Multiomics Analysis Reveals Distinct Immunogenomic Features of Lung Cancer with Ground-Glass Opacity

Kezhong Chen¹*, Jing Bai²*, Alexandre Reuben³, Heng Zhao¹, Guannan Kang¹, Chunliu Zhang², Qingyi Qi⁴, Yaping Xu², Shawna Hubert^{3,5}, Lianpeng Chang², Yanfang Guan^{2,6}, Lin Feng⁷, Kai Zhang⁸, Kaitai Zhang⁷, Xin Yi^{2,6}, Xuefeng Xia², Shujun Cheng⁷, Fan Yang^{1‡}, Jianjun Zhang^{3,5‡}, and Jun Wang^{1‡}

¹Department of Thoracic Surgery and ⁴Department of Radiology, Peking University People's Hospital, China; ²Geneplus-Beijing Institute, Beijing China; ³Department of Thoracic/Head and Neck Medical Oncology and ⁵Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas; ⁶Department of Computer Science and Technology, School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China; and ⁷State Key Laboratory of Molecular Oncology, Department of Etiology and Carcinogenesis, Cancer Institute and Hospital, and ⁸Department of Cancer Prevention, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China

ORCID IDs: 0000-0002-9723-6153 (K.Z.); 0000-0001-7872-3477 (J.Z.); 0000-0002-7033-5012 (J.W.).

Abstract

Rationale: Ground-glass opacity (GGO)-associated lung cancers are common and radiologically distinct clinical entities known to have an indolent clinical course and superior survival, implying a unique underlying biology. However, the molecular and immune characteristics of GGO-associated lung nodules have not been systemically studied.

Objectives: To provide mechanistic insights for the treatment of these radiologically distinct clinical entities.

Methods: We initiated a prospective cohort study to collect and characterize pulmonary nodules with GGO components (nonsolid and part-solid nodules) or without GGO components, as precisely quantified by using three-dimensional image reconstruction to delineate the molecular and immune features associated with GGO. Multiomics assessment conducted by using targeted gene panel sequencing, RNA sequencing, TCR (T-cell receptor) sequencing, and circulating tumor DNA detection was performed.

Measurements and Main Results: GGO-associated lung cancers exhibited a lower tumor mutation burden than solid

nodules. Transcriptomic analysis revealed a less active immune environment in GGO components and immune pathways, decreased expression of immune activation markers, and less infiltration of most immune-cell subsets, which was confirmed by using multiplex immunofluorescence. Furthermore, T-cell repertoire sequencing revealed lower T-cell expansion in GGO-associated lung cancers. HLA loss of heterozygosity was significantly less common in lung adenocarcinomas with GGO components than in those without. Circulating tumor DNA analysis suggested that the release of tumor DNA to the peripheral blood was correlated with the tumor size of non-GGO components.

Conclusions: Compared with lung cancers presenting with solid lung nodules, GGO-associated lung cancers are characterized by a less active metabolism and a less active immune microenvironment, which may be the mechanisms underlying their indolent clinical course.

Clinical trial registered with www.clinicaltrials.gov (NCT 03320044).

Keywords: ground-glass opacity; genomics; T-cell repertoire; immune infiltration; circulating tumor DNA

(Received in original form January 16, 2021; accepted in final form September 1, 2021)

*These authors contributed equally to this work.

[‡]Co-senior authors.

Supported by the National Natural Science Foundation of China (82072566 and 81602001 [K.C.]) and Peking University People's Hospital Research and Development Funds (RS2019-01 [K.C.]). J.Z. was supported by a Khalifa Scholar Award, the National Cancer Institute of the National Institutes of Health Research Project (grant R01CA234629-01), the AACR–Johnson & Johnson Lung Cancer Innovation Science Grant (18-90-52-ZHAN), the MD Anderson Physician Scientist Program, the Sabin Family Foundation Award, and the Duncan Family Institute Cancer Prevention Research Seed Funding Program.

Author Contributions: K.C., X.X., S.C., J.Z., and J.W. conceived and designed the study. K.C., H.Z., G.K., and Q.Q. collected samples as well as clinical information. J.B., C.Z., Y.X., L.C., and Y.G. performed the experiments. K.C., J.B., C.Z., Y.X., L.C., and J.Z. analyzed the data. K.C., J.B., A.R., C.Z., Y.X., S.H., Y.G., X.X., and J.Z. wrote the manuscript. K.C., J.B., Y.G., L.F., Kai Zhang, Kaitai Zhang, and S.C. provided intellectual discussions and ideas regarding the content of manuscript. X.Y., S.C., F.Y., and J.W. supervised the study. All authors read and approved the final version to be published.

Correspondence and requests for reprints should be addressed to Jun Wang, M.D., Department of Thoracic Surgery, Peking University People's Hospital, Xi Zhi Men South Avenue, No. 11, Beijing 100044, China. E-mail: wangjun@pkuph.edu.cn.

Am J Respir Crit Care Med Vol 204, Iss 10, pp 1180–1191, Nov 15, 2021 Copyright © 2021 by the American Thoracic Society Originally Published in Press as DOI: 10.1164/rccm.202101-0119OC on September 2, 2021 Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the

Subject: Ground-glass opacity (GGO)associated lung cancers are common and radiologically distinct clinical entities known to have an indolent clinical course and superior survival, implying an underlying unique biology. However, the molecular characteristics of GGO-associated lung nodules have not been systemically studied.

What This Study Adds to the Field:

Multiomics analysis revealed GGOassociated lung cancers with a lower mutational burden, a less active immune environment, and less circulating tumor DNA shedding, highlighting the intrinsic biological features that GGO represents early during lung adenocarcinoma carcinogenesis, when cancer cells and the antitumor immune response may be at an equilibrium. This study provides mechanistic insights for the diagnosis and treatment of these radiologically distinct clinical entities.

With the increasing implementation of low-dose computed tomography (CT)-guided lung cancer screening and the use of high-resolution diagnostic CT scans, there has been a drastic increase in the detection of pulmonary nodules (1). The radiological texture of pulmonary nodules can be solid (solid nodules [SNs]), partly solid (part-solid nodules [PSNs]) with a mixture of solid and ground-glass opacity (GGO) components, or nonsolid (nonsolid nodules [NSNs]) with pure GGO (2).

