JCI insight

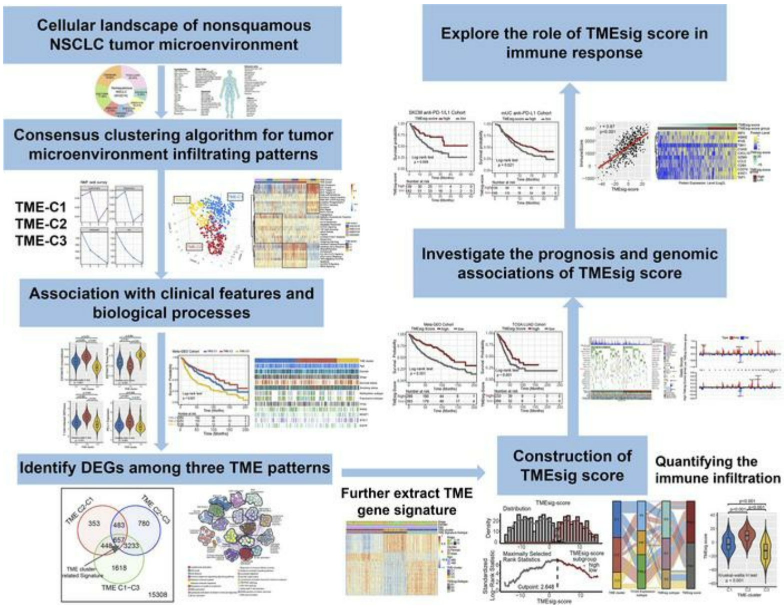
Deciphering the tumor microenvironment cell–infiltrating landscape reveals microenvironment subtypes and therapeutic potentials for nonsquamous NSCLC

Hao Chen, Tongchao Zhang, Yuan Zhang, Hao Wu, Zhen Fang, Yang Liu, Yang Chen, Zhe Wang, Shengtao Jia, Xingzhao Ji, Liang Shang, Fengying Du, Jin Liu, Ming Lu, Wei Chong

JCI Insight. 2022;7(12):e152815. https://doi.org/10.1172/jci.insight.152815.

Research Article **Immunology** **Oncology**

Graphical abstract



JCI insight

Deciphering the tumor microenvironment cell–infiltrating landscape reveals microenvironment subtypes and therapeutic potentials for nonsquamous NSCLC

Hao Chen,[1] Tongchao Zhang,[1] Yuan Zhang,[1] Hao Wu,[2,3,4] Zhen Fang,[2,3,4] Yang Liu,[2,3,4] Yang Chen,[5] Zhe Wang,[6] Shengtao Jia,[6] Xingzhao Ji,[4,7] Liang Shang,[2,3,4] Fengying Du,[2,3,4] Jin Liu,[8] Ming Lu,[1] and Wei Chong[2,3,4]

[1]Clinical Research Center of Shandong University, Clinical Epidemiology Unit, Qilu Hospital of Shandong University, Jinan, Shandong, China. [2]Department of Gastrointestinal Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong, China. [3]Key Laboratory of Engineering of Shandong Province, Shandong Provincial Hospital, Jinan, Shandong, China. [4]Medical Science and Technology Innovation Center, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, Shandong, China. [5]Department of Radiation Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China. [6]Department of Tumor Cell Biology, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China. [7]Department of Pulmonary Medicine, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong, China. [8]Research Center for Experimental Nuclear Medicine, School of Basic Medical Sciences, Shandong University, Jinan, Shandong, China.

Recent studies highlighted the clinicopathologic importance of the tumor microenvironment (TME) in delineating molecular attributes and therapeutic potentials. However, the overall TME cell infiltration landscape in nonsquamous non–small cell lung cancer (NSCLC) has not been comprehensively characterized. In this study, we used consensus non-negative matrix factorization molecular subtyping to determine TME cell infiltration patterns and identified 3 TME clusters (TME-C1, -C2, -C3) characterized by distinct clinicopathologic features, infiltrating cells, and biological processes. Proteomics analyses revealed that cyclic GMP-AMP–stimulator of interferon genes immune signaling–mediated protein and phosphorylation levels were significantly upregulated in inflammation-related TME-C2 clusters. The score extracted from the TME-related signature (TMEsig-score) divided patients with NSCLC into high- and low-score subgroups, where a high score was associated with favorable prognosis and immune infiltration. The genomic landscape revealed that patients with low TMEsig-score harbored more somatic copy number alterations and higher mutation frequency of driver genes involving *STK11*, *KEAP1*, *SMARCA4*, and others. Drug sensitivity analyses suggested that tumors with high TMEsig-score were responsible for favorable clinical response to immune checkpoint inhibitor treatment. In summary, this study highlights that comprehensive recognizing of the TME cell infiltration landscape will contribute to enhancing our understanding of TME immune regulation and promote effectiveness of precision biotherapy strategies.

Conflict of interest: The authors have declared that no conflict of interest exists.

Submitted: July 1, 2021
Accepted: May 3, 2022
Published: June 22, 2022

Reference information: *JCI Insight*. 2022;7(12):e152815. https://doi.org/10.1172/jci.insight.152815.

Introduction

Nonsquamous non–small cell lung cancer (NSCLC) is the most common cancer worldwide and accounts for approximately 55%–60% of lung cancer deaths (1). Despite great advances in the treatment of nonsquamous NSCLC, the prognosis remains poor due to the presence of locally advanced or widely metastatic tumors in the majority of patients at the time of diagnosis (2). Thus, reliable markers that can precisely estimate clinical prognosis and therapeutic response would have tremendous value in guiding the management of lung cancer (3). The critical role of the tumor immune microenvironment on cancer development and metastasis has been recently recognized, suggesting that infiltration of various types of immune and stroma cells in tumor parenchyma might be a promising source of drug target and prognostic biomarkers (4). Indeed, assessment of the extent of tumor-infiltrating lymphocytes has been confirmed to be an important supplemental marker to the TNM staging system for relapse and mortality prediction (5). Therefore, understanding of the tumor microenvironment (TME) in NSCLC is critical and remains to be investigated.

A growing appreciation of the role of the TME in promoting lung carcinogenesis has also driven the development of anticancer therapies that target the TME (6). Recent immunotherapies targeting specific immune checkpoint molecules, such as programmed cell death 1 (PD-1) (e.g., pembrolizumab and nivolumab) or programmed cell death ligand 1 (PD-L1) (e.g., atezolizumab), have demonstrated a remarkable clinical benefit in NSCLC (7, 8). Moreover, drugs targeting other components of the TME, including transforming growth factor–β (TGF-β), vascular endothelial growth factor (VEGF), and aromatase, have already been used in clinical practice or trials (2, 4). Besides immune cells, tumors are also surrounded by vasculature, cancer-associated fibroblasts (CAFs), extracellular matrix (ECM), and more (9), which are considered to have a complex impact on tumorigenesis and immune surveillance. Therefore, understanding the heterotypic reciprocal crosstalk among cancer cells, various stromal cells, and ECM is an area of active research.

However, conventional methods of measuring the tumor immune and stromal cells, such as immunohistochemistry or flow cytometry, are not capable of comprehensively assessing the immune effects of different cell types or do not show effective discriminating power among closely related cell populations, which is largely due to the limitation in the number of immune markers that can be simultaneously measured with current techniques (10). As an alternative, continuously accumulating multiomics data provided an ideal resource for large-scale analysis of the TME landscape, and various computational approaches (e.g., xCell, CIBERSORT, MCP-counter) (11) have been developed and utilized to dissect the TME. Therefore, a comprehensive dissection of the heterogeneity and complexity of TME landscape would contribute to identifying different tumor immune phenotypes and guiding patients into personalized therapeutic regimens (12). Certain transcriptomic and genomic signatures, such as T cell–inflamed gene expression profile (GEP), ImmuneScore, aneuploidy, and tumor mutation burden (TMB), have also been associated with immune modulation and immunotherapy benefits (13). Additionally, phosphoproteomics provided abundant resources to investigate the potential immune modulation mechanism and therapeutic vulnerabilities in NSCLC. Therefore, promising biomarkers could be revealed, which will prove highly effective in recognizing patients' response to immunotherapy and will benefit the search for new therapeutic targets.

In this study, we systematically characterized the fractions of 64 immune and stromal cell types based on the xCell-annotated nonsquamous NSCLC immune profile. Three distinct TME cell–infiltrating patterns (termed as TME-C1, -C2, and -C3) with unsupervised consensus clustering were identified, and the distribution of cell subsets was corroborated by CIBERSORT and MCP-counter algorithms. The molecular characteristics and clinicopathologic features of these 3 TME clusters were closely linked to 3 previously reported immunophenotypes: immune excluded, inflamed, and desert (14, 15). Proteomics analyses revealed that cyclic GMP-AMP–stimulator of interferon genes (cCAMP-STING) immune signaling–mediated protein and phosphorylation levels were significantly augmented in the TME-C2 subtype. Moreover, we constructed a scoring scheme to quantify the immune infiltration level of individual tumors and guide patients' immune checkpoint inhibitor (ICI) therapy regimens. These findings will contribute to enhancing our cognition of TME infiltration characterization and developing more effective immunotherapy strategies for NSCLC.

Results

Cellular landscape and immune profile of TME cell infiltration patterns in NSCLC. We first summarized the workflow of our study design to illustrate the construction scheme of nonsquamous NSCLC TME cell–infiltrating patterns and TME signatures (Figure 1). The meta–National Center for Biotechnology Information Gene Expression Omnibus (meta-GEO) cohort of 681 NSCLC tumors with matched TME cell profiles were stratified into 3 distinct clusters by unsupervised hierarchical clustering analysis (Supplemental Figure 1, A and B; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.152815DS1). The 3 TME cluster pattern was dominated by different TME cell infiltration profiles, respectively termed TME-C1, TME-C2, and TME-C3 (Figure 2A). The 3D projection of tumors per TME clusters (subnetworks) based on uniform manifold approximation and projection (UMAP) unsupervised clustering corroborated the classification effect of non–negative matrix factorization (non-NMF) (Figure 2B). Based on the box plot of xCell-annotated cell enrichment score, most of the cell subpopulations had a significant difference among the 3 TME clusters (Supplemental Figure 1C). TME-C1 was characterized by upregulation in the infiltration of endothelial cells, hepatocytes, fibroblasts, monocytes, epithelial cells, and others; TME-C2 was distinguished by $CD8^+$ T cells, $CD4^+$ T cells, macrophage M1 cells, NK cells, and others; and TME-C3 was enriched for Th2 cells, multipotent progenitors, smooth muscle cells, basophils, and others (Figure 2A and Supplemental Figure 1C). Consistent with the above results, we also classified the TME cell profiles of The Cancer Genome Atlas lung adenocarcinoma (TCGA-LUAD) tumors into 3 clusters and noticed the similar distribution of cell subpopulations among the 3 TME clusters (Supplemental Figure 1D). We further used the ESTIMATE algorithm to quantify the overall immune infiltration (Estimate-ImmuneScore) and tumor cell purity (Estimate-tumor purity) among the 3 TME patterns. TME-C2 exhibited the highest immune scores, followed by TME-C1 and TME-C3 in meta-GEO and TCGA cohorts (Kruskal-Wallis H test, both $P < 0.001$, Figure 2C and Supplemental Figure 2A). Conversely, TME-C3 had a higher tumor purity than TME-C1 and TME-C2, suggesting that TME-C3 subtype tumors were surrounded by fewer nontumor components (e.g., immune cells and stromal cells) (Kruskal-Wallis H test, $P < 0.001$, Figure 2D). Moreover, we separately investigated the distribution of T cell–inflamed GEP score and PD-L1 (CD274) expression among the 3 TME clusters and found that TME-C2 exhibited the highest T cell–inflamed score and PD-L1 level compared with TME-C1 and TME-C3 (Kruskal-Wallis H test, meta-GEO cohort, Figure 2, E and F; TCGA-LUAD, Supplemental Figure 2, B and C; both $P < 0.001$). We also employed the CIBERSORT and MCP-counter algorithms to infer the microenvironment cell subsets and noticed a significant difference of immune and stroma cell distributions within the 3 TME clusters (Supplemental Figure 2, D and E).

Furthermore, prognosis analyses revealed significant differences between the 3 TME cell–infiltrating subgroups, in which TME-C2 exhibited a prominent survival advantage, whereas TME-C3 represented the worst prognosis in the meta-GEO cohort ($P < 0.001$, log-rank test, Figure 2G). Multivariate Cox proportional hazards regression analysis further demonstrated that the TME clustering model was associated with patients' survival outcome after adjusting for clinicopathologic factors in these 2 cohorts (meta-GEO cohort: TME-C1 vs. TME-C2, HR 0.66 [95% CI, 0.50 to 0.86], $P = 0.002$; Figure 2H). We also performed identical analyses in TCGA-LUAD cohort, and similar results were ascertained ($P = 0.016$, log-rank test, Figure 2I; TME-C1 vs. TME-C2, HR 0.69 [95% CI, 0.49 to 0.99], $P = 0.042$, Figure 2J).

