented IHC which retains the advantage of revealing protein subcellular localization and cellular organization and heterogeneity.

Keratin Structure and Function and Clinical Relevance. Among the most abundant proteins detected by MS analysis were epithelial keratins, and they were among the most strikingly highly differentially expressed. For example, six KRTs (5, 15, 6A, 16, 4, 13) were detected exclusively in SCC (Table 5). As expected, the ADC xenografts were predominantly characterized by keratins typically associated with simple

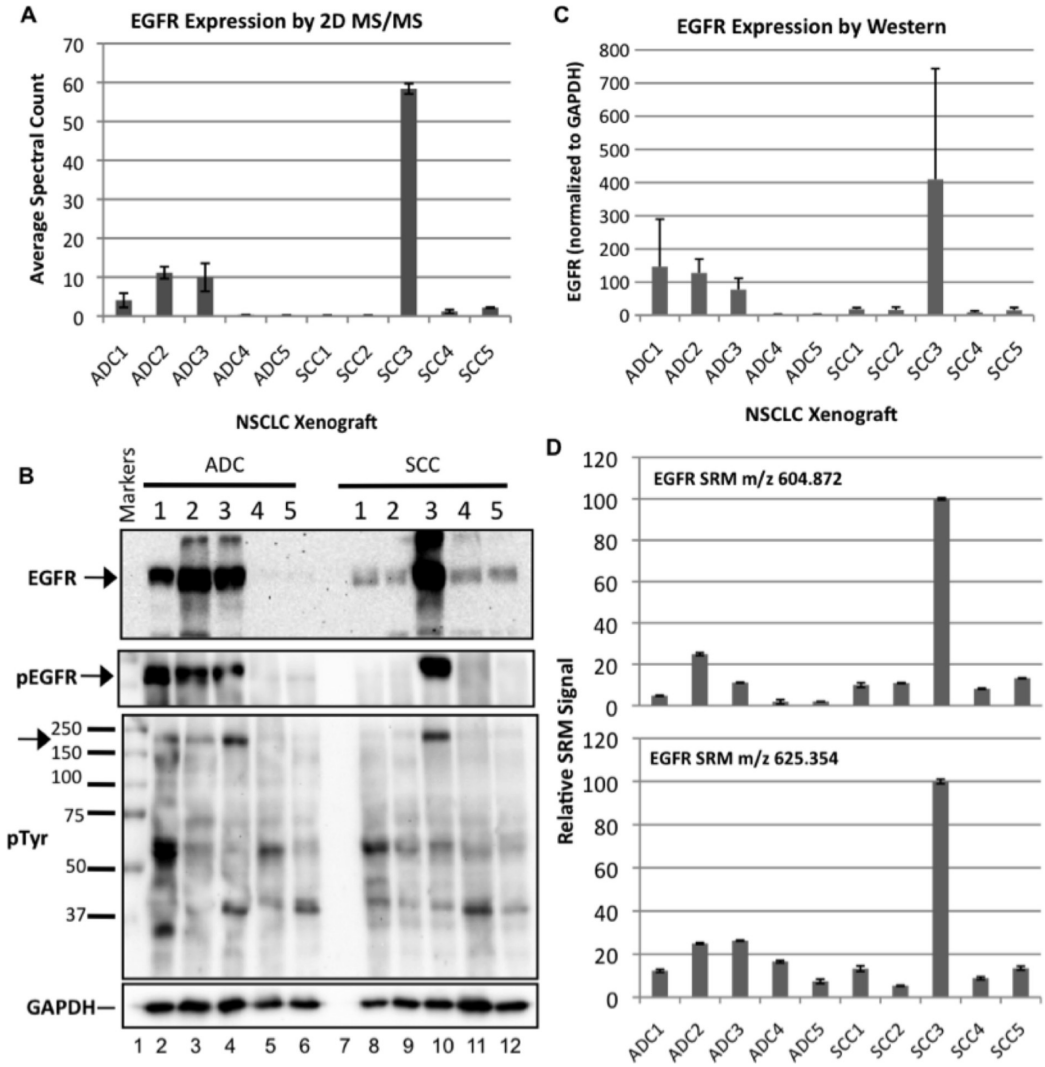

Figure 9. Analysis of EGFR expression and activation in NSCLC xenografts. (A) EGFR protein measured by spectral counting ($n = 3$, ± SD). (B,C) Anti-EGFR Western blotting (for the histogram, $n = 3$, ± SD). Receptor phosphorylation at Y1068 was imaged by Western analysis as indicated (pEGFR) and compared with total cellular antiphosphotyrosine (pTyr) staining. Arrows indicate migration of EGFR proteins. The lower panel (in B) is a Western blot of GAPDH as a loading control. (D) SRM measurement of two EGFR peptides, as listed in Table 6 ($n = 2$, error bars denote range).

epithelia; SCC also expressed these to some extent, but were notable for their expression of KRTs associated with stratified epithelium. Some but not all ADC also expressed very low levels of the squamous-type (i.e., stratified epithelial) keratins KRT14 and KRT17. The less characterized KRT80, which was detected previously in lung[27] was identified in one SCC tumor (SCC1), and three SCC xenografts expressed the rare KRT78 (refer to Table 5). SCC3 was unique both in its high level expression of EGFR, and as the only SCC found to express KRT7, which was otherwise only seen in ADC. The KRT7 expression in this instance may be an indication SCC3 arose through squamous metaplasia. The keratins are a key structural component of the 3-dimensional epithelial barrier.