Studies have shown that GGOassociated lung cancers are relatively indolent, with nearly 100% 5-year overall survival after resection being shown, suggesting that GGO-associated lung cancers may have a biology that is distinct from that of SN lung cancers (3). Recent studies have revealed the molecular features of early-stage lung cancers (4–7). However, little is known about the molecular and immune features of these radiologically distinct GGO-associated lung nodules (8–10). Current treatment strategies are solely based on clinical features without an understanding of the biology underlying these unique entities. Management of GGO-associated lung nodules varies greatly, with clinicians having to rely on their individual experience of pulmonary GGO and often encountering cases that are inconsistent with available guidelines (11). Interestingly, clinical management is much more aggressive in East Asian countries than in Western countries (12–14).

We initiated a prospective study (clinicaltrials.gov identifier NCT 03320044) and collected and analyzed blood and tissue specimens from patients presenting with pulmonary nodules (with or without GGO components as based on three-dimensional [3D] image reconstruction) suspicious for lung cancers to delineate the molecular and immune features of tumors with GGO components versus tumors without GGO components. Some of the results of this study have been previously reported in the form of abstracts (15, 16).

Methods

Study Design and Participants

Patients presenting with solidary pulmonary nodules who were suspected of having resectable lung cancers were enrolled prospectively. Eligible patients were >18 years of age and had no distant metastasis on preoperative scans (clinical stages I-IIIA) and no personal history of malignancy 5 years before surgery. Patients who 1) received neoadjuvant therapy, 2) had multiple primary lung cancers, 3) had R1 or R2 resection, or 4) had no blood samples were excluded. Written informed consent was obtained from each patient, and the study was approved by the Peking University People's Hospital Medical Ethics Committee (2017PHB105-01) in accordance with the Declaration of Helsinki and was registered with clinicaltrials.gov (NCT 03320044).

3D Radiologic Evaluation

All patients underwent high-resolution CT scans to ensure accurate volumetric analysis.

The total tumor volume and GGO components of each lesion were determined by using 3D reconstructions and were automatically obtained by using the CT lung analysis workstation Lung VCAR (GE Healthcare) (Figure 1). For cases in which segmentation and nodules were visually mismatched, the margins of the whole nodule and solid component were manually adjusted.

Targeted Next-Generation Sequencing

Please see the online supplement for specimen collection, processing, and DNA/ RNA extraction. Both cell-free DNA (cfDNA) from blood and genomic DNA from tissues were subjected to next-generation sequencing of 1,021 frequently mutated cancer genes and exons 2–3 of HLA genes (*see* Table E1 in the online supplement).

RNA Sequencing and Gene Expression Profiling

DESeq2 software (17) was used for gene differential expression analysis between tumors with GGO components and tumors without GGO components. Significantly differentially expressed genes were identified by using thresholds of a twofold change and a false discovery rate less than 0.05. Enrichment analysis of gene function and pathways was conducted by using gene set enrichment analysis with Benjamini-Hochberg correction for multiple hypothesis testing (Q < 0.05) (18). The "GOplot" R package (R Foundation for Statistical Computing) was employed to visualize gene ontology analysis.

Tumor-Infiltrating Lymphocyte Subpopulation Analysis

Single-sample gene set enrichment analysis (19) was used to calculate the enrichment scores of immune-cell types in the tumor microenvironment. Gene signatures of 28 immune-cell types associated with innate and adaptive immunity were derived as previously described (20). Tumors were further subclassified into different immune groups by using the Euclidean distance and "ward.D" clustering.

Multiplexed Immunofluorescence and Image Analysis

Tissue multiplex immunofluorescence staining was performed by using the Opal

This article has a related editorial.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Polaris five-color immunohistochemistry staining kit (Akoya Biosciences). Image Analysis was performed by using inForm 2.4.8 image analysis software (Akoya Biosciences).

High-Throughput Sequencing of T-Cell Receptor β Genes

TCR (T-cell receptor) sequencing was conducted as previously described (see the online supplement). Shannon entropy was calculated for the clonal abundance of all productive TCR sequences. Normalized Shannon entropy was determined by dividing the Shannon entropy by the natural logarithm of the number of unique productive TCR sequences as a surrogate for TCR diversity (21). Clonality is defined as "1 - normalized entropy," with values ranging from 0 to 1. The Morisita overlap index, which takes into account the specific T-cell rearrangements and their respective frequencies, was used to measure the similarity and overlap of TCR repertoires between different specimens (22).

HLA Typing and Loss of Heterozygosity

The HLA type was determined by using the OptiType algorithm (23). Polymorphisms in exons 2 and 3 of class I genes (HLA-A, HLA-B, HLA-C) were considered for HLA typing. HLA loss of heterozygosity (LOH) was assessed by using the LOHHLA computational tool (24).

Statistical Analysis

Comparisons of proportions and variables between different groups were performed by using the Wilcoxon rank sum test. Fisher exact or a chi-square tests were used to compare categorical variables. Spearman correlation was used to analyze the correlation between univariates. All statistical analyses were performed in the R statistical environment (version 3.3.4).

Additional details are provided in the online supplement.

Results

Low-Grade Lung Adenocarcinomas Were Enriched for GGO

From June 2017 to July 2018, 101 patients presenting with solitary pulmonary nodules suspicious for lung cancers who underwent surgical resection with curative intent were subjected to multiomics analyses (Figures 1 and E1). The final histologic results included lung adenocarcinoma (LUAD) (n = 82), squamous-cell carcinoma (SCC) (n = 4), small-cell lung carcinoma (SCLC) (n = 1), and nonmalignant entities (n = 14; Tables 1 and E2). Of the 101 patients, 31 patients had GGO components (NSN/PSN components) on preoperative CT scans, including 1 patient with primary alveolar epithelial hyperplasia, 11 patients with low-grade LUADs (n = 4 for lepidic-predominant invasive adenocarcinoma and n = 7 for minimally invasive adenocarcinoma), and 19 with intermediate/high-grade LUADs. All four SCCs and one SCLC presented as pure SNs. Compared with intermediate/highgrade LUADs, low-grade LUADs were significantly enriched for tumors with GGO components (100% of low-grade LUADs vs. 26.7% of intermediate/high-grade LUADs; $P = 1.3 \times 10^{-5}$, chi-square test).

