ORIGINAL ARTICLE

Lung Adenocarcinoma and Squamous Cell Carcinoma Gene Expression Subtypes Demonstrate Significant Differences in Tumor Immune Landscape

Hawazin Faruki, DrPH,[a,*] Gregory M. Mayhew, DrPH,[a] Jonathan S. Serody, MD,[b] D. Neil Hayes, MD,[b] Charles M. Perou, PhD,[b,c] Myla Lai-Goldman, MD[a]

[a]*GeneCentric Diagnostics, Durham, North Carolina*
[b]*Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina*
[c]*Department of Genetics, University of North Carolina, Chapel Hill, North Carolina*

Received 30 January 2017; revised 13 March 2017; accepted 14 March 2017
Available online - 21 March 2017

ABSTRACT

Introduction: Molecular subtyping of lung adenocarcinoma (AD) and lung squamous cell carcinoma (SCC) reveal biologically diverse tumors that vary in their genomic and clinical attributes.

Methods: Published immune cell signatures and several lung AD and SCC gene expression data sets, including The Cancer Genome Atlas, were used to examine immune response in relation to AD and SCC expression subtypes. Expression of immune cell populations and other immune related genes, including CD274 molecule gene (*CD274*) (programmed death ligand 1), was investigated in the tumor microenvironment relative to the expression subtypes of the AD (terminal respiratory unit, proximal proliferative, and proximal inflammatory) and SCC (primitive, classical, secretory, and basal) subtypes.

Results: Lung AD and SCC expression subtypes demonstrated significant differences in tumor immune landscape. The proximal proliferative subtype of AD demonstrated low immune cell expression among ADs whereas the secretory subtype showed elevated immune cell expression among SCCs. Tumor expression subtype was a better predictor of immune cell expression than *CD274* (programmed death ligand 1) in SCC tumors but was a comparable predictor in AD tumors. Nonsilent mutation burden was not correlated with immune cell expression across subtypes; however, major histocompatibility complex class II gene expression was highly correlated with immune cell expression. Increased immune and major histocompatibility complex II gene expression was associated with improved survival in the terminal respiratory unit and proximal inflammatory subtypes of AD and in the primitive subtype of SCC.

Conclusions: Molecular expression subtypes of lung AD and SCC demonstrate key and reproducible differences in immune host response. Evaluation of tumor expression subtypes as potential biomarkers for immunotherapy should be investigated.

Keywords: Lung cancer; Adenocarcinoma; Squamous cell carcinoma; Immune response; PD-L1; Gene expression

Introduction

NSCLC is a heterogeneous disease that is typically classified into two broad subtypes, adenocarcinoma (AD) and squamous cell carcinoma (SCC), by using standard pathology methods. Beyond the morphologic differentiation, multiple gene expression subtypes that differ in their prognosis, underlying genomic alterations, and

***Corresponding author.**

***Disclosure:** Drs. Faruki, Mayhew, and Lai-Goldman are employees of GeneCentric Diagnostics and are named in three pending patents for lung subtyping. Drs. Perou and Hayes are board of directors members, equity stock holders, and consultants for GeneCentric Diagnostics and are each also named as inventors on an issued U.S. patent and three other pending patents for lung cancer subtyping. Dr. Serody is a consultant for GeneCentric Diagnostics.*

Address for correspondence: Hawazin Faruki, DrPH, GeneCentric Diagnostics, 280 South Mangum St., Suite 350, Durham, NC 27701. **E-mail:** Hawazin@genecentric.com

ISSN: 1556-0864

http://dx.doi.org/10.1016/j.jtho.2017.03.010

potential response to treatment have been identified within AD and SCC tumors.[1–4] The three gene expression AD subtypes include terminal respiratory unit (TRU), proximal proliferative (PP), and proximal Inflammatory (PI),[2] which were formerly referred to as the *bronchioid*, *magnoid,* and *squamoid* subtypes, respectively.[4] SCC includes four subtypes: primitive, classical, basal, and secretory.[1,3] Lung AD and SCC expression subtypes are not discernable by standard morphologically based diagnoses; however, they demonstrate significant differences in key genomic, genetic, and clinical characteristics, including tumor differentiation, stage-specific survival, underlying drivers, and likely response to various therapies.[1–4] The AD subtype TRU is characterized by enrichment for alterations in *EGFR* and anaplastic lymphoma receptor tyrosine kinase gene (*ALK*), nonsmokers, and a better prognosis. Serine/threonine kinase 11 gene (*STK11*) deletion, high proliferation, brain metastases, and poor prognosis characterize the PP subtype, and tumor protein p53 gene (*TP53*) mutations are more characteristic of the PI subtype.[2,4,5] The SCC subtypes are enriched for retinoblastoma 1 gene (*RB1*) loss in the primitive subtype, kelch like ECH associated protein gene (*KEAP*)/nuclear factor erythroid 2, like 2 gene (*NFE2L2*) oxidative stress alterations in the classical subtype, greater inflammatory response in the secretory subtype, and neurofibromin 1 gene (*NF1*) alterations in the basal subtype.[1,3]

Investigation of the tumor immune response across multiple tumor types by using genomic data has been valuable in identifying the role of tumor-infiltrating lymphocytes in prognosis and response to immunotherapy.[6–8] Bremnes et al. recently reviewed the role of tumor-infiltrating lymphocytes and the associated composition of the immune microenvironment as key determinants of the outcomes of patients with NSCLC.[9] Immune checkpoint inhibitors targeting the programmed cell death 1/programmed death ligand 1 (PD-L1) interaction have been shown to reverse the lung tumor–induced immunosuppressive microenvironment, releasing an effective host antitumor immune response and leading to remarkable improvements in survival.[10–12]

Biomarkers for predicting response to immunotherapy have included a variety of anti–PD-L1 antibodies assayed by using immunohistochemistry (IHC). In melanoma and lung AD, increase in tumor cells expressing PD-L1 has been associated with an increase in response[10,12,13]; however, variable cutoffs, multiple antibodies with differing affinities, and lack of method standardization have resulted in conflicting findings regarding the value of PD-L1 IHC testing.[11,12,14] Other predictors of response, including nonsilent mutation and neoantigen burden[15] and gene expression of a variety of immune response genes,[16] have been studied but remain investigational tools at this time. In this study, we explore the immune landscape of genomic subtypes of lung AD and SCC to characterize the immune microenvironment by using publicly available genomic data sets. Given the intrinsic biologic differences of gene expression subtypes of lung AD and SCC tumors, we were interested in investigating subtype-specific immune characteristics and associated immune cell/marker expression differences that might contribute to our understanding of prognosis and/or response to immunotherapy.

Materials and Methods

Sample Data Sets

Multiple publicly available data sets were assembled to evaluate tumor subtype infiltration of immune cells and associated survival differences. The AD and SCC data sets included several publicly available lung cancer gene expression data sets. Data sources and sample numbers and characteristics are provided in Supplementary Table 1.

