

RESEARCH

Open Access



# The intratumoral landscape of T cell receptor repertoire in esophageal squamous cell carcinoma

Meng Yang<sup>1</sup>, Dan He<sup>1</sup>, Yu Sun<sup>2</sup>, Yunquan Guo<sup>1</sup>, Yu Ma<sup>3</sup>, Lin Feng<sup>2</sup> and Meng Liu<sup>1\*</sup>

## Abstract

**Background** Esophageal squamous cell carcinoma (ESCC) is a malignant neoplasm with detrimental implications for human health. The landscape of ESCC therapy has been revolutionized by the introduction of immunotherapy, specifically involving immune checkpoint inhibitors (ICIs). A number of studies have documented the prognostic significance of T-cell receptor (TCR) repertoire and its association with many tumors. Nevertheless, the TCR repertoire landscape and its significance in ESCC still need to be explored.

**Methods** In this study, we conducted RNA-Seq analysis to investigate the characteristics of the TCR repertoire in 90 patients. Moreover, high-throughput TCR sequencing was performed on tumor tissues from 41 patients who received immunotherapy. Additionally, a comprehensive analysis of the T-cell receptor repertoire landscape within ESCC tumors was carried out through immunohistochemical staining on all patient samples.

**Results** We noticed a diminished diversity of TCR repertoire within the tumor compared to its adjacent normal tissue. In terms of immunotherapy responses, non-responsive patients exhibited higher TCR repertoire diversity indices and an increased frequency of common V and J genes. Additionally, elevated TCR repertoire diversity correlated with improved overall survival rates. Lastly, immunohistochemical staining results indicated a correlation between TCR repertoire diversity and the tumor immune microenvironment (TIME).

**Conclusions** Our study primarily describes the landscape of TCR repertoires in ESCC through three aspects: differences in tumor tissues, immune response to immunotherapy, and survival prognosis of patients. These results emphasize the importance of TCR repertoire characteristics as unique and relevant biomarkers for ESCC immunotherapy.

**Keywords** T cell receptor repertoire, Immunotherapy response, Esophageal squamous cell carcinoma, Overall survival

\*Correspondence:

Meng Liu

liuxiaomeng198154@163.com

<sup>1</sup>Affiliated Tumor Hospital of Xinjiang Medical University, Xinjiang Uygur Autonomous Region, Urumqi 830011, PR China

<sup>2</sup>State Key Laboratory of Molecular Oncology, Department of Etiology and Carcinogenesis, National Clinical Research Center for Cancer/Cancer Hospital, National Cancer Center, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100021, PR China

<sup>3</sup>The Fourth People's Hospital of Urumqi, Xinjiang Uygur Autonomous Region, Urumqi 830002, PR China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Introduction

Esophageal cancer is a malignant neoplasm originating from the esophageal epithelial cells, representing one of the most prevalent and lethal malignancies within the gastrointestinal tract, posing a significant threat to human well-being [1]. According to the Global Burden of Disease (GBD) statistics, the global incidence of esophageal cancer cases rose from 310,000 in 1990 to 640,000 in 2020, indicating an 88.79% increase [2]. In China, esophageal squamous cell carcinoma (ESCC) is one of the most common types of esophageal cancer, accounting for over 90% of all cases [3]. Due to the elusive and highly invasive nature of ESCC, many patients are already in advanced stages of ESCC at the time of diagnosis, missing the optimal treatment window, which results in a five-year survival rate of only 30% [4]. In recent years, the development of immune checkpoint inhibitors (ICIs) has made significant strides and fundamentally revolutionized the landscape of cancer treatment; however, only 20–40% of patients experience benefits from ICIs and even fewer achieve long-term disease control [5–9]. The composition of tumor-infiltrating lymphocytes (TILs) in ESCC is related to patient survival prognosis, and a higher proportion of TILs infiltrating the tumor microenvironment indicates a more favorable response to treatment [10, 11]. Previous research has demonstrated the strong relationship between T-cell immunity and the formation of tumors, as well as the clinical outcomes of various tumor types, such as breast cancer, melanoma, stomach, bladder, lung, and pancreatic cancer [12–17]. Consequently, in order to more accurately identify individuals who can potentially benefit from immunotherapy, a deeper comprehension of the relationship between T cells and tumors is required.