GGO Was Associated with a Lower Tumor Mutation Burden

DNA from resected pulmonary nodules and DNA from paired peripheral blood mononuclear cells (PBMCs) (as the germline DNA control) were subjected to nextgeneration sequencing of 1,021 cancerassociated and HLA genes (Table E1). No mutations were identified in any of the nonmalignant pulmonary nodules, and at least one nonsilent mutation or structural variant was detected in every tumor specimen (Table E3). In the 82 LUADs, the predominant histologic type in our cohort, the median tumor mutation burden (TMB), defined as the number of nonsynonymous single-nucleotide variants and indels per megabase, was 4 mutations (Mut)/Mb, which was slightly lower than that of The Cancer Genome Atlas LUAD cohort (median TMB, 4.52 Mut/Mb; P = 0.027). The TMB was significantly higher in intermediate/highgrade LUADs than in low-grade LUADs (Figures 2A and 2B and Table E4), which is consistent with previous reports (6). In addition, the TMB was significantly higher in LUADs presenting as SNs than in those with GGO components (6.71 Mut/Mb for SNs vs. 3.08 Mut/Mb for PSNs and 1.85 Mut/Mb for NSNs; Figure 2C). This remained true when only the 67 stage I LUADs were analyzed (4.41 Mut/Mb for SNs vs. 3.10 Mut/Mb for PSNs and 1.86 Mut/Mb for NSNs; Figure E2A). Furthermore, the TMB was negatively associated with the proportion of GGO components (Figure E2B).

We next investigated whether certain cancer gene mutations were associated with

GGOs. Among 82 LUADs, EGFR (63.4%), TP53 (34.1%), KRAS (15.9%), and PIK3CA (11%) were among the top mutated cancer genes. GGO was not significantly associated with mutations of any cancer genes. However, the incidence of KRAS mutations appeared to progressively increase as the GGO component decreased (Figures 2D and E2C). PIK3CA mutations and EML4-ALK fusions were also exclusively detected in SNs (Figures 2D and E2C). Pathway analysis of these mutations demonstrated that the RAS pathway was the most frequently altered pathway in tumors with GGO components, followed by the TP53 pathway, which is consistent with a previous report (Figure E2D) (10).

GGO Was Associated with a Less Active Metabolism and Less Immune Infiltration

Fifty-one LUADs (37 SNs and 14 NSNs/ PSNs) with sufficient remaining tissue were subjected to RNA sequencing (RNA-Seq). A total of 682 significantly differently expressed genes between SNs and NSNs/PSNs (Table E5) were identified. Pathway analysis revealed 10 significantly upregulated pathways in SNs, 9 of which were related to metabolism (Figure E3A), indicating a more active overall metabolism in LUADs presenting as SNs. The only non-metabolism-related pathway was the cytokine-cytokine receptor interaction pathway (false discovery rate, 0.02; Figures E3A and E3B). Similarly, enrichment analysis of gene ontology also demonstrated metabolism- and immune-associated processes being upregulated in SNs (Figure E3C).

Unsupervised clustering based on immune infiltration was used to classify the 51 LUADs into two clusters (Figure 3A). Tumors from the immune-high group demonstrated upregulation of genes associated with immune activation, including granzyme, perforin, CD8, and cytolytic activity (CYT), among others (25). Of note, the 14 NSNs/PSNs were enriched in the immune-low group (10 of 22 immune-low tumors vs. 4 of 29 immune-high tumors; P = 0.025, Fisher exact test). Compared with SNs, NSNs/PSNs had significantly less expression of genes involved in immune activation (e.g., CD8A, IFN-y, CYT, etc.) and immunomodulation, including several critical chemokines known to promote the recruitment of immune cells (CXCL5, CXCL6, CXCL8, CCL20, CXCL13, CXCL1, CXCL9) (26) (Figures 3B, 3C, and E3D).



Figure 1. Research strategy. (A) Description of sample collection and multiomics analysis. (B) Example axial CT images, 3D vascular reconstruction, and volumetric measurement for one nonsolid nodule (NSN) and one part-solid nodule (PSN) (50%). The ground-glass opacity

Moreover, the expression of some of these critical immune genes was negatively correlated with the proportion of GGOs (Figure E4A). We next inferred the major immune-cell subtypes from RNA-Seq data and observed an overall lower degree of immune infiltration in NSNs/PSNs than in SNs (Figures 3D and E4B). In results similar to those for gene expression, the infiltration of certain immune cells such as activated CD4 T cells, activated CD8⁺ T cells, and regulatory T cells was negatively associated with the proportion of GGO components (Figure E4C). Interestingly, unlike all other immunecell types, eosinophils, which are known to play important roles in T-helper cell type 2 (Th2) recruitment related to allergic reactions and inflammatory responses (27), demonstrated a higher degree of infiltration in NSNs/PSNs than in SNs (Figure 3D). Similarly, DAPK2 and CSF3R, two genes involved in granulocytic maturation, were upregulated in the GGO-associated LUADs, whereas IFN- γ and TNF-a, two cytokines associated with Th1, were lower in GGOassociated LUADs. These results imply a shift in the Th1/Th2 balance toward Th2 dominance in GGO-associated LUADs (28), which has been credited for cancer immune tolerance. Furthermore, multiplex immunofluorescence staining using T-cell markers in a subset of LUADs (5 NSNs/PSNs and 6 SNs) confirmed lower amounts of CD4⁺, CD8⁺, and regulatory T-cell (Treg) infiltration in NSNs/PSNs, which is consistent with RNA-Seq deconvolution (Figures 3E, E4D, and E4E). Taken together, these results are suggestive of an overall less active immune landscape in GGO-associated LUADs as compared with their SN-associated counterparts.

GGO Was Associated with Distinct T-Cell Repertoire Metrics

TCR repertoire reflects the breadth and strength of the T-cell response (22, 29), which plays a central role in antitumor immune responses across different cancer types, including lung cancer (30). TCR sequencing was performed on 84 tumors and 81 PBMC specimens from 87 patients with cancer as well as on 14 tissue specimens and 12 PBMC specimens from 14 patients with nonmalignant nodules who had sufficient remaining DNA. The median number of unique TCR clonotypes per patient was 1,456 for tissue and 6,840 for PBMCs. TCR diversity (Shannon entropy), the proportion of top 10 T-cell clones, and T-cell clonality were not statistically different between nonmalignant nodules and LUADs (Figures E5A and E5B). Interestingly, when GGO components were considered, TCR clonality and the proportion of top 10 TCR clones were significantly lower in NSNs/PSNs than in SNs, which is consistent with lower T-cell expansion (Figure 4A and Table E4).

TCR parameters from PBMCs were not different across clinical parameters (Figures 4B, E5C, and E5D). However, compared with PBMCs from 94 age- and sex-matched healthy Chinese individuals, PBMCs from patients with pulmonary nodules had higher clonality and lower diversity, indicating antigen-specific T-cell expansion in these patients (Figures E5E and E5F). As expected, the TCR similarity as assessed by using the Morisita index of overlap between tissues and PBMCs from the same patients exceeded that across different patients between tissues, between PBMCs, and between PBMCs and tissues (Figures E5G and E5H). Importantly, compared with patients with LUADs presenting with SNs, patients who had NSN/PSN-associated LUADs shared more T-cell clones between PBMCs and paired tumor tissues (*P* = 0.036; Figures 4C, E5I, and E5J). Furthermore, we applied the TRUST4 algorithm (31) to RNA-Seq data from the 51 tumors to infer the TCR repertoire attributes for validation purposes. As shown in Figure E5K, the clonality estimated by using the TRUST4 algorithm was also higher in SNs than in NSNs/ PSNs, although it did not reach statistical significance, which is consistent with TCRsequencing data, supporting more potent T-cell expansion in SN-associated LUADs.