Publicly Available Expression Data Sets and Subtype Assignments

We used four previously published AD data sets with a total of 1190 patient samples. The published data included the data sets in The Cancer Genome Atlas (TCGA)[2] and the articles by the Director's Challenge Consortium for the Molecular Classification of Lung A et al.,[17] Tomida et al.,[18] and Wilkerson et al.,[4] all of which were derived from fresh frozen specimens. The published TCGA data included expression profiles from 58 tumor-adjacent normal lung tissue samples. For the TCGA, upper quantile normalized RNA sequencing (RNA-Seq) by expectation-maximization data were downloaded from Firehose[19] and log2-transformed. Affymetrix Cel files from the article by the Director's Challenge Consortium for the Molecular Classification of Lung A et al.[17] were downloaded from the National Cancer Institute website (https://wiki.nci.nih.gov/display/caIntegrator/caIntegrator+Directory), and robust multi-array average expression measures were generated by using the Affy package in R (Affymetrix, Santa Clara, CA). Normalized Agilent array data (Agilent, Santa Clara, CA) were downloaded from the Gene Expression Omnibus (GEO) website for Tomida et al.[18] (GSE13213) and Wilkerson et al.[4] (GSE26939).

We used four published gene expression data sets of lung SCC samples representing a total of 761 patients, including the TCGA[1] and the articles by Lee et al.,[16] Raponi et al.,[20] and Wilkerson et al.[3] The published TCGA data included expression profiles from 51 tumor-adjacent normal lung tissue samples. For the TCGA,

upper quantile normalized RNA-Seq by expectation-maximization data were downloaded from Firehose[19] and log2-transformed. Normalized Affy array data were downloaded from GEO for Lee et al.[16] (GSE8894) and Raponi et al.[20] (GSE4573), and normalized Agilent array data were downloaded from GEO for Wilkerson et al.[3] (GSE17710). To determine AD subtype (TRU, PP, or PI) and SCC subtype (basal, classical, primitive, or secretory), we applied the published AD 506-gene nearest centroid classifier and the SCC 208 classifier as described previously in Wilkerson et al.[3,4] A full list of data sets used is included in Supplementary Table 1.

Gene Sets

Our investigation of immune differences by subtype used the 24 immune cell gene signatures from Bindea et al.,[6] each of which had a varying number of genes and was classified as an adaptive or innate immunity cell signature. Adaptive immune cell (AIC) signatures included T cells, central memory T (Tcm) cells, effector memory T cells, type 1 helper (Th1) T cells, type 2 helper T (Th2) cells, T follicular helper (TFH) cells, T helper 17 (Th17) cells, regulatory T (Treg) cells, $\gamma\delta$ T (T$\gamma\delta$) cells, CD8 T cells, cytotoxic T cells, and B cells, and innate immune cell signatures included natural killer (NK) cells, NK CD56dim cells, NK CD56bright cells, dendritic cells (DCs), immature DCs, plasmacytoid DCs, activated DCs, mast cells, eosinophils, macrophages, and neutrophils. For each signature, we assigned a score to each sample by calculating the average expression value of all genes in the list. The interferon (IFN) gene signature was a new signature developed by us for this analysis and included 13 interferon signaling pathway genes selected by the authors to investigate IFN pathway expression; the IFN gene list can be found in Supplementary Table 2. A 13-gene major histocompatibility complex class II (MHC II) signature score,[21] was also included along with four additional individual gene immunity markers: programmed cell death 1 gene (*PDCD1*), PD-L1 (CD274 molecule gene [*CD274*]), cytotoxic T-lymphocyte associated protein 4 gene (*CTLA4*), and programmed death ligand 2 (programmed cell death 1 ligand 2 gene [*PDCD1LG2*]). For all signatures, a summary of missing genes in various data sets is included in Supplementary Table 2.

Immune Cell Genomic Evaluations

Using the TCGA lung AD and SQ data to investigate overall immunity marker trends by subtype, we plotted expression heatmaps separately for AD and SCC in which samples were arranged by subtype and markers were grouped according to their ordering in Bindea et al.,[6] Marker-subtype association test p values (Kruskal-Wallis test) and Mann-Whitney test p values for marker distribution differences between each pair of subtypes were calculated. Correlations among the 30 immune markers were made by plotting matrices of pairwise Spearman rank correlation coefficients in which markers were ordered by hierarchical clustering. We compared the MHC II signature in tumor versus in tumor-adjacent normal lung tissue by using the Mann-Whitney test. To evaluate the reproducibility of immunity marker differences among the subtypes, we plotted normalized T-cell signatures by subtype for each data set.

Prediction Strength

To assess the prediction strength of subtype as a predictor of immune markers relative to that of PD-L1, a linear regression model of each signature with subtype as the sole predictor, and again with PD-L1 as the sole predictor, was fit by using the TCGA data set. PD-L1 expression was treated as a low/high categorical variable with equal proportions in each group. Scatter plots of adjusted R^2 when subtype was the predictor against adjusted R^2 when PD-L1 was the predictor were inspected for overall trends.

Genomic Associations with T-Cell Expression

Several genetic and genomic alterations characteristic of AD and SCC subtypes, including *EGFR*, *TP53*, and *STK11* inactivation in AD and *RB1*, *NFE2L2,* and *NF1* expression in SCC, were examined for association with T-cell expression with and without adjustment for subtype by using linear regression. Mutation and copy number variation data were downloaded from Firehose[19] and for *STK11* samples were called inactive when reported as deleted and/or mutated. When association with T cells was strong, we plotted the marker distribution by subtype and evaluated association evidence by using Fisher's exact test and the Kruskal-Wallis test for binary and continuous markers, respectively, and compared every pair of subtypes for marker distribution differences by using Fisher's exact test or the Mann-Whitney test.

With data on nonsilent mutation burden per megabase available in the supplementary TCGA information,[1,2] we investigated association with T-cell expression by using linear regression. Association between mutation burden and subtype was evaluated overall by using the Kruskal-Wallis test and between each pair of subtypes by using the Mann-Whitney test.

Subtype and immune signature associations with a 13-gene MHC II signature,[21] which were calculated as an average of all genes in the list (the gene list in Supplementary Table 2), were investigated by using the Kruskal-Wallis test for overall differences and the Mann-Whitney test for comparing two subtypes. For immune

signature–MHC II associations in both tumor and tumor-adjacent normal tissue, Spearman correlation coefficients and *p* values were calculated.

Survival Analysis

We tested for immune marker–survival associations in the TCGA data sets, overall and separately within each subtype, by using Cox proportional hazards models. Immune markers were left as continuous variables after being centered and scaled to have a mean of 0 and variance of 1. Stage IV patients were excluded from the analysis because of heterogeneity in their clinical management and poor representation (i.e., small sample size) in the genomic data sets. Evaluations within a specific subtype were adjusted for stage, and overall evaluations were adjusted for both stage and subtype. Forest plots showing hazard ratios and confidence intervals for significant signatures were made. All statistical analyses were conducted with R 3.2.0 software (http://www.R-project.org).