TME infiltration patterns characterized by specific clinical features and molecular processes. Next, we explored the relationship between TME clustering patterns and conventional clinical characteristics in nonsquamous NSCLC (Figure 3A). The TCGA genomic analysis revealed an overall molecular characterization of LUAD and suggested LUAD was classified into 3 transcriptional subtypes (proximal-inflammatory [PI], proximal-proliferative [PP], and terminal respiratory unit [TRU]) or 3 methylation subtypes (significantly altered CpG island methylator phenotype-high [CIMP-high], CIMP-low, CIMP-intermediate) (16). In this study, TME-C2 tumors were predominantly clustered into prognostically favorable subtypes of TRU and CIMP-intermediate, whereas TME-C3 tumors were mainly concentrated within the worse outcome subtypes of PP and CIMP-low (Supplemental Figure 3, A and B). Cigarette smoking–caused genomic instability is a common factor that contributes to TME remodeling. We observed that the proportion of current smokers in the TME-C2 subtype was significantly higher than other subtypes (Supplemental Figure 3C), which suggested that tobacco smoking may promote the formation of a proinflammatory TME. Emerging evidence has shown that oncogenic driver mutations (e.g., *TP53*, *KRAS*, *KEAP1*, *STK11*, and *EGFR*) are associated with discrete immune phenotypes in LUADs (17, 18). We also investigated the association between gene mutation and TME clusters and noticed that *KEAP1*- and *STK11*-mutated tumors exhibited a significantly lower proportion of TME-C2 cluster than wild-type tumors (Figure 3B). However, marginal differences were observed in terms of *TP53*, *EGFR*, and *KRAS* mutations (Supplemental Figure 3D). Recent immunogenomic analysis also identified 6 immune subtypes across cancer types (19). Subsequent studies showed that the identified immune subtypes of macrophage regulation, lymphocyte infiltration signature, and IFN-γ response were markedly elevated in the TME-C2 subgroup, yet wound healing and proliferation were enhanced in TME-C1 and TME-C3 subgroups (Supplemental Figure 3E). We further compared the somatic copy

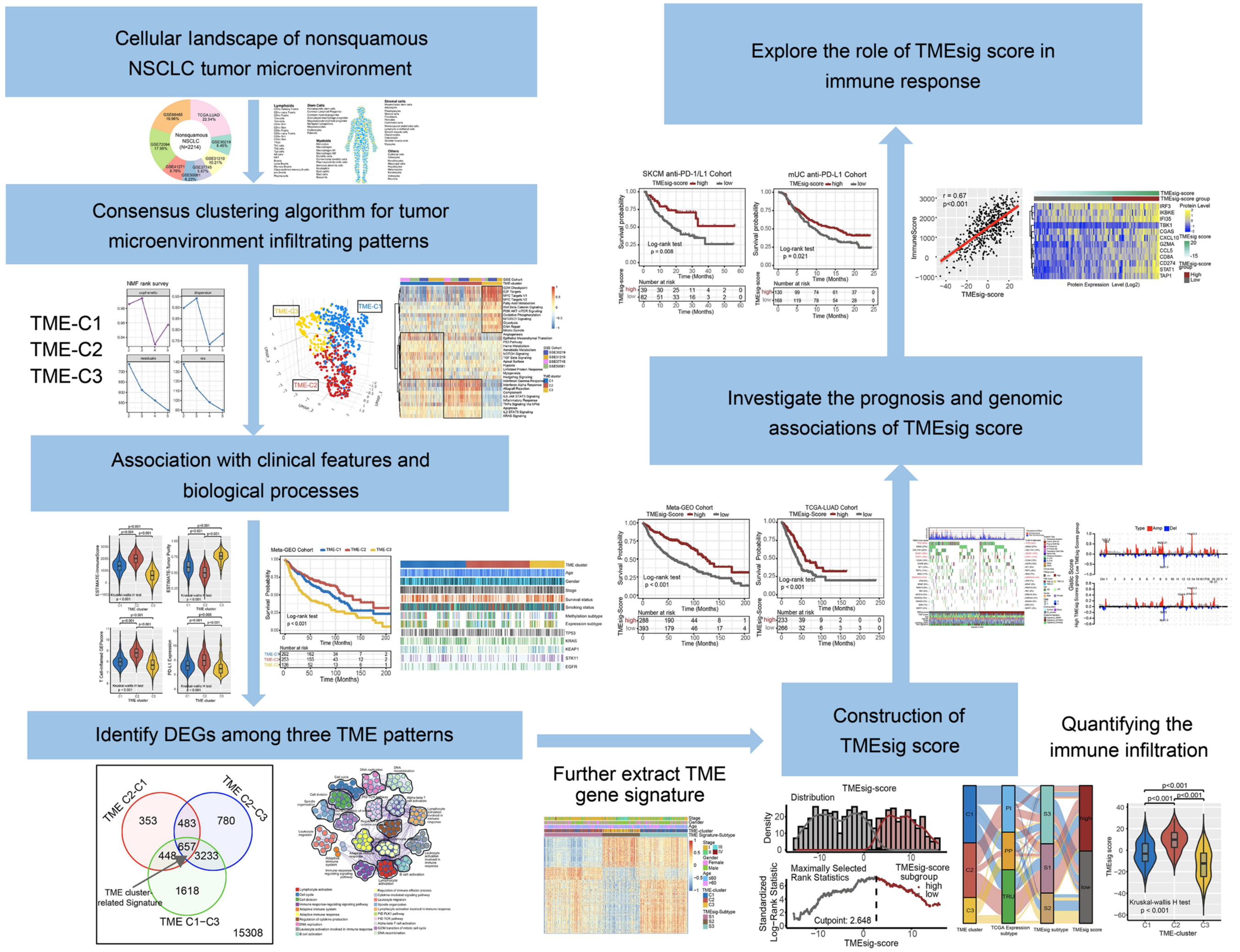


Figure 1. The overview of study design to evaluate the construction scheme of TME cell–infiltrating patterns and TMEsig-score in nonsquamous NSCLC.

number alterations (SCNAs) and aneuploidy score with TME clusters and observed a lower SCNA and aneuploidy level in TME-C2 (Figure 3, C and D), consistent with the previous conclusions that SCNA and aneuploidy score correlated positively with immune evasion and tumor cell proliferation (20).

To explore the biological molecular changes underlying 3 distinct TME phenotypes, we performed gene set variation analysis (GSVA) against the Hallmark gene set in NSCLC tumors (Figure 3E and Supplemental Figure 3F). GSVA results indicated that TME-C2 was significantly enriched in immune infiltration–related circuits, including IFN-γ/α response, allograft rejection, IL-6/JAK/STAT3 signaling, inflammatory response, and so on; TME-C1 was highly enriched in stroma-related signaling pathways, including TGF-β pathway, epithelial-mesenchymal transition (EMT), P53 pathway, and so on; TME-C3 presented enrichment pathways prominently associated with carcinogenic activation and highly proliferative features, such as fatty acid metabolism, PI3K/AKT/mTOR, G2M checkpoint, E2F targets, MYC target V1/V2, mitotic spindle, and so on. Interestingly, moderate immune activation and immune cell infiltration were also observed in the TME-C1 cluster and consistent with the previously reported immunophenotype of immune excluded, in which immune cells are retained in the stroma surrounding tumor cell nests rather than penetrating their parenchyma (14, 15). Subsequent analyses by using the Mariathasan et al. curated signaling signature set (21) demonstrated that stroma and angiogenesis activity were significantly enhanced in TME-C1 while immune activation was markedly elevated in TME-C2, which confirmed our speculation (Supplemental Figure 3G).

We further investigated the enriched pathways between TME-C2 and TME-C3 against a REACTOME gene set with gene set enrichment analysis (GSEA). Strikingly, STING-mediated innate immune response to cytosolic DNA was significantly enriched in TME-C2 clusters (Figure 3F).

