Prognostic value of the baseline circulating T cell receptor β chain diversity in advanced lung cancer

Jiaqi Wang§, Zhixin Bie§, Yajing Zhang§, Lin Li§, Yujie Zhu§, Yi Zhang§, Xin Nie§, Ping Zhang§, Gang Cheng§, Xuebing Di§, Xiaoguang Li§, Shujun Cheng§, and Lin Feng§*

§State Key Laboratory of Molecular Oncology, Department of Etiology and Carcinogenesis, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; *Minimally Invasive Tumor Therapies Center, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China; Department of Medical Oncology, Beijing Hospital, National Center of Gerontology; Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

ABSTRACT
An indicator for systemic evaluation of the adaptive immune status is lacking. Peripheral blood is important in antitumour immunity, and the T-cell receptor (TCR) repertoire diversity is key for effective immunity. This study aimed to investigate changes in the circulating T cell receptor β chain (TCRB) diversity during the first few (1 ~ 4) treatment cycles and its clinical value in patients with advanced lung cancer. TCRB-enriched sequencing data combined with transcriptomic RNA sequencing data of peripheral blood leukocytes were obtained from 72 patients with advanced lung cancer before and after targeted therapy or chemotherapy. Changes in the circulating TCRB diversity during treatment and the relationship of the baseline circulating TCRB diversity with prognosis and therapeutic effects were evaluated. We found that targeted therapy or chemotherapy did not significantly affect the T lymphocyte composition or circulating TCRB diversity (3.83 vs 3.74, T-test, p = .16) in patients with advanced lung adenocarcinoma (LUAD) during the first few treatment cycles. The higher circulating TCRB diversity was linked to improved therapeutic effects (T-test, p = .0003) in LUAD patients receiving targeted therapy. Higher baseline circulating TCRB diversity was associated with better prognosis. In addition, a five-factor prognostic risk score model was built for more accurate prognosis prediction for LUAD patients. The chemotherapeutic agents for advanced lung cancer do not significantly affect adaptive immune function over the first few treatment cycles. The circulating TCRB diversity reflects the adaptive immunological repertoire and might be a convenient indicator for evaluating the adaptive immune status and prognosis.

Introduction
The treatment options for patients with advanced lung cancer were developed according to the clinical guidelines and are based on the molecular characteristics and histology of the tumors.1, 2 These treatments not only induce tumor apoptosis or inhibit tumor proliferation but also enhance tumor immunogenicity and promote the antitumour immune response, particularly adaptive immunity. Accumulating evidence suggests that the innate and adaptive immune systems play a crucial role in the antitumour effects of conventional chemotherapy-based and targeted therapy-based treatments.3-5 A powerful antitumour response requires durable antitumour immunity, and the activation of adaptive immunity is indispensable for killing tumor cells.6 Consequently, the antitumour ability of the host immune system might considerably influence the therapeutic effect and prognosis of patients with advanced lung cancer.

The T cell recognition of tumor antigens presented by major histocompatibility complex (MHC) molecules via the T cell receptor (TCR) is crucial for the activation of antitumour immune responses. The TCR diversity is a prerequisite for the recognition of tumor antigens.7 The TCR consists of αβ or γδ heterodimers, and the vast majority of mature T cells in the peripheral blood are αβ T cells (more than 90%).7,8 The diversity of αβ T cells is mainly concentrated in both the Vα and Vβ complementarity determining region 3 (CDR3) loops,9 particularly in CDR3 of the TCRβ chain. The recombination of V(D)J regions and random nucleotide insertions and deletions contribute to the superior variability of CDR3.10 Focusing on the TCR repertoire might allow a deeper understanding of the interactions between treatment and the host immune status in patients with advanced lung cancer.

In non-small cell lung cancer (NSCLC), a higher T cell density in the blood indicates a better outcome following
surgery.\(^\text{11}\) Tumor-specific T cells recognize tumor-associated antigens and undergo intratumor clonal expansion. The TCR sequences expressed by these tumor-specific T cells are detectable in peripheral blood and remain for the long-term.\(^\text{12}\) Zhang, J. et al. demonstrated that the peripheral blood TCR repertoire acts as a vital compartment for rejuvenating anti-tumour immunity.\(^\text{13}\) These results indicate that the peripheral blood TCR repertoire might reflect the antitumour ability of the host. Several studies have suggested the prognostic value of the circulating T cell receptor \(\beta\) chain (TCRB) diversity in immunotherapy.\(^\text{14,15}\) However, the relationship between the baseline circulating TCRB diversity and prognosis remains unclear in patients with advanced lung cancer receiving conventional therapy.\(^\text{16}\) Other studies on the TCR repertoire of lung cancer mainly recruited NSCLC patients but did not pay attention to small cell lung cancer (SCLC).\(^\text{11,14,16}\) Previous studies have shown that EGFR-TKI therapy can enhance tumor-specific T cell infiltration and reactivation and trigger tumor lysis mediated by NK cells and tumor-specific T cells.\(^\text{4,17}\) Although chemotherapeutic drugs that exert cytotoxic effects induce systemic immunosuppressive effects, some chemotherapeutic compounds can enhance the immunogenicity of tumor cells and the activity of cytotoxic T cells and NK cells and thus ultimately promote antitumour immunity.\(^\text{3,18,19}\) Furthermore, the effects of cytotoxic chemotherapy and targeted therapy on the circulating TCRB diversity in patients with advanced lung cancer remain unknown.

