**Title:** Immunogenomic profiling of lung adenocarcinoma reveals high-grade growth patterns are associated with an immunogenic tumor microenvironment

**Authors:** Neal Akhave<sup>1</sup>, Erin M. Bayley<sup>6</sup>, Meredith Frank<sup>1</sup>, Carmen Behrens<sup>1,4</sup>, Jiexin Zhang<sup>2</sup>, Runzhe Chen<sup>1</sup>, Xin Hu<sup>3</sup>, Edwin Roger Parra<sup>4</sup>, Won-Chul Lee<sup>,3</sup>, Stephen Swisher<sup>6</sup>, Luisa Solis<sup>4</sup>, Annikka Weissferdt<sup>7</sup>, Cesar Moran<sup>7</sup>, Neda Kalhor<sup>7</sup>, Jianhua Zhang<sup>3</sup>, Paul Scheet<sup>5</sup>, Ara A. Vaporciyan<sup>6</sup>, Boris Sepesi<sup>6</sup>, Don L. Gibbons<sup>1</sup>, John V. Heymach<sup>1</sup>, Jack J. Lee<sup>8</sup>, Ignacio I. Wistuba<sup>1,4</sup>, P. Andrew Futreal<sup>3</sup>, Jianjun Zhang<sup>1,3,\*</sup>, Junya Fujimoto<sup>3,\*</sup> and Alexandre Reuben<sup>1,\*</sup>

Affiliations: <sup>1</sup>Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA. <sup>2</sup>Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA. <sup>3</sup>Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA. <sup>4</sup>Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA. <sup>5</sup>Department of Epidemiology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA. <sup>6</sup>Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA. <sup>7</sup>Department of Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA. <sup>8</sup>Department of Biostatistics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA.

**Keywords:** NSCLC; Lung Adenocarcinoma histologic subtypes; Translational Pathology

Running title: LUAD growth pattern associated with immune microenvironment

**Word count:** 4538 (including methods); 3626 (excluding methods)

Abstract word count: 112

Figure/Table count: 17 (4 main; 13 in supplemental)

**Correspondence should be addressed to:** Jianjun Zhang (<u>izhang20@mdanderson.org</u>), Junya Fujimoto (<u>ifujimot@mdanderson.org</u>) or Alexandre Reuben (<u>areuben@mdanderson.org</u>)

## Abstract

Lung cancer is the leading cause of cancer-related mortality in the United States. Lung adenocarcinoma (LUAD) is the most common subtype and the most epidemiologically and genetically heterogeneous. Pathologists have routinely observed phenotypic heterogeneity among LUAD primary tumors as reflected by distinct patterns of tumor growth. However, despite prior implications on the association of immune-genomic environment and prognosis, this information is not utilized clinically. Herein, applying multiplatform immune-genomic analysis, we investigate two distinct classification systems and demonstrate that high-grade patterns of growth are associated with a distinct immunogenic tumor microenvironment that is predicted with a favorable response to immunotherapy, a finding with growing importance in the era of adjuvant and neoadjuvant immunotherapy.

## Introduction

Lung cancer remains the leading cause of cancer-related mortality in men and women in the United States<sup>1,2</sup>. However, over the course of the last century, and particularly with the advent of genomic sequencing, there have been significant advances in the understanding of lung cancer risk factors, genetic basis, and treatments. Small cell and non-small cell lung cancer (NSCLC) are regarded as the most clinically meaningful subtypes of lung cancer<sup>3,4</sup> with NSCLC further sub-classified as lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC) due to distinct genetic basis and disparate responses to systemic therapy<sup>5-7</sup>.

LUAD is the most common subtype and the most epidemiologically and genetically heterogeneous<sup>5</sup>. Tobacco exposure is the most prevalent risk factor for LUAD; however, nearly 20-30% of cases are identified in never smokers<sup>2</sup>. Further, LUAD in smokers is enriched for activating mutations in *KRAS*, high tumor mutation burden, and increased T-cell infiltration and clonality<sup>5,8</sup>. In contrast, LUAD in never smokers is rarely associated with *KRAS* alterations and predicts for the presence of a diverse set of oncogenic mutations or fusions, each with distinct disease biology and treatment<sup>3,5,8</sup>.

However, pathologists have routinely observed phenotypic heterogeneity among LUAD primary tumors as reflected by distinct patterns of tumor growth which prompted the International Association for the Study of Lung Cancer (IASLC)/American Thoracic Society (ATS)/European Respiratory Society (ERS)<sup>9</sup> to develop a multidisciplinary classification that was subsequently incorporated into the World Health Organization (WHO) classification of LUAD<sup>10</sup>. This classification system identified five distinct patterns

of growth in LUAD: acinar, lepidic, papillary, micropapillary, and solid<sup>11</sup>, which were classified by predominant subtype.

Multiple clinical studies have attempted to determine the relationship of patterns of growth and genomic driver of prognosis and have suggested that solid or micropapillary tumors are high-grade and associated with worse prognosis<sup>12-22</sup>. However, due to large-scale differences in sample size and clinical variables such as completeness of surgical resection or pathologic stage, findings have not been uniformly consistent. Thus, at this time, LUAD patterns of growth have had limited use in clinical decision-making. However, particularly in the immunotherapy and cellular therapy era, further study is warranted to understand the relationship of LUAD pattern of growth and immune activation and senescence within the tumor microenvironment that could affect efficacy of these novel agents.