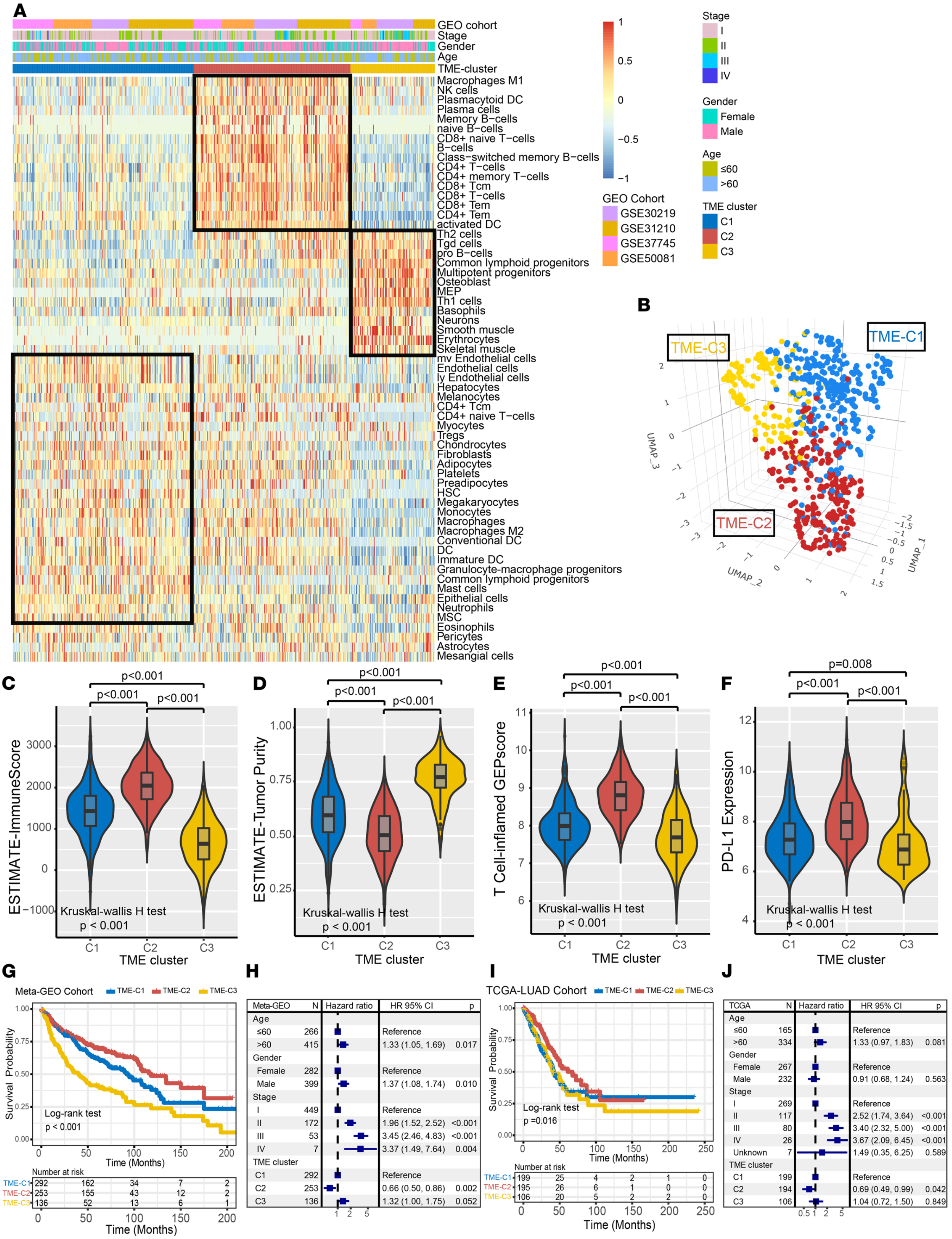

A
GEO cohort
Stage
Gender
Age
TME-cluster
Macrophages M1
NK cells
Plasmacytoid DC
Plasma cells
Memory B−cells
naive B−cells
CD8+ naive T−cells
B−cells
Class−switched memory B−cells
CD4+ T−cells
CD4+ memory T−cells
CD8+ Tcm
CD8+ T−cells
CD8+ Tem
CD4+ Tem
activated DC
Th2 cells
Tgd cells
pro B−cells
Common lymphoid progenitors
Multipotent progenitors
Osteoblast
MEP
Th1 cells
Basophils
Neurons
Smooth muscle
Erythrocytes
Skeletal muscle
mv Endothelial cells
Endothelial cells
ly Endothelial cells
Hepatocytes
Melanocytes
CD4+ Tcm
CD4+ naive T−cells
Myocytes
Tregs
Chondrocytes
Fibroblasts
Adipocytes
Platelets
Preadipocytes
HSC
Megakaryocytes
Monocytes
Macrophages
Macrophages M2
Conventional DC
DC
Immature DC
Granulocyte-macrophage progenitors
Common lymphoid progenitors
Mast cells
Epithelial cells
Neutrophils
MSC
Eosinophils
Pericytes
Astrocytes
Mesangial cells
1
0.5
0
−0.5
−1
GEO Cohort
GSE30219
GSE31210
GSE37745
GSE50081
Stage
I
II
III
IV
Gender
Female
Male
Age
≤60
>60
TME cluster
C1
C2
C3
B
TME-C3
TME-C1
TME-C2
UMAP_3
UMAP_2
UMAP_1
C
p<0.001
p<0.001
p<0.001
ESTIMATE-ImmuneScore
Kruskal-wallis H test
p < 0.001
C1 C2 C3
TME cluster
D
p<0.001
p<0.001
p<0.001
ESTIMATE-Tumor Purity
Kruskal-wallis H test
p < 0.001
E
p<0.001
p<0.001
p<0.001
T Cell-inflamed GEPscore
Kruskal-wallis H test
p < 0.001
F
p=0.008
p<0.001
p<0.001
PD-L1 Expression
Kruskal-wallis H test
p < 0.001
G
Meta-GEO Cohort
TME-C1 TME-C2 TME-C3
Survival Probability
Log-rank test
p < 0.001
Time (Months)
Number at risk
TME-C1 292 162 34 7 2
TME-C2 253 155 43 12 2
TME-C3 136 52 13 6 1
H
Meta-GEO | N | Hazard ratio | HR 95% CI | p
Age
≤60 266 Reference
>60 415 1.33 (1.05, 1.69) 0.017
Gender
Female 282 Reference
Male 399 1.37 (1.08, 1.74) 0.010
Stage
I 449 Reference
II 172 1.96 (1.52, 2.52) <0.001
III 53 3.45 (2.46, 4.83) <0.001
IV 7 3.37 (1.49, 7.64) 0.004
TME cluster
C1 292 Reference
C2 253 0.66 (0.50, 0.86) 0.002
C3 136 1.32 (1.00, 1.75) 0.052
I
TCGA-LUAD Cohort
TME-C1 TME-C2 TME-C3
Survival Probability
Log-rank test
p =0.016
Time (Months)
Number at risk
TME-C1 199 25 4 2 1 0
TME-C2 195 26 6 1 0 0
TME-C3 106 20 5 2 2 0
J
TCGA | N | Hazard ratio | HR 95% CI | p
Age
≤60 165 Reference
>60 334 1.33 (0.97, 1.83) 0.081
Gender
Female 267 Reference
Male 232 0.91 (0.68, 1.24) 0.563
Stage
I 269 Reference
II 117 2.52 (1.74, 3.64) <0.001
III 80 3.40 (2.32, 5.00) <0.001
IV 26 3.67 (2.09, 6.45) <0.001
Unknown 7 1.49 (0.35, 6.25) 0.589
TME cluster
C1 199 Reference
C2 194 0.69 (0.49, 0.99) 0.042
C3 106 1.04 (0.72, 1.50) 0.849

Figure 2. Association of TME cell–infiltrating patterns with immune signature and prognosis in NSCLC. (**A**) Unsupervised clustering of TME cell landscape for 681 lung cancer patients in meta-GEO cohort. Stage, sex, age, GEO database cohort, and TME cluster are shown as patient annotations. The numbers of patients with TME-C1, TME-C2, and TME-C3 phenotypes are 292, 253, and 136, respectively. (**B**) 3D UMAP projection of NSCLC tumors per TME subtype (subnetworks) based on unsupervised clustering. (**C**–**F**) Indicator of immune infiltration level was compared by Kruskal-Wallis test among 3 TME clusters, including immune score (**C**), tumor purity (**D**), T cell immune GEP score (**E**), and PD-L1 expression (**F**). When $P < 0.05$, pairwise comparisons were made using Dunn's test with Benjamini-Hochberg adjustment for multiple comparisons. (**G**) Kaplan-Meier curves for overall survival of 3 TME phenotypes in NSCLC meta-GEO cohort. (Log-rank test, $P < 0.001$.) (**H**) Forest plot representation of multivariate Cox model depicted association between TME clusters and overall survival (OS) after being adjusted for age, sex, and stage. Square data markers indicate estimated hazard ratios (HRs) and the length of the horizontal line represented the 95% confidence interval for each variable. (**I**) Kaplan-Meier curves for OS of 3 TME clusters in TCGA-LUAD cohort; log-rank test, $P = 0.016$. The numbers of patients in TME-C1, TME-C2, and TME-C3 clusters are 199, 195, and 106, respectively. (**J**) Forest plot representation of multivariate Cox model–calculated association between OS and TME cluster with other clinical factors taken into account.

Considering the activation of cGAS-STING pathway resulted in *TBK1* and *IRF3* protein phosphorylation and further induced the production of IFNs and inflammatory cytokines (*CXCL10* and *CCL5*), we downloaded and curated the Clinical Proteomic Tumor Analysis Consortium LUAD (CPTAC-LUAD) multiomics data to validate the potential regulation mechanism. Inferred TME cell infiltration patterns based on the NMF algorithm were compared with CPTAC-LUAD–defined immune clusters (hot tumor enriched [HTE], cold tumor enriched [CTE], and normal adjacent tissue–enriched [NAT-enriched]) and noticed that TME-C2 was significantly dominated by HTE, whereas TME-C1 and -C3 were mainly clustered in CTE ($P < 0.001$, Figure 3G). Similarly, TME-C2 exhibited the highest Estimate-ImmuneScore, followed by TME-C1 and TME-C3 in the CPTAC-LUAD cohort (Kruskal-Wallis H test, $P < 0.001$, Supplemental Figure 3H). In order to comprehensively characterize the cGAS-STING milieu across 3 clusters, we analyzed the expression of crucial STING-related mRNAs and proteins (including *TMEM173*, *cGAS*, *TBK1*, *IRF3*, *NFKB1*, *NFKB2*, *RELB*, *CCL5*, *CCL10*, and *STAT1*) in the CPTAC samples. A significant elevation of STING-related mRNA and protein levels was observed in the TME-C2 cluster (Figure 3H and Supplemental Figure 3I). We further compared the *IRF3* and *NFKBIE* protein phosphorylation level between the 3 TME clusters and noticed a pronounced upregulation in the TME-C2 cluster (Figure 3, I and J).

Based on these findings, we speculated that the identified 3 TME cell infiltration patterns were characterized by distinct immune phenotypes. TME-C2 was recognized as an immune-inflamed phenotype and characterized by immune pathway activation and abundant immune cell infiltration, as well as downregulation of SCNAs and aneuploidy. TME-C3 was recognized as an immune-desert phenotype and represented by tumor cell proliferation, glycolysis, and suppression of immunity, coupled with a higher mutation rate of *KEAP1* and *STK11*. TME-C1 was recognized as an immune-excluded phenotype and distinguished by stromal activation and exclusion of immune cell infiltration.

TME phenotype-related DEGs in NSCLC. To identify the underlying genetic alterations and expression perturbations of each TME phenotype, we applied the empirical Bayes algorithm to determine the overlapping DEGs between the 3 TME phenotypes. A total of 657 DEGs that represented the critical distinguishing index of the 3 TME patterns were considered TME signature genes and illustrated in a Venn diagram (Figure 4A). Based on the representative TME phenotype-related signature genes, we performed unsupervised consensus clustering analysis and obtained 3 stable TME transcriptomic subtypes. These stratifications divided NSCLC patients into 3 distinct TME gene signature subgroups characterized by distinct clinicopathologic features, which were defined as TMEsig-S1, TMEsig-S2, and TMEsig-S3, respectively (meta-GEO cohort, Figure 4B; TCGA cohort, Supplemental Figure 4A). We found that patients in the TMEsig-S2 subgroup were dominated by a higher proportion of male sex and advanced tumor stage, and patients with TME-C3 infiltration pattern were mainly concentrated in the TMEsig-S2 subgroup. Gene ontology (GO) enrichment and Metascape analyses of these signature genes revealed that immune-related biological processes were significantly overrepresented (Figure 4, C and D). Further prognosis analysis indicated that the survival outcomes of the 3 TME signature subtypes were significantly different in NSCLC samples. The TMEsig-S1 signature was proved to be associated with better prognosis, while TMEsig-S2 was associated with a worse survival outcome (meta-GEO cohort: $P < 0.001$; TCGA-LUAD cohort: $P = 0.001$, log-rank test, Figure 4, E and F). Similarly, the TME signature subtype could stratify patients into subgroups with significantly different prognoses in the independent validation cohorts (GEO GSE72094, GSE68465, GSE41271, all $P < 0.05$; Supplemental Figure 4, B–D). The association between the TME gene signatures and survival remained statistically significant after taking into account age, sex,

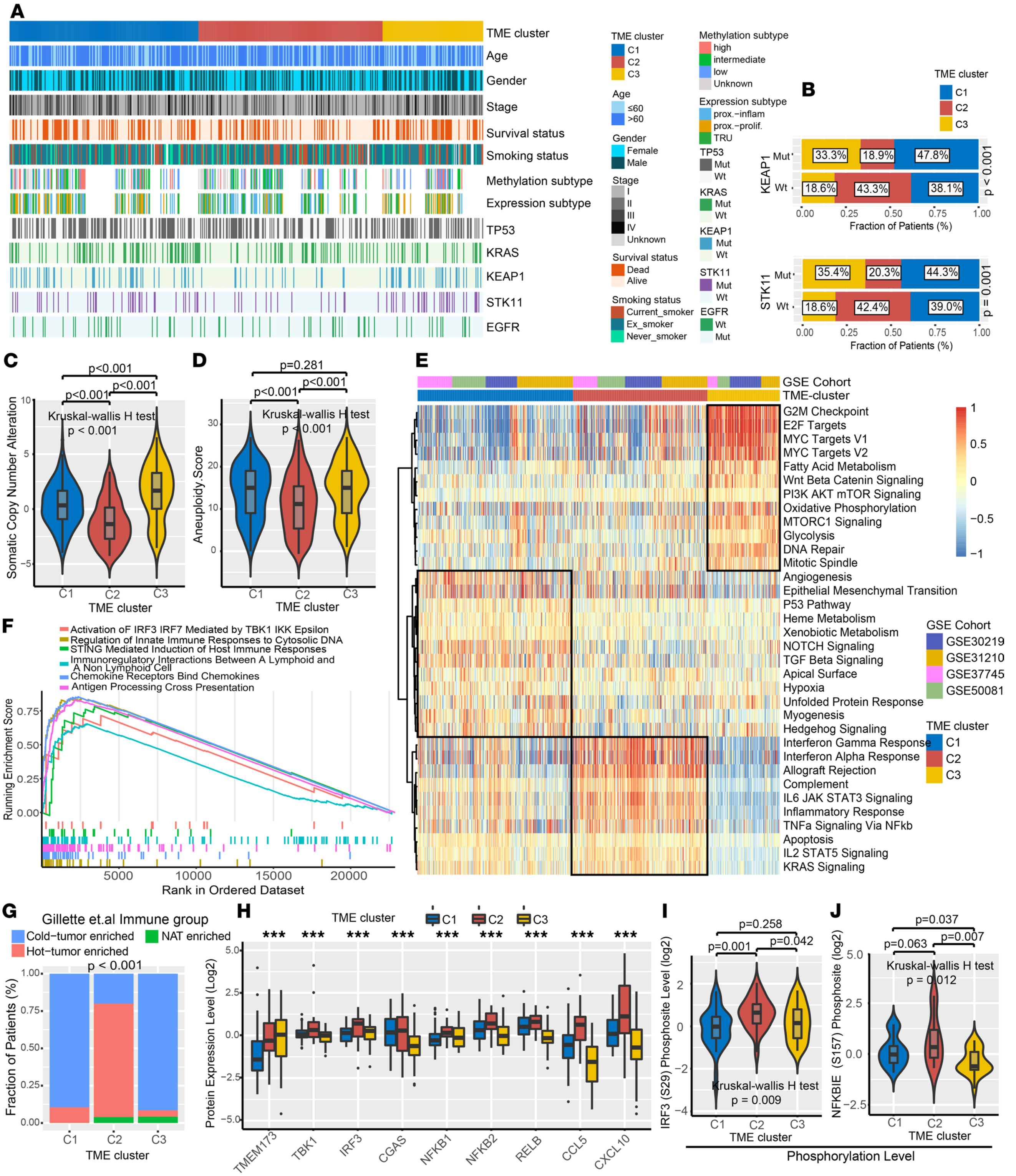


Figure 3. Clinical characteristics and biological processes among 3 TME cell infiltration patterns. (**A**) An overview of the association between TME clusters and nonsquamous NSCLC clinical and molecular characteristics in TCGA cohort. (**B**) The mutational proportion of *STK11* and *KEAP1* among 3 TME clusters (Fisher's exact test). (**C** and **D**) SCNA (**C**) and aneuploidy score (**D**) was compared between 3 TME clusters (Kruskal-Wallis *H* test followed by Dunn's test for pairwise comparisons). (**E**) Heatmap shows the top enriched biological pathways calculated by GSVA algorithm in distinct TME phenotypes. Hallmark gene set (h.all.v7.0) curated from Molecular Signatures Database (MSigDB) was regarded as the reference gene signatures. (**F**) GSEA plots showing the cGAS-STING-IRF3–mediated gene sets were enriched in TME-C2. (**G**) Comparison of TME clusters with CPTAC annotated immune groups in CPTAC-LUAD proteomics cohort (Fisher's exact test). (**H**) Comparison of the protein expression levels of STING pathway-related molecules among 3 TME clusters in CPTAC cohort. (**I** and **J**) Violin plot shows the phospho-IRF3 (S157) and phospho-NFKBIE (S157) protein levels among different TME clusters. The bottom and top of the boxes were the 25th and 75th percentiles (interquartile range). The whiskers encompassed

1.5 times the interquartile range. The statistical difference of 3 gene clusters was compared through the Kruskal-Wallis *H* test. $^{***}P < 0.001$. PI, proximal-inflamed; PP, proximal-proliferation; TRU, terminal respiratory unit.

and tumor stage in the meta-GEO cohort (Cox proportional hazards model, TMEsig-S2 vs. TMEsig-S3, HR 0.78 [95% CI, 0.60 to 1.00], $P = 0.054$; TMEsig-S2 vs. TMEsig-S1, HR 0.48 [95% CI, 0.36 to 0.65], $P < 0.001$; Supplemental Figure 4E). Consistent results were obtained in TCGA-LUAD cohort (Cox model, TMEsig-S2 vs. TMEsig-S3, HR 0.64 [95% CI, 0.44 to 0.93], $P = 0.019$; TMEsig-S2 vs. TMEsig-S1, HR 0.49 [95% CI, 0.33 to 0.73], $P < 0.001$; Supplemental Figure 4F).