Results

Immune Cell Evaluations by Subtype

Examination of immune cell gene signatures including both AICs and innate immune cells as well as individual immune gene markers revealed clear differences among the AD and SCC subtypes (Fig. 1). In AD, immune expression was consistently lower in the PP subtype for most cell types examined. Expression was similar in the TRU and PI subtypes for most T cells but could be differentiated between the TRU and PI subtypes by greater expression of some innate immune cells (DCs, NK CD56bright cells, mast cells, and eosinophils) and several AICs (B cells, TFH cells, Tcm cells, Th17 cells, and CD8 T cells) in the TRU subtype, whereas the PI subtype showed higher expression of Th1 and Th2 cells, Treg cells, cytotoxic T cells, and NKCD56dim cells. Box plots of all the immune cells and markers by AD and SCC subtype can be found in Supplementary Figure 1. Marker-subtype association test *p* values (Kruskal-Wallis test) and Mann-Whitney test *p* values for marker distribution differences between each pair of subtypes can be found in Supplementary Table 3.

Immunotherapy targets, *CTLA4* and *CD274 (*PD-L1), demonstrated consistently higher expression in the PI subtype across multiple data sets (see Supplementary Fig. 1). In the PP tumors, both AICs and innate immune cell expression as well as immunotherapy target expression were low relative to those in other ADs (see Supplementary Fig. 1). Among the SCC subtypes, the secretory subtype showed consistently higher immune cell expression of both innate and AICs with one exception, the Th2 signature, in which case both primitive and

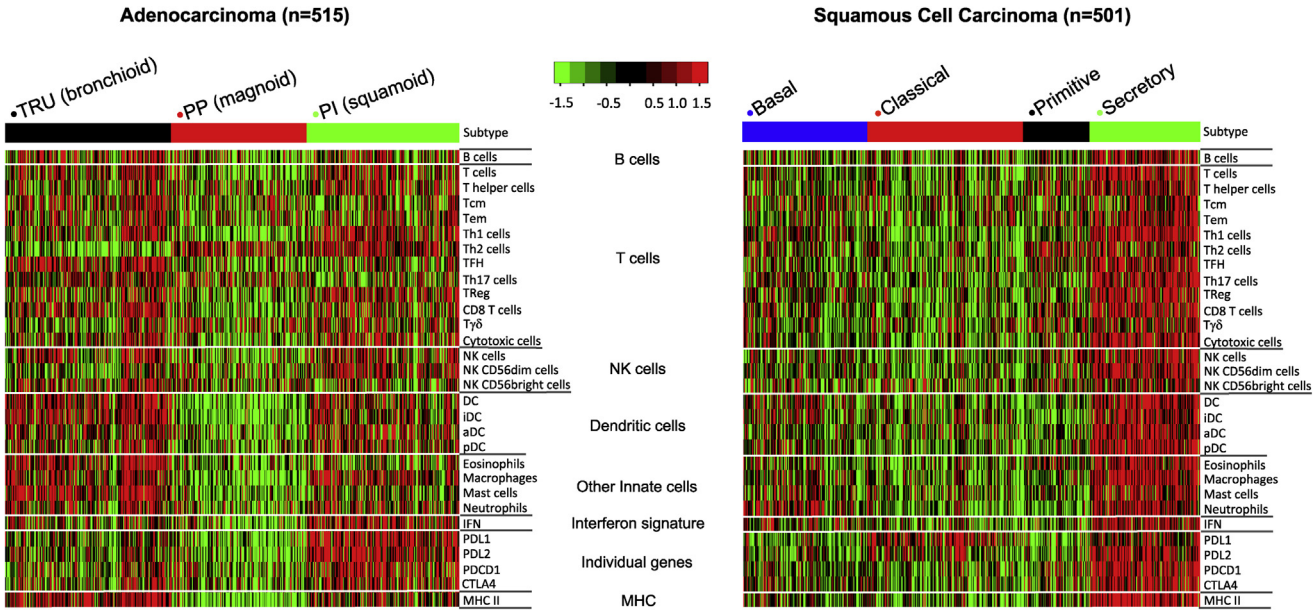

Figure 1. Heatmaps of Bindea et al.[6] showing immune cell signature expression, other immune signatures, and individual immune markers in lung adenocarcinoma and squamous cell carcinoma gene expression data sets.[1,2] TRU, terminal respiratory unit; PP, proximal proliferative; PI, proximal inflammatory; Tcm, central memory T cell; Tem, effector memory T cell; Th1, type 1 T helper cell; Th2, type 2 T helper cell; TFH, T follicular helper cell; Th17, T-helper 17 cell; Treg, regulatory T cell; T$\gamma\delta$, $\gamma\delta$ T cell; NK, natural killer; DC, dendritic cell; iDC, immature dendritic cell; aDC, activated dendritic cell; pDC, plasmacytoid dendritic cell; IFN, interferon; PD-L1, programmed death ligand 1; PD-L2, programmed death ligand 2; PDCD1, programmed cell death 1; CTLA4, cytotoxic T-lymphocyte associated protein 4; MHC II, major histocompatibility complex class II.

secretory had comparable expression (see Supplementary Fig. 1). The classical subtype demonstrated the lowest immune cell expression of all the SCC subtypes. Unlike in the case of AD subtypes, *CD274 (*PD-L1) expression did not correlate with other immune cell expression in SCC subtypes. This is especially obvious in the classical subtype where *CD274* (PD-L1) expression was high despite relatively low expression of other immune cells (see Fig. 1 and Supplementary Fig. 1).

Hierarchical clustering analysis grouped AICs together and innate immune cells together. In AD, AICs such as T cells, cytotoxic cells, CD8 cells, Th1 cells, *PDCD1*, *CTLA4,* and Treg cells had high pairwise correlations (Spearman correlation coefficient >0.53), and similarly for innate immune cells, immature DCs, DCs, macrophages, neutrophils, mast cells, and eosinophils had high pairwise correlations (Spearman correlation coefficient >0.52). In SCC, patterns were similar (Spearman correlation coefficient >0.68 for AICs and >0.55 for innate immune cells). NK cells were not consistently correlated with innate immune cells. In particular, NK CD56dim cells (cytolytic activity) were more strongly correlated with AICs than with innate immune cells (Spearman correlation coefficient >0.50 for AICs versus >0.09 for innate cells in AD and >0.70 for AICs versus >0.42 for innate cells in SCC). Hierarchical clustering is shown in Supplementary Figure 2, and all correlation coefficients and *p* values for each pair of immune markers are included in Supplementary Table 4.