Previous studies applied multiplex PCR with specific primers against each V and J gene and genomic DNA or cDNA templates to construct TCR repertoire libraries;\(^\text{16,20–22}\) however, these results are compromised by the accumulation of PCR mismatch errors. Many studies have demonstrated that a high TCRB diversity is related to better prognosis in lung cancer and melanoma.\(^\text{14,15}\) However, Liu et al. did not find the direct prognostic value of the baseline circulating TCRB diversity in advanced lung cancer,\(^\text{16}\) and this finding might be because their TCR library construction method was based on multiplex PCR, which has high error rate and provides inaccurate quantification. A novel TCR repertoire library construction method based on the switch mechanism at the 5’-end of RNA templates (SMART) has been developed,\(^\text{23}\) and this method allows nearly absolute error correction. We modified this method to obtain the SMART-based UMI-corrected TCR library construction method, which involves the use of two universal TCR C gene primers and cDNA templates to ensure specificity during reverse transcription and PCR amplification.\(^\text{24}\)

In our study, we combined TCRB sequencing with transcriptome sequencing of peripheral blood leukocytes of patients with advanced lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC) or small cell lung cancer (SCLC) from the same sample pre- and posttreatment to explore the clinical value of the circulating TCRB diversity and the impact of treatment on the T lymphocyte composition and circulating TCRB diversity in these patients. We were committed to developing an accessible, noninvasive and unbiased approach for evaluating the global immune status of patients with advanced lung cancer and predicting their treatment response and prognosis.

### Material and methods

#### Study design and patient population

This study involved 72 patients with advanced lung cancer, including 39 patients with lung adenocarcinoma, 10 patients with lung squamous cell carcinoma and 23 patients with small cell lung cancer. We included patients who received initial therapy or who relapsed at least 6 months after their last therapy and excluded patients who had a prior history of autoimmune disease. These patients were divided into four groups depending on their pathologic classifications and treatment regimens. The LUAD target therapy group included 23 LUAD patients receiving EGFR-TKI targeted therapy. The LUAD chemotherapy group included 16 LUAD patients who were treated with pemetrexed alone or combined with platinum. The LUSC chemotherapy group included 10 LUSC patients who were treated with platinum with gemcitabine or paclitaxel. The LUSC chemotherapy group included 23 SCLC patients who were treated with platinum with etoposide. Peripheral blood leukocytes obtained before (baseline) and after treatment (1 ~ 2 cycles of treatment: 87.5%; 3 ~ 4 cycles of treatment: 12.5%) were collected for RNA extraction and for performing paired TCRB sequencing and transcriptomic RNA sequencing (RNAseq). Table 1 lists the baseline characteristics of the patients, and details of each patient are available in Supplementary Table S1. The study was approved by the Ethics Review Committee of Beijing Hospital (No.2017BJYYEC-007–01&2019BJYYEC-148–02). All patients were informed about the study and provided written informed consent.

#### Treatment evaluation

The treatment efficacy was evaluated by two independent clinicians according to the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1. Progression-free survival (PFS) was calculated as the time from the start of targeted therapy or chemotherapy to either tumor progression or death. Detailed follow-up information for the patients is shown in Supplementary Table S1.

### Table 1. Baseline characteristics of patients.

| Pathology     | Number of cases (%) |
|---------------|---------------------|
| Gender        |                     |
| Female        | 21(54)              |
| Male          | 18(46)              |
| Age ≤60       | 17(44)              |
| Age >60       | 22(56)              |
| Smoking history |                   |
| Current/Former | 11(28)            |
| Never/Light   | 28(72)              |
| Stage         |                     |
| III           | 3(8)                |
| IV            | 36(92)              |
| Treatment     |                     |
| Targeted therapy | 23(59)           |
| Chemotherapy  | 16(41)              |
| Therapeutic effect |            |
| PR            | 16(41)              |
| non-PR        | 23(59)              |
| Total         | 39                  |

LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; SCLC: small cell lung cancer; PR: partial response.
TCRB sequencing and transcriptome sequencing

We performed TCRB sequencing of 144 samples (all 72 patients before and after treatment) and transcriptomic RNA-seq of 138 samples. Transcriptome sequencing were not obtained for three patients due to limited RNA. SMART-based UMI-corrected TCRB libraries were constructed using two universal TCR C gene primers and cDNA templates, as detailed in our previous work. RNA-seq and TCRB sequencing were performed using an Illumina HiSeq X ten platform. The raw TCRB sequencing and transcriptome sequencing data reported in this study have been deposited in the Genome Sequence Archive in National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number HRA000211 [http://bigd.big.ac.cn/gsa-human].

The preprocessing for TCRB sequencing data

We first extracted and merged UMIs and effective TCRB sequences from the raw TCRB sequencing data using MiGEC. The TCRB sequences were aligned to the human TCR reference sequences to extract and quantify corresponding CDR3 sequences or clonotypes using MIXCR. The log10-transformed Shannon-Weiner index (ShannonDI), which takes into account both the number of TCR clonotypes and their relative abundances, was used to indicate the diversity of the TCR repertoire, and this index was calculated using VDJtools. Unique clonotypes were defined as TCRB clonotypes with unique CDR3 amino acid sequences. The overlap of the T cell repertoire was measured using the ‘cosine’ method, and the shared clones were calculated using the R package ‘immunarch’.

Estimation of relative abundance of lymphocyte subpopulations

We inferred the constituent ratios of lymphocyte subsets by inputting gene signature and bulk RNA-seq data into CIBERSORTx. The gene signature of the lymphocyte subsets was established based on the single-cell RNA sequencing of peripheral blood mononuclear cells (PBMCs) from cancer patients (unpublished data), which is listed in Supplementary Table S2. We also used ABSolute Immune Signal (ABIS) deconvolution to evaluate the relative abundance of the lymphocyte subpopulations.

Establishment of the five-factor prognostic risk score model

Based on a ratio of 3:1, patients with LUAD were divided into a training dataset (N = 27) and a testing dataset (N = 10) by stratified sampling based on pharmacotherapeutic regimens. To further select the best factors related to patient prognosis, we used the Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression method and leave-one-out cross-validation to filter effective factors with the R package ‘glmnet’. Among the factors, the platelet-to-lymphocyte ratio (PLR) was the ratio of platelet count to lymphocyte count, and the change in circulating TCRB diversity was calculated using the following formula: log10(posttreatment ShannonDI—pretreatment ShannonDI). Regression coefficients (β) were used to calculate a weighted sum of the effective factors, which was used as the prognostic risk score. Using the training, testing and entire datasets, Kaplan-Meier survival analysis and Harrell’s concordance index (C-index) were used to evaluate the prognostic value and predictive ability of the model, respectively.