Herein, we describe the most comprehensive examination of the relationship between LUAD pattern of growth with genomic-immunologic disease correlates and survival. Using the PROSPECT (Profiling of Resistance patterns and Oncogenic Signaling Pathways in Evaluation of Cancers of the Thorax; LAB07-0233) patient cohort from the University of Texas M.D. Anderson Cancer Center, a robust dataset of stage I-III NSCLC tumors collected from 1996 to 2008, we applied two distinct classification methodologies to understand the relationship between LUAD patterns of growth and their relationship with baseline clinical variables, prognosis, and immunogenomic environment.

Our study reaffirms conclusions of prior studies that high-grade patterns of growth are associated with worse relapse-free and overall survival; however, in contrast to prior studies, this relationship is the result of an association of high-grade patterns with higher

stage of disease at time of surgery. Additionally, we demonstrate that when tumors are reclassified by a risk-based tiered fashion, the presence of a high-grade pattern of growth of any percentage is associated with prolonged tobacco exposure, elevated tumor mutation burden, and an immune infiltrated tumor microenvironment with increased cytotoxic CD8 T-cell infiltration, PD-L1 expression, and decreased T-cell repertoire homology with tumor-adjacent lung tissue, which may predict for differences in therapeutic response to novel immunotherapies.

## Methods

### Patient Cohort and Sample Collection

After informed consent, enrollment of study participants and collection of study samples was performed via the PROSPECT (Profiling of Resistance patterns and Oncogenic Signaling Pathways in Evaluation of Cancers of the Thorax; LAB07-0233) study approved by the University of Texas M.D. Anderson Cancer Center's Institutional Review Board (IRB). Between 1996 and 2008, fresh-frozen and formalin-fixed paraffin-embedded (FFPE), peripheral blood mononuclear cells (PBMC), serum, tumor, and tumor-adjacent lung tissue samples were obtained from 174 LUAD patients at time of diagnosis and surgery. Histological subtype was defined and pathologic classification was performed on all samples and patients were treatment-naïve at the time of surgery. As part of pathologic sub-classification, patterns of growth were quantified within tumors such that the sum total of all patterns within an individual tumor was 100%. LUAD tumors were classified using two distinct classification systems: a predominant-pattern method, as per IASLC-ATS-ERS and WHO, and a risk-based tiered method. In the predominant-pattern method,

LUAD tumors were classified by the largest single pattern of growth and invasion (*i.e* acinar, lepidic, papillary, or solid) as defined by relative percentage. Conversely, in the risk-based tiered method, tumors were simply classified by the presence or absence of high- (*i.e.* solid) or low-grade (*i.e.* lepidic) patterns of growth, independent of relative percent. Tumors with papillary and/or acinar patterns of growth were classified as intermediate-grade and were classified by the predominant pattern. For the purposes of comparison, independent of classification method, solid tumors were designated as high-grade LUAD group, and all lepidic, papillary, and acinar tumors were combined to form the low/intermediate-grade LUAD group.

#### TCR variable beta chain sequencing

Sequencing of the CDR3 regions of the human TCR- $\beta$  chains was performed using the immunoSEQ Assay (Adaptive Biotechnologies, Seattle, WA)<sup>23-25</sup>. T cell density, richness, clonality, and similarity were calculated as previously described<sup>8,26</sup>. Briefly, TCR $\beta$  counts were normalized to the total amount of DNA usable for TCR sequencing as determined by PCR-amplification and sequencing of housekeeping genes before calculating T cell density. The preseqR package was used to calculated richness by extrapolating to 400,000 templates for PBMCs and 120,000 templates for tissue. To allow comparison of samples with unequal numbers of T cells, both richness and clonality normalize for sampling depth. Clonality was defined as 1-Peilou's eveness<sup>27</sup>. As described previously, differential abundance analyses were used to identify TCRs that were enriched in one sample over another<sup>28</sup>. Parameters were as follows: minTotal = 5, productiveOnly = True, alpha = 0.1, count = aminoAcid. Statistical analysis was performed in R version 3.2.

## Whole Exome Sequencing

Whole exome sequencing (WES) was performed on tumors and spatially segregated tumor-adjacent lung tissues to determine somatic point mutations in a prior study<sup>29,30</sup>, where single nucleotide variants (SNV) and small insertions and deletions (indels) were detected using MuTect<sup>31</sup> and Pindel<sup>32</sup>, respectively., followed by variants annotation and filter<sup>33</sup>. In addition, genomic DNA from 96 available matched peripheral blood samples was sequenced as germline control to identify the mutations in the tumor-adjacent lung tissues. Blood DNA was analyzed to identify mutations related to clonal hematopoiesis of indeterminate potential (CHIP) based on annotation specified previously<sup>34</sup>. WES data is available in the EGA (EGAS00001004026).

## Gene Expression Profiling

RNA microarray was performed in a prior study on 141 patients<sup>35,36</sup> using the Illumina HumanWG-6 v3.0 expression bead chip. Then an extended robust multi-array analysis (RMA) background correction model<sup>37</sup> was applied to obtain normalized gene expression profiles for individual samples. Gene expression data is available in the GEO repository (GSE42127).

## CIBERSORT

To estimate the abundances of different cell types in mixed cell population, we ran CIBERSORT <sup>38</sup> (<u>https://cibersort.stanford.edu</u>) using leukocyte gene signature matrix (LM22) containing 22 functionally defined human immune subsets to deconvolute our normalized gene expression profiles with default settings.