[5] In reference to the role of keratins in epithelia, Moll et al.[5] stated "this main cytoskeletal function transcends the single cell level." Hence, the measurement of KRT proteins in the primary xenograft model provides insight into a key structural component of three-dimensional lung tumors. van Dorst et al.[28] examined by IHC a limited set of keratins in adenocarcinomas and squamous cell carcinomas, including 16 from lung, and noted the difficulty in classifying the tumors by this method. The ability to comprehensively measure KRTs as demonstrated in this study suggests that efficient classification of ADC and SCC subtypes may be achieved, if not assisted by multiplexed-SRM-mediated, comprehensive KRT profiles.

Plakophilin-1 (PKP1) was also found highly differentially expressed in SCC, and functions in the linkage of intermediate filaments to desmosomes.[29] Interestingly, PKP1 overexpression correlates with increased cell proliferation and size, and regulates translation through interaction with eIF4A1.[30] The related protein PKP3 is up-regulated and oncogenic in NSCLC.[31] While our study was not aimed at identifying target proteins differentially expressed between tumor and normal tissue, this example illustrates the potential to discover and link tumor molecular markers with the cancer phenotype. The discovery of differential expression of the urea cycle enzyme CPS1 illustrates the potential monitoring of metabolic profiles by proteomics. Elevated expression of another urea cycle protein, ARG2, in NSCLC and in large cell carcinoma in particular, was observed previously.[32] Mutationally activated forms of EGFR are recognized drug targets in NSCLC, but how EGFR markers, such as EGFR protein expression, gene copy