Strength of Association with AIC Expression

Strength of association between *CD274* (PD-L1) expression and AIC signatures as compared with AD or SCC subtype was investigated. Figure 2 shows that CD274-signature associations (measured by adjusted r^2) were stronger than AD subtype for some cells (T, Th1, Treg, cytotoxic, T helper, effector memory T, and T$\gamma\delta$ cells) but not for others (TFH, Th2, CD8, Th17, and Tcm cells). In AD, median *F*-test *p* value and adjusted R^2 of the predictive strength of gene expression subtype versus *CD274* expression for association with AICs was 5.97E−13 and 0.10 for gene expression subtype versus 1.18E−10 and 0.08 for *CD274* expression. In SCC tumors, subtype was a better predictor of AIC expression for all cells examined and the median *F*-test *p* value and adjusted R^2 were 2.16E−24 and 0.20 for subtype versus 4.36E−5 and 0.03 for *CD274* (see Fig. 2).

Evaluation of T-Cell Signature in Multiple Data Sets

To evaluate the reproducibility of our findings, we examined individual signatures across each data set separately, with typical results shown in Figure 3. T-cell immune signature expression subtype patterns by subtype in AD and SCC were remarkably reproducible across a variety of gene expression data sets and a variety of gene expression platforms, including RNA-Seq (Illumina, San Diego, CA) and microarrays from both Affymetrix and Agilent (see Fig. 3). AD and SCC subtypes showed similar T-cell expression patterns independent of platform. Similar observations were made for other immune cell types and expression biomarkers, including B cells, NK cells, DCs, macrophages, *PDCD1* (programmed cell death 1), and *CD274* (PD-L1), which can be seen in Supplementary Figure 3.

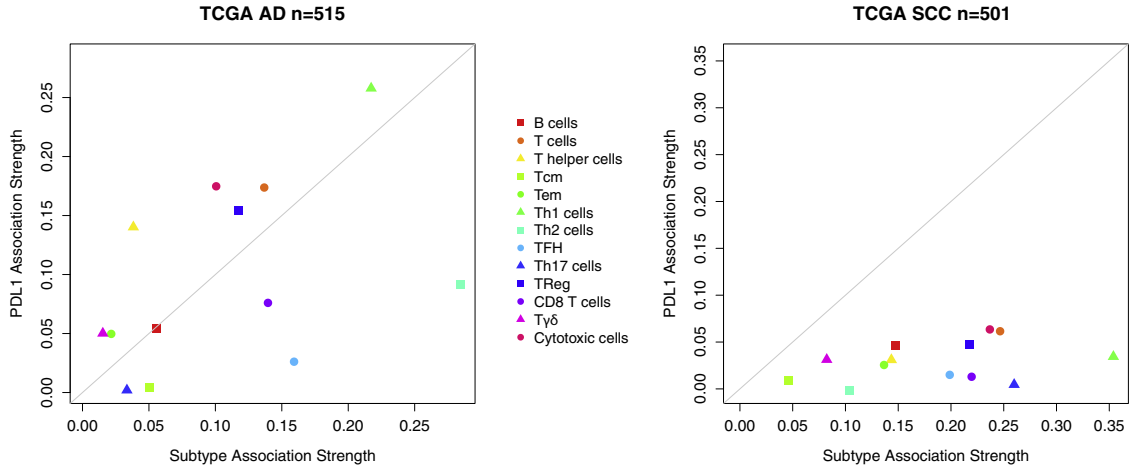

Figure 2. Association strength (adjusted r^2) between CD274 molecule gene (*CD274*) (programmed death ligand 1 [PD-L1]) expression and immune signature versus strength between subtype and immune signature for 13 adaptive immune cell (AIC) expression signatures in adenocarcinoma (AD) and squamous cell carcinoma (SCC) data sets. Association between subtype and AIC was greater for some AICs in AD and for all AICs tested in SCC. Tcm, central memory T cell; Tem, effector memory T cell; Th1, type 1 T helper cell; Th2, type 2 T helper cell; TFH, T follicular helper cell; Th17, T helper 17 cell; Treg, regulatory T cell; T$\gamma\delta$, $\gamma\delta$ T cell.

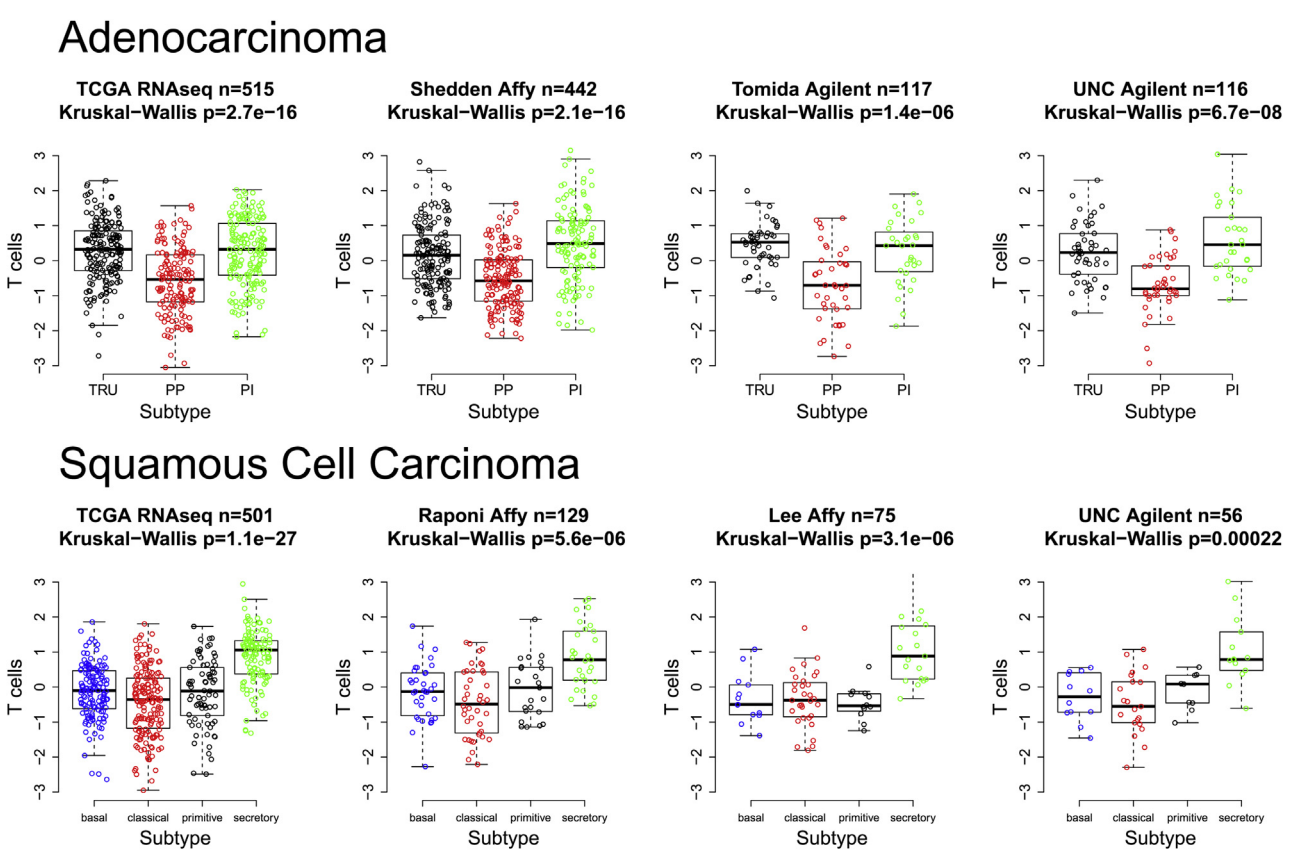

Figure 3. Reproducibility of T-cell signature gene expression subtype patterns across multiple adenocarcinoma data sets[2,4,17,18] and squamous cell carcinoma data sets.[1,3,16,20] RNA sequencing (RNA-Seq Kit [Illumina, San Diego, CA]) and microarrays from both Affymetrix (Santa Clara, CA) and Agilent (Santa Clara, CA). TCGA, The Cancer Genome Atlas; TRU, terminal respiratory unit; PP, proximal proliferative; PI, proximal inflammatory.