Statistical analysis

The unpaired Student’s t test and Wilcoxon rank-sum test were used to compare the differences between two unpaired groups. The paired Student’s t test and Wilcoxon signed-rank test were used to compare the differences between paired groups. Pearson correlation and Spearman rank correlation were used to measure the associations among variables. A multivariate Cox proportional hazards regression model was used to identify independent prognostic factors. The survival analysis was performed through Kaplan–Meier curve analysis and the log-rank test. The GSEA analysis was conducted using the R packages ‘DESeq2’ and ‘clusterProfiler’. R (version 3.6.2) was used for the statistical analyses and for the preparation of the figures.

Results

Baseline clinical information and characteristics of the TCRB repertoire of the recruited patients

A total of 144 blood samples from 72 patients with advanced lung cancer (64 treatment-naïve patients and eight patients who relapsed) were recruited into the study. Table 1 and Supplementary Table S1 show the baseline characteristics of the recruited patients. All patients received treatment according to standard guidelines, and the follow-up time ranged from 14 to 1307 days (median follow-up, 197 days). The TCRB repertoires of each patient before and after treatment contained 104,913 ± 44,727 (mean ± SD) and 108,327 ± 44,681 (mean ± SD) TCRB clonotypes, respectively.

Patients with advanced lung adenocarcinoma retain a relatively stable T lymphocyte composition and circulating TCRB diversity after the first few treatment cycles

To explore the effect of targeted therapy and chemotherapy on systemic immunity in patients with advanced lung adenocarcinoma, we separately compared the composition of peripheral blood immune cells in the LUAD targeted therapy group and the LUAD chemotherapy group before and after treatment. In terms of absolute numbers, both targeted therapy and chemotherapy did not significantly alter the neutrophil count or the lymphocyte count (Supplementary Figure S1b-c). Furthermore, we inferred the constituent ratios of peripheral blood mononuclear cell subpopulations using the CIBERSORTx algorithm. The mean constituent ratios before and after treatment are presented in Supplementary Figure S1a. No significant changes were detected after the first few targeted
therapy or chemotherapy cycles in the main T cell subtypes including CD4 naïve T cells (CD4 naïve T), CD8 naïve T cells (CD8 naïve T), CD4+ central memory T cells (CD4 Tcm), CD4+ effector T cells and effector memory T cells (CD4 Teff and Tem), CD8+ effector T cells (CD8 Teff) or CD4+ regulatory T cells (CD4 Treg) (Figure 1).

We also analyzed the changes in leukocyte components in the LUSC chemotherapy group and the SCLC chemotherapy group. The chemotherapy significantly reduced the neutrophil count in these patients (Supplementary Figure S1b), while the lymphocyte count and constituent ratios of most T lymphocyte subsets remained stable during chemotherapy (Figure 1, Supplementary Figure S1c). After chemotherapy, patients with LUSC had a much higher proportion of CD4 Tcm (Figure 1b) and a much lower proportion of CD4 Treg (Figure 1d). Similar observations have been made in recent studies: Julia EP et al. revealed that an increased frequency of CD4+ central memory T cells was associated with the immunotherapy response and Tseng et al. indicated that cisplatin significantly decreased T regulatory cells in the peripheral blood of TC-1 tumor-bearing mice. 

For patients with SCLC, the constituent ratio of CD8 naïve T cells was significantly elevated after chemotherapy (Figure 1e), which suggests that more CD8 naïve T cells were mobilized to initiate immune responses. This phenomenon was previously less reported in SCLC.

To investigate whether targeted therapy or chemotherapy affects the T cell repertoire, we subsequently compared the circulating TCRB diversity in the four groups before and after treatment. No significant changes in the circulating TCRB diversity were found after the first few treatment cycles (Figure 2a). We found that the pretreatment circulating TCRB diversity was significantly correlated with the posttreatment circulating TCRB diversity in each of the four groups (Supplementary Figure S2). Neither chemotherapy nor targeted therapy can significantly alter the immune status of patients with advanced lung cancer. Therefore, over the first few treatment cycles, both targeted therapy and chemotherapy for patients with advanced lung adenocarcinoma did not have a marked effect on the peripheral blood T lymphocyte composition.

**The circulating TCRB diversity is linked to therapeutic effects in the LUAD targeted therapy group**

Next, we horizontally compared the T cell composition of patients with different therapeutic effects in the LUAD targeted therapy group. The LUAD targeted therapy group was divided into two groups based on the RECIST (version 1.1): PR (partial response) group and non-PR group. The PR group had a larger proportion of CD4 naïve T cells before and after treatment (Figure 2c) compared with the non-PR group, which indicated that the group that exhibited good efficacy reserved more CD4 naïve T cells to respond against multiple antigens. The patients in the PR group had more CD4 naïve T cells that might be activated by the antigen-MHC complex and differentiate into specific subtypes. These CD4+ T cells, especially Th1 cells, can equip CD8+ cytotoxic T lymphocytes to eliminate tumor cells. CD4+ T-cell mediated antitumour responses in the peripheral blood are gradually being valued. We further explored the relationship between the circulating TCRB diversity and therapeutic effect: the PR group had a higher circulating TCRB diversity than the non-PR group both before and after treatment (Figure 2b). This result is consistent with the changes in constituent ratios of CD4 naïve T cells and reminded us that the circulating TCRB diversity might quantitatively reflect the changes in the T lymphocyte composition. Besides, the circulating TCRB diversity and CD4 naïve T cells were not significantly altered during the first few targeted therapy cycles both in PR group and non-PR group (Figure 2b-c).