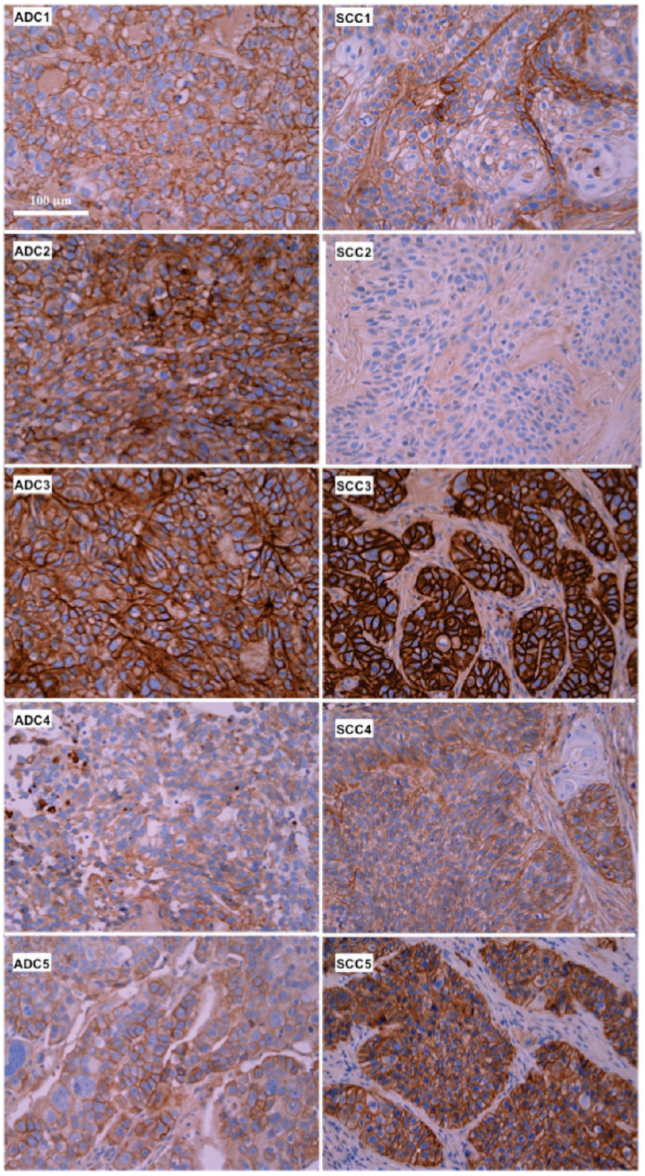

Figure 10. Immunohistochemistry of EGFR in 10 NSCLC xenografts. EGFR immunohistochemistry of representative sections in the indicated 10 NSCLC xenograft tumors.

number, and mutation status, should be incorporated into clinical decision making remains an evolving and contentious issue.[33] EGFR expression was expected in both ADC and SCC, and was reproducibly relatively quantified by spectral counting and SRM. These results complement well the measurement of EGFR by IHC, which informs of positive cells, and subcellular localization (e.g., peripheral staining corresponding to plasma membrane localization). However, it has been recognized that various EGFR mutations, exposure to ligands, and therapeutic treatments can differentially effect EGFR protein stability and subcellular localization.[34,35] Therefore, the accurate quantification of EGFR protein levels by SRM may enable the further stratification of NSCLC in terms of EGFR levels, beyond what has been achieved by IHC, measures of gene copy number, and mutations. In the 10 xenografts analyzed in this study, the range in EGFR expression was approximately 50-fold. In experiments not shown, SRM measurement of a spiked-in, stable-isotope-labeled EGFR peptide (identical in sequence to the m/z 604.87 peptide, Table 6) allowed us to estimate the level of EGFR at approximately 6×10^5 copies per cell in SCC3, which expressed the highest amount of EGFR. These examples illustrate the potential application of SRM to quantify drug target protein levels with greater precision than is achieved by current methods such as IHC. This may facilitate a better assessment of correlations in EGFR protein levels and responsiveness to EGFR-directed drugs.

The sensitivity and versatility (i.e., multiplexing) of SRM enabled the assembly of a single assay to measure keratins, the target EGFR, and examples of metabolic enzymes. In preliminary experiments we were able to monitor mouse-only peptides representing stroma and blood cells/vasculature. A more detailed examination of human/tumor and murine/stroma material is under investigation. In conclusion, we envision the development of SRM-based assays to measure NSCLC subtypes, tumor/stroma content, and the levels of expression and activation of validated drug targets such as the EGFR (and phosphorylated EGFR) and metabolic enzymes. Additional clinical utility will be realized as these methods are further adapted for the analysis of FFPE patient tissue specimens (e.g., ref 36). This information and strategic approach has the potential to improve the recognition and treatment of NSCLC and other cancers.