T cells play a crucial role in immune responses, and as key participants in the adaptive immune process, their surface T cell receptors (TCRs) also play a vital role. By recognizing novel antigens through the major histocompatibility complex (MHC), T cells specifically activate and expand via their TCRs to combat diverse pathogens [18, 19]. The specificity and diversity of TCRs are mainly determined by the highly variable complementarity-determining region 3 (CDR3); CDR3 is encoded in the V-D-J or V-J gene rearrangement regions and directly contacts antigen peptides when presented in the MHC binding groove [20, 21]. Therefore, CDR3 can provide abundant information about TCR specificity and is commonly targeted for TCR sequencing. It is worth noting that the TCR repertoire often co-evolves with new antigens, indicating that it can reflect self-immune status and disease progression. Many studies have also revealed the characteristics of the TCR repertoire in different types of cancer, and by monitoring the clonality and diversity of the TCR repertoire, the response to treatment can

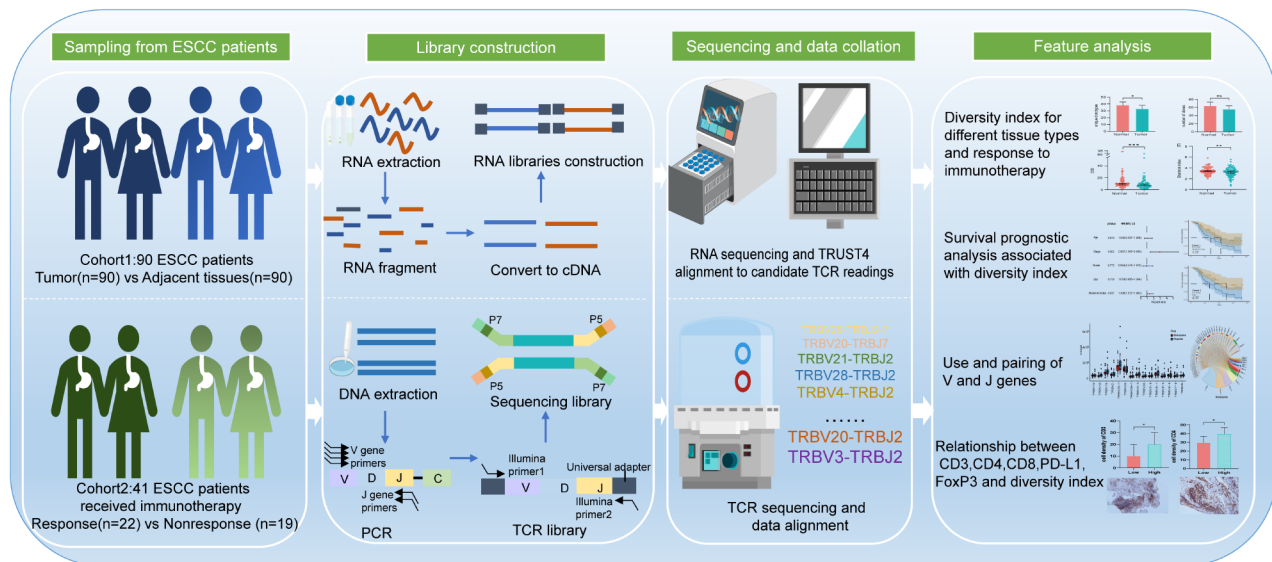
be dynamically evaluated [12, 22]. Currently, there is a paucity of research on TCR repertoire in ESCC. Cihui Yan's study revealed the diversity of tumor-infiltrating TCRs and their interplay between tumor and peripheral blood when treating ESCC patients with radiotherapy combined with Camrelizumab [23]. However, further exploration is needed to investigate the TCR repertoire characteristics of ESCC patients and the relationship between TCR repertoire and clinical features as well as prognosis.

In order to investigate the TCR repertoire landscape of ESCC, we conducted an analysis of ESCC patients from two cohorts. These analyses included tumor tissue and its adjacent normal tissue, immune therapy response, and patient survival prognosis. Additionally, we conducted immunohistochemical staining to investigate the expression of immune cell-related biomarkers, with a primary focus on T-cell subtypes (CD3, CD4, CD8, PD-L1, and FoxP3). This was performed to elucidate the correlation between TCR repertoire diversity and the tumor immune microenvironment (TIME). Please refer to Fig. 1 for specific collection and analysis strategies. In summary, our data reveal differences in the TCR repertoire within the tumor tissues of ESCC patients, which may have potential clinical value in terms of immune therapy and survival prognosis.