## Immunohistochemistry

Tumor tissue was fixed in formalin and embedded in paraffin. For immunohistochemical staining, tissue was cut and mounted at a thickness of 4µm per slide. Slides were then stained with CD3 polyclonal (1:100, DAKO), CD4 clone 4B12 (1:80, Leica Biosystems), CD8 clone C8/144B (1:25, Thermo Scientific), PD-L1 clone E1L3N (1:100, Cell Signaling Technology), PD-1 clone EPR4877-2 (1:250, Abcam), CD45RO clone UCHL1 (ready-touse, Leica Biosystems), FoxP3 clone 206D (1:50, BioLegend), and Granzyme B clone F1 (ready-to-use, Leica Biosystems)<sup>39</sup> antibodies. Slides were then stained using diaminobenzidine as chromogen and the Leica Bond Polymer refine detection kit (Leica Biosystems). Slides were then counterstained with hematoxylin and scanned using an Aperio AT2 automated slide scanner (Leica Biosystems). Quantification was performed on 5 x 1mm2 regions per tumor sample within the tumor center and measuring the average density of positive cells per region as a count of positive cells/mm2. For PD-L1, H-score was calculated by multiplying the proportion of positive cells in the sample (0-100%) by the intensity of staining (1+, 2+, or 3+) to obtain a score ranging between 1 and 300.

## Statistical Analysis

All plots were generated using GraphPad Prism 8.0 (La Jolla, CA). For baseline characteristics, for continuous variables, t-test was applied. Further, for categorical variables, based upon sample size of < or >5, chi-square or Fisher's exact test was applied. Because not all distributions of TCR variables met the normality assumption, Mann-Whitney U test or Kruskal-Wallis test (two-sided) was applied for assessing

differences among groups. Wilcoxon signed-rank test was used to compare matched samples. Spearman's rank correlation (two-sided) was used to assess monotonic relationships between two continuous variables. For survival analysis, we first performed Log-rank (Mantel-Cox) test. Finally, despite the exploratory nature of the study, p-values were adjusted for multiple hypothesis testing (p-value < 0.05).

### Results

# Risk-based tiered system reclassifies tumors by presence of any amount of highgrade pattern of growth and demonstrates relationship with smoking and stage

One hundred and seventy-four patients with LUAD enrolled onto the PROSPECT study were included in our analysis. All patients underwent curative-intent surgery followed by classification of histologic pattern of growth by expert thoracic pathologist and comprehensive multi-platform profiling. Dual classification systems were utilized to ensure that the commonly used predominant-pattern (IASLC/ATS/ERS) methodology would not disregard smaller percentage tumor patterns, particularly high-grade growth patterns, whose molecular contribution to the tumor microenvironment would be masked within the context of an otherwise-classified low/intermediate grade tumor. Thus, we devised a risk-based tiered classification method to stratify tumors for the presence or absence of high-risk tumor components and then assessed relationship with tumor microenvironment. Finally, before comparative analysis, independent of classification method, solid tumors were placed in the high-grade group and lepidic, papillary, and acinar tumors were combined to form the low/intermediate grade group (**Supplementary**)

**Figure 1a-b**). With this, as shown in **Supplementary Figure 1b-c**, when comparing the predominant pattern to risk-based tiered classification method, 40 (23%) LUAD tumors classified as low/intermediate grade by predominant pattern classification (i.e. containing  $\leq$  49% solid pattern of growth) were reclassified as high-grade using the risk-based tiered classification.

Comparison of baseline characteristics, as shown in **Supplementary Table 1** (predominant pattern classification) and **Supplementary Table 2** (risk-based tiered classification), identified high-grade tumors as enriched for association with active and heavy tobacco use, as defined by current smoking and smoking for >20 pack-years, respectively. Notably, differences were particularly pronounced when utilizing the risk-based tiered method (**Supplementary Table 2**) with 54% and 80% of high-grade group patients being either current active smokers or having >20 pack-year history of smoking, relative to only 26% (p=0.001) or 58% (p=0.002) of low-intermediate grade group patients.

Further, comparison of pathologic stage using both classification systems (**Supplementary Table 1; Supplementary Table 2**) demonstrated a significant association of high-grade tumors with advanced stage by the predominant pattern classification with 64% of high-grade tumors being stage 2 or 3, in contrast to only 37% of low/intermediate grade tumors (p=0.001). However, this difference was not observed when using the risk-based tiered classification (p=0.59). Of note, as shown in **Figure 1a-c**, no association was observed between percentage of high-grade growth pattern and primary tumor size (T-stage) (**Figure 1b**; r=0.18; p=0.08; Spearman rank correlation). However, increasing percentage of high-grade growth pattern was strongly associated

with presence of nodal metastases (N-stage) (**Figure 1c**; p<0.001). Altogether, these findings suggest that when high-grade patterns of growth are dominant, disease is likely of advanced stage with increased risk of nodal metastases, which may reflect a more invasive tumor biology and warrants evaluation about the need for more thorough lymph node evaluation pre- or intra-operatively.

# High-grade pattern, when predominant, *may* be associated with worsened diseasefree & overall survival

Next, we examined the relationship between LUAD pattern of growth and prognosis. As in prior studies<sup>12-14</sup>, we found that the presence of high-grade growth patterns, when defined by predominant pattern, was associated with worse relapse-free (RFS) and overall survival (OS) (**Figure 1d; Figure 1e**). However, unlike prior studies which suggested that this difference was independent of pathologic stage, our findings were tied to pathologic stage with loss of survival disadvantage when stratified by stage 1 or stage 2 and 3 cohorts (**Figure 1f-i**). Nonetheless, in the stage 1 population, although not statistically significant, there was a trend towards worsened RFS and OS (**Figure 1f-g**), as has been shown previously<sup>17,19,40</sup> and supports the use of the predominant pattern (IASLC-ATS-ERS) classification as the preferred method to stratify tumors to understand disease prognosis.

Use of the risk-based tiered method to assess RFS and OS (**Supplementary Figure 2af**) was not associated with differences in RFS or OS, even among stage 1 adenocarcinomas. This finding suggests that the impact of high-grade patterns of growth

upon prognosis is present only when it is the predominant component, at least in the preimmunotherapy era.