The immune cell subsets and gene sets were also compared among the 3 TME gene signature subtypes. Here, we utilized Charoentong et al.'s (22) curated 28 representative immune cells to explore the immune infiltration level. We observed a significant elevation in cytotoxic T lymphocytes, B cells, and dendritic cells yet a reduction of neutrophils, regulatory T cells, and Th cells in the TMEsig-S2 subgroup (Supplemental Figure 4G). Then we analyzed the expression of inflammatory, stroma-related mRNAs in the NSCLC samples to explore the relationship between the 3 gene clusters and the molecular perturbation milieu (Supplemental Figure 4H). *CXCL9*, *GZMA*, *PRF1*, *CD8A*, *TNF*, *PDCD1*, *LAG3*, and *CTLA4* were considered immune-activated-related transcripts; *TGFB1*, *ACTA2*, *COL4A1*, *TWIST1*, *ADAM12*, *FSTL3*, *SMAD9*, and *TPM1* were considered to be TGF-β/EMT pathway-relevant transcripts; *CCNE1*, *RFC3*, *MKI67*, *POLD2*, *LIG1*, *BRCA1*, *FANCA*, *FANCD2*, *CDK2*, and *POLE* were considered to be cell cycle/proliferation–related transcripts. We found the mRNAs relevant to the immune-activated pathway were substantially upregulated in TME-S1, which suggested this subtype was deemed as an immune-activated group, while TME-S2 and TME-S3 showed high expression of mRNAs related to stroma-activated and cell proliferation–related transcripts.

Construction of the TMEsig-score and exploration of its clinical and molecular relevance. The aforementioned results validated the role of TME phenotype in regulation of prognosis and immune infiltration, but these analyses were only based on patient populations and cannot accurately predict the pattern of TME phenotype in individual tumors. Therefore, we developed a scoring scheme termed TMEsig-score (Methods), which based on specified TME signature genes related to prognosis, quantifies the TME immune infiltration of individuals with NSCLC. Multivariate Cox regression model analysis considering patient age, sex, and tumor stage showed the TMEsig-score served as an independent prognostic biomarker for evaluating patient outcomes in NSCLC (meta-GEO cohort, HR 0.98 [95% CI, 0.97 to 0.99], $P < 0.001$; TCGA-LUAD cohort, HR 0.98 [95% CI, 0.96 to 0.99], $P < 0.001$, Supplemental Figure 5, A and B). Considering the stability of the developed scoring system, we compared the prognostic value of our TMEsig-score with previous identified risk score models (23–26) in lung cancer: the Harrell concordance index (C-index) for the overall survival of TMEsig-score was 0.696 [95% CI 0.657–0.734] and marginally larger than the other 4 score models (risk model by Li et al. 0.650 [0.611–0.690], risk score by Zhao et al. 0.630 [0.590–0.678], immune risk score by Zhang et al. 0.609 [0.569–0.649], immune signature by Song et al. 0.627 [0.586−0.668]; Supplemental Figure 5C). We further stratified the nonsquamous NSCLC tumors into high versus low TMEsig-score subgroup based on the optimal cutoff point of the meta-GEO cohort (Methods, Figure 5A). Kaplan-Meier plot showed that the high TMEsig-score subgroup was significantly associated with a better prognosis (HR, 0.44 [95% CI, 0.35 to 0.57], $P < 0.001$, Figure 5B). This cutoff value was also employed in TCGA-LUAD cohort and revealed a similar result (HR, 0.54 [95% CI, 0.40 to 0.73], $P < 0.001$, log-rank test, Figure 5C).

Furthermore, we also compared the TMEsig-score with TME cell infiltration patterns and TMEsig subtype and noticed that TME-C3 and TMEsig-S2 exhibited a lower TMEsig-score, whereas TME-C2 and TMEsig-S1 were linked to a higher TMEsig-score (Figure 5D and Supplemental Figure 5D). We also explored the relationship between the TMEsig-score and TCGA molecular subtype to find that the TRU subtype was linked to the highest TMEsig-score, and the PP subtype was linked to a lower TMEsig-score than the PI subtype (Figure 5E). An alluvial graph was used to further show the dynamic change of specimens' flow in different TME phenotypes (Supplemental Figure 5E). Next, we examined the relationship between ESTIMATE algorithm–inferred overall immune infiltration (ImmuneScore) and the TMEsig-score through Spearman analysis to find a significantly positive correlation between the TMEsig-score and the ImmuneScore ($r = 0.67$, $P < 0.001$, Figure 5F). Meanwhile, it was found that the high TMEsig-score subgroup had a higher level of immune subtype of lymphocyte infiltration signature, macrophage regulation, and IFN-γ response but lower level of proliferation and wound healing compared with patients with a low TMEsig-score (Supplemental Figure 5F).

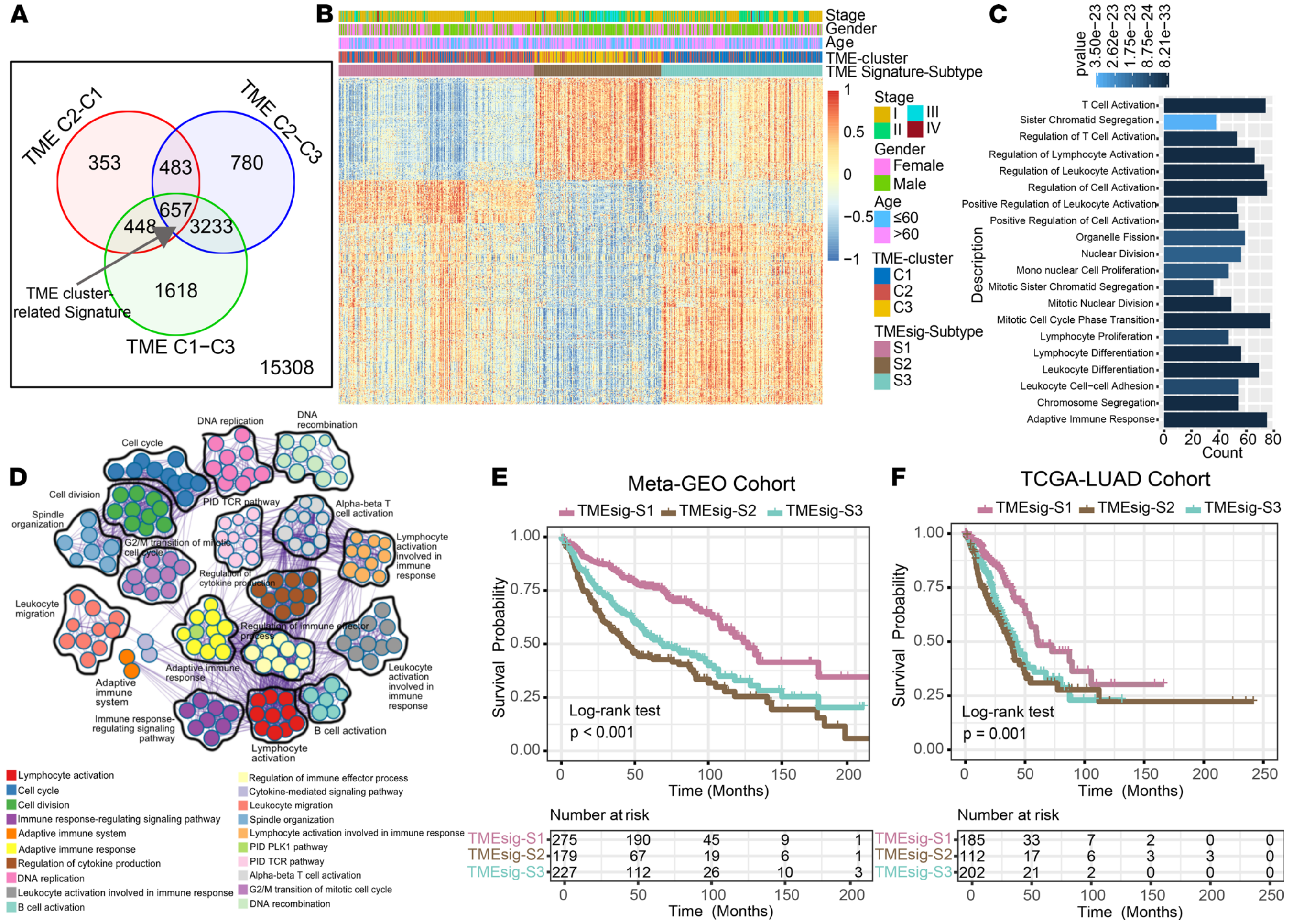


Figure 4. TME phenotype-related DEGs in nonsquamous NSCLC. (**A**) A total of 657 overlapped differential expressed genes of the 3 TME clusters were recognized as the TME phenotype-related gene signature and shown in a Venn diagram. (**B**) Unsupervised clustering of TME phenotype-related gene signatures to classify patients into different transcriptomic subtypes, termed as TME gene S1–S3, respectively. The stage, sex, age, TME clusters, and TME signature subtype were used as patient annotations. (**C**) Functional annotation for TME phenotype-related genes using GO enrichment analysis. The color depth of the bar plots represented the statistical significance of enriched pathways. (**D**) Metascape enrichment network visualization showed the intracluster and intercluster similarities of enriched terms, up to 20 terms per cluster. Cluster annotations are shown in the color key. (**E** and **F**) The survival curves of the TME gene signature subtypes were estimated by the Kaplan-Meier plotter in Meta-GEO and TCGA-LUAD cohort (meta-GEO cohort, $P < 0.001$; TCGA-LUAD cohort, $P = 0.001$; log-rank test).

We further constructed the TMEsig-score model in the CPTAC cohort and found that the majority of HTE tumors were concentrated in the high TMEsig-score subgroup (χ^2 test, $P < 0.001$, Figure 5G). ImmuneScore was significantly upregulated in the high TMEsig-Score group likewise (Supplemental Figure 5G). Aforementioned results revealed the enrichment of STING-mediated innate immune response in TME-C2. We also investigated the association between TMEsig-score and cGAS-STING–related proteins in the CPTAC-LUAD cohort. A protein expression heatmap illustrated the STING-related molecules (*IRF3*, *TBK1*, *cGAS*), inflammatory cytokines (*CCL5*, *CXCL10*, *CD8A*, *GZMA*), and immune checkpoint molecules (*CD274*, *TAP1*), and others were significantly positively correlated with TMEsig-score (all Spearman $r > 0.25$, $P < 0.001$, Figure 5H). We further collected the expression profile of nonsquamous NSCLC cell lines ($n = 100$) from the Cancer Cell Line Encyclopedia (CCLE) project (27) and divided them into 2 TMEsig-score subgroups (Supplemental Table 1). The 10 highest and lowest TMEsig-score cell lines were separately illustrated in Figure 5I. We selected the top-ranked and ATCC-sourced cell lines (high TMEsig-score: NCI-H23, NCI-H1299, NCI-H358, NCI-H838; low TMEsig-score: Calu-3, NCI-H1395, NCI-H2126) and performed Western blot analysis on cGAS-STING pathway–related molecules. Cell lines of high TMEsig-score subtype showed an obviously upregulated expression of STING signaling molecules (*cGAS*, *STING*, phosphorylated [p-] *STING*, *TBK1*, NF-κB, p–NF-κB, *IRF3*, *CCL5*, and *CXCL10*) than low TMEsig-score subgroup (Figure 5, J and K; see complete unedited blots in the supplemental material).