**Circulating TCRB diversity partly reflects adaptive immunological repertoire**

Each clone of the naïve T cells and central memory T cells (CD62L+ CCR7+) contains only a few cells. They recognize foreign antigens, activate, differentiate into effector or memory T cells and present the primary source to mount effective immune responses. They represent the ability of the immune system to recognize and clear nonself-antigens and they present high TCRB diversity. Effector T cells or effector memory T cells (CD62L± CCR7-) experienced antigen-specific clonal expansion, and each clone contains many cells. They present low TCRB diversity. Based on the relative abundance of lymphocyte subpopulations evaluated by the CIBERSORTx, the percentages of naïve T cells and CD8 Teff in total T cells were 29.18 ± 11.98(mean ± SD, %) and 14.5 ± 12.96(mean ± SD, %), respectively. Previous studies have shown that naïve T cells and CD4 Tcm display negligible clonal expansion while CD8 Teff is one of the most clonally expanded subgroups of T cells. Thus, circulating TCRB diversity depends strongly on the number of naïve T cells and central memory T cells and their relative proportions to effector T cells.

In our study, to explore whether the circulating TCRB diversity can reflect the host immune status and adaptive immunological repertoire, we performed Spearman correlation analyses between the circulating TCRB diversity and the constituent ratios of lymphocyte subsets (evaluated with CIBERSORTx) in each of the four groups and all patients before and after treatment. According to the results of the correlation analysis of the four groups, the circulating TCRB diversity tended to be positively correlated with the proportion of CD62L+ CCR7+ T cells and negatively correlated with the proportion of CD62L± CCR7- T cells proportion (Supplementary Figure S3a). The result of combining all patients is more significant due to increased sample size. The circulating TCRB diversity was positively correlated with the CD4 naïve T cell proportion before (rho = 0.39, p = .00096, Figure 3a) and after treatment (rho = 0.52, p < .0001, Figure 3b) and the CD8 naïve T cell proportion before (rho = 0.26, p = .029, Figure 3c) and after treatment (rho = 0.35, p = .0029, Figure 3d). In contrast, the circulating TCRB diversity was negatively correlated with the CD8 Teff proportion before (rho = −0.50, p < .0001, Figure 3e) and after treatment (rho = −0.58, p < .0001, Figure 3f). In addition, we found that the circulating TCRB diversity showed
Figure 1. Changes in the constituent ratios of peripheral blood major T lymphocyte subpopulations in the four groups before and after treatment. Changes in the constituent ratios of CD4 naïve T (a), CD4 Tcm (b), CD4 Teff and Tem (c), CD4 Treg (d), CD8 naïve T (e), CD8 Teff (f), and CD8 Tem (g) in the LUAD targeted therapy group (N = 21), the LUAD chemotherapy group (N = 16), the LUSC chemotherapy group (N = 10), and the SCLC chemotherapy group (N = 22) before and after treatment (paired Wilcoxon signed-rank test, *p < .05; **p < .01; ***p < .001; NS, P > .05 not significant). LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small cell lung cancer. CD4 naïve T, CD4 naïve T cells; CD4 Tcm, CD4-positive central memory T cells; CD4 Teff and Tem, CD4-positive effector T cells and effector memory T cells; CD4 Treg, CD4-positive regulatory T cells; CD8 naïve T, CD8 naïve T cells; CD8 Teff, CD8-positive effector T cells; CD8 Tem, CD8-positive effector memory T cells.
a positive correlation with the CD4 Tcm proportion (rho = 0.36, \( p = 0.0022 \), Figure 3g) and a negative correlation with the CD4 Teff and Tem proportion (rho = -0.31, \( p = 0.0084 \), Figure 3i) before treatment, but no similar correlation was found after treatment (Figure 3h, 3j). We obtained similar correlations with Absolute Immune Signal (ABIS) deconvolution (Supplementary Figure S3b-d).

We then defined a new parameter, the immune reserve index, to quantify the adaptive immunological repertoire using bulk RNA sequencing data. The immune reserve index was calculated using the following formula: 2(CD4 naïve T cell proportion + CD8 naïve T cell proportion + CD4 Tcm proportion)/2(CD8 Teff proportion + CD4 Teff and Tem proportion). Strong positive correlations were found between the circulating TCRB diversity and the immune reserve index in each of the four groups before (Supplementary Figure S4a-4d) and after treatment (Supplementary Figure S4e, S4f, S4h) except for the LUSC chemotherapy group after treatment (Supplementary Figure S4g). Although not statistically significant, there was a trend of a positive correlation between the immune reserve index and the circulating TCRB diversity in the LUSC chemotherapy group after treatment. Accordingly, the circulating TCRB diversity quantitatively reflects the relative proportions of CD62L+ CCR7 + T cells and CD62L± CCR7- T cells. Thus, the circulating TCRB diversity partly reflects the adaptive immunological repertoire and might be related to prognosis.

**Baseline circulating TCRB diversity is an independent prognostic factor for progression-free survival**

The presence of a preexisting immune response is often linked to a more favorable prognosis.\(^{42}\) To further assess the prognostic value of the baseline circulating TCRB diversity, we constructed two multivariate Cox proportional hazards regression models that included the following clinical parameters: the baseline circulating TCRB diversity, gender, age, smoking status, and pathological type or pharmacotherapeutic regimens.
Figure 3. Correlation between the circulating TCRβ diversity and the constituent ratios of lymphocyte subsets. Spearman correlation analysis between the circulating TCRβ (T cell receptor β chain) diversity and the constituent ratios of CD4 naïve T cells before (a) and after treatment (b), between the circulating TCRβ diversity and the constituent ratios of CD8 naïve T cells before (c) and after treatment (d), between the circulating TCRβ diversity and the constituent ratios of CD8 Teff before (e) and after treatment (f), between the circulating TCRβ diversity and the constituent ratios of CD4 Tcm before (g) and after treatment (h), and between the circulating TCRβ diversity and the constituent ratios of CD4 Teff and Tem before (i) and after treatment (j). For the Spearman correlation analysis, samples with matched RNAseq data and TCR sequencing data were included (N = 69). CD4 naïve T, CD4 naïve T cells; CD4 Tcm, CD4-positive central memory T cells; CD4 Teff and Tem, CD4-positive effector T cells and effector memory T cells; CD4 Treg, CD4-positive regulatory T cells; CD8 naïve T, CD8 naïve T cells; CD8 Teff, CD8-positive effector T cells; CD8 Tem, CD8-positive effector memory T cells.
The baseline circulating TCRB diversity, when treated as a continuous variable, was found to be an independent prognostic factor for progression-free survival in advanced lung cancer in the two models (Figure 4a, S5). A log10-transformed Shannon-Weiner index of 3.91 was determined as the optimal cutoff value of the baseline circulating TCRB diversity using the R package ‘survminer’. Because of the significant difference in PFS between the LUSC chemotherapy group and other groups, we specifically calculated a cutoff value of the baseline circulating TCRB diversity (3.535) for this group. We then split all the patients into a high-diversity group and a low-diversity group based on the corresponding optimal cutoff value. The Kaplan–Meier curves suggest that a high baseline circulating TCRB diversity predicts a better prognosis in patients with LUAD (log-rank analysis, \( p = .02 \), Figure 4b), LUSC (log-rank analysis, \( p = .03 \), Figure 4c) or SCLC (log-rank analysis, \( p = .005 \), Figure 4d). Hence, patients with a good baseline adaptive immunological repertoire exhibited a relatively good prognosis.