GGO Was Associated with a Lower Incidence of HLA LOH

We next explored whether the presence of GGO was associated with HLA heterozygous

states or LOH, which has been reported to impact antigen presentation, immune escape, and the response to immunotherapy (24, 32). Overall, HLA LOH was identified in 31.0% (27 of 87) of lung cancers, which is consistent with previous reports (24), but was identified in none of the nonmalignant pulmonary nodules. The incidence of lost alleles was similar between different HLA genes (Figure 4D). Of note, HLA LOH was significantly more common in LUADs presenting with SNs than in those presenting with NSNs/ PSNs (40.4% [21 of 52] vs. 16.7% [5 of 30]; P = 0.029, chi-square test) (Figures 4E and 4F). The HLA homozygous ratio was not different in malignant versus nonmalignant nodules (17 of 82 vs. 2 of 14; P = 0.73, chisquare test) or in LUADs with GGO components (NSNs/PSNs) versus LUADs without GGO components (10 of 52 vs. 7 of 30; P = 0.87, chi-square test) (Figure 4E).

Release of Tumor DNA to Peripheral Blood May Be Associated with GGO Status

To assess whether tumor DNA shedding to blood is impacted by GGO status, cfDNA was subjected to ultra-deep sequencing of the same 1,021 cancer genes to an average depth of 25,984× (Figure 5A). The concentration of cfDNA from NSNs/PSNs was 12.90 ng/ml, which was similar to the concentration among nonmalignant nodules (10.46 ng/ml; P = 0.45), but was significantly lower than that among SNs (14.89 ng/ml; P = 0.048) (Figure 5B). In LUADs, the cfDNA concentration was higher in later-stage disease, although the difference was not statistically significant (P = 0.69, Kruskal-Wallis test) (Figure E6A).

By using a tumor-informed strategy (identical mutations in paired tumor DNA), mutations were detected in cfDNA from 35 of 85 patients with lung cancer, including 27 of 55 (49%) patients with SNs and 8 of 30 (27%) patients with NSNs/PSNs (P = 0.076, chi-square test) (Figure 5C). Mutations were identified in cfDNA from all four patients with SCC and from one patient with SCLC but were identified in 30 of 80 (38%) patients with LUAD (P = 0.01, chi-square test) (Figure 5D and Table E6), which is consistent with

Figure 1. (*Continued*). components are equal to the NS divided by the TV. In *B*, the histologic features of nodules, including (*a*) an NSN of minimally invasive adenocarcinoma, (*b*) a PSN of minimally invasive adenocarcinoma, and (*c*) an SN of invasive adenocarcinoma, are shown Scale bars, 100 μ m. 3D = three-dimensional; CT = computed tomography; NS = volume size of NSN; pts. = patients; S = volume size of solid component; SN = solid nodule; TCR = T-cell receptor; TV = total volume of lesion.

Table 1. Clinical Characteristics

Age at diagnosis, yr, mean (range) 60 (25 to 85) Sex 52 Female 49 Stage (version 8 TNM criteria) 1 Nonmalignant 14 Stage I 69 Stage III 10 Stage IV 1 Histologic type 1 Nonmalignant 14 SCC 4 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 7 Yes 24 No 70 Tumor size, cm 5 ≤1 5 1-2 52 2-3 28 3-4 10	Characteristic	Total (N = 101)
Sex 52 Male 49 Stage (version 8 TNM criteria) 14 Nonmalignant 14 Stage I 69 Stage II 10 Stage IV 1 Histologic type 1 Nonmalignant 14 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Sono 77 GGO status 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	Age at diagnosis, yr, mean (range)	60 (25 to 85)
Male 49 Stage (version 8 TNM criteria) 14 Nonmalignant 14 Stage I 69 Stage II 10 Stage III 7 Stage IV 1 Histologic type 1 Nonmalignant 14 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 4 PPA 6 SPA 4 IMA 15 PEAC 1 Stage SPA 24 No 77 GGO status 70 Tumor size, cm 5 ≤1 5 1-2 52 2-3 28 3-4 10	Female	52
Stage (version 8 TNM criteria) 14 Nonmalignant 10 Stage I 10 Stage II 7 Stage IV 1 Histologic type 14 Nonmalignant 14 SCC 4 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 7 Yes 24 NSN 70 Tumor size, cm 5	Male	49
Stage I 69 Stage II 10 Stage III 7 Stage IV 1 Histologic type 14 Nonmalignant 14 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 7 Yes 24 NSN 7 SN 70 Tumor size, cm 5 ≤1 52 2-3 28 3-4 10	Stage (version 8 TNM criteria)	1/
Stage II 10 Stage III 7 Stage IV 1 Histologic type 14 Nonmalignant 14 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 45 IMA 15 PEAC 1 Smoking 7 Yes 24 No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 ≤1 52 2-3 28 3-4 10	Stage I	69
Stage III 7 Stage IV 1 Histologic type 1 Nonmalignant 14 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 24 Yes 24 No 7 GGO status 7 SN 70 Tumor size, cm 5 ≤1 52 2-3 28 3-4 10	Stage II	10
Stage IV 1 Histologic type 14 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 82 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 77 GGO status 77 NSN 7 SN 77 GGO status 7 NSN 7 SN 70 Tumor size, cm 5 ≤1 5 1-2 52 2-3 28 3-4 10	Stage III	7
Nonmalignant 14 SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 7 Yes 24 No 77 GGO status 7 NSN 7 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	Histologic type	I
SCC 4 SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 77 Yes 24 No 77 GGO status 7 NSN 7 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	Nonmalignant	14
SCLC 1 LUAD 82 Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 24 No 77 GGO status 7 NSN 7 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	SCC	4
Pathologic subtype 7 MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 24 No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10		82
MIA 7 LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 24 No 77 GGO status 77 NSN 7 PSN 24 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	Pathologic subtype	62
LPA 4 APA 45 PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 24 No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	MIA	7
PPA 6 SPA 4 IMA 15 PEAC 1 Smoking 24 No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	ΔΡΔ	4
SPA 4 IMA 15 PEAC 1 Smoking 24 No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	PPA	
$\begin{array}{cccc} \text{IMA} & 15 \\ \text{PEAC} & 1 \\ \text{Smoking} & & & \\ \text{Yes} & & 24 \\ \text{No} & & 77 \\ \text{GGO status} & & & \\ \text{NSN} & & & 7 \\ \text{PSN} & & & 24 \\ \text{SN} & & & 70 \\ \text{Tumor size, cm} & & & \\ \hline & \leqslant 1 & & 5 \\ 1-2 & & & 52 \\ 2-3 & & & 28 \\ 3-4 & & & 10 \\ \hline & & & & & 10 \\ \hline \end{array}$	SPA	4
FEAC 1 Smoking 24 No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 $1-2$ 52 $2-3$ 28 $3-4$ 10	IMA	15
Yes 24 No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 ≤1 5 1-2 52 2-3 28 3-4 10	Smoking	I
No 77 GGO status 7 NSN 7 PSN 24 SN 70 Tumor size, cm 5 1-2 52 2-3 28 3-4 10	Yes	24
NSN 7 PSN 24 SN 70 Tumor size, cm ≤ 1 ≤ 1 5 1-2 52 2-3 28 3-4 10	No CCO atatua	77
PSN 24 SN 70 Tumor size, cm 5 ≤1 5 1-2 52 2-3 28 3-4 10	NSN	7
SN 70 Tumor size, cm 5 ≤1 5 1-2 52 2-3 28 3-4 10	PSN	24
≤ 1 5 $1-2$ 52 $2-3$ 28 $3-4$ 10	SN	70
1-2 52 2-3 28 3-4 10	lumor size, cm ≤1	5
2–3 28 3–4 10	1–2	52
3–4 10	2–3	28
>4	3–4 >4	10