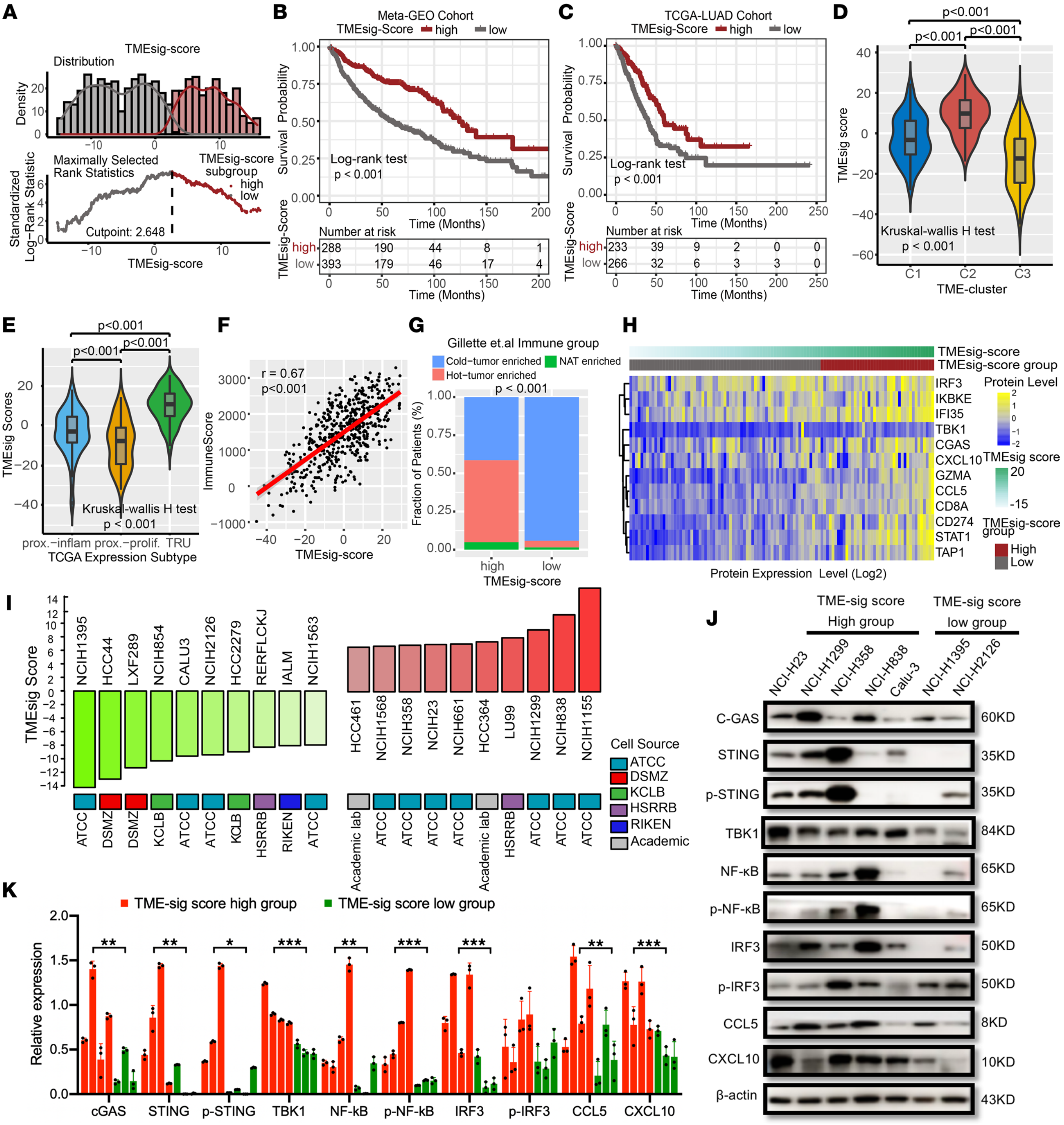


Figure 5. Construction of the TME score and exploration of its biological relevance. (**A**) Identification of the cutoff point of the TMEsig-score subgroup in nonsquamous NSCLC. TMEsig-score with the highest standardized log-rank statistics was regarded as the optimal cutoff point. (**B**) Kaplan-Meier curves for high versus low TMEsig-score subgroups in meta-GEO cohort (log-rank test, $P < 0.001$). (**C**) Kaplan-Meier curves for high versus low TMEsig-score patient subgroups in TCGA-LUAD cohort (log-rank test, $P < 0.001$). (**D** and **E**) Violin plot showing the distribution of TMEsig-score in different TME clusters (**D**) and TCGA expression subtypes (**E**). The differences between the 3 groups were compared through the Kruskal-Wallis test (both $P < 0.001$). (**F**) Spearman's correlation between TMEsig-score and ImmuneScore in TCGA-LUAD cohort ($r = 0.67$). (**G**) Comparison of TMEsig-score group with CPTAC-annotated immune group in CPTAC-LUAD proteomics cohort (Fisher's exact test). (**H**) Heatmap shows correlation of TMEsig-score with STING-related molecules (IRF3, TBK1, cGAS), inflammatory cytokines (CCL5, CXCL10, CD8A, GZMA), and immune checkpoint molecules (CD274, TAP1) in the CPTAC-LUAD cohort. (**I**) The nonsquamous NSCLC cell subsets of 10 highest and lowest TMEsig-score in CCLE data set. (**J**) Western blot analyses of cGAS-STING pathway–related molecules in selected ATCC-sourced cell lines of different TMEsig-score subgroups. (**K**) Comparison of relative level of cGAS-STING pathway–related molecules in different TMEsig-score cell subsets by Western blot. Data represented with mean ± SD. The differences between the 2 groups were compared through the Student's *t* test ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

Genomic alterations between high and low TMEsig-score subgroups. Mounting evidence has demonstrated that there is a correlation between tumor genomic mutations and responsiveness to immunotherapy (28, 29). Therefore, to gain further insights into the mutational processes operative in nonsquamous NSCLC tumors, we extracted 3 mutational signatures with varying numbers of somatic mutations from the mutational profiles (Figure 6A). The extracted mutational signatures included *APOBEC* somatic mutations (signature 2, 30888 of 208063), smoking-related mutations (signature 4, 158725 of 208063), and DNA mismatch repair deficiency-related (dMMR) mutations (signature 6, 18450 of 208063). Samples with high TMEsig-score had a higher frequency of signature 6 and a lower frequency of signature 2 (P = 0.003, χ^2 test, Figure 6B). We next investigated the distributions of somatic alterations and performed significantly mutated gene (SMG) analysis for nonsquamous NSCLC samples between the TMEsig high-score and low-score subgroups. The mutational landscapes showed that *TP53* (60% vs. 34%), *KEAP1* (25% vs. 10%), *STK11* (19% vs. 11%), and *SMARCA4* (13% vs. 4%) had higher somatic mutation rates in the TMEsig low-score subgroup, compared with high-score subgroup (Fisher's exact test, $P < 0.05$, Figure 6C and Supplemental Figure 5H). Consequently, we also noticed a higher level of SCNAs in the TME low-score subgroup (Figure 6D). We further analyzed the Genomic Identification of Significant Targets in Cancer–derived (GISTIC-derived) scores and copy number gain/loss frequencies in the high and low TMEsig-score subtypes. The results showed that the low TMEsig-score subgroup had higher GISTIC scores and specific copy number variation (CNV) regions than the high-score subgroup (1q21.3, 1q22, and 8q24.21 in low subgroup; 9p23, 12q14.1, and 14q13.1 in high subgroup; Figure 6E). Then, we compared the differences in fraction genome altered (FGA) and segments among different subtypes and observed a significant augmentation in low TMEsig-score group (Wilcoxon rank-sum test, $P < 0.001$, Figure 6, F and G). Above results enabled us to more comprehensively assess the association between the TMEsig-score and genomic variation and facilitated revealing the underlying complex relationship between individual somatic mutations and immune regulation.

Identification of potential immunotherapy response of TMEsig-score. ICI treatment represented by PD-1/PD-L1 inhibitors has undoubtedly made a major breakthrough in NSCLC antitumor therapy. Recent evidence demonstrated that newly identified molecular markers (such as Tumor Immune Dysfunction and Exclusion [TIDE] and T cell–inflamed GEP) are widely and strongly recommended to evaluate the response to anti–PD-1/PD-L1 immunotherapy. We found that the TIDE was significantly decreased in the TMEsig high-score group, and T cell–inflamed GEP was significantly elevated in the high TMEsig-score subgroup (TIDE and T cell–inflamed GEP distribution in meta-GEO and TCGA-LUAD, all $P < 0.001$, Figure 7, A–D). Due to the strong connection of the TME signatures with the immune response, we further investigated whether the TMEsig-score can predict patients' response to ICI therapy in 2 independent immunotherapy cohorts. In the anti–PD-1 cohort, the significant therapeutic benefits and immune response were confirmed in patients with high TMEsig-score compared with those with low TMEsig-score (HR 0.44 [95% CI, 0.24 to 0.80], P = 0.008, Figure 7E; response rate: 51.3% vs. 32.9%, Figure 7F). Consistent results were also validated in the anti–PD-L1 cohort (HR 0.71 [95% CI, 0.53 to 0.95], P = 0.021, Figure 7G; response rate: 31.5% vs. 16.1%, Figure 7H). The distribution of TMEsig-score in groups with different immunotherapy clinical response is shown in Figure 7I, and it was found that patients with higher TMEsig-score were more likely to benefit from immunotherapy (Kruskal-Wallis H test, $P < 0.001$). Tumor tissue is genomically heterogeneous, with varying tumor mutation burden (TMB), and presents a correlation with durable clinical response to anti–PD-1/PD-L1 immunotherapy (30, 31). A significant positive correlation between TMEsig-score and TMB was identified among the anti–PD-L1 data set (Spearman r = 0.24, $P < 0.001$, Figure 7J). Therefore, we divided the overall population into 4 subgroups according to the TMEsig-score and TMB distribution, TMEsig-score-H+TMB-H, TMEsig-score-H+TMB-L, TMEsig-score-L+TMB-H, and TMEsig-score-L+TMB-L. Finally, the TMEsig-score-H+TMB-H subgroup exhibited the best survival compared with the other 3 subgroups (log-rank test, P = 0.008, Figure 7K).

TMEsig-score model in clinical practice. We further validated the relationship between TMEsig-score and prognosis in 3 independent cohorts of lung cancer (GEO GSE72094, GSE68465, and GSE41271). TMEsig-score was significantly elevated in TMEsig-S1 and -S3 subtypes (Supplemental Figure 6, A–C). Patients with a higher TMEsig-score exhibited significant clinical benefits and survival advantage (GSE72094, HR 0.40 [95% CI, 0.27 to 0.60], $P < 0.001$, Supplemental Figure 6D; GSE68465, HR 0.63 [95% CI, 0.48 to 0.83], P = 0.001, Supplemental Figure 6E; GSE41271, HR 0.41 [95% CI, 0.26 to 0.64], $P < 0.001$, Supplemental Figure 6F).