Somatic Genetic and Genomic Associations with T-Cell Expression

Evaluation of mutation burden and its association with T-cell expression in the AD and SCC subtypes was conducted. Nonsilent mutation burden in the TCGA AD data differed by subtype, with the PI subtype showing the highest burden and the TRU subtype showing the lowest burden (Fig. 4*A*). The PI subtype, which is enriched for *TP53* mutations[2] and had greater mutation burden, was associated with higher immune cell expression. However, the TRU subtype had the lowest mutation burden (see Fig. 4*A*) and rarely harbored *TP53* mutations,[2] yet it demonstrated high immune cell expression features (see Fig. 1 and Supplementary Fig. 1). In SCC, nonsilent mutation burden was not significantly different across subtypes ($p = 0.54$ [Fig. 4*B*]) despite significant differences in T-cell expression (see Fig. 1 and Supplementary Fig. 1). When linear regression was used, mutation burden was not found to be correlated with T-cell immune cell expression in either AD or SCC data sets ($p = 0.24$ in AD and $p = 0.9$ in SCC [Supplementary Fig. 4*D–H*]).

Several AD and SCC gene expression subtype–enriched genetic and genomic alterations were investigated for their association with T-cell expression, including *EGFR*, *TP53*, and *STK11* inactivation in AD and *RB1*, *NF1*, and *NFE2L2* expression in SCC. In AD, only *STK11* inactivation was markedly associated with T-cell gene expression signatures ($p = 0.0007$ [see Supplementary Fig. 4*A*]), but the association was lost after adjustment for expression subtype ($p = 0.43$). In SCC, *NFE2L2* expression was associated with T cells ($p = 1.2E{-}07$ [see Supplementary Fig. 4*G*]), as was *NF1* expression ($p = 0.01$ [see Supplementary Fig. 4*F*]), but adjustment for expression subtype eliminated significance to $p = 0.47$ and $p = 0.26$, respectively. Loss of *STK11* in AD[22–24] and *KEAP/NFE2L2* alterations in SCC[25] have been associated with reduced immune response in NSCLC. Our evaluation showed enrichment of *STK11* inactivation in the low–immune response AD PP subtype (see Fig. 4*C*) and *KEAP/NFE2L2* alterations affecting the oxidative stress pathway in the low–immune response SCC classical subtype (see Fig. 4*D*); however, neither *STK11* nor *NFE2L2* was a significant predictor after adjustment for expression subtype (see Supplementary Fig. 4*A* and *G*).

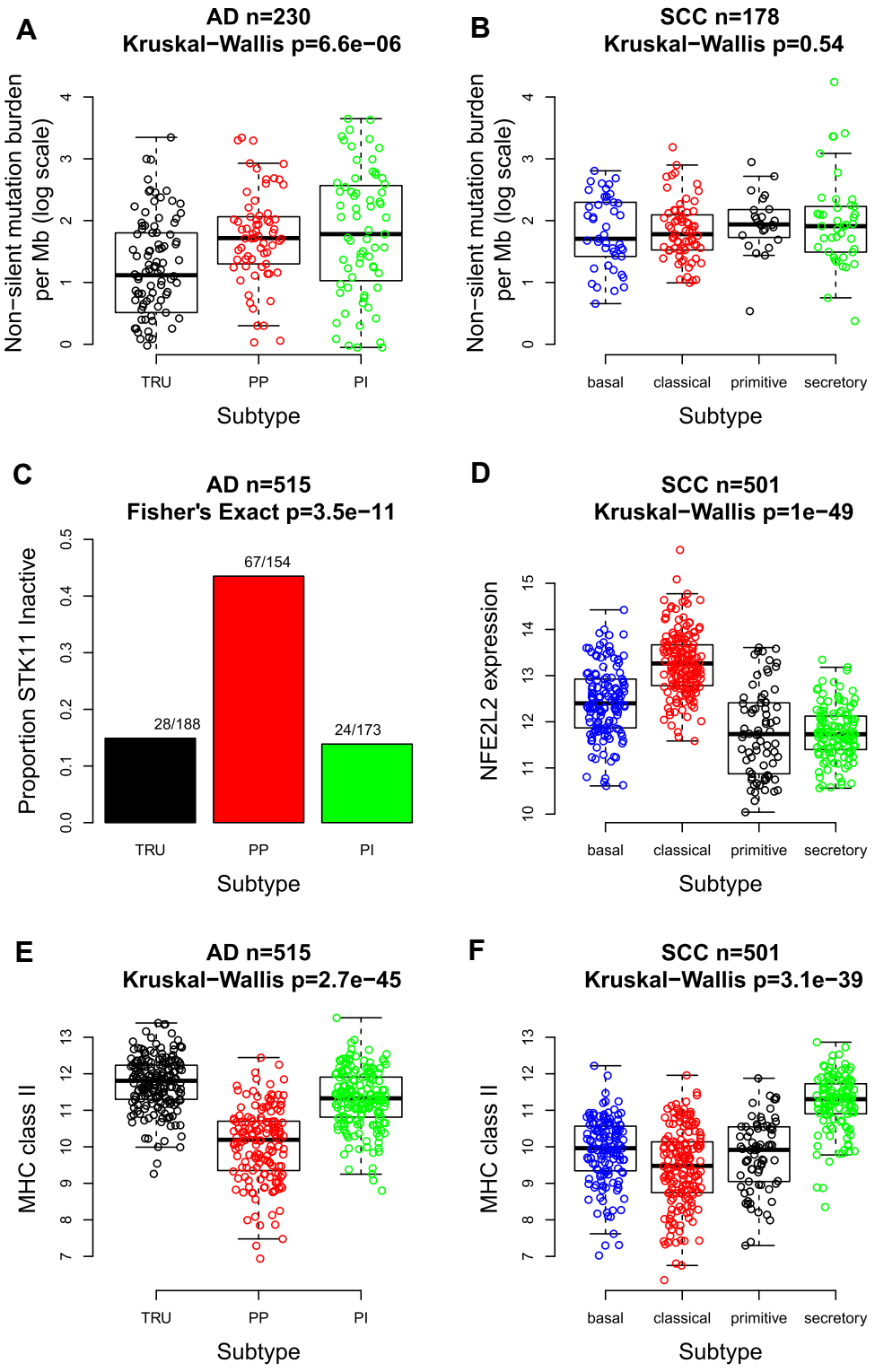

Figure 4. Adenocarcinoma (AD) and squamous cell carcinoma (SCC) subtype nonsilent mutation burden (*A* and *B*), serine/threonine kinase 11 gene (*STK11*) inactivation (mutation and/or deletion) in AD (*C*), nuclear factor erythroid 2, like 2 gene (*NFE2L2*) expression in SCC (*D*), and major histocompatibility complex class II (MHC II) signature (*E* and *F*) with association test *p* values. Mb, megabase; TRU, terminal respiratory unit; PP, proximal proliferative; PI, proximal inflammatory.