### High-grade pattern is associated with increased tumor mutation burden

Due to the burden of smokers in the high-grade group, we next assessed the relationship of pattern of growth and tumor mutation burden (TMB), which has been recurrently shown to be increased in lung cancer, particularly smoking-associated lung cancer, relative to other cancer types<sup>41,42</sup>. Additionally, TMB has been shown to be correlated with increased T-cell clonality and T-cell activation<sup>8</sup>. Finally, TMB of greater than 10 mutations per megabase has been shown to be an independent predictor of response to immunotherapy across human cancer<sup>43</sup>, including LUAD<sup>44</sup>.

First, to validate findings of prior studies in our dataset, we assessed the relationship of TMB, tobacco use and T-cell clonality. TMB demonstrated a significant correlation with pack-years of smoking (**Figure 2a**; *r=0.37*; *p=0.0001*; *Spearman rank correlation*) and T-cell clonality (**Figure 2b**; *r=0.28*; *p=0.02*; *Spearman rank correlation*). Next, we compared TMB between high-grade and low-intermediate grade tumors by total non-synonymous exonic mutations (NSEM) and by number of tumors with >10 mutations/megabase (Mut/Mb) or <10 Mut/Mb. Utilizing the predominant-pattern classification, high-grade tumors showed no association with TMB by total NSEM (**Supplementary Figure 3a**; *p=0.52*). However, high-grade tumors were significantly enriched for tumors with >10 Mut/Mb (**Supplementary Figure 3b**; *p=0.02*). Further, when comparing TMB by risk-based tiered methodology, differences in TMB were even more pronounced with high-grade tumors showing significant increase in TMB relative to low/intermediate grade

tumors, both in total NSEM and number of tumors with >10 Mut/Mb. Specifically, highgrade tumors were found to have 273 (ranging from 8 to 857) NSEM per tumor compared to only 116 (ranging from 3 to 653) NSEM per tumor in low-intermediate grade tumors (**Supplementary Figure 4a**; p=0.046). Additionally, high-grade tumors showed significant enrichment for tumors with >10 Mut/Mb (**Figure 2c**; p=0.008) with 12 (86%) of tumors with >10 Mut/Mb identified as high-grade.

Altogether, these findings suggest that high-grade patterns, as identified by both classification systems, are associated with increased TMB. However, the risk-based tiered classification system more clearly stratifies tumors by TMB and suggests that high-grade pattern at any percentage within a tumor may predict for increased responsiveness to immunotherapy, in part due to its association with increased TMB.

## High-grade pattern is associated with mutations in TP53

We next assessed the relationship between high and low/intermediate grade pattern and known LUAD oncogenic drivers, *KRAS* and *EGFR*, and tumor suppressor genes, *TP53*, *STK11*, and *KEAP1*<sup>45</sup>. Notably, although whole exome sequencing was used in this dataset, fusions were poorly captured and therefore we were unable to assess for presence of fusion-specific oncogenic drivers in LUAD involving *ALK*, *ROS1*, and *RET* genes. Neither predominant-pattern classification nor risk-based tiered method showed an enrichment of activating mutations in *KRAS* or *EGFR* (**Supplementary Figure 3c**; **Figure 2d**) or loss of function mutations in *STK11 or KEAP1* (**Supplementary Figure 3d**; **Figure 2e**). However, both classification methods demonstrated significant enrichment for mutations in *TP53* (**Figure 2d**; *p=0.002*; **Supplementary Figure 3d**; *p=0.003*). As

with TMB, *TP53* mutations have been suggested to also predict responsiveness to checkpoint inhibitors,<sup>46,47</sup> and create for an environment that allows for cancer cell proliferation<sup>48</sup> and future genomic doubling, a finding associated with poor prognosis<sup>48</sup>.

# High-grade pattern is associated with increased immune infiltration and T-cell activation

As prior studies have suggested relationship of *TP53* mutations and increased TMB with increased infiltration of cytotoxic and activated CD8 T-cells and M0/M1 macrophages<sup>49</sup>, we next assessed the relationship between high- and low/intermediate grade tumors and tumor microenvironment. As our group has previously published upon the immune characteristics of this patient cohort using TCRseq, IHC, and GEP<sup>8,26,29,30,35,36,39</sup>, we reanalyzed the data to interrogate the composition of the high-grade and low/intermediate grade tumors.

First, when utilizing the predominant-pattern classification, IHC revealed no differences in the density of T-cell populations within the tumor (**Supplementary Figure 5a**). However, the ratio of CD4 to CD8 T-cells was significantly elevated in low/intermediate grade tumors relative to high-grade tumors with ratio of 1.71 (ranging from 0.34 to 5.18) and 1.32 (ranging from 0.42 to 4.04), respectively (**Supplementary Figure 5b**; p=0.01). Moreover, using CIBERSORT, intra-tumoral CD4 T-cells were more likely to be of activated memory CD4 T-cell programming (**Supplementary Figure 6b**; p=0.003). However, no differences were found among other cell types (**Supplementary Figure 6b**; p=0.003).

In the TAL, high-grade tumors were notable for significant increase in density of CD4 and CD8 T-cells and Granzyme B (GZMB) relative to low/intermediate-grade tumors (**Supplementary Figure 5c**). Further, although there was no difference in CD4:CD8 ratio within TAL (**Supplementary Figure 5b**; p=0.17), the ratio of CD4 T-cells between tumor and TAL was significantly increased in low/intermediate grade tumors relative to high-grade tumors with Tumor:TAL-CD4 ratio of 0.96 (ranging from 0.22 to 2.31) and 0.78 (ranging from 0.11 to 2.0), respectively (**Supplementary Figure 5d**; p=0.002). Altogether, our findings suggest that high-grade patterns may drive a more active T-cell population both within LUAD tumors and in the tumor-adjacent lung.