**Construction and validation of the five-factor prognostic risk score model**

We constructed a five-factor prognostic risk score model to achieve a more accurate evaluation of prognosis for patients with advanced lung adenocarcinoma. Using the LASSO Cox regression model and leave-one-out cross-validation, five factors were identified as being associated with PFS in the training dataset: smoke status, overlap index, pretreatment clonality, pretreatment platelet-lymphocyte ratio (PLR) index, and change of the circulating TCRB diversity. The risk score was obtained according to the following formula:

\[ \text{Risk score} = (2.114 \times \text{value of smoke status}) + (-10.26 \times \text{value of overlap index}) + (-0.4046 \times \text{value of pretreatment clonality}) + (8.304 \times \text{value of pretreatment PLR index}) + (1.655 \times \text{value of change of circulating TCRB diversity}) \]

We divided the patients into a high-risk group and a low-risk group according to the median value of the risk score. Kaplan–Meier survival curves obtained using the training dataset illustrate that patients in the low-risk group have longer PFS than those in the high-risk group (log-rank analysis, \( p = .003 \), Figure 5a). The same results were obtained with the testing dataset (log-rank analysis, \( p = .03 \), Figure 5b) and the entire dataset (log-rank analysis, \( p = .003 \), Figure 5c). The predictive accuracy of the five-factor prognostic risk score model was assessed by the C-indexes obtained for the training, testing and entire datasets: 0.736, 0.778, and 0.681, respectively (Figure 5d). In addition, we constructed a nomogram (Figure 5e) based on the LASSO Cox regression model that would help...
Clinicians predict the 6-month, 1-year and 3-year PFS probabilities for patients with advanced lung adenocarcinoma.

**Discussion**

We found that the EGFR-TKI targeted therapy or chemotherapy did not significantly affect the T lymphocyte composition or circulating TCRB diversity in patients with advanced lung adenocarcinoma over the first few treatment cycles. A higher circulating TCRB diversity was found to be associated with improved therapeutic effects in the LUAD targeted therapy group. In addition, the circulating TCRB diversity partly reflects the adaptive immunological repertoire. The baseline circulating TCRB diversity is an independent prognostic factor for progression-free survival. Moreover, the five-factor prognostic risk score model provided a good prediction of PFS in patients with advanced lung adenocarcinoma. These results suggested that the baseline circulating TCRB diversity might
serve as a predictive biomarker for patients with advanced lung cancer receiving targeted therapy or chemotherapy.

Recent studies have revealed that host systemic immunity is compromised across tumor types and that changes in the immune macroenvironment largely depend on the tumor burden.\(^{43}\) Tumor resection, which markedly decreases the tumor burden, can revert systemic immune disruptions and improve the global immune status in mouse breast cancer and colorectal cancer models and patients with renal cell carcinoma.\(^ {22,43}\) In our study, targeted therapy or chemotherapy did not strongly affect adaptive immune function over the first few treatment cycles. Other studies have drawn similar conclusions.\(^{22,44}\) In fact, we also explored whether the changes in the circulating TCRB diversity during therapy in the SCLC patients treated with etoposide with cisplatin are different from those treated with etoposide with carboplatin. We found that the difference in platinum salt did not affect the change of the circulating TCRB diversity (Supplementary Figure S6), and this finding was limited to the sample size. For patients with advanced lung cancer under heavy tumor burden, targeted therapy or chemotherapy is unable to reduce the tumor burden as fast as surgery. At present, immunotherapy, which strengthens the host immune system to fight cancer cells, has attracted growing attention. Han et al. examined the peripheral blood immunity of NSCLC patients receiving immunotherapy and revealed that the pre-therapy TCR diversity of peripheral PD-1 + CD8 + T cells predicts the clinical response to immunotherapy.\(^ {14}\) Although PD-1+ CD8 + T cells account for a relatively small percentage of T cells and might be insufficiently representative of systemic immunity, we speculate that the circulating TCRB diversity can monitor the efficacy of immunotherapy and accurately reflect changes in the immune macroenvironment. These findings suggest that the baseline global immune status is particularly important for patients with advanced lung cancer.

Tumor-infiltrating lymphocytes have attracted the attention of many researchers.\(^ {45}\) However, effective systemic immune responses are necessary for tumor eradication. In effective immunotherapy, changes in the T cell population in peripheral blood parallel those in secondary lymphoid organs, which are critical sites of T lymphocyte activation and proliferation.\(^ {46}\) In addition, researchers have observed the exchange of T-cell clones between tumors and peripheral blood during neoadjuvant immunotherapy.\(^ {17}\) Peripheral T cell cytotoxicity is closely correlated with the cytotoxicity of tumor-infiltrated T cells.\(^ {17}\) Peripheral blood, tissues and organs throughout the body and secondary lymphoid organs form a complete closed loop in which the three individual components interact with each other. Hence, the number, phenotype, function, and repertoire of peripheral blood T cells indirectly indicate the strength of the antitumour immune response. In the present study, the circulating TCRB diversity exhibited positive correlations with naïve T cell subsets and negative correlations with CD8 T effector before and after treatment. The correlations between the circulating TCRB diversity and the immune reserve index demonstrated that the circulating TCRB diversity reflects the relative proportions of CD62L+ CCR7 + T cells and CD62L+ CCR7- T cells. We can conclude that the circulating TCRB diversity reflects the flexibility of the adaptive immune response, the adaptive immunological repertoire of the host and the potential of antitumour immunity.