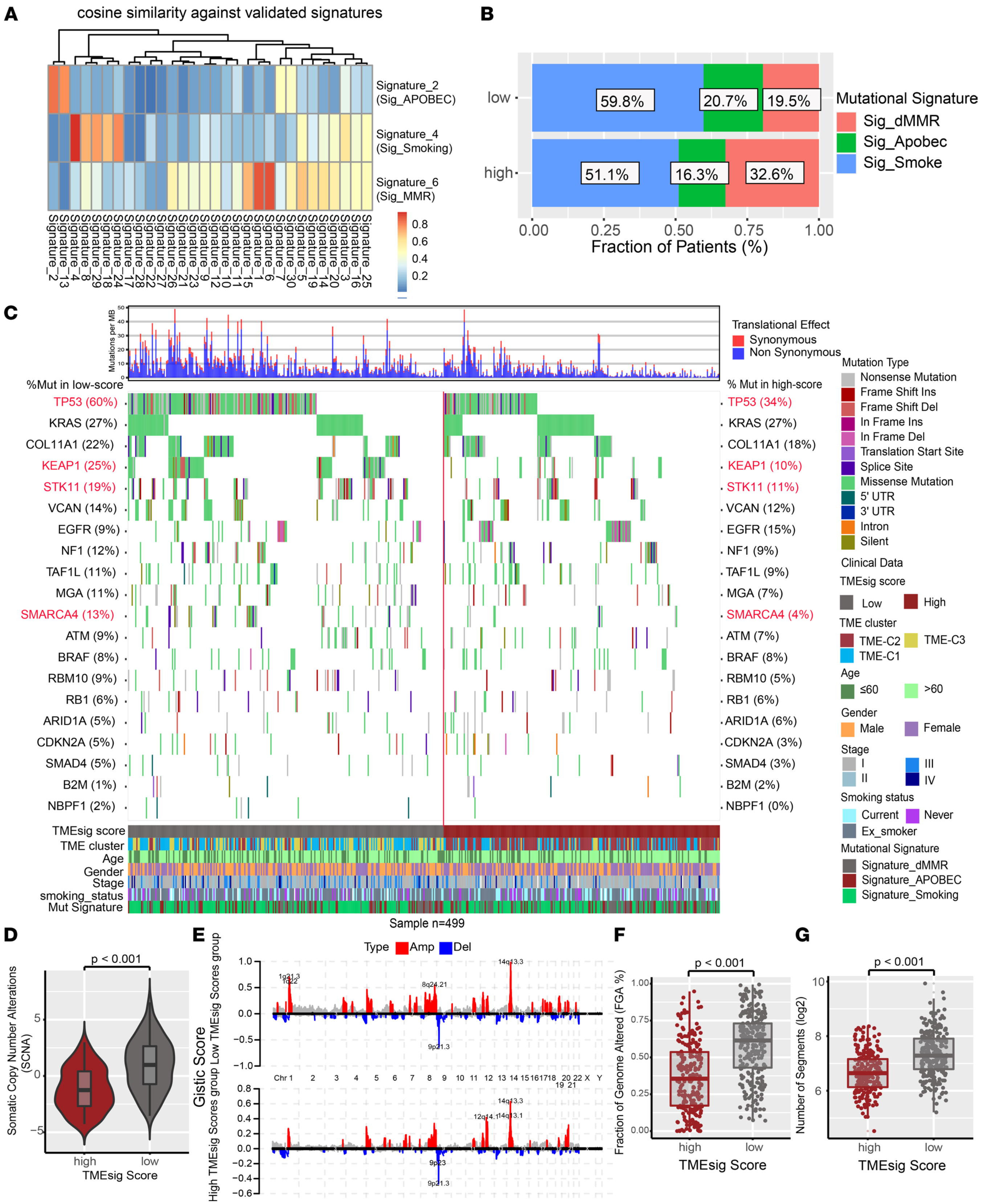


Figure 6. Genetic alternations between high and low TMEsig-score subgroups based on TCGA-LUAD cohort. (**A**) Cosine similarity analysis of extracted 3 mutational signatures against the 30 identified signatures in Catalogue of Somatic Mutations in Cancer (COSMIC, v2) with heatmap illustration. (**B**) The proportion of extracted TCGA-LUAD mutational signatures (smoking-, APOBEC-, and MMR-related signature) across different TMEsig-score subgroups ($P = 0.003$, χ^2 test). (**C**) Mutational landscape of SMGs in TCGA-LUAD stratified by low (left panel) versus high TMEsig-score (right panel) subgroups. Individual patients are represented in each column. The upper bar plot showed TMB. Mutational frequencies of SMGs in different TMEsig subtypes were depicted

in 2 sides of the plot and highlighted in red for those statistically significant. TME cluster, age, stage, sex, smoking status, and mutational signatures were shown as patient annotations. (**D**) Relative distribution of SCNA in TMEsig-score high versus low subgroups ($P < 0.001$, Wilcoxon rank-sum test). (**E**) Significant amplifications and deletions of CNVs were detected and compared between the TMEsig-score low and high subgroups. (**F** and **G**) Differences of fraction genome altered (FGA) and segment numbers values in different TMEsig-score groups ($P < 0.001$, Wilcoxon's rank-sum test).

We also integrated all the collected nonsquamous NSCLC data sets into a combined cohort and explored the prognostic value of the TMEsig-score. Concurrently, OS and relapse-free survival were significantly associated with TMEsig-score (Supplemental Figure 6, G and H), and the results of the receiver operating characteristic (ROC) curves analysis validated the predictive advantage of the established risk model (AUC = 0.739, Supplemental Figure 6I). To better predict the lung cancer prognosis, a nomogram was developed based on the multivariate Cox proportional hazards analysis to calculate the 3- and 5- year survival probability for an individual patient by the points associated with 4 risk factors (age, sex, TNM stage, TMEsig-score; Supplemental Figure 6J). The TMEsig-score showed the largest range of risk points in the nomogram, followed by the age and tumor stage. Based on these results, we concluded that TMEsig-score was supported as a favorable prognostic biomarker and a higher score indicated a better survival outcome.

Discussion

A comprehensive understanding of the molecular and biological features of the TME will extend our knowledge of immune contexture in tumorigenesis and provide potential markers for prognosis assessment and therapeutic vulnerabilities for nonsquamous NSCLC. Our findings highlight the value of integrating abundant clinical information with molecular tumor characterization and the need to generate such multimodal data. In this study, we elucidated the interactions between the clinical characteristics of nonsquamous NSCLC and TME cell–infiltrating patterns and identified 3 TME clusters with distinct immune phenotypes and survival outcomes. Stroma activation, angiogenesis, and moderate immunity, similar to immune-excluded phenotypes, were predominately in TME-C1. Activation of STING innate immunity and IFN signaling, similar to immune-inflamed phenotypes, was predominately in TME-C2. Highly proliferative features and paucity of T cell infiltration, similar to the immune-desert phenotype, were predominately in TME-C3. With the integration of TME phenotype-related gene signature, a scoring framework that adopted multiple algorithms (termed TMEsig-score) was established to quantify the TME infiltration pattern of individual tumors. Integrated analysis revealed that the TMEsig-score was an effective biomarker for NSCLC prognosis estimation and was markedly associated with immune signaling enrichment and inflammation infiltration, implying its molecular insights into characterization of the TME landscape. Furthermore, TMEsig-score was proved to be a crucial factor that predicted the genomic alterations and clinical response to anti–PD-1/PD-L1 immunotherapy.

Human lung cancer often exhibits desmoplasia, which is characterized by the presence of CAFs, ECM, and immune cells (32, 33). Previous studies reported that preexisting immunity in the immune-excluded phenotype was rendered ineffective by a block in tumor penetration through the stroma or by the retention of immune cells in the stroma (14, 34). Therefore, the surrounding stroma impeded the antitumor response of immune-excluded tumors and restricted clinical benefits from ICI agents. Specific molecular inhibitors targeting TGF-β or VEGF have been shown to reprogram the TME and restore antitumor immunity (21, 35). Based on these findings, we speculated that NSCLC patients with the TME-C1 pattern may benefit from combination treatment with ICI agents and TGF-β/VEGF blockade. TME-C3 tumors were devoid of T cell infiltration and primarily presented with highly proliferative features with increased Wnt/β-catenin, cell mitosis, and fatty acid metabolism activation. Therefore, a synergic therapy of chemotherapy, radiotherapy, or target therapy was more suitable for such a TME pattern. Platinum-based alkylating agents and pemetrexed-based antimetabolite agents are frequently used in nonsquamous NSCLC treatment (36) and are proposed to combine with immune-activated drugs to treat the TME-C3 subtype. Therefore, combination of the specific anticancer agents based on the identified TME subtype will strengthen the potential efficacy of identified patient subgroups and provide new insights into tumor precision therapy. Of course, further prospective clinical studies are needed to verify the treatment options of patients with different TME clusters or TMEsig-scores.

In addition, we observed that the TMEsig-score was strongly associated with biomarkers of response to ICI treatment, including TMB, T cell–inflamed GEP, and TIDE, implying this score could predict immunotherapy efficacy. Actually, we identified the robust prediction ability of the TMEsig-score in response

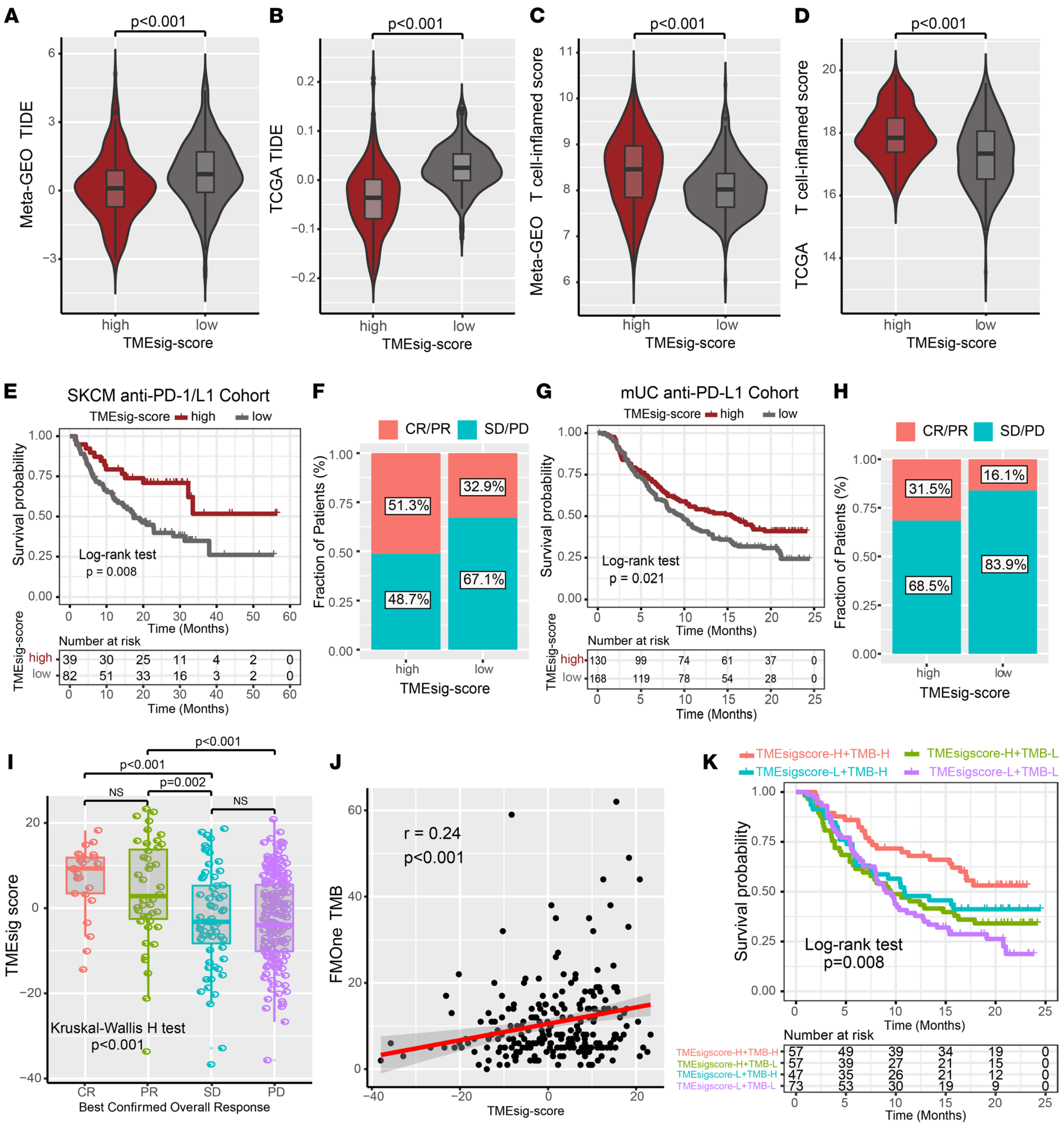


Figure 7. High TMEsig-score predicts immunotherapeutic benefits. (**A** and **B**) The relative distribution of TIDE was compared between TMEsig-score high versus low subgroups in meta-GEO (**A**) and TCGA-LUAD (**B**) cohort, respectively (Wilcoxon's rank-sum test). (**C** and **D**) Relative distribution of T cell–inflamed GEP score was also compared between TMEsig-score high versus low groups in meta-GEO (**C**) and TCGA-LUAD (**D**) cohort (Wilcoxon's rank-sum test). (**E**) Kaplan-Meier curves for high versus low TMEsig-score subgroups in the anti–PD-1/PD-L1 immunotherapy cohort. Log-rank test, *P* = 0.008. (**F**) The fraction of patients with clinical response to anti–PD-1 therapy in high or low TMEsig-score subgroups. CR/PR vs. SD/PD: 51.3% vs. 48.7% in the high TMEsig-score groups, 32.9% vs. 67.1% in the low TMEsig-score groups. (**G**) Kaplan-Meier curves for high versus low TMEsig-score patient in the anti–PD-L1 immunotherapy cohort. Log-rank test, *P* = 0.021. (**H**) The fraction of patients with clinical response to anti–PD-L1 immunotherapy in low or high TMEsig-score groups. CR/PR vs. SD/PD: 31.5% vs. 68.5% in the high TMEsig-score groups, 16.1% vs. 83.9% in the low TMEsig-score groups. CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease. (**I**) The distribution of TMEsig-score in groups with different clinical response to immunotherapy were shown (Kruskal-Wallis *H* test, *P* < 0.001). (**J**) A significant positive correlation between TMEsig-score and TMB was identified within the anti–PD-L1 cohort (Spearman *r* = 0.24, *P* < 0.001). (**K**) The overall population were divided into 4 subgroups according to the TMEsig-score and TMB level, and the TMEsig-score-H+TMB-H subgroup exhibited the best survival compared with the other 3 subgroups (log-rank test, *P* = 0.008).

to ICI treatment via 2 independent ICI cohorts. These findings verified our hypothesis that the expression profile–based TME cell infiltration pattern could be applied in clinical practice to determine immune phenotypes and guide therapeutic regimens.