Acknowledgment. This work was supported through funding provided by the Canada Research Chairs Program (M.F.M., T.K.), the Ontario Research Fund—Research Excellence Award (M.-S.T., I.J., F.A.S., M.F.M., T.K.), the Ontario Institute for Cancer Research (M.F.M., M.-S.T., I.J.), and the Terry Fox Foundation Training Program in Molecular Pathology of Cancer at CIHR (STP 53912) (N.Y.). We thank Ming Li, Devang Panchal, James Ho, and Jing Xu for expert technical assistance related to xenograft procurement and IHC.

Supporting Information Available: MS data will be deposited to the Tranche repository and made freely available. A set of tables designated SI Tables 1−7 are available and provide details on the following: SI Table 1, 1D Protein Report; SI Table 2, 1D Spectral Counts; SI Table 3, 1D Differential Proteins; SI Table 4, 2D Human Protein Report; SI Table 5, 2D Murine Protein Report; SI Table 6, 2D Differential Proteins; and SI Table 7, IHC Reagents. SI Figure 1 is a Venn Diagram of Technical and Biological Reproducibility associated with the 1D LC−MS/MS analyses. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Rosell, R.; Felip, E.; Garcia-Campelo, R.; Balana, C. The biology of non-small-cell lung cancer: identifying new targets for rational therapy. *Lung Cancer* **2004**, *46* (2), 135–48.
(2) Scagliotti, G.; Hanna, N.; Fossella, F.; Sugarman, K.; Blatter, J.; Peterson, P.; Simms, L.; Shepherd, F. A. The differential efficacy of pemetrexed according to NSCLC histology: a review of two Phase III studies. *Oncologist* **2009**, *14* (3), 253–63.
(3) Coate, L. E.; John, T.; Tsao, M. S.; Shepherd, F. A. Molecular predictive and prognostic markers in non-small-cell lung cancer. *Lancet Oncol.* **2009**, *10* (10), 1001–10.
(4) Tsao, M. S.; Sakurada, A.; Cutz, J. C.; Zhu, C. Q.; Kamel-Reid, S.; Squire, J.; Lorimer, I.; Zhang, T.; Liu, N.; Daneshmand, M.; Marrano, P.; da Cunha Santos, G.; Lagarde, A.; Richardson, F.; Seymour, L.; Whitehead, M.; Ding, K.; Pater, J.; Shepherd, F. A. Erlotinib in lung cancer - molecular and clinical predictors of outcome. *N. Engl. J. Med.* **2005**, *353* (2), 133–44.
(5) Moll, R.; Divo, M.; Langbein, L. The human keratins: biology and pathology. *Histochem. Cell Biol.* **2008**, *129* (6), 705–33.
(6) Chen, F.; Luo, X.; Zhang, J.; Lu, Y.; Luo, R. Elevated serum levels of TPS and CYFRA 21−1 predict poor prognosis in advanced non-