Prior research has shown that patients with advanced lung cancer exhibit a significantly lower TCRB diversity than healthy people.\(^ {16}\) Elimination, which is the first process of tumor immunoediting,\(^ {48}\) is markedly suppressed by damaging the ability of T cells to recognize tumor antigens. Effective tumor treatment requires restarting the elimination process. Our study proved that patients with a higher circulating TCRB diversity tend to exhibit improved therapeutic effects in the LUAD targeted therapy group. Furthermore, the baseline circulating TCRB diversity is an independent prognostic factor for the PFS of patients with advanced lung cancer. Patients with a higher baseline circulating TCRB diversity enjoyed longer PFS. This finding applies to patients with LUAD, patients with LUSC, and patients with SCLC. These results further proved that the circulating TCRB diversity represents the capacity of the antitumour immunity of the host. Patients with a high baseline circulating TCRB diversity not only have more naïve T cells to recognize tumor antigens and thus restart the elimination process and prevent immune escape but also provide a material basis for treatment-induced immune responses. In addition, a good prediction of PFS was obtained with the five-factor prognostic risk score model in patients with advanced lung adenocarcinoma.

The therapeutic efficacy of chemotherapeutic agents largely relies on the host immune response.\(^ {14,19}\) The treatment indirectly activates antitumour immune responses in advanced lung cancer. Several studies have reported that EGFR-TKI targeted therapy can trigger innate and adaptive immunity by enhancing NK cell-mediated tumor cell killing, dendritic cell-mediated antigen-presenting capabilities and cytotoxic CD8 + T cell-mediated tumor lysis.\(^ {4,5,17,49,50}\) In the current study, we also detected the activation of NK cells, CD8 + T cells and CD4 + T cells (Supplementary Figure S7a) in peripheral blood. Chemotherapy induces immunogenic tumor cell death that increases T cell recognition and subverts the immunosuppressive state through transient lymphodepletion or stimulatory effects on immune effectors.\(^ {3}\) In our GSEA, CD8 + and CD4 + T cells were activated in patients with NSCLC or SCLC after chemotherapy (Supplementary Figure S7b, c). This result provides a theoretical basis for combination immunotherapy with chemotherapy or targeted therapy. These effective immune responses are based on certain immunological repertoires. Overall, the global immune status is critical for patients with advanced lung cancer: patients with a good baseline global immune status are more likely to respond to treatment and achieve improved therapeutic effects. The baseline circulating TCRB diversity is an accessible, noninvasive and unbiased approach for evaluating the immune status and predicting prognosis.

Siminca, D. et al observed rarely shared T cell clones between paired pre- and posttreatment blood samples and theorized that rebound thymic activity reconstitutes the T cell repertoire after chemotherapy.\(^ {21}\) However, our analysis of the overlap in the T cell repertoire showed a clear clonal overlap before and after treatment in most patients (Supplementary Figure S8). The cumulative frequency of shared clones accounted for over 60% of the total frequency in our data. In
addition, chemotherapy-induced thymic involution might further reduce the output of naïve T cells.\textsuperscript{31} During treatment, homeostatic proliferation might be an important mechanism for the maintenance of the naïve T cell pool.\textsuperscript{38} The relative contribution of homeostatic proliferation and rebound thymic activity to maintaining T cell repertoire homeostasis requires further exploration.

As a preliminary study, some deficiencies were observed in our study. This study included patients with advanced lung cancer with different pathological types who treated with different treatment regimens. This is indeed our shortcoming. Moreover, a further significant limitation was the small sample size of the specific pathological type. To further verify our conclusions, we will enlarge the sample size in the following independent studies and explore changes and prognostic value of the circulating TCRB diversity in one specific pathological type lung cancer treated with one specific drug or combination during therapy. Besides, another limitation of our study was that we did not isolate PBMC subpopulations from blood. However, our method is simple to operate and specific for clinical application. We analyzed TCRB sequencing data in combination with transcriptome data. Without sampling at multiple time points, we were unable to capture the developmental trajectories of the circulating TCRB diversity throughout the treatment process. We did not further explore the relationship between T cell clones in peripheral blood and in tumors after treatment due the unavailability of tumor tissue from advanced patients.

In conclusion, therapy with chemotherapeutic agents cannot significantly impact the systemic immune over the first few treatment cycles, and a good global immune status is needed to restart the elimination process. The baseline circulating TCRB diversity is associated with the prognosis of patients with advanced lung cancer. The circulating TCRB diversity should be an important reference indicator for clinical antitumour treatment. The five-factor prognostic risk score model might be a worthy prognostic indicator in clinical applications. The clinician is supposed to combine the antitumour immune ability of the host and antitumour therapy organically on the basis of a comprehensive consideration, and this approach might make it possible to live with tumors.

**Acknowledgments**

We are grateful to the all the participants and their families. We would like to thank Dr. Liping Guo for help with data analysis and for help with preparing the manuscript.

**Funding**

This work was supported by the The National Key R&D Program of China under Grant [2018YFC1705104] and CAMS Innovation Fund for Medical Sciences (CIFMS) under Grant [2016-12M-3-005].

**ORCID**

Lin Feng http://orcid.org/0000-0003-0951-713X

**Declaration of interest**

The authors have declared no conflicts of interest.

**References**

1. Ferrara MG, Di Noia V, D’Argento E, Vita E, Damiano P, Cannella A, Ribelli M, Pilotto S, Milella M, Tortora G, et al. Oncogene-Addicted Non-Small-Cell Lung Cancer: treatment Opportunities and Future Perspectives. Cancers. 2020;12(5):1196. doi:10.3390/cancers12051196.