Definition of abbreviations: APA = acinar-predominant invasive adenocarcinoma; GGO = ground-glass opacity; IMA = invasive mucinous adenocarcinoma; LPA = lepidicpredominant invasive adenocarcinoma; LUAD = lung adenocarcinoma; MIA = minimally invasive adenocarcinoma; NSN = nonsolid nodule; PEAC = pulmonary enteric adenocarcinoma; PPA = papillary-predominant invasive adenocarcinoma; PSN = part-solid nodule; SCC = lung squamous-cell carcinoma; SCLC = small-cell lung carcinoma; SN = solid nodule; SPA = solidpredominant adenocarcinoma.

previous studies (33). Furthermore, the detection rate of circulating tumor DNA, defined as cfDNA with positive mutations identical to those in paired tumor tissues, was positively correlated with the solid tumor volume (P = 0.011), but was not positively correlated with the total tumor volume incorporating GGO components, further supporting that GGO was associated with less tumor DNA shedding (Figures 5E, E6B, and E6C). Among seven patients with LUAD and NSNs, mutations were detected in only one patient with a pulmonary vein running through the lesion (Figure E6D), highlighting the impact of vascularization on tumor DNA shedding.

Discussion

The GGO component has been speculated to represent early development of lung cancers (34), and the molecular features associated with GGO may reflect the molecular events associated with early lung carcinogenesis. The initiation and progression of cancers result from the accumulation of mutations offering survival and growth advantages (5), in line with our findings that GGO components are associated with a lower TMB. These results are consistent with findings from our previous studies regarding the progressive increase of the TMB with the progression from precancers to invasive LUADs (6). In addition, RNA-Seq data revealed upregulation of several metabolic pathways in SN-associated LUADs compared with GGO-associated LUADs, indicating increasing metabolic activities together with cancer progression, which is consistent with the slow growth of GGOassociated lung cancers (35) and the known fact that quickly growing tumors tend to have a more active metabolism (36).

In addition to accumulating molecular alterations, cancer evolution is constantly shaped by the dynamic interaction between cancer cells and host factors, particularly immune surveillance (37), which has become an intense research focus (9, 38-42). However, immune profiling results are often inconsistent between studies, reflecting the profound heterogeneity of LUADs. Furthermore, these studies were either not focused on GGO-associated LUADs or were limited by a small sample size; therefore, our understanding of the immune features associated with GGO remains rudimentary. In the current study, we applied multiomics analyses to a relatively large cohort of lung nodules with a focus on GGO. Our results revealed a generalized less active immune microenvironment in GGO-associated LUADs compared with SN-associated LUADs, including less expression of immune-related genes, downregulation of immune pathways, less infiltration of immune cells, and a less expanded TCR repertoire.

We integrated the genomic and immune features in Figure 6 to provide a global view of this cohort of LUADs with or without GGO components. The parallel trends in genomic and immune features after the incremental changes of GGO components are consistent with the concept of cancer immunoediting composed of three phases: elimination, equilibrium, and escape (43). During elimination, tumor cells are actively cleared by both innate and acquired immune responses. After elimination comes the second phase, equilibrium, when the antigenicity of tumor cells is weakened and tumors are not readily recognized and cleared by the immune system. However, tumor cells are constantly under pressure from the immune system and therefore cannot grow excessively. The generalized lower inflammation/immune contexture and lower incidence of HLA LOH in GGOassociated LUADs are in line with the notion that the cancer cells and antitumor immune response may be in the equilibrium phase of



Figure 2. Genomic landscape of lung adenocarcinoma with or without ground-glass opacity (GGO) components. (*A*) Top mutated cancer genes in at least three patients. Clinical features are listed at the bottom of the figure, and the number of mutated genes in each patient is shown at the top of the figure as red bars. The prevalence of mutated cancer genes in this cohort is represented as blue bars to the left of the figure. Patient PTHOG0047 with the highest tumor mutation burden in this cohort is detailed in the online supplement. (*B*) The association of the tumor mutation burden with low-grade ($n = 11, 2.27 \pm 1.19$ Mut/Mb) versus intermediate or high-grade histology ($n = 71, 5.74 \pm 13.29$ Mut/Mb) or (*C*) GGO status (NSN: $n = 7, 1.85 \pm 1.06$ Mut/Mb; PSN: $n = 23, 3.08 \pm 1.27$ Mut/Mb; SN: $n = 52, 6.71 \pm 15.44$ Mut/Mb). *N* represents the sample size in each group, and the mean \pm SD is also shown in brackets. (*D*) The incidence of commonly mutated cancer genes in lung adenocarcinomas with different proportions of GGO components (grouped as quartiles). For box plots within violin plots, each box indicates the first quartile (Q1) and third quartile (Q3), and the black horizontal line represents the median; the upper whisker is the min[max(x), Q3 + 1.5 × IQR], and the lower whisker is the max[min(x), Q1 - 1.5 × IQR], where x represents the data, Q3 is the 75th percentile, Q1 is the 25th percentile, and IQR = Q3 - Q1. The widths of the violin plots indicate the kernel density of the data. CNV = copy number variation; IQR = interquartile range; LPA = lepidic-predominant invasive adenocarcinoma; max = maximum; MIA = minimally invasive adenocarcinoma; min = minimum; Mut = mutations; NSN = nonsolid nodule; PSN = part-solid nodule; SN = solid nodule; SV = structural variation.