small-cell lung cancer patients treated with gefitinib. *Med. Oncol.* **2009**, *27* (3), 950–7.
(7) de Hoog, C. L.; Mann, M. Proteomics. *Annu. Rev. Genomics Hum. Genet.* **2004**, *5*, 267–93.
(8) Troiani, T.; Schettino, C.; Martinelli, E.; Morgillo, F.; Tortora, G.; Ciardiello, F. The use of xenograft models for the selection of cancer treatments with the EGFR as an example. *Crit. Rev. Oncol. Hematol.* **2008**, *65* (3), 200–11.
(9) Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **2008**, *4*, 222.
(10) Tong, J.; Taylor, P.; Jovceva, E.; St-Germain, J.; Jin, L.; Nikolic, A.; Gu, X.; Li, Z.; Trudel, S.; Moran, M. Tandem Immunoprecipitation of phosphotyrosine-mass spectrometry (TIPY-MS) indicates C19orf19 becomes tyrosine phosphorylated and associated with activated epidermal growth factor receptor. *J. Proteome Res.* **2008**, *7* (3), 1067–77.
(11) St-Germain, J. R.; Taylor, P.; Tong, J.; Jin, L. L.; Nikolic, A.; Stewart, I. I.; Ewing, R. M.; Dharsee, M.; Li, Z. H.; Trudel, S.; Moran, M. F. Multiple Myeloma Phosphotyrosine Proteomic Profile Associated with FGFR3 expression, ligand activation, and drug inhibition. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (47), 20127–32.
(12) Taylor, P.; Nielsen, P. A.; Trelle, M. B.; Horning, O. B.; Andersen, M. B.; Vorm, O.; Moran, M. F.; Kislinger, T. Automated 2D peptide separation on a 1D nano-LC−MS system. *J. Proteome Res.* **2009**, *8* (3), 1610–6.
(13) Prakash, A.; Tomazela, D. M.; Frewen, B.; Maclean, B.; Merrihew, G.; Peterman, S.; Maccoss, M. J. Expediting the development of targeted SRM assays: using data from shotgun proteomics to automate method development. *J. Proteome Res.* **2009**, *8* (6), 2733–9.
(14) Cox, B.; Kotlyar, M.; Evangelou, A. I.; Ignatchenko, V.; Ignatchenko, A.; Whiteley, K.; Jurisica, I.; Adamson, S. L.; Rossant, J.; Kislinger, T. Comparative systems biology of human and mouse as a tool to guide the modeling of human placental pathology. *Mol. Syst. Biol.* **2009**, *5*, 279.
(15) Sodek, K. L.; Evangelou, A. I.; Ignatchenko, A.; Agochiya, M.; Brown, T. J.; Ringuette, M. J.; Jurisica, I.; Kislinger, T. Identification of pathways associated with invasive behavior by ovarian cancer cells using multidimensional protein identification technology (MudPIT). *Mol. Biosyst.* **2008**, *4* (7), 762–73.
(16) Zybailov, B.; Coleman, M. K.; Florens, L.; Washburn, M. P. Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal. Chem.* **2005**, *77* (19), 6218–24.
(17) Zybailov, B.; Mosley, A. L.; Sardiu, M. E.; Coleman, M. K.; Florens, L.; Washburn, M. P. Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **2006**, *5* (9), 2339–47.
(18) Drake, R. R.; Elschenbroich, S.; Lopez-Perez, O.; Kim, Y.; Ignatchenko, V.; Ignatchenko, A.; Nyalwidhe, J. O.; Basu, G.; Wilkins, C. E.; Gjurich, B.; Lance, R. S.; Semmes, O. J.; Medin, J. A.; Kislinger, T. In-depth proteomic analyses of direct expressed prostatic secretions. *J. Proteome Res.* **2010**, *9* (5), 2109–16.
(19) Kislinger, T.; Rahman, K.; Radulovic, D.; Cox, B.; Rossant, J.; Emili, A. PRISM, a generic large scale proteomic investigation strategy for mammals. *Mol. Cell. Proteomics* **2003**, *2* (2), 96–106.
(20) Elschenbroich, S.; Ignatchenko, V.; Sharma, P.; Schmitt-Ulms, G.; Gramolini, A. O.; Kislinger, T. Peptide separations by on-line MudPIT compared to isoelectric focusing in an off-gel format: application to a membrane-enriched fraction from C2C12 mouse skeletal muscle cells. *J. Proteome Res.* **2009**, *8* (10), 4860–9.