Recent proteogenomic characterization delineated tumor heterogeneity and therapeutic vulnerabilities in LUAD (37, 38). We also integrated multiomics data and revealed the cGAS-STING–related mRNA and proteins were significantly altered in different TME modification subgroups. Activation of the STING innate immune pathway has been demonstrated to promote IFN secretion and lymphocyte infiltration and enhance immunotherapy response in tumors (39, 40). Here, we found that STING-related molecules (including *cGAS*, *TBK1*, *STING*, p-*STING*, and *IRF3*) and IFN-I–targeted genes (*CCL5*, *CXCL10*, and *IFI35*) were obviously enhanced in TME-C2 and high TMEsig-score subgroups, which indicated its sustainable benefit from cancer immunotherapy. Considering the low or defective STING pathway activity in TME-C2/C3 and low TMEsig-score subgroups, STING agonist coadministrated with ICI agents would hold a promise of treating such patients. However, the specific application of STING agonist combined with immunotherapy needs further prospective clinical trials.

Moreover, our data revealed that TMEsig-score had a crucial role in genomic alteration and was associated with SCNAs, mutational burden, and dMMR mutational signatures. Recent studies demonstrated that SCNAs and aneuploidy were widespread in cancer and correlated with immune evasion and cancer cell proliferation (20, 41). In our study, the low TMEsig-score subgroup was characterized by heightened SCNAs, FGA, and segments, which indicated the immune-repressive role on the low TMEsig-score subtype. Evaluation of the mutated driver genes/signatures underlying human tumors is also a critical foundation for cancer diagnostics, prognosis, and targeted therapy selection (29, 42). Recent studies reported that variants in *STK11* (*LKB1*) suppress the STING pathway and T cell inactivation and are associated with a lack of benefit from anti–PD-1/PD-L1 treatment (43, 44). In this study, we found that patients who harbored mutations of *STK11* had a higher proportion in TME-C1/-C3 and low TMEsig-score subgroups, which implied the *STK11* mutation played a role in induction of immune suppression. Moreover, on the basis of *KEAP1* mutational co-occurrences, Marinelli et al. identified mutations in 4 genes (*KEAP1*, *STK11*, *SMARCA4*, *PBRM1*) that were potentially associated with reduced efficacy of immunotherapy in LUAD (45). We also observed a higher mutation frequency of *KEAP1*, *STK11*, and *SMARCA4* in the low TMEsig-score subgroup, manifesting the molecular significance of tumor driver mutation in reprograming of immune phenotype. In addition, dMMR mutational signature was significantly associated ($P = 0.003$) with the high TMEsig-score subtype, which indicated again that genomic instability plays a crucial role in mediating tumor immunogenicity.

Tissue digital parsing methods (xCell, CIBERSORT, and so on) are an emerging tool for large-scale characterization and dissection of tumor cellular heterogeneity. Although we utilized the canonical gene signature algorithm to portray the cellular landscape, the actual TME was not completely recognized because of the technical and method limitations. Furthermore, heterogeneity in sequencing approaches and data normalization between cohorts hindered our ability to develop more standardized and precise immune features to improve the accuracy and stability of the prediction model. The clinical relevance of the identified TME molecular signatures for subtyping was validated mostly on the bulk RNA database, but not single-cell, level, because of limited single-cell transcriptome-sequencing resources in nonsquamous NSCLC clinical practice. In order to evaluate the distribution of cell subpopulations in TME subtyping more accurately, single-cell sequencing and spatial transcriptome sequencing are worthy of consideration. We also realized that the self-owned, large-scale, independent sequencing database was absent from this study and limited conclusion reliability and strength. Hence, we are collecting relevant cases from a hospital alliance to complete the verification of the findings in nonsquamous NSCLC, and it has been listed as one of our subsequent priority tasks. Besides, the TME infiltration patterns and TMEsig-score were identified by using retrospective data sets; thus, a prospective cohort of NSCLC patients receiving immunotherapy is needed to validate clinical utility. In addition, patient-derived xenograft model validation that can simulate the authentic in vivo tumor microenvironment is also necessary for accurately evaluating the effectiveness of this finding.

In summary, comprehending the TME cell infiltration landscape mediated by multiple cell subsets will contribute to enhancing our understanding of TME immune regulation. Our findings provided ideas for predicting clinical outcome and guiding treatment strategies based on TME infiltration pattern in nonsquamous NSCLC. Meanwhile, the constructed TMEsig-score quantified the TME immune infiltration of individual tumors and promoted personalized cancer immunotherapy in the future.

Methods

Collection and preprocessing of publicly attainable data sets. Multiomics sequencing data and clinical annotation of nonsquamous NSCLC samples were retrospectively collected from publicly available data sets of the NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/), TCGA (https://cancergenome.nih.gov/), and the CPTAC databases (https://cptac-data-portal.georgetown.edu/cptac). The selection criteria of nonsquamous NSCLC data sets were adopted from the workflow of Li et al. (46) in which a sample size less than 100 was excluded. Finally, a total of 2324 nonsquamous NSCLC patients were enrolled for this analysis, including those from the GSE30219 (n = 187), GSE31210 (n = 226), GSE37745 (n = 130), GSE50081 (n = 138), GSE41271 (n = 194), GSE68465 (n = 442), GSE72094 (n = 398), TCGA-LUAD (n = 499), and CPTAC-LUAD (n = 110) (37) (Supplemental Table 2). GSE42127 (47) was excluded from this analysis because its probe cell intensity (CEL) files extensively overlapped with the GSE41271 series. Since the GSE30219, GSE31210, GSE37745, and GSE5008 data sets adopted the same microarray sequencing platform (Affymetrix HG-U133 plus 2.0), we downloaded the raw CEL files and performed identical background adjustment and quantile normalization by affy and simpleaffy packages. We then further integrated them into a meta-cohort by using ComBat function from R package sva for batch removal. GSE41271, GSE68465, GSE72094, TCGA-LUAD, and CPTAC-LUAD were recognized as the independent validation cohorts. TCGA RNA-sequencing data (FPKM format) were downloaded from the UCSC Xena (https://gdc.xenahubs.net/download/TCGA-LUAD.htseq_fpkm.tsv.gz) and transformed into transcripts per kilobase million (TPM) format. The clinical information and cluster subtyping of the meta-GEO and TCGA data sets are listed in Supplemental Tables 3 and 4 In addition, gene-level proteomics and phosphoproteomic data were downloaded from the CPTAC website, and protein abundances were log2-transformed and median-centered.

Comparison of genomic alterations. The somatic mutation (MuTect2) and CNV data of TCGA-LUAD were downloaded from TCGA database (https://portal.gdc.cancer.gov/). Nonsynonymous mutation (including frameshift mutation, inflame mutation, missense mutation, nonsense mutation and splice site mutation) counts were recognized as TMB. The GISTIC score and gene copy number amplification and deletion data for each sample were analyzed by GISTIC 2.0 software and plotted by maftools package. The FGA, SCNA, and aneuploidy score of each lung cancer sample was determined and curated from previous studies (20).

Inference of infiltrating cells in the TME. To fully portray the cellular heterogeneity of the TME landscape in the NSCLC, we utilized the gene signature–based xCell algorithm (48) to infer 64 immune and stromal cell types, spanning multiple adaptive and innate immunity cells, hematopoietic progenitors, epithelial cells, and ECM cells. The xCell employed a curve fitting approach for linear comparison of cell types and introduced a novel spillover compensation technique to reduce dependencies between closely related cell types. Gene expression profiles were prepared using standard annotation files, and data were uploaded to the xCell web portal (https://xcell.ucsf.edu/), with the algorithm run using the xCell signature. Moreover, we also employed the CIBERSORT and MCP-counter algorithms to quantify the immune cell subsets and validate the reliability of xCell-identified 3 TME clusters of NSCLC. CIBERSORT (49) is a deconvolution algorithm that uses a set of reference gene expression values (a signature with 547 genes) considered a minimal representation for each cell type and, based on those values, infers cell type proportions in data from bulk tumor samples with mixed cell types using support vector regression. MCP-counter (50) is a transcriptomic marker-based approach, which allows the robust quantification of the absolute abundance of 8 immune and 2 stromal cell populations in heterogeneous tissues.

Consensus molecular clustering for TME-infiltrating cells. We performed hierarchical consensus clustering with non-NMF algorithm to identify distinct TME immune patterns based on the infiltration level of 64 cell subpopulations. The xCell-derived cell profiles were factorized into 2 non-negative matrices, W and H (i.e., $A \approx WH$). Repeated factorization of matrix A was performed, and its outputs were aggregated to obtain consensus clustering of NSCLC samples. The optimal number of clusters was selected according to cophenetic, dispersion, and silhouette coefficients. The R package NMF (version 0.22.0) with the Frobenius algorithm and 200 nruns was used to perform the consensus clustering in both the meta-GEO data set and TCGA cohort. Besides, 3D projection on NSCLC tumor clustering was performed by UMAP-based dimension reduction and visualized by plotly package.

GSVA and GO annotation. We utilized the GSVA algorithm within the R package GSVA to investigate the variation in biological processes among different TME cluster patterns. The well-defined biological signatures were derived from the Hallmark gene set (downloaded from MSigDB database v7.1), Mariathasan et al. constructed gene set (21) (curated from IMvigor210CoreBiologies packages), and Thorsson et al. constructed gene sets (19). GO annotation for TME phenotype-related genes was performed in the R package clusterProfiler with the cutoff value of $FDR < 0.01$.

Quantifying the predictors of immune response: T cell GEP score, TIDE, and ESTIMATE. T cell–inflamed GEP is a superior predictor of response to anti–PD-1 regimens, which contains IFN-γ–responsive genes related to antigen presentation, cytotoxic activity, and adaptive immune resistance (51). The T cell–inflamed scores were calculated and weighted by averaging of the included genes for the IFN-γ (6-gene) and expanded immune (18-gene) signatures. The TIDE algorithm proposed by Jiang et al. was utilized to model distinct tumor immune evasion mechanisms (52), including dysfunction of tumor-infiltrating cytotoxic T lymphocytes (CTLs) and exclusion of CTLs by immunosuppressive factors. A higher TIDE score indicated cells more likely to induce immune escape, thus indicating a lower response rate to ICI treatment. The ESTIMATE algorithm (53), which takes advantage of the unique properties of the transcriptional profiles, was used to infer the infiltration level of immune and stromal cell score in tumor tissues and to estimate tumor purity. Tumor tissues with abundant immune cell infiltration represented a higher immune score and lower level of tumor purity.

SMGs and tumor mutational signatures. We utilized the MutSigCV algorithm described in a prior study and filtration criteria to identify SMGs (54, 55). The mutational landscape of identified SMGs in TCGA cohort was depicted by the waterfall function of the R GenVisR package. Mutational signatures extracted from TCGA genomic data adopted the maftools package. ExtractSignatures function, based on Bayesian variant non-NMF, factorized the mutation portrait matrix into 2 non-negative matrix signatures and contributions, where signatures represent mutational processes and contributions represent the corresponding mutational activities. The SignatureEnrichment function can automatically determine the optimal number of extracted mutational signatures and assign them to each sample based on the mutational activities. The extracted mutational portrait of lung cancer was compared and annotated by cosine similarity analysis against the COSMIC.

Identification of DEGs between distinct TME modification phenotypes. Our consensus clustering algorithm divided patients into 3 distinct TME infiltration patterns, and we next determined TME infiltration pattern–related DEGs between distinct TME phenotypes. The R package limma was used to evaluate DEGs in lung cancer samples between different TME clusters. Specifically, gene expression data were normalized by voom and then fed to lmFit and eBayes functions to calculate the differentially expressed statistics. The significance filtering criteria of DEGs was set as an adjusted P value less than 0.001.