immunoediting. During the equilibrium phase, cancers may acquire a small number of somatic mutations, and tumor cells continue to evolve under a constant but less active immune pressure; this balance may remain for years, mirroring the indolent clinical course of GGO-associated LUADs. Cancer cells continue accumulating molecular alterations (e.g., progressively increase the TMB) until the acquisition of certain molecular changes that empower the cancer cells to evade the immune response (e.g., HLA LOH), at which point the equilibrium is broken and invasive cancer development is promoted.



Figure 3. The association of the immune contexture derived from RNA sequencing with different ground-glass opacity (GGO) statuses in lung adenocarcinomas. (*A*) Heatmap of normalized enrichment scores for infiltration of 28 immune cells. The upper panel represents the expression of cytolytic activity, PRF1, GZMA, CD8, and IFNG. Patients' features, including smoking, stage, GGO status (red, solid nodule [SN] with no GGO; blue, nonsolid nodule [NSN] + part-solid nodule [PSN]), and the GGO-component percentage (indicated as color scales), are listed. (*B*) Expression of immune markers as determined by using RNA-sequencing data (14 NSN/PSN cases vs. 37 SN cases). The expression levels were 3.43 ± 0.73 versus $4.18 \pm 0.76 \log_2$ (transcripts per million [TPM] + 1) for CD8A, 0.59 ± 0.44 versus $1.31 \pm 0.97 \log_2$ (TPM + 1) for IFNG, and 4.23 ± 0.58 versus $4.80 \pm 0.72 \log_2$ (TPM + 1) for cytolytic activity. (*C*) Volcano plot of differentially expressed genes comparing SNs with NSNs/PSNs. Red dots and green dots indicate significantly upregulated and downregulated genes in NSN and PSNs, respectively. (*D*) Comparisons of the NSNs/PSNs with SNs for different immune cells (14 NSN/PSN cases vs. 37 SN cases). The measures of infiltration were



Figure 4. T-cell clonality (left), diversity (middle), and the proportion of the top 10 clones (right) in (*A*) nonmalignant nodules versus lung adenocarcinoma (LUAD) tissues with ground-glass opacity (GGO) components (NSNs/PSNs) or without GGO components (SNs) and (*B*) peripheral blood mononuclear cell (PBMCs). (*C*) Shared T-cell clones in LUAD specimens and PBMCs as quantified by using the MOI. (*D*) For each patient, the HLA heterozygote status and data on whether HLA-A/B/C kept or lost alleles are listed. The right bar represents the number of patients in whom HLA loss of heterozygosity (LOH) events on related HLA genes occurred. (*E*) Numbers of patients with HLA LOH events (left panel) and an HLA homozygous status (right) among patients with LUAD with a GGO component versus without a GGO component. *P* values as calculated by using the chi-square test and ORs are shown. (*F*) The incidence of HLA LOH events in LUADs with different proportions of GGO components (grouped as quartiles). MOI = Morisita overlap index; NSN = nonsolid nodule; OR = odds ratio; PSN = part-solid nodule; SN = solid nodule.

Immune checkpoint blockade (ICB) strategies that target T cells by using anti–PD-1/PD-L1 have revolutionized the therapeutic landscape in cancers (44). The use of ICB is also being tested in early-stage lung cancers, including GGO-associated LUADs (clinicaltrials.gov identifier NCT 04047186) and even precancers (clinicaltrials.gov identifier NCT 03634241). As data from metastatic cancer have suggested that cancers' high TMB and active immune infiltration may be associated with a superior benefit being gained from the use of ICB (45), a relatively lower TMB and degree of immune infiltration in GGO-associated LUADs in the current study may raise concerns for these trials. A recent study reported a response rate of only 8.1% in synchronous GGO-associated

Figure 3. (*Continued*). 0.20 ± 0.85 versus 0.30 ± 0.09 for activated CD4 T cells, 0.46 ± 0.06 versus 0.51 ± 0.06 for activated CD8 T cells, 0.40 ± 0.06 versus 0.45 ± 0.09 for regulatory T cells, and 0.01 ± 0.06 versus -0.03 ± 0.07 for eosinophils. (*E*) Representative images from patients with NSNs and SNs for CD3⁺ T cells (magenta), CD4⁺ T cells (green), CD8⁺ T cells (red), FoxP3⁺ regulatory T cells (cyan), and DAPI (blue), as determined by using multiplex immunofluorescence staining. Scale bars, 100 μ m. FDR = false discovery rate; MDSC = myeloid-derived suppressor cell.



Figure 5. Cell-free DNA (cfDNA) mutations in patients with lung adenocarcinoma with clinicopathologic features. (*A*) The number of variations from tissue and plasma, the VAF of driver genes detected from plasma, and clinical features in 82 patients with lung adenocarcinoma are listed. (*B*) The cfDNA concentration of nonmalignant lung nodules versus lung cancers presenting as SNs or NSNs/PSNs. Numbers of patients (*C*) with versus without GGO component and (*D*) with adenocarcinoma versus with nonadenocarcinoma. (*E*) The solid volume in circulating tumor DNA (ctDNA)-positive versus ctDNA-negative nodules. For all box plots or box plots within violin plots, each box indicates the first quartile (Q1) and third quartile (Q3), and the black horizontal line represents the median; the upper whisker is the min[max(*x*), Q3 + 1.5 × IQR], and the lower whisker is the max[min(*x*), Q1 - 1.5 × IQR], where *x* represents the data, Q3 is the 75th percentile, Q1 is the 25th percentile, and the IQR = Q3 - Q1. The widths of violin plots indicate the kernel density of the data. GGO = ground-glass opacity; IQR = interquartile range; LPA = lepidic-predominant invasive adenocarcinoma; max = maximum; MIA = minimally invasive adenocarcinoma; min = minimum; NSN = nonsolid nodule; PSN = part-solid nodule; SN = solid nodule; SNV = single-nucleotide variant; VAF = variant allele frequency.

lung nodules in a cohort of patients with lung cancer receiving anti–PD-1/PD-L1 therapy (42). However, the pathologic nature of these lung nodules was not known, and the benefit from ICB in patients with GGO-associated, early-stage LUADs or precancers has yet to be determined.