MHC II

Given recent interest in MHC II gene expression association with immune infiltration in triple-negative breast cancer[21] and in colon cancer,[6] in combination with our own observations of differential expression of several MHC II genes across the subtypes, we

investigated the association of immune cell expression in AD and SCC lung cancer with MHC II genes by using a published 13-gene MHC II signature.[21] MHC II gene expression varied significantly across tumor subtypes (Figures 4*E* and *F*) ($p = 2.7E{-}45$ and 3.1E–39 in AD and SCC, respectively). We next examined tumor-adjacent normal lung tissue expression profiles from the AD and SCC data sets (n = 58 and 51, respectively) to evaluate whether MHC II expression was lung tissue–specific versus tumor specific. Higher MHC II expression in tumor-adjacent normal lung tissue as compared with in lung tumor tissue suggested tumor-specific decreased expression in the tumor microenvironment as compared with in normal lung tissue taken from patients with lung cancer ($p = 4.8E{-}16$ and $p < 2.2E{-}16$ in AD and SCC, respectively). MHC II gene expression was strongly correlated with several immune cells in both AD and SCC, including T-cell expression (Spearman correlation of 0.66 in AD and 0.86 in SCC), B-cell expression (Spearman correlation of 0.5 in AD and 0.69 in SCC), and DC expression (Spearman correlation of 0.69 in AD and 0.76 in SCC). Scatter plots of MHC II and immune cell expression in AD and SCC subtypes with correlation coefficients and associated *p* values can be found in Supplementary Figure 5.

Tumor-adjacent normal MHC II expression was uniformly high relative to tumor samples (see Supplementary Fig. 5), and correlations with immune cell expression were lower. In contrast with the results regarding other subtype-specific genomic alterations, MHC II expression remained a significant predictor of T-cell immune cell expression in linear regression models after adjustment for subtype in the AD and SCC data sets ($p < 1E{-}50$).

Survival Analysis

Immune infiltrates have been associated with improved survival in NSCLC[9,26–29]; however, analyses have typically not considered the subtypes of AD or SCC. Using Cox proportional hazard models, we calculated subtype-specific hazard ratios per unit increase in normalized expression. Hazard ratios and confidence intervals for markers that were significant (nominal *p* value <0.05) for at least one subtype after adjustment for pathologic stage are shown in Figure 5. For AD subtypes, a unit increase in expression for many innate and AICs, including *CD274* (PD-L1), *CTLA4*, and MHC II signature, was significantly associated with improved survival in the PI subtype of AD but not in other AD subtypes (Fig. 5*A*).

Among the SCC subtypes, a unit increase in expression for Th1 cells, Th2 cells, TFH cells, DCs, macrophages, and MHC II was significantly associated with improved survival in the primitive subtype (Fig. 5*B*). We demonstrated that increased immune cell expression is not consistently associated with improved survival and appears to be expression subtype–dependent. Curiously, the inflammatory SCC secretory subtype expected to demonstrate improved survival with increased expression of immune cells did not show significant associations with survival (except for mast cells). The secretory SCC subtype demonstrated uniformly high expression of immune cells, which may have prevented detection of a significant survival benefit. Alternatively, immune expression in the SCC secretory subtype may not be associated with improved survival. Subtype-specific and overall hazard ratios and confidence intervals (adjusted for subtype and pathologic stage) for all the signatures and genes evaluated are included in Supplementary Table 5.

Discussion

The mortality rate associated with lung cancer remains high, and despite the presence of several potentially targetable mutations in lung AD, many lung tumors do not carry a known driver mutation that can be targeted with existing therapies[30]; therefore, there is a significant need to improve management for this large group of patients with lung cancer. Recent evidence from the Cancer Genome Atlas Network confirms earlier findings that genomic analysis of lung AD and SCC tumors defines unique biologic subtypes with potential to inform treatment and management decisions.[1–4] Given the lack of a reproducible biomarker for immune therapy response in NSCLC and the potential for clinical differentiation by gene expression subtypes to inform prognosis and treatment outcomes, the immune landscape of lung AD and SCC molecular subtypes was explored.

We have demonstrated consistent but variable immune features of lung AD and SCC expression subtypes, with decreased immune cell expression in the PP subtype of AD and elevated expression in the secretory and (to a lesser extent) primitive subtypes of SCC. Results of *CD274* (PD-L1) gene expression were not always correlated with immune cell expression, particularly in SCC, in which gene expression subtypes were more strongly predictive of T-cell gene expression than *CD274*. This is consistent with earlier observations demonstrating lack of association of PD-L1 IHC expression with response to checkpoint inhibitors in patients with SCC[11] but is potentially in conflict with another study suggesting association of PD-L1 expression and treatment response in patients with NSCLC.[12]

Evaluation of mutation burden did not confirm a strong independent association of mutation burden with immune cell expression in either AD or SCC despite earlier observations of mutation burden and neoantigen associations with response to checkpoint inhibitors in NSCLC.[15] In AD, the PI subtype, enriched for *TP53*