Construction of the TME signature score. We constructed a TMEsig scoring scheme to quantify the relative TME infiltration level of individual patients by using principal component analysis (PCA). Specifically, the overlapping DEGs identified from different TME clusters were curated and employed to perform prognostic analysis for each gene using a univariate Cox regression model (Supplemental Table 5). The genes with a significant prognostic impact were extracted for further feature selection by using recursive feature elimination with random forest and the 10-fold cross-validation method in the caret R package. We then curated the expression profile of the final determined genes to perform PCA, and principal components 1 and 2 were extracted and served as the signature score. This method mainly focuses on the score on the set with the largest block of well-correlated (or inverse-correlated) genes in the set, while downweighting contributions from genes that do not track with other set members. We then adopted a formula similar to previous study to define the TMEsig-score (56): TMEsig-score = $\sum(PC1_i+PC2_i)$, where $_i$ is the expression of final determined TME phenotype-related genes.

Cell lines and culture conditions and reagents. Calu-3 cells, NCI-H23 cells, NCI-H1299 cells, NCI-H358 cells, NCI-H838 cells, NCI-H1395, and NCI-H2126 cells were purchased from ATCC. Authentication of the cells was performed by short tandem repeat analysis at Beijing Microread Genetics Co., Ltd. NCI-H23 cells, NCI-H1299 cells, NCI-H358 cells, NCI-H838 cells, NCI-H1395 cells, and NCI-H2126 cells were cultured in 1640 medium (Gibco) supplemented with 10% fetal bovine serum (Gibco) in a 5% CO_2 incubator at 37°C, while Calu-3 cells were in DMEM (Gibco).

Western blot analysis and antibodies. In brief, total cell lysates were prepared with cell lysis buffer. After denaturing via boiling, total protein was quantified using a BCA protein assay kit (Solarbio). Equivalent amounts of protein were separated by SDS-PAGE at 80 V for 2.5 hours and transfected to PVDF membranes for 1.5 hours. The membranes were washed using 1% TBST by 3 cycles of 5 minutes after incubation with primary antibodies targeting *cGAS* (Santa Cruz Biotechnology, SC-515777), *STING* (Cell Signaling Technology, CST13647S), p-*STING* (Cell Signaling Technology, CST19781S), *TBK1* (Cell Signaling Technology, CST3504S), NF-κB (Cell Signaling Technology, CST4764S), p-NF-κB (Cell Signaling Technology, CST3033S), *IRF3* (Santa Cruz Biotechnology, SC-33641), p-*IRF3* (Cell Signaling Technology, CST29047S), *CCL5* (Santa Cruz Biotechnology, SC-514019), *CXCL10* (Santa Cruz Biotechnology, SC-374092), and β-actin (Proteintech, 20536-1-AP) at 4°C overnight. Then membranes were treated with secondary antibodies (Proteintech, SA00001-1 and SA00001-2).

Collection of transcriptomic and clinical information of the ICI-based cohort. We systematically searched the gene expression profiles in ICIs, which could be publicly obtained and coupled with detailed clinical information. Two immunotherapeutic cohorts were finally included in our study: metastatic melanoma treated with nivolumab/pembrolizumab (anti–PD-1 monoclonal Ab) (57) and metastatic urothelial cancer treated with atezolizumab (anti–PD-L1 mcAb) (21). The gene expression profiles of pretherapy biopsy samples were curated and transformed into the TPM format for further analysis.

Statistics. The statistical analyses in this study were generated by R 3.6.1. For quantitative data, statistical significance for normally distributed variables was estimated by Student's 2-tailed *t* tests, and non-normally distributed variables were analyzed by the Wilcoxon's rank-sum test. For comparisons of more than 2 groups, Kruskal-Wallis tests and 1-way ANOVA were used as nonparametric and parametric methods, respectively. The χ^2 test and Fisher's exact test were used to analyze contingency tables depending on specific grouping condition. Cox proportional hazards model was used to analyze the association between the TME modification pattern and prognosis with the R package forestmodel. The surv-cutpoint function from the survminer package utilized the maximally selected rank statistics to determine the optimal cutpoint, and was applied to stratify samples into high and low TMEsig-score subgroups. The C-index proposed by Harrell et al. (58) and the ROC curve were used to assess the prognosis classification performance of the TMEsig-score model, and the C-index and AUC were calculated by using survcomp and timeROC packages, respectively. Nomogram model combined with TMEsig and detailed clinical information was utilized for calculating the survival probability by using rms package. All comparisons were 2 sided with an α level of 0.05, and the Benjamini-Hochberg method was applied to control the FDR for multiple-hypothesis testing.

Study approval. This study was approved by the Shandong Provincial Hospital Institutional Review Board (Jinan, China), which waived additional informed consent because all data used in this study were obtained from public databases. Participants in the original studies have provided informed consent.

Author contributions

HC, WC, and ML conceived and designed the study; HC, WC, ML, and FD developed methodology; HC, JL, LS, YC, TZ, YZ, XJ, ZW, and SJ acquired data (provided data, provided facilities, etc.); HC, WC, FD, HW, ZF, and YL analyzed and interpreted data (e.g., statistical analysis, biostatistics, computational analysis); HC, WC, TZ, and ML wrote, reviewed, and/or revised the manuscript; HC, WC, JL, and FD provided administrative, technical, or material support (i.e., reporting or organizing data, constructing databases); and HC, WC, JL, and ML supervised the study.

Acknowledgments

We sincerely thank Xiangchun Li and Yichen Yang from Tianjin Medical University Cancer Institute and Hospital for analysis advice and also thank Liming Chen from Harvard University for meaningful discussion and revising the manuscript. This study was supported by grants from National Scientific Foundation of China (82103322, 82102702), Natural Science Foundation of Shandong Province of China (ZR2020QH180, ZR2021QH141), Key Research and Development Program of Shandong Province (No. 2021CXGC011104; No. 2019JZZY010104; No. 2019GSF108146), academic promotion programme of Shandong First Medical University (2019QL021), Special Foundation for Taishan Scholars Program of Shandong Province (No. ts20190978), and The Research Incubation Funding of Shandong Provincial Hospital (2020FY035).

Address correspondence to: Wei Chong, Department of Gastrointestinal Surgery, Key Laboratory of Engineering of Shandong Province, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong 250021, China. Email: chongwei@sdfmu.edu.cn. Or to: Ming Lu, Clinical Research Center of Shandong University, Clinical Epidemiology Unit, Qilu Hospital of Shandong University, Jinan, Shandong 250021, China. Email: lvming@sdu.edu.cn. Or to: Jin Liu, Research Center for Experimental Nuclear Medicine, School of Basic Medical Sciences, Shandong University, Jinan 250012, China. Email: miracle713@163.com.

1. Hirsch FR, et al. Lung cancer: current therapies and new targeted treatments. *Lancet*. 2017;389(10066):299–311.
2. Herbst RS, et al. The biology and management of non-small cell lung cancer. *Nature*. 2018;553(7689):446–454.
3. Yuan M, et al. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduct Target Ther*. 2019;4:61.
4. Altorki NK, et al. The lung microenvironment: an important regulator of tumour growth and metastasis. *Nat Rev Cancer*. 2019;19(1):9–31.
5. Li X, et al. Identification of an immune signature predicting prognosis risk and lymphocyte infiltration in colon cancer. *Front Immunol*. 2020;11:1678.
6. Jin MZ, Jin WL. The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct Target Ther*. 2020;5(1):166.
7. Reck M, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med*. 2016;375(19):1823–1833.
8. West H, et al. Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol*. 2019;20(7):924–937.
9. Salmon H, et al. Host tissue determinants of tumour immunity. *Nat Rev Cancer*. 2019;19(4):215–227.
10. Zhou R, et al. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer. *Cancer Immunol Immunother*. 2019;68(3):433–442.
11. Sturm G, et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics*. 2019;35(14):i436–i445.
12. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discov*. 2019;18(3):197–218.
13. Litchfield K, et al. Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell*. 2021;184(3):596–614.
14. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. 2017;541(7637):321–330.
15. Hegde PS, Chen DS. Top 10 challenges in cancer immunotherapy. *Immunity*. 2020;52(1):17–35.
16. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7511):543–550.
17. Busch SE, et al. Lung cancer subtypes generate unique immune responses. *J Immunol*. 2016;197(11):4493–4503.
18. Spranger S, Gajewski TF. Impact of oncogenic pathways on evasion of antitumour immune responses. *Nat Rev Cancer*. 2018;18(3):139–147.
19. Thorsson V, et al. The immune landscape of cancer. *Immunity*. 2018;48(4):812–830.
20. Davoli T, et al. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*. 2017;355(6322):eaaf8399.
21. Mariathasan S, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature*. 2018;554(7693):544–548.
22. Charoentong P, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18(1):248–262.
23. Li Z, et al. Establishment of a gene signature to predict prognosis for patients with lung adenocarcinoma. *Int J Mol Sci*. 2020;21(22):8479.
24. Zhao J, et al. Identification of a novel gene expression signature associated with overall survival in patients with lung adenocarcinoma: a comprehensive analysis based on TCGA and GEO databases. *Lung Cancer*. 2020;149:90–96.
25. Zhang M, et al. An immune-related signature predicts survival in patients with lung adenocarcinoma. *Front Oncol*. 2019;9:1314.
26. Song Q, et al. Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma. *J Transl Med*. 2019;17(1):70.
27. Ghandi M, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature*. 2019;569(7757):503–508.
28. Chen H, et al. The immune response-related mutational signatures and driver genes in non-small-cell lung cancer. *Cancer Sci*. 2019;110(8):2348–2356.
29. Chong W, et al. Association of clock-like mutational signature with immune checkpoint inhibitor outcome in patients with melanoma and NSCLC. *Mol Ther Nucleic Acids*. 2021;23:89–100.
30. Rizvi NA, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348(6230):124–128.
31. Chen H, et al. Association of LRP1B mutation with tumor mutation burden and outcomes in melanoma and non-small cell lung cancer patients treated with immune check-point blockades. *Front Immunol*. 2019;10:1113.
32. Bremnes RM, et al. The role of tumor stroma in cancer progression and prognosis: emphasis on carcinoma-associated fibroblasts and non-small cell lung cancer. *J Thorac Oncol*. 2011;6(1):209–217.
33. Kalluri R. The biology and function of fibroblasts in cancer. *Nat Rev Cancer*. 2016;16(9):582–598.
34. Lambrechts D, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med*. 2018;24(8):1277–1289.
35. Hack SP, et al. Augmenting anticancer immunity through combined targeting of angiogenic and PD-1/PD-L1 pathways: challenges and opportunities. *Front Immunol*. 2020;11:598877.
36. Gadgeel S, et al. Updated analysis from KEYNOTE-189: pembrolizumab or placebo plus pemetrexed and platinum for previously untreated metastatic nonsquamous non-small-cell lung cancer. *J Clin Oncol*. 2020;38(14):1505–1517.
37. Gillette MA, et al. Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. *Cell*. 2020;182(1):200–225.
38. Xu JY, et al. Integrative proteomic characterization of human lung adenocarcinoma. *Cell*. 2020;182(1):245–261.
39. Kwon J, Bakhoum SF. The cytosolic DNA-sensing cGAS-STING pathway in cancer. *Cancer Discov*. 2020;10(1):26–39.
40. Hopfner KP, Hornung V. Molecular mechanisms and cellular functions of cGAS-STING signalling. *Nat Rev Mol Cell Biol*. 2020;21(9):501–521.

41. Bareche Y, et al. Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. *Ann Oncol*. 2018;29(4):895–902.
42. Chen H, et al. The new identified biomarkers determine sensitivity to immune check-point blockade therapies in melanoma. *Oncoimmunology*. 2019;8(8):1608132.
43. Koyama S, et al. STK11/LKB1 deficiency promotes neutrophil recruitment and proinflammatory cytokine production to suppress T-cell activity in the lung tumor microenvironment. *Cancer Res*. 2016;76(5):999–1008.
44. Kitajima S, et al. Suppression of STING associated with LKB1 loss in KRAS-driven lung cancer. *Cancer Discov*. 2019;9(1):34–45.
45. Marinelli D, et al. KEAP1-driven co-mutations in lung adenocarcinoma unresponsive to immunotherapy despite high tumor mutational burden. *Ann Oncol*. 2020;31(12):1746–1754.
46. Li B, et al. Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer. *JAMA Oncol*. 2017;3(11):1529–1537.
47. Tang H, et al. A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small cell lung cancer patients. *Clin Cancer Res*. 2013;19(6):1577–1586.
48. Aran D, et al. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017;18(1):220.
49. Newman AM, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–457.
50. Becht E, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17(1):218.
51. Ayers M, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest*. 2017;127(8):2930–2940.
52. Jiang P, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*. 2018;24(10):1550–1558.
53. Yoshihara K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
54. Lawrence MS, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214–218.
55. Chen H, et al. Age-related mutational signature negatively associated with immune activity and survival outcome in triple-negative breast cancer. *Oncoimmunology*. 2020;9(1):1788252.
56. Chong W, et al. m^{6}A regulator-based methylation modification patterns characterized by distinct tumor microenvironment immune profiles in colon cancer. *Theranostics*. 2021;11(5):2201–2217.
57. Liu D, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med*. 2019;25(12):1916–1927.
58. Harrell FE Jr., et al. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361–387.