Another point worth discussing is the impact of GGO components on the mutation detection sensitivity of liquid biopsy. Liquid biopsy has been widely adapted in the diagnosis, prognosis, and disease monitoring of lung cancers (46, 47). Several studies have demonstrated that the sensitivity of cfDNA analysis depends on the tumor volume (33, 48), but little attention has been paid to the impact of the GGO component. Our results suggest that the proportion of GGO components is also a phenotype associated with detection sensitivity.

One limitation of our study is that the majority of patients were never smokers, which is different from many other previous studies in the Western population. On the other hand, a large proportion of neversmokers was also reported in studies on early-stage lung cancers presenting as GGOpredominant nodules in an East Asian population (10, 49), which may reflect the ethnic and etiologic difference in East Asian patients with lung cancer. As shown in Figures E6E–E6G, we assessed the potential impact of smoking on our findings and observed similar trends regardless of smoking status, although we were not able to reach statistical significance, as we were likely limited by the small sample size of the smoking group. Nevertheless, studies on other ethnic groups and smokers are warranted before these intriguing findings are generalized. In addition, tumor evolutionary heterogeneity should be taken into consideration. Some GGO-associated lung nodules gradually grow, whereas some remain unchanged for years, and a



Figure 6. Integrated genomic and immune features across patients with lung adenocarcinoma with different GGO components. ctDNA = circulating tumor DNA; CYT = cytologic activity; GGO = ground-glass opacity; LOH = loss of heterozygosity; MUT = mutation; NSN = nonsolid nodule; PSN = part-solid nodule; SN = solid nodule; TMB = tumor mutation burden.

considerable proportion of LUADs do not appear to have a GGO phase on the basis of data from the National Lung Cancer Screening Trial (50). Further studies on large cohorts of lung cancers with or without GGO components that use long-term, longitudinal CT scans may be able to identify different subgroups to determine whether radiologic changes over time reflect the dynamic changes of the immune microenvironment. More comprehensive profiling approaches, such as spatial transcriptomic single-cell sequencing technologies, which could leverage gene expression profiles while retaining information about the spatial tissue, are needed to reveal the cellular ecosystem and its precise proportions of activated and exhausted immune-cell types and to elucidate the interaction between early changes in the immune microenvironment and tumor cells in GGO-associated lung cancers.

In summary, we initiated a prospective study to collect and characterize pulmonary nodules with or without GGO components, which were precisely quantified by using 3D image reconstruction to delineate immunogenomic features associated with GGO. To the best of our knowledge, this is the first prospective study to focus on the GGO components of pulmonary nodules by using a multiomics approach. The use of precise 3D volumetric analysis increased the accuracy of GGO-component quantification compared with the traditionally adopted twodimensional tumor measurement (consolidation tumor ratio). Intrinsic biological features may have contributed to the indolent clinical course of GGOassociated lung cancers, supporting the hypothesis that GGO may represent early carcinogenesis of a subset of LUADs when cancer cells and the antitumor immune response are at an equilibrium and providing mechanistic insights into the diagnosis and treatment of these radiologically distinct clinical entities.

<u>Author disclosures</u> are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank Yun Wang and Yanyan Hou of Peking University People's Hospital for assistance with sample collection and preservation.

References

- Yankelevitz DF, Yip R, Smith JP, Liang M, Liu Y, Xu DM, et al.; International Early Lung Cancer Action Program Investigators Group. CT screening for lung cancer: nonsolid nodules in baseline and annual repeat rounds. *Radiology* 2015;277:555–564.
- Naidich DP, Bankier AA, MacMahon H, Schaefer-Prokop CM, Pistolesi M, Goo JM, et al. Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the Fleischner Society. Radiology 2013;266:304–317.
- Kozower BD, Larner JM, Detterbeck FC, Jones DR. Special treatment issues in non-small cell lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest* 2013;143:e369S–e399S.
- Chen J, Yang H, Teo ASM, Amer LB, Sherbaf FG, Tan CQ, et al. Genomic landscape of lung adenocarcinoma in East Asians. Nat Genet 2020;52: 177–186.
- Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al.; TRACERx Consortium. Tracking the evolution of non-small-cell lung cancer. N Engl J Med 2017;376: 2109–2121.
- Hu X, Fujimoto J, Ying L, Fukuoka J, Ashizawa K, Sun W, et al. Multiregion exome sequencing reveals genomic evolution from preneoplasia to lung adenocarcinoma. *Nat Commun* 2019;10:2978.
- Kadara H, Choi M, Zhang J, Parra ER, Rodriguez-Canales J, Gaffney SG, 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.

- Kobayashi Y, Mitsudomi T, Sakao Y, Yatabe Y. Genetic features of pulmonary adenocarcinoma presenting with ground-glass nodules: the differences between nodules with and without growth. *Ann Oncol* 2015;26:156–161.
- Zhang C, Zhang J, Xu FP, Wang YG, Xie Z, Su J, et al. Genomic landscape and immune microenvironment features of preinvasive and early invasive lung adenocarcinoma. J Thorac Oncol 2019;14:1912–1923.
- Li Y, Li X, Li H, Zhao Y, Liu Z, Sun K, et al. Genomic characterisation of pulmonary subsolid nodules: mutational landscape and radiological features. Eur Respir J 2020;55:1901409.
- 11. Li R, Li X, Xue R, Yang F, Wang S, Li Y, *et al.* Early metastasis detected in patients with multifocal pulmonary ground-glass opacities (GGOs). *Thorax* 2018;73:290–292.
- MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology* 2017;284: 228–243.
- Pedersen JH, Rzyman W, Veronesi G, D'Amico TA, Van Schil P, Molins L, et al. Recommendations from the European Society of Thoracic Surgeons (ESTS) regarding computed tomography screening for lung cancer in Europe. Eur J Cardiothorac Surg 2017;51:411–420.
- Bai C, Choi CM, Chu CM, Anantham D, Chung-Man Ho J, Khan AZ, et al. Evaluation of pulmonary nodules: clinical practice consensus guidelines for Asia. *Chest* 2016;150:877–893.
- Chen K, Zhao H, Bai J, Yang F, Chang L, Guan Y, et al. A prospective cohort study of TMB and determinants of ctDNA detection by comprehensive genomic profiling in stage I lung adenocarcinomas [abstract]. J Thorac Oncol 2018;13:S726.
- Chen K, Bai J, Zhao H, Yang F, Zhang C, Wang Y, et al. Comprehensive profiling of genomic and TCR repertoire in localized stage lung adenocarcinomas from a prospective cohort study [abstract]. Ann Oncol 2019;30:ii26.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15: 550.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545–15550.
- Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
- Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotypeimmunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* 2017;18:248–262.
- 21. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, et al. Comprehensive assessment of T-cell receptor β -chain diversity in $\alpha\beta$ T cells. Blood 2009;114:4099–4107.
- 22. Reuben A, Gittelman R, Gao J, Zhang J, Yusko EC, Wu CJ, et al. TCR repertoire intratumor heterogeneity in localized lung adenocarcinomas: an association with predicted neoantigen heterogeneity and postsurgical recurrence. *Cancer Discov* 2017;7:1088–1097.
- Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* 2014;30:3310–3316.
- 24. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, *et al.*; TRACERx Consortium. Allele-specific HLA loss and immune escape in lung cancer evolution. *Cell* 2017;171:1259–1271, e11.
- 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.
- Griffith JW, Sokol CL, Luster AD. Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annu Rev Immunol* 2014;32:659–702.
- 27. Maggi E. The TH1/TH2 paradigm in allergy. *Immunotechnology* 1998;3: 233–244.
- Campbell JD, HayGlass KT. T cell chemokine receptor expression in human Th1- and Th2-associated diseases. Arch Immunol Ther Exp (Warsz) 2000;48:451–456.
- Joshi K, de Massy MR, Ismail M, Reading JL, Uddin I, Woolston A, et al.; TRACERx Consortium. Spatial heterogeneity of the T cell receptor repertoire reflects the mutational landscape in lung cancer. Nat Med 2019;25:1549–1559.