Reclassification of tumors by risk-based tiered methodology was able to identify additional differences in immune composition between high-grade and low/intermediate grade tumors. Firstly, IHC was now able to identify a significantly increased density of infiltrating CD8 T-cells (p=0.02) and GZMB-expressing cells (p=0.01) (**Figure 3a**) in high-grade tumors, a difference not found with other T-cell populations (**Figure 3a**; **Supplementary Figure 7a**). Further, although all T-cell markers were positively correlated with GZMB, among high-grade tumors the density of GZMB was most positively correlated with CD8 T-cells (**Figure 3b**; p=<0.0001; r=0.67; Spearman rank correlation) as opposed to CD4 T-cells (**Supplementary Figure 7b**; p=<0.0001; r=0.35; Spearman rank correlation), highlighting the cytotoxic nature of infiltrating CD8 T-cells.

Further, functional assessment by CIBERSORT revealed that CD4 T-cells within low/intermediate grade tumors favored a resting CD4 memory cell phenotype (**Figure 3c**; p=0.03), whereas CD4 T-cells within high-grade tumors favored an activated CD4 memory cell phenotype (**Figure 3c**; p=0.005). Additionally, although no differences

among nearly all other adaptive and innate immune cells were observed (**Supplementary Figure 8a-b**), high-grade tumors demonstrated slight increase in prevalence of nonactivated (M0) (**Figure 3d**; p=0.03) and pro-inflammatory (M1) macrophages (**Figure 3d**; p=0.05), but not anti-inflammatory (M2) macrophages (**Figure 3d**; p=0.79). This finding suggests that high-grade tumors may be more primed to foster an anti-tumor immune response<sup>50</sup> and supports the idea of a distinct immunogenic TME phenotype within highgrade tumors. Finally, the lack of increase in activated NK cell programming among highgrade tumors further supports that the observed increase in granzyme B expression is likely solely due to increased CD8 T cell infiltration (**Supplementary Figure 8b**).

Further comparison of intra-tumoral CD4:CD8 ratio revealed a further intensification of differences among high-grade and low-intermediate grade tumors. Specifically, among low/intermediate grade tumors, intra-tumoral CD4:CD8 ratio was 1.88 (ranging from 0.58 to 5.18) compared to high-grade tumors with CD4:CD8 ratio of 1.37 (ranging from 0.34 to 4.04) (**Figure 3e**; *p*=0.001), again suggesting enrichment of CD8+ T cells within high-grade tumors, something not seen in the TAL (**Figure 3e**; *p*=0.08).

Finally, review of T-cell populations within TAL revealed significantly increased T-cell density across all markers tested among high-grade tumors (**Supplementary Figure 8c**). However, notably, ratio of Tumor:TAL CD3 (p=0.005) and CD4 (p=0.01), but not CD8 (p=0.82), was higher among low-intermediate grade tumors suggesting that CD8 T cells may be less capable to infiltrate or excluded from these tumors (**Figure 3f**).

High-grade pattern is associated with increased expression of immune checkpoints, particularly PD-L1

Considering the increased T-cell infiltration seen in high-grade tumors, we next assessed the expression of checkpoints used by malignant tumors, including LUAD, to induce immune senescence.

First, using the predominant-pattern classification, we compared PD-L1 H-score between high-grade and low/intermediate grade tumors. As shown in **Supplementary Figure 9a**, PD-L1 H-score was significantly elevated in high-grade tumors (p=0.008), however, gene expression profiling of other immune checkpoints, CTLA-4, TIGIT, and LAG-3 did not demonstrate significant differences (**Supplementary Figure 9b**).

However, analysis using risk-based tiered classification revealed not only significant enrichment of PD-L1 intensity among high-grade relative to low/intermediate grade tumors (**Figure 4a**; *p*=0.032) but also increase in LAG-3 expression, an immune checkpoint molecule for which checkpoint inhibitors are in development for NSCLC and already showing success in combination with PD-1 inhibitors in metastatic melanoma<sup>51</sup> (**Figure 4b**; p=0.04). Altogether, these findings reinforce that even small percentages of high-grade pattern, as defined by risk-based tiered classification, unveil key differences in immunogenic potential of LUAD tumors, which may have clinical implications.

# High-grade pattern is associated with decreased T-cell repertoire homology compared to TAL

Finally, to further investigate the nature of the infiltrating T-cell repertoire between highgrade and low/intermediate grade tumors, we analyzed TCR sequencing data generated from tumor and TAL repertoires.

First, when applying the predominant-pattern classification, we identified no differences in T-cell clonality between high-grade and low-intermediate grade tumors (**Supplementary Figure 10a**; p=0.10). Further, comparison of homology between tumor and TAL by Jaccard (**Supplementary Figure 10b**; p=0.58) and Morisita indices (**Supplementary Figure 10c**; p=0.30), and total number of shared T-cell clones among top 100 clones (**Supplementary Figure 10d**; p=0.27) showed no differences between high-grade and low/intermediate grade tumors.