2. Arbour KC, Riedl GJ. Systemic Therapy for Locally Advanced and Metastatic Non–Small Cell Lung Cancer. Jama. 2019;322(8):764–774. doi:10.1001/jama.2019.11058.

3. Zitvogel L, Apetoh L, Ghiringhelli F, Kroemer G. Immunological aspects of cancer chemotherapy. Nat Rev Immunol. 2008;8(1):59–73. doi:10.1038/nri2216.

4. Liu Z, Han C, Dong C, Shen A, Hsu E, Ren Z, Lu C, Liu Z, Zhang A, Timmerman C, et al. Hypofractionated EGFR tyrosine kinase inhibitor limits tumor relapse through triggering innate and adaptive immunity. Sci Immunol. 2019;4(38):eaav6473. doi:10.1126/sciimmunol.aaav6473.

5. Jia Y, Li X, Jiang T, Zhao S, Zhao C, Zhang L, Liu X, Shi J, Qiao M, Luo J, et al. EGFR-targeted therapy alters the tumor microenvironment in EGFR-driven lung tumors: implications for combination therapies. Int J Cancer. 2019;145(5):1432–1444. doi:10.1002/ ijc.32191.

6. Prendergast GC, Jaffee EM. Cancer Immunologists and Cancer Biologists: why We Didn’t Talk Then but Need to Now. Cancer Res. 2007;67(8):3500–3504. doi:10.1158/0008-5472.CAN-06-4626.

7. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. Nature. 1988;334(6181):395–402. doi:10.1038/334359a0.

8. Girardi M. Immunosurveillance and Immunoregulation by γδ T Cells. J Invest Dermatol. 2006;126(1):25–31. doi:10.1038/jid.5700003.

9. Davis MM, Boniface JJ, Reich Z, Lyons D, Hamp1 J, Arden B, Chien Y-H. LIGAND RECOGNITION BY αβ T CELL RECEPTORS. Annu Rev Immunol. 1998;16(1):523–544. doi:10.1146/annurev.immunol.16.1.523.

10. Garcia KC, Teyton L, Wilson IA. Structural basis of T cell recognition. Annu Rev Immunol. 1999;17(1):369–397. doi:10.1146/annurev.immunol.17.1.369.

11. Reuben A, Zhang J, Chioo S-H, Gittelman RM, Li J, Lee W-C, Fujimoto J, Behrens C, Liu X, Wang F, et al. Comprehensive T cell repertoire characterization of non-small cell lung cancer. Nat Commun. 2020;11(1):603. doi:10.1038/s41467-019-14273-0.

12. Joshi K, De Massy MR, Ismail M, Reading JL, Uddin I, Woolston A, Hatipoglu E, Oakes T, Rosenthal R, Peacock T, et al. Spatial heterogeneity of the T-cell receptor repertoire reflects the mutational landscape in lung cancer. Nat Med. 2019;25(10):1549–1559. doi:10.1038/s41591-019-0592-2.

13. Zhang J, Ji Z, Caushi JK, El Asmar M, Anagnostou V, Cottrell TR, Chan HY, Suri P, Guo H, Merghoub T, et al. Compartimental Analysis of T-cell Clonal Dynamics as a Function of Pathologic Response to Neoadjuvant PD-1 Blockade in Resectable Non–Small Cell Lung Cancer. Am J Cancer Res. 2020;26(6):1327–1337. doi:10.1158/1078-0432.CCR-19-2931.

14. Han J, Duan J, Bai H, Wang Y, Wan R, Wang X, Chen S, Tian Y, Di W, Fei K, et al. TCR Repertoire Diversity of Peripheral PD-1 + CD8 + T Cells Predicts Clinical Outcomes after Immunotherapy in Patients with Non–Small Cell Lung Cancer. Cancer Immunol Res. 2020;8(11):146–154. doi:10.1158/2326-6066.CIR-19-0398.

15. Charles J, Mouret S, Challende I, Leccia M-T, De Fraipont F, Perez S, Plantier N, Plumas J, Manuel M, Chaperot L, et al. T-cell receptor diversity as a prognostic biomarker in melanoma patients. Pigment Cell Melanoma Res. 2020;33(4):612–624. doi:10.1111/pcm.12866.
16. Liu -Y, Yang Q-F, Yang J-S, Cao R-B, Liang J-Y, Liu Y-T, Zeng Y-L, Chen S, Xia X-F, Zhang K, et al. Characteristics and prognostic significance of profiling the peripheral blood T-cell receptor repertoire in patients with advanced lung cancer. Int J Cancer. 2019;145(5):1423–1431. doi:10.1002/ijc.32145.

17. Dominguez C, Tsang K-Y K-Y, Falena C. Short-term EGFR blockade enhances immune-mediated cytotoxicity of EGFR mutant lung cancer cells: rationale for combination therapies. Cell Death Dis. 2016;7(9):e2380–e. doi:10.1038/cddis.2016.297.

18. Liu W, Zhang L, Xiu Z, Guo J, Wang L, Zhou Y, Jiao Y, Sun M, Cai J. cDNA:Combination of Immune Checkpoint Inhibitors with Chemotherapy in Lung Cancer. [cdata(oncotargets and Therapy)]. 2020;7:2299–2321. doi:10.2147/OTTT.S255491.

19. Galluzzi L, Humeau J, Buqué A, Zitvogel L, Kroemer G. Immunostimulation with chemotherapy in the era of immune checkpoint inhibitors. Nat Rev Clin Oncol. 2020;17(12):725–741. doi:10.1038/s41571-020-0413-z.

20. Zhang C, Palashati H, Tan Q, Ku W, Miao Y, Xiong H, Lu Z. Immediate and substantial evolution of T-cell repertoire in peripheral blood and tumor microenvironment of patients with esophageal squamous cell carcinoma treated with preoperative chemotherapy. Carcinogenesis. 2018;39(11):1389–1398. doi:10.1093/carcin/bgy116.