(21) Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242–7.
(22) Wolters, D. A.; Washburn, M. P.; Yates, J. R. 3rd, An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **2001**, *73* (23), 5683–90.
(23) Hatzfeld, M.; Franke, W. W. Pair formation and promiscuity of cytokeratins: formation in vitro of heterotypic complexes and intermediate-sized filaments by homologous and heterologous recombinations of purified polypeptides. *J. Cell Biol.* **1985**, *101* (5 Pt 1), 1826–41.
(24) Pawson, T. Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* **2004**, *116* (2), 191–203.
(25) Whiteaker, J. R.; Zhang, H.; Zhao, L.; Wang, P.; Kelly-Spratt, K. S.; Ivey, R. G.; Piening, B. D.; Feng, L. C.; Kasarda, E.; Gurley, K. E.; Eng, J. K.; Chodosh, L. A.; Kemp, C. J.; McIntosh, M. W.; Paulovich, A. G. Integrated pipeline for mass spectrometry-based discovery and confirmation of biomarkers demonstrated in a mouse model of breast cancer. *J. Proteome Res.* **2007**, *6* (10), 3962–75.
(26) Yanagisawa, K.; Shyr, Y.; Xu, B. J.; Massion, P. P.; Larsen, P. H.; White, B. C.; Roberts, J. R.; Edgerton, M.; Gonzalez, A.; Nadaf, S.; Moore, J. H.; Caprioli, R. M.; Carbone, D. P. Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* **2003**, *362* (9382), 433–9.
(27) Hesse, M.; Zimek, A.; Weber, K.; Magin, T. M. Comprehensive analysis of keratin gene clusters in humans and rodents. *Eur. J. Cell Biol.* **2004**, *83* (1), 19–26.
(28) van Dorst, E. B.; van Muijen, G. N.; Litvinov, S. V.; Fleuren, G. J. The limited difference between keratin patterns of squamous cell carcinomas and adenocarcinomas is explicable by both cell lineage and state of differentiation of tumour cells. *J. Clin. Pathol.* **1998**, *51* (9), 679–84.
(29) Bass-Zubek, A. E.; Godsel, L. M.; Delmar, M.; Green, K. J. Plakophilins: multifunctional scaffolds for adhesion and signaling. *Curr. Opin. Cell Biol.* **2009**, *21* (5), 708–16.
(30) Wolf, A.; Krause-Gruszczynska, M.; Birkenmeier, O.; Ostareck-Lederer, A.; Huttelmaier, S.; Hatzfeld, M. Plakophilin 1 stimulates translation by promoting eIF4A1 activity. *J. Cell Biol.* **2010**, *188* (4), 463–71.
(31) Furukawa, C.; Daigo, Y.; Ishikawa, N.; Kato, T.; Ito, T.; Tsuchiya, E.; Sone, S.; Nakamura, Y. Plakophilin 3 oncogene as prognostic marker and therapeutic target for lung cancer. *Cancer Res.* **2005**, *65* (16), 7102–10.
(32) Rotondo, R.; Mastracci, L.; Piazza, T.; Barisione, G.; Fabbi, M.; Cassanello, M.; Costa, R.; Morandi, B.; Astigiano, S.; Cesario, A.; Sormani, M. P.; Ferlazzo, G.; Grossi, F.; Ratto, G. B.; Ferrini, S.; Frumento, G. Arginase 2 is expressed by human lung cancer, but it neither induces immune suppression, nor affects disease progression. *Int. J. Cancer* **2008**, *123* (5), 1108–16.
(33) Shepherd, F. A.; Tsao, M. S. Epidermal growth factor receptor biomarkers in non-small-cell lung cancer: a riddle, wrapped in a mystery, inside an enigma. *J. Clin. Oncol.* **2010**, *28* (6), 903–5.
(34) Grandal, M. V.; Madshus, I. H. Epidermal Growth Factor Receptor and Cancer: Control of Oncogenic Signalling by Endocytosis. *J Cell Mol Med* **2008**, *12* (5A), 1527–34.
(35) Hynes, N. E.; Lane, H. A. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat. Rev. Cancer* **2005**, *5* (5), 341–54.
(36) Taylor, P.; Tong, J.; Shih, W.; Darfler, M.; Tsao, M.; Krizman, D.; Eitner, C.; Moran, M. F. Detection and quantification of EGF receptor phosphorylation in formalin-fixed tumor sections by selected/multiple reaction monitoring mass spectrometry. *Eur. J. Cancer* **2009**, *7* (4), 31.

PR100491E