However, utilization of risk-based tiered classification, although again not demonstrating a difference in T-cell clonality between high-grade and low-intermediate grade tumors (Figure 4c; *p*=0.12), did show distinct differences in T-cell repertoire homology between tumor and TAL. Specifically, evaluation by Jaccard index revealed significantly greater homology of T-cell clones between tumor and TAL in low-intermediate grade tumors relative to high-grade tumors (**Figure 4d**; p=0.048), a finding that approached but did not reach significance when using the Morisita Index (**Supplementary Figure 11a**; p=0.051), suggesting differences in clonal expansion and distribution between low/intermediate and high-grade tumors. Furthermore, among the top 100 most prevalent T-cell clones identified in individual tumors, low/intermediate grade tumors had an average of 34 among the top 100 in the TAL (Supplementary Figure 11b). In contrast, high-grade tumors had an average of only 27 among the top 100 in the TAL (Supplementary Figure **11b**; *p*=0.028). Additionally, a negative correlation was seen between increasing highgrade pattern and shared Top 100 T-cell clones (Figure 4e; r=-0.33; p=0.0002; Spearman rank correlation). Shared T-cell repertoire homology has been suggested to be non-tumor specific by a recent publication by our group<sup>8</sup>. Therefore, results suggest that the T-cell repertoire in high-grade tumors, as identified by risk-based tiered classification, may reflect more tumor-specific clonal expansion that may further predict and augment immunotherapy responses.

## Discussion

This study represents the most comprehensive analysis of the genomic and immune underpinnings of LUAD morphologic patterns of growth and invasion. Utilizing two distinct methods of classification, the predominant-pattern (WHO) classification and risk-based tiered classification, our study highlights the key clinical, prognostic, genomic, and immunologic differences between tumors with high-grade and low/intermediate grade patterns of growth.

We demonstrate that, when classified by the predominant-pattern method, high-grade tumors are associated with worsened relapse-free and overall survival. However, unlike prior studies which suggest that pattern of growth is an independent predictor of survival, we demonstrate that prognosis is more closely linked to pathological stage rather than pattern of growth. Further, our findings suggest that increasing percentage of high-grade growth patterns within LUAD primary tumors are associated with increased risks of nodal metastases, suggesting a more invasive phenotype among high-grade tumors. Current standard of care for mediastinal lymph node dissection is lymphadenectomy of stations 2R, 4R, 7, 8, and 9 for right-sided tumors and stations 4L, 5, 6, 7, 8, 9 for left-sided tumors<sup>52</sup>. However, our results suggest that in tumors with high-grade pattern of growth,

even in small percentage, expanded lymph node dissection to capture occult metastases may be necessary.

Further, we suggest that the predominant-pattern classification is ill equipped to identify meaningful differences in immune microenvironment of LUAD tumors. Although recent studies have also demonstrated distinct epidemiologic and genomic differences between high and low/intermediate grade tumors<sup>53</sup>, our study is the first to demonstrate in detail the distinct immune differences, as generated by the risk-based tiered classification.

When using risk-based tiered classification, we demonstrate that high-grade tumors appear to represent the morphologic phenotype of an active, heavy smoker, particularly those with greater than 20 pack-years of smoking. Furthermore, our multiplatform analysis reveals that these high-grade tumors have increased TMB, loss of mutations in TP53, and increased expression of PD-L1 relative to low/intermediate grade tumors, all of which are independent predictors of response to checkpoint blockade<sup>44,46,47,54</sup>. Finally, we describe the novel finding that tumor microenvironments of high-grade tumors, when classified by risk-based tiered classification, possess a distinct immunogenic phenotype with significant infiltration of cytotoxic CD8<sup>+</sup> T-cells, increased activated CD4 memory Tcells, and polarization of macrophages toward pro-inflammatory M1. Further, with less shared homology with TAL, infiltrating T-cells are likely more tumor specific and thus more primed for anti-tumor immune response. Altogether, these findings define an early-stage LUAD subset that would be predicted for favorable response to immunotherapy, including standard-of-care checkpoint inhibitors, a finding which is increasingly important in the era of neoadjuvant and adjuvant immunotherapy. It may also define a subset predicted to respond to novel checkpoint inhibitors in development, particularly against LAG-3, for

which there is an ongoing clinical trial in NSCLC (NCT03516981) and tumor-infiltrating lymphocyte (TIL) therapy.

Notably, our study does have certain limitations. First, as our cohort of patients underwent surgical resection between 1996 and 2008, pre-surgical staging with PET/CT and endobronchial ultrasound (EBUS) was not performed and could have influenced the difference in clinical stage between groups. Therefore, further study is warranted to determine if our findings would hold true in the current era of pre-surgical staging. Moreover, although we controlled for pathologic stage in our comparison of prognosis, we did not assess for other variables that have been suggested to impact prognosis in early-stage NSCLC such as level of tumor differentiation and pleural or vascular invasion<sup>55,56</sup>. Additionally, for the sake of simplicity, although nearly all LUAD tumors consist of multiple different co-existing patterns of growth and invasion, our analyses evaluated tumors as a whole as opposed to specific analysis of individual growth patterns within a tumor. Therefore, correlations between an individual growth pattern within a tumor and genotypic or immune trends is not possible. However, as stated before, we feel that the design of the risk-based tiered classification compensates for this by accounting for and unmasking contribution of even small high-grade patterns within a tumor.

In conclusion, we suggest that high-grade patterns of growth, of any percentage within a LUAD tumor, possess a distinct immunogenic tumor phenotype that should be considered at the time of surgery and in clinical decision-making regarding adjuvant therapy.

## Acknowledgements

This study was supported by the NIH CCSG Award (CA016672 (Institutional Tissue Bank (ITB) and Research Histology Core Laboratory (RHCL)) and the Translational Molecular Pathology-Immunoprofiling Moonshot Platform (TMP-IL) at the Department Translational Molecular Pathology, the University of Texas MD Anderson Cancer Center; the Lung Specialized Programs of Research Excellence grant 5 (P50 CA070907).