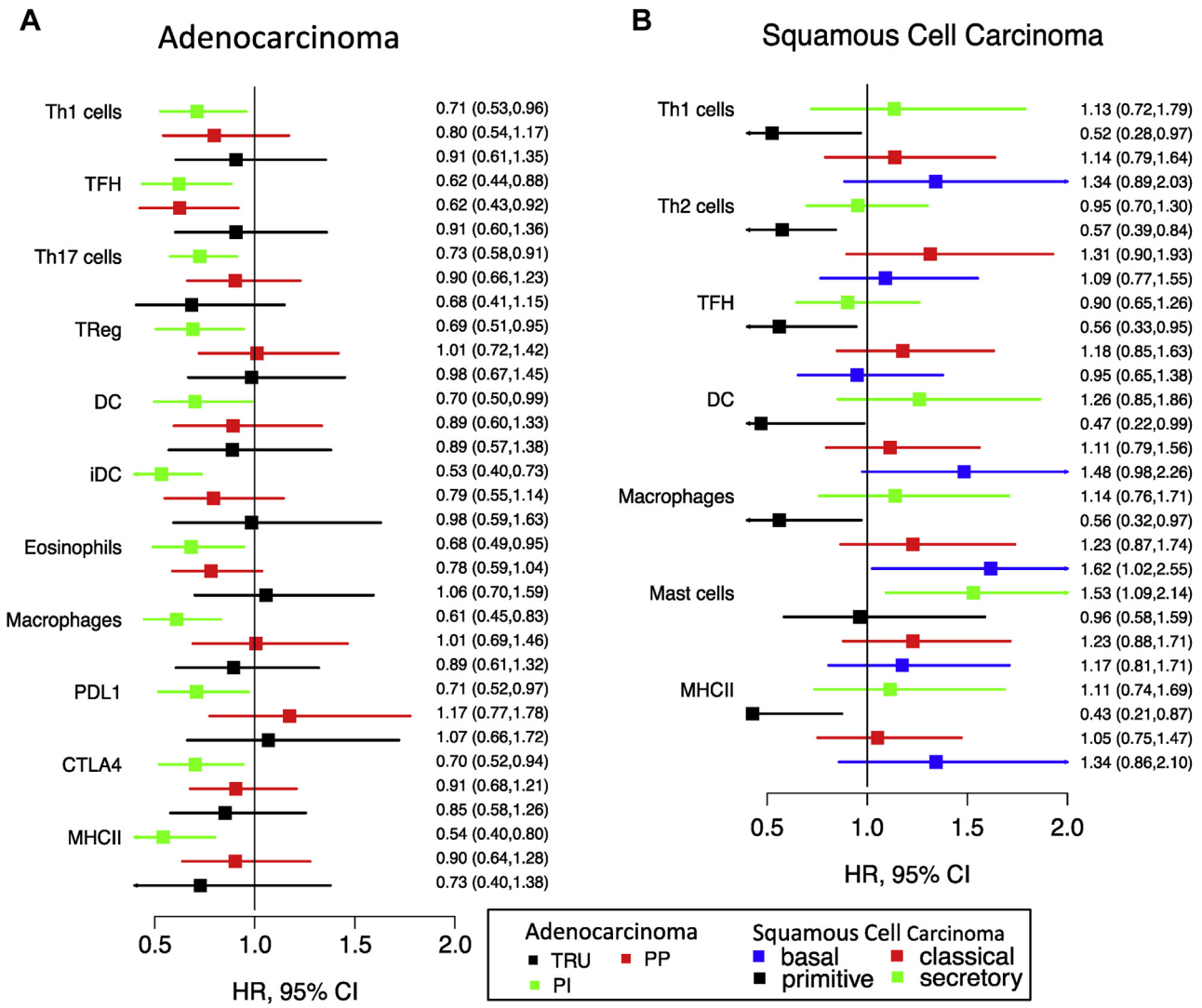

Figure 5. Subtype-specific immune marker hazard ratios and 95% confidence intervals (CIs) for 5-year overall survival in stage I to III adenocarcinoma (*A*) and stage I to III squamous cell carcinoma (*B*). Hazard ratios (HRs) correspond to a unit increase in the normalized immune marker and were adjusted for pathological stage using Cox models. Only markers that were significant (nominal p value <0.05) for at least one subtype are shown. Th1, type 1 T helper cells; Th2, type 2 T helper cells; TFH, T follicular helper cell; Th17, T helper 17 cell; Treg, regulatory T cell; DC, dendritic cell; iDC, immature dendritic cell; PD-L1, programmed death ligand 1; CTLA4, cytotoxic T-lymphocyte associated protein 4; MHC II, major histocompatibility complex class II; TRU, terminal respiratory unit; PP, proximal proliferative; PI, proximal inflammatory.

mutations, was associated with elevated immune cell expression and a higher mutation burden; however, the TRU subtype was associated with the lowest mutation burden despite relatively high immune expression. As mutation burden association with immune response is investigated, it will be important to consider distribution of AD subtypes in the sample population, as the relative proportion of various subtypes within a given cohort could bias conclusions regarding mutation burden association with immune response for AD as a whole. This study did not investigate neoantigens, and it is possible that neoantigen burden results might yield different results; however, previous observations have demonstrated a high correlation of neoantigens with nonsilent mutations.[15] In SCC, there was no association of silent mutation burden and T-cell immune expression, which is in agreement with an earlier study of immune cytolytic activity across tumor types.[8]

Other potential genomic associations with lack of immune cell expression, such as *STK11* loss/ inactivation[22–24] in the AD or *KEAP1/NFE2L2* oxidative stress pathway alterations in SCC,[25] were reflective of differences in enrichment across the subtypes, but the alteration association with immune cell expression was not subtype independent. Association of T-cell immune expression with MHC II genes and with improved

survival in the AD PI subtype and in the primitive subtype of SCC was observed. Expression of MHC II genes with increased immune cell expression may be reflective of ongoing immune cell infiltration or may be tumor associated, as suggested recently in triple-negative breast cancer.[21] More importantly, it is associated with improved survival in some of the subtypes but not others, thus providing potentially valuable clinical justification for additional investigation of expression subtypes of AD and SCC. This adds to a growing body of literature linking improved prognosis to infiltrating immune cells in NSCLC[9,26–29] and in a variety of other tumors[7] but is unique in identifying prognostic differences of immune cell expression across expression subtypes that may not be discernable when evaluating AD or SCC tumors as a whole.

Recent studies suggest that adaptive immune responses in the lung may be initiated in the tumor environment in nodelike "tertiary lymphoid structures,"[31] thus providing an explanation for the relatively high expression of MHC II genes in the normal lung as compared with in the tumor microenvironment. Despite earlier observations of loss of MHC class I gene mutations on chromosome 6 in SCC as a potential mechanism of immune suppression in lung SCC tumors,[1] preliminary evaluation of chromosome 6 deletions in our data set did not confirm an association of loss of MHC II with lack of immune expression in AD nor in SCC. Nonetheless, downregulation of MHC II genes by other means may be an important mechanism in evading host response, and further investigation of MHC II genes and the role they play in variable immune infiltration in the tumor microenvironment of AD and SCC subtypes is warranted.

This work is significant for presenting the differential immune response associated with intrinsic lung AD and SCC gene expression subtypes. Reproducible differences in immune response and associated survival implications likely reflect the underlying tumor biology of the gene expression subtypes with potential to inform immunotherapy response and reduce selection bias in drug trials. This study is limited by the use of genomic analyses in publicly available gene expression data sets. The immune signatures used[6] have been applied in multiple tumor types to characterize the tumor immune response[7]; to our knowledge, however, they have not been specifically evaluated in various normal lung anatomic sites (proximal versus distal airway) where immune differences may exist and where future evaluations will be needed to further elucidate the lung tumor immune landscape. Publicly available gene expression data sets are disproportionately derived from early-stage fresh frozen surgical resection samples with sufficient residual tissue for genomic study. Genomic advances have recently established the reliability of gene expression analysis in formalin-fixed paraffin-embedded tumor tissues, facilitating the use of archived tissue for such analyses, and preliminary data suggest similar immune landscape findings. Advanced tumors, however, particularly stage IV tumors, are rarely well represented in these archives despite frequent diagnoses of such tumors. We anticipate that this limitation will be overcome as improvements in genomic technology permit reliable analyses from more limited samples, and/or from circulating nucleic acids.