21. Siminca D, Akıyüz N, Schlipfke S, Mohme M, V.wenserski L, Mährle T, Fanchi LF, Lamskus K, Binder M, wenserski L, Mährle T, et al. T-cell receptor next-generation sequencing reveals cancer-associated repertoire metrics and reconstitution after chemotherapy in patients with hematologic and solid tumors. OncoImmunology. 2019;8(11):e1644110. doi:10.1080/2162402X.2019.1644110.

22. Gustafsson CE, Jadhav R, Cao W, Qi Q, Pegramp M, Tian L, Weyand CM, Goronzy JJ. Immune cell repertoires in breast cancer patients after adjuvant chemotherapy. JCI Insight. 2020;5(4). doi:10.1172/jci.insight.134569.

23. Shugay M, Britanova OV, Merzlyak EM, Turchaninova MA, Mamedov IZ, Tuganbaev TR, Bolotin DA, Staroverov DB, Putintseva EV, Plevkova K, et al. Towards error-free profiling of immune repertoires. Nat Methods. 2014;11(6):653–655. doi:10.1038/nmeth.2960.

24. Guo L, Bi X, Li Y, Wen L, Zhang W, Jiang W, Ma J, Feng L, Zhang K, Shou J, et al. Characteristics, dynamic changes, and prognostic significance of TCR repertoire profiling in patients with renal cell carcinoma. J Pathol. 2020;251(1):26–37. doi:10.1002/path.5396.

25. Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, Tang B, Dong L, Ding N, Zhang Q, et al. GSA: genome Sequence Archive. Genomics Proteomics Bioinformatics. 2017;15(1):14–18. doi:10.1006/jjp.2017.01.001.

26. Database Resources of the National Genomics Data Center in. Database Resources of the National Genomics Data Center in 2020.. Nucleic Acids Res. 2020;48(D1):D24–D33. doi:10.1093/nar/gkz913.

27. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, Chudakov DM. MIXCR: software for comprehensive adaptive immunity profiling. Nat Methods. 2015;12(5):380–381. doi:10.1038/nmeth.3364.

28. Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, Pogorelyy MV, Nazarov VI, Zvyagina IV, Kirgizova VL, et al. VDJtools: unifying Post-analysis of T Cell Receptor Repertoires. PLoS Comput Biol. 2015;11(11): e1004503. doi:10.1371/journal.pcbi.1004503.

29. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773–782. doi:10.1038/s41587-019-0114-2.

30. Monaco G, Lee B, Xu W, Mustafah S, Huang YY, Carré C, Burdin N, Visan L, Ceccarelli M, Poidinger M, et al. RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute Deconvolution of Human Immune Cell Types. Cell Rep. 2019;26(6):1627–1640. doi:10.1016/j.celrep.2019.01.041.

31. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16(5):284–287. doi:10.1089/omi.2011.0118.

32. Julia EP, Mandò P, Rizzo MM, Cueto GR, Tsou F, Luca R, Papareli C, Bravo AI, Astorino W, Mordoh J, et al. Peripheral changes in immune cell populations and soluble mediators after anti-PD-1 therapy in non-small cell lung cancer and renal cell carcinoma patients. Cancer Immunol Immunother. 2019;68(10):1585–1596. doi:10.1007/s00262-019-02391-z.

33. Tseng C-W, Hung C-F, Alvarez RD, Trimble C, Huh WK, Kim D, Chuang C-M, Lin C-T, Tsai Y-C, He L, et al. Pretreatment with Cisplatin Enhances E7-Specific CD8+ T-Cell–Mediated Antitumor Immunity Induced by DNA Vaccination. Am J Cancer Res. 2008;14(10):3185–3192. doi:10.1172/jci.insight.134569.

34. Zheng W, Zhang Z. Understanding T-cell responses to tumor antigens. Nat Rev Immunol. 2010;10(4):245–485. doi:10.1038/nri2770.

35. Zitvogel L, Kroemer G. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol. 2017;14(12):717–734. doi:10.1038/nrclinonc.2017.101.

36. Allen BM, Hiam KJ, Burnett CE, Venida A, DeBarge R, Tovooren O, Marquez I, Cho MW, Carmi Y, Spitzen MH, et al. Systemic dysfunction and plasticity of the immune macroenvironment in cancer models. Nat Med. 2020;26(7):1125–1134. doi:10.1038/s41591-020-0892-6.

37. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, Chudakov DM. MIXCR: software for comprehensive adaptive immunity profiling. Nat Methods. 2015;12(5):380–381. doi:10.1038/nmeth.3364.

38. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773–782. doi:10.1038/s41587-019-0114-2.
47. Iwahori K, Shintani Y, Funaki S, Yamamoto Y, Matsumoto M, Yoshida T, Morimoto-Okazawa A, Kawashima A, Sato E, Gottschalk S, et al. Peripheral T cell cytotoxicity predicts T cell function in the tumor microenvironment. Sci Rep. 2019;9(1):2636. doi:10.1038/s41598-019-39345-5.

48. Miller JAP, Sadelain M. The journey from discoveries in fundamental immunology to cancer immunotherapy. Cancer Cell. 2015;27(4):439–449. doi:10.1016/j.ccell.2015.03.007.

49. Mallmann-Gottschalk N, Sax Y, Kimmig R, Lang S, Brandau SEGFR. EGFR-Specific Tyrosine Kinase Inhibitor Modifies NK Cell-Mediated Antitumoral Activity against Ovarian Cancer Cells. Int J Mol Sci. 2019;20(19):4693. doi:10.3390/ijms20194693.

50. Chu C-L, Lee Y-P, Pang C-Y, Lin H-R, Chen C-S, You R-I. Tyrosine kinase inhibitors modulate dendritic cell activity via confining c-Kit signaling and tryptophan metabolism. Int Immunopharmacol. 2020;82:106357. doi:10.1016/j.intimp.2020.106357.

51. Wang W, Thomas R, Sizova O, Su D-M. Thymic Function Associated With Cancer Development, Relapse, and Antitumor Immunity – a Mini-Review. Front Immunol. 2020;11:773. doi:10.3389/fimmu.2020.00773.