AR is supported by the Exon20 Group, Rexanna's Foundation for Fighting Lung Cancer, the Happy Lungs Project, the Waun Ki Hong Lung Cancer Research Fund, MD Anderson's Lung Cancer Moon Shot, the Petrin Fund, the University Cancer Foundation via the Institutional Research Grant program at the University of Texas MD Anderson Cancer Center, and the Cancer Prevention and Research Institute of Texas (CPRIT - RP210137). All other authors declare no disclosures.

## Disclosures

AR serves on the Scientific Advisory Board and has received honoraria from Adaptive Biotechnologies. J.V.H serves on the Scientific Advisory Committees of AstraZeneca, EMD Serono, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Hengrui Therapeutics, Eli Lilly, Spectrum, Sanofi, Takeda, Mirati Therapeutics, BMS, BrightPath Biotherapeutics, Janssen Global Services, Nexus Health Systems, Pneuma Respiratory, Roche, Leads Biolabs, and RefleXion and provides research support to AstraZeneca, Bristol-Myers Squibb, and Spectrum and has received royalties and licensing fees from Spectrum.

## **Author contributions**

N.A., M.L.F., and A.R. wrote the manuscript. N.A., J.Z., J.F., and A.R. designed the study.

A.R., C.B., X. H., H.K., R.C., W-C.L., Jianh.Z, P.S., I.W., P.A.F generated and analyzed

the data. N.A., A.R., M.L.F, E.M.B., J.Z., and J.F. interpreted the data. JieZ, J.L., and P.S.

performed statistical analysis. A.V., B.S., S.S., C.M., N.K, J.H, D.G. J.F., E.R.P, A.W. and

I.W. obtained clinical samples and data. J.F., E.R.P, I.W. and A.W. performed pathologic

analysis of all samples. JZ, JF, and AR jointly supervised the study. All authors reviewed

and edited the manuscript.

## References

1. Henley SJ, Ward EM, Scott S, et al. Annual report to the nation on the status of cancer, part I: National cancer statistics. Cancer 2020;126:2225-49.

2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin 2021;71:7-33.

3. Ettinger DS, Wood DE, Aisner DL, et al. NCCN Guidelines Insights: Non-Small Cell Lung Cancer, Version 2.2021. J Natl Compr Canc Netw 2021;19:254-66.

4. Kalemkerian GP, Loo BW, Akerley W, et al. NCCN Guidelines Insights: Small Cell Lung Cancer, Version 2.2018. J Natl Compr Canc Netw 2018;16:1171-82.

5. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543-50.

6. Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012;489:519-25.

7. Scagliotti GV, Parikh P, von Pawel J, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 2008;26:3543-51.

8. Reuben A, Zhang J, Chiou SH, et al. Comprehensive T cell repertoire characterization of non-small cell lung cancer. Nat Commun 2020;11:603.

9. Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011;6:244-85.

10. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol 2015;10:1243-60.

11. Travis WD, Brambilla E, Noguchi M, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. Proc Am Thorac Soc 2011;8:381-5.

12. Warth A, Muley T, Meister M, et al. The novel histologic International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society

classification system of lung adenocarcinoma is a stage-independent predictor of survival. J Clin Oncol 2012;30:1438-46.

13. Tsuta K, Kawago M, Inoue E, et al. The utility of the proposed IASLC/ATS/ERS lung adenocarcinoma subtypes for disease prognosis and correlation of driver gene alterations. Lung Cancer 2013;81:371-6.

14. Russell PA, Wainer Z, Wright GM, Daniels M, Conron M, Williams RA. Does lung adenocarcinoma subtype predict patient survival?: A clinicopathologic study based on the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary lung adenocarcinoma classification. J Thorac Oncol 2011;6:1496-504.

15. Gu J, Lu C, Guo J, et al. Prognostic significance of the IASLC/ATS/ERS classification in Chinese patients-A single institution retrospective study of 292 lung adenocarcinoma. J Surg Oncol 2013;107:474-80.

16. Murakami S, Ito H, Tsubokawa N, et al. Prognostic value of the new IASLC/ATS/ERS classification of clinical stage IA lung adenocarcinoma. Lung Cancer 2015;90:199-204.

17. Hung JJ, Jeng WJ, Chou TY, et al. Prognostic value of the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society lung adenocarcinoma classification on death and recurrence in completely resected stage I lung adenocarcinoma. Ann Surg 2013;258:1079-86.

18. Tsao MS, Marguet S, Le Teuff G, et al. Subtype Classification of Lung Adenocarcinoma Predicts Benefit From Adjuvant Chemotherapy in Patients Undergoing Complete Resection. J Clin Oncol 2015;33:3439-46.

19. Yoshizawa A, Motoi N, Riely GJ, et al. Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. Mod Pathol 2011;24:653-64.

20. Borczuk AC. Prognostic considerations of the new World Health Organization classification of lung adenocarcinoma. Eur Respir Rev 2016;25:364-71.

21. Butnor KJ. Controversies and challenges in the histologic subtyping of lung adenocarcinoma. Transl Lung Cancer Res 2020;9:839-46.

22. Yoshizawa A, Sumiyoshi S, Sonobe M, et al. Validation of the IASLC/ATS/ERS lung adenocarcinoma classification for prognosis and association with EGFR and KRAS gene mutations: analysis of 440 Japanese patients. J Thorac Oncol 2013;8:52-61.

23. Robins HS, Campregher PV, Srivastava SK, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. Blood 2009;114:4099-107.
24. Robins H, Desmarais C, Matthis J, et al. Ultra-sensitive detection of rare T cell clones. J Immunol Methods 2012;375:14-9.