This study reflects an analysis at a single point in time before the clinical introduction of immunotherapy. Changes in the immune landscape after initiation of immunotherapy have been observed, at least in melanoma,[32] and suggest new areas for exploration of the tumor immune microenvironment in serial longitudinal samples, which could not be evaluated here. In addition, acquired resistance involving alterations in IFN receptor signaling and/or antigen presentation, as noted recently in patients with melanoma in whom resistance to checkpoint inhibitors develops,[33] will likely reveal new insights into the immune landscape. Nonetheless, at a time when reliable immune oncology biomarkers are so urgently needed, we believe that this work is significant for identifying and characterizing these immune differences by expression subtype, while acknowledging the need for additional studies in immunotherapy-treated patient cohorts as such data become available.

In conclusion, expanded use of molecular testing to better characterize lung tumors is likely, and we believe desirable, as genomics drives improved therapeutics and more personalized oncology treatment plans. Gene expression–based tumor subtyping of lung AD and SCC provides valuable information regarding differential immune features of NSCLC tumors, with potential to inform immunotherapy drug development, treatment selection, and improved outcomes.

Acknowledgements

We thank the many patients who donated their tissue for medical research that has become part of data sets in the public domain.

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *Journal of Thoracic Oncology* at www.jto.org and at http://dx.doi.org/10.1016/j.jtho.2017.03.010.

References

1. Cancer Genome Atlas Research Network, Hammerman PS, Lawrence MS, et al. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012;489:519-525.

2. Cancer Genome Atlas Research Network, Collisson EA, Campbell JD, et al. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511:543–550.
3. Wilkerson MD, Yin X, Hoadley KA, et al. Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. *Clin Cancer Res*. 2010;16:4864–4875.
4. Wilkerson MD, Yin X, Walter V, et al. Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLoS One*. 2012;7:e36530.
5. Hayes DN, Monti S, Parmigiani G, et al. Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J Clin Oncol*. 2006;24:5079–5090.
6. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 2013;39:782–795.
7. Iglesia MD, Parker JS, Hoadley KA, Serody JS, Perou CM, Vincent BG. Genomic analysis of immune cell infiltrates across 11 tumor types. *J Natl Cancer Inst*. 2016;108:djw144.
8. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160:48–61.
9. Bremnes RM, Busund LT, Kilvaer TL, et al. The role of tumor-infiltrating lymphocytes in development, progression, and prognosis of non-small cell lung cancer. *J Thorac Oncol*. 2016;11:789–800.
10. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med*. 2015;373:1627–1639.
11. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N Engl J Med*. 2015;373:123–135.
12. Garon EB, Rizvi NA, Hui R, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. 2015;372:2018–2028.
13. Gettinger SN, Horn L, Gandhi L, et al. Overall survival and long-term safety of nivolumab (anti-programmed death 1 antibody, BMS-936558, ONO-4538) in patients with previously treated advanced non-small-cell lung cancer. *J Clin Oncol*. 2015;33:2004–2012.
14. Shukuya T, Carbone DP. Predictive markers for the efficacy of anti-PD-1/PD-L1 antibodies in lung cancer. *J Thorac Oncol*. 2016;11:976–988.
15. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348:124–128.
16. Lee HJ, Lee JJ, Song IH, et al. Prognostic and predictive value of NanoString-based immune-related gene signatures in a neoadjuvant setting of triple-negative breast cancer: relationship to tumor-infiltrating lymphocytes. *Breast Cancer Res Treat*. 2015;151:619–627.
17. Director's Challenge Consortium for the Molecular Classification of Lung A, Shedden K, Taylor JM, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med*. 2008;14:822–827.
18. Tomida S, Takeuchi T, Shimada Y, et al. Relapse-related molecular signature in lung adenocarcinomas identifies patients with dismal prognosis. *J Clin Oncol*. 2009;27:2793–2799.
19. Broad Institute TCGA Genome Data Analysis Center (2015): Analysis-ready standardized TCGA data from Broad GDAC Firehose stddata__2015_02_04 run. http://gdac.broadinstitute.org/runs/stddata__2015_02_04/data/. Accessed February 26, 2015.
20. Raponi M, Zhang Y, Yu J, et al. Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res*. 2006;66:7466–7472.
21. Forero A, Li Y, Chen D, et al. Expression of the MHC class II pathway in triple-negative breast cancer tumor cells is associated with a good prognosis and infiltrating lymphocytes. *Cancer Immunol Res*. 2016;4:390399.
22. Cao C, Gao R, Zhang M, et al. Role of LKB1-CRTC1 on glycosylated COX-2 and response to COX-2 inhibition in lung cancer. *J Natl Cancer Inst*. 2015;107:358.
23. Koyama S, Akbay EA, Li YY, et al. STK11/LKB1 deficiency promotes neutrophil recruitment and proinflammatory cytokine production to suppress T-cell activity in the lung tumor microenvironment. *Cancer Res*. 2016;76:999–1008.
24. Schabath MB, Welsh EA, Fulp WJ, et al. Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma. *Oncogene*. 2016;35:3209–3216.
25. Hast BE, Cloer EW, Goldfarb D, et al. Cancer-derived mutations in KEAP1 impair NRF2 degradation but not ubiquitination. *Cancer Res*. 2014;74:808–817.
26. Kayser G, Schulte-Uentrop L, Sienel W, et al. Stromal CD4/CD25 positive T-cells are a strong and independent prognostic factor in non-small cell lung cancer patients, especially with adenocarcinomas. *Lung Cancer*. 2012;76:445–451.
27. Liu H, Zhang T, Ye J, et al. Tumor-infiltrating lymphocytes predict response to chemotherapy in patients with advance non-small cell lung cancer. *Cancer Immunol Immunother*. 2012;61:1849–1856.
28. Schalper KA, Brown J, Carvajal-Hausdorf D, et al. Objective measurement and clinical significance of TILs in non-small cell lung cancer. *J Natl Cancer Inst*. 2015;107:dju435.
29. Yuan A, Hsiao YJ, Chen HY, et al. Opposite effects of M1 and M2 macrophage subtypes on lung cancer progression. *Sci Rep*. 2015;5:14273.
30. National Comprehensive Cancer Network. Clinical practice guideline in oncology. Non-small cell lung cancer. https://www.nccn.org/professionals/physician_gls/f_guidelines.asp. Accessed September 15, 2016.
31. Remark R, Becker C, Gomez JE, et al. The non-small cell lung cancer immune contexture. A major determinant of tumor characteristics and patient outcome. *Am J Respir Crit Care Med*. 2015;191:377–390.
32. Chen PL, Roh W, Reuben A, et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov*. 2016;6:827–837.
33. Zaretsky JM, Garcia-Diaz A, Shin DS, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N Engl J Med*. 2016;375:819–829.