- Reuben A, Zhang J, Chiou SH, Gittelman RM, Li J, Lee WC, et al. Comprehensive T cell repertoire characterization of non-small cell lung cancer. Nat Commun 2020;11:603.
- Song L, Cohen D, Ouyang Z, Cao Y, Hu X, Liu XS. TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. *Nat Methods* 2021;18:627–630.
- 32. Negrao MV, Lam VK, Reuben A, Rubin ML, Landry LL, Roarty EB, et al. PD-L1 expression, tumor mutational burden, and cancer gene mutations are stronger predictors of benefit from immune checkpoint blockade than HLA class I genotype in non-small cell lung cancer. J Thorac Oncol 2019;14:1021–1031.
- Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, et al.; TRACERx Consortium; PEACE Consortium. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature 2017; 545:446–451.
- 34. Travis WD, Asamura H, Bankier AA, Beasley MB, Detterbeck F, Flieder DB, et al. The IASLC Lung Cancer Staging Project: proposals for coding T categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the TNM classification of lung cancer. J Thorac Oncol 2016;11:1204–1223.
- Kakinuma R, Noguchi M, Ashizawa K, Kuriyama K, Maeshima AM, Koizumi N, et al. natural history of pulmonary subsolid nodules: a prospective multicenter study. J Thorac Oncol 2016;11:1012–1028.
- Rosenthal R, Cadieux EL, Salgado R, Bakir MA, Moore DA, Hiley CT, et al.; TRACERx Consortium. Neoantigen-directed immune escape in lung cancer evolution. *Nature* 2019;567:479–485.
- Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. Nat Med 2018;24:541–550.
- Naito M, Aokage K, Saruwatari K, Hisakane K, Miyoshi T, Hishida T, et al. Microenvironmental changes in the progression from adenocarcinoma in situ to minimally invasive adenocarcinoma and invasive lepidic predominant adenocarcinoma of the lung. *Lung Cancer* 2016;100:53–62.
- Chen H, Carrot-Zhang J, Zhao Y, Hu H, Freeman SS, Yu S, *et al.* Genomic and immune profiling of pre-invasive lung adenocarcinoma. *Nat Commun* 2019;10:5472.
- 40. Dejima H, Hu X, Chen R, Zhang J, Fujimoto J, Parra ER, *et al.* Immune evolution from preneoplasia to invasive lung adenocarcinomas and underlying molecular features. *Nat Commun* 2021;12:2722.
- Hu X, Estecio MR, Chen R, Reuben A, Wang L, Fujimoto J, et al. Evolution of DNA methylome from precancerous lesions to invasive lung adenocarcinomas. *Nat Commun* 2021;12:687.
- Wu F, Li W, Zhao W, Zhou F, Xie H, Shi J, et al. Synchronous groundglass nodules showed limited response to anti-PD-1/PD-L1 therapy in patients with advanced lung adenocarcinoma. *Clin Transl Med* 2020; 10:e149.
- O'Donnell JS, Teng MWL, Smyth MJ. Cancer immunoediting and resistance to T cell-based immunotherapy. *Nat Rev Clin Oncol* 2019;16:151–167.
- Allinson JP, Brown J, Gibb K, Navani N. Immunotherapy in non-small cell lung cancer: which patients and at which stage? *Am J Respir Crit Care Med* 2019;199:1277–1279.
- 45. Beattie J, Yarmus L, Wahidi M, Rivera MP, Gilbert C, Maldonado F, et al.; American Thoracic Society Assembly on Thoracic Oncology. The Immune landscape of non-small-cell lung cancer: utility of cytologic and histologic samples obtained through minimally invasive pulmonary procedures. Am J Respir Crit Care Med 2018;198:24–38.
- Chen K, Sun J, Zhao H, Jiang R, Zheng J, Li Z, *et al.* Non-invasive lung cancer diagnosis and prognosis based on multi-analyte liquid biopsy. *Mol Cancer* 2021;20:23.
- Yamauchi Y, Safi S, Blattner C, Rathinasamy A, Umansky L, Juenger S, et al. Circulating and tumor myeloid-derived suppressor cells in resectable non-small cell lung cancer. Am J Respir Crit Care Med 2018;198:777–787.
- Chen K, Zhao H, Shi Y, Yang F, Wang LT, Kang G, et al. Perioperative dynamic changes in circulating tumor DNA in patients with lung cancer (DYNAMIC). *Clin Cancer Res* 2019;25:7058–7067.
- 49. Fu F, Zhang Y, Wen Z, Zheng D, Gao Z, Han H, et al. Distinct prognostic factors in patients with stage I non-small cell lung cancer with radiologic part-solid or solid lesions. J Thorac Oncol 2019;14:2133–2142.
- Yip R, Henschke CI, Xu DM, Li K, Jirapatnakul A, Yankelevitz DF. Lung cancers manifesting as part-solid nodules in the National Lung Screening Trial. *AJR Am J Roentgenol* 2017;208:1011–1021.