25. Carlson CS, Emerson RO, Sherwood AM, et al. Using synthetic templates to design an unbiased multiplex PCR assay. Nat Commun 2013;4:2680.

26. Reuben A, Gittelman R, Gao J, et al. TCR Repertoire Intratumor Heterogeneity in Localized Lung Adenocarcinomas: An Association with Predicted Neoantigen Heterogeneity and Postsurgical Recurrence. Cancer Discov 2017;7:1088-97.

27. Kirsch I, Vignali M, Robins H. T-cell receptor profiling in cancer. Mol Oncol 2015;9:2063-70.

28. DeWitt WS, Emerson RO, Lindau P, et al. Dynamics of the cytotoxic T cell response to a model of acute viral infection. J Virol 2015;89:4517-26.

29. Choi M, Kadara H, Zhang J, et al. Mutation profiles in early-stage lung squamous cell carcinoma with clinical follow-up and correlation with markers of immune function. Ann Oncol 2017;28:83-9.

30. Kadara H, Choi M, Zhang J, et al. Whole-exome sequencing and immune profiling of early-stage lung adenocarcinoma with fully annotated clinical follow-up. Ann Oncol 2017;28:75-82.

31. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31:213-9.
32. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach

to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics 2009;25:2865-71.

33. Takahashi K, Wang F, Kantarjian H, et al. Preleukaemic clonal haemopoiesis and risk of therapy-related myeloid neoplasms: a case-control study. Lancet Oncol 2017;18:100-11.

34. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med 2014;371:2488-98.

35. Tang H, Xiao G, Behrens C, et al. A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small cell lung cancer patients. Clin Cancer Res 2013;19:1577-86.

36. Lou Y, Diao L, Cuentas ER, et al. Epithelial-Mesenchymal Transition Is Associated with a Distinct Tumor Microenvironment Including Elevation of Inflammatory Signals and Multiple Immune Checkpoints in Lung Adenocarcinoma. Clin Cancer Res 2016;22:3630-42.

37. Xie Y, Wang X, Story M. Statistical methods of background correction for Illumina BeadArray data. Bioinformatics 2009;25:751-7.

38. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12:453-7.

39. Parra ER, Behrens C, Rodriguez-Canales J, et al. Image Analysis-based Assessment of PD-L1 and Tumor-Associated Immune Cells Density Supports Distinct Intratumoral Microenvironment Groups in Non-small Cell Lung Carcinoma Patients. Clin Cancer Res 2016;22:6278-89.

40. Yoshiya T, Mimae T, Tsutani Y, et al. Prognostic Role of Subtype Classification in Small-Sized Pathologic N0 Invasive Lung Adenocarcinoma. Ann Thorac Surg 2016;102:1668-73.

41. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415-21.

42. Chae YK, Davis AA, Raparia K, et al. Association of Tumor Mutational Burden With DNA Repair Mutations and Response to Anti-PD-1/PD-L1 Therapy in Non-Small-Cell Lung Cancer. Clin Lung Cancer 2019;20:88-96 e6.

43. Marabelle A, Fakih M, Lopez J, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. Lancet Oncol 2020;21:1353-65.

44. Hellmann MD, Ciuleanu TE, Pluzanski A, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. N Engl J Med 2018;378:2093-104.

45. Skoulidis F, Heymach JV. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. Nat Rev Cancer 2019;19:495-509.

46. Hellmann MD, Nathanson T, Rizvi H, et al. Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. Cancer Cell 2018;33:843-52 e4.

47. Dong ZY, Zhong WZ, Zhang XC, et al. Potential Predictive Value of TP53 and KRAS Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. Clin Cancer Res 2017;23:3012-24.

48. Bielski CM, Zehir A, Penson AV, et al. Genome doubling shapes the evolution and prognosis of advanced cancers. Nat Genet 2018;50:1189-95.

49. Lin X, Wang L, Xie X, et al. Prognostic Biomarker TP53 Mutations for Immune Checkpoint Blockade Therapy and Its Association With Tumor Microenvironment of Lung Adenocarcinoma. Front Mol Biosci 2020;7:602328.

50. Lin Y, Xu J, Lan H. Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications. J Hematol Oncol 2019;12:76.

51. LAG3-PD-1 Combo Impresses in Melanoma. Cancer Discov 2021;11:1605-6.

52. Darling GE, Allen MS, Decker PA, et al. Randomized trial of mediastinal lymph node sampling versus complete lymphadenectomy during pulmonary resection in the patient with N0 or N1 (less than hilar) non-small cell carcinoma: results of the American College of Surgery Oncology Group Z0030 Trial. J Thorac Cardiovasc Surg 2011;141:662-70.

53. Caso R, Sanchez-Vega F, Tan KS, et al. The Underlying Tumor Genomics of Predominant Histologic Subtypes in Lung Adenocarcinoma. J Thorac Oncol 2020;15:1844-56.

54. Gandhi L, Rodriguez-Abreu D, Gadgeel S, et al. Pembrolizumab plus Chemotherapy in Metastatic Non-Small-Cell Lung Cancer. N Engl J Med 2018;378:2078-92.

55. Strauss GM, Herndon JE, 2nd, Maddaus MA, et al. Adjuvant paclitaxel plus carboplatin compared with observation in stage IB non-small-cell lung cancer: CALGB 9633 with the Cancer and Leukemia Group B, Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups. J Clin Oncol 2008;26:5043-51. 56. Park SY, Lee JG, Kim J, et al. Efficacy of platinum-based adjuvant chemotherapy

in T2aN0 stage IB non-small cell lung cancer. J Cardiothorac Surg 2013;8:151.



Figure 2: Akhave et al.



Figure 3: Akhave et al.









Figure 4: Akhave et al.