## Materials and methods

### Study cohort

Cohort 1: The transcriptome data and clinical features of 90 patients with tumor tissues and adjacent normal tissues were obtained from the Affiliated Tumor Hospital of Xinjiang Medical University and are accessible through the National Genomics Data Center (GSA-Human: HRA006108) [24]. Cohort 2: This cohort comprised paraffin-embedded tissue specimens collected from 41 ESCC patients who underwent endoscopic biopsy at the Affiliated Tumor Hospital of Xinjiang Medical University between 2021 and 2023. All 41 patients in Cohort 2 received their initial immunotherapy, which involved a treatment regimen combining anti-PD1 drugs with chemotherapy. The dosage was administered based on individual body surface area, and efficacy was evaluated after a continuous period of 2–3 cycles. Based on the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 standard, the clinical response level was evaluated, with partial response (PR) or complete response (CR) defined as responsive patients, while patients experiencing disease progression (PD) or stable disease (SD) were defined as non-responsive patients [25]. Representative imaging is depicted in Figure S1 (Figure S1A–B). The tumor-node-metastasis (TNM) classification, based on the 8th edition of The American Joint Committee on Cancer (AJCC) guidelines, was confirmed by independent examination



**Fig. 1** Overview of the study design. The study utilized RNA expression profile data from ESCC patient samples at the Affiliated Tumor Hospital of Xinjiang Medical University (cohort 1,  $n = 90$ ), as well as TCR sequencing data from samples of ESCC patients who received immunotherapy for the first time (cohort 2,  $n = 41$ ). The analysis focused on TCR repertoire features, including unique clone types, total clone numbers, top clone frequencies, diversity index, and differences in CDR3 aa motifs between tumor tissue and adjacent normal tissue, as well as between patients who responded or did not respond to immunotherapy. Immunohistochemical staining confirmed the association of the TCR repertoire with TIME and suggested that the TCR repertoire diversity index could potentially predict survival prognosis in ESCC patients

by two pathologists for all samples. Overall survival (OS) was defined as the duration in months from ESCC diagnosis to death due to any cause. Complete clinical information of patients in both cohorts, including age, gender, alcohol consumption, smoking history, TNM stage, clinical stage, and tumor grade (Table 1), was available. The present study obtained approval from the Ethics Committee of Xinjiang Medical University Affiliated Tumor Hospital, with approval number K-2,023,006, and strictly adhered to applicable standards and regulations in accordance with the Helsinki Declaration (revised in 2013). Written informed consent was obtained from all participating patients.

### High-throughput TCR sequencing

The genomic DNA was extracted from formalin-fixed paraffin-embedded (FFPE) specimens of the 41 ESCC patients in cohort 2, using a DNA extraction kit (abs60301, Absin, Shanghai). Subsequently, the concentration of DNA was measured using a NanoDrop One ultraviolet spectrophotometer (NanoDrop One, Thermo Fisher Scientific, America). Next, deep TCR  $\beta$  chain (TRB) high-throughput sequencing analysis was performed on DNA samples with qualified concentrations. In brief, a multiplex PCR amplification strategy targeting the CDR3 of TRB was employed. Firstly, DNA underwent the first round of PCR (PCR-1) using 31 forward TRBV primers and 12 reverse TRBJ primers from the Multiplex PCR Kit (QIAGEN, Germany) to amplify

the CDR3 region of TRB. During PCR-1, amplification was achieved through a cycling reaction under different temperature conditions: 30 s at 94 °C, 30 s at 68 °C, and 60 s at 72 °C, with a final extension of 10 min at 72 °C. Subsequently, Agenecourt AMPure XP beads (A63882, Beckman Coulter Inc) were used to purify and remove impurities and unamplified DNA molecules from the amplified TRB products. The purified products then underwent a second round of PCR (PCR-2) with Illumina sequencing indexes, followed by another purification step using Agenecourt AMPure XP beads. Finally, the purified PCR products were diluted to appropriate concentrations for sequencing sample preparation and sequenced on an Illumina NovaSeq system, following standard protocols with paired-end reads of 150 bp in length.

### Reprocessing of TCR data

**Cohort 1:** From a large RNA sequencing dataset, the TCR repertoire was extracted using the TRUST4 program. Appropriate criteria were chosen in accordance with TRUST4's requirements. Raw sequence files were used as input, and TRUST4 applied strict overlap criteria to each read or read pair to find possible reads [26]. During the assembly of candidate segments into immune receptor sequences, TRUST4 employed a consensus-based assembly approach using read overlaps. Reads with somatic hypermutations were clustered within the same overlapping group. The TRUST4 tool aligns assembled overlapping clusters with sequences from the International

**Table 1** Table of cohort characteristics. Clinical and pathological characteristics of ESCC patients divided by different cohorts.

|                      | Cohort 1<br>N=90 | Cohort 2<br>N=41 |
|----------------------|------------------|------------------|
| Gender:              |                  |                  |
| Female               | 28 (31.1%)       | 18 (43.9%)       |
| Male                 | 62 (68.9%)       | 23 (56.1%)       |
| Age(years):          |                  |                  |
| <=60                 | 45 (50.0%)       | 19 (46.3%)       |
| >60                  | 45 (50.0%)       | 22 (53.7%)       |
| Race:                |                  |                  |
| Chinese Han          | 31 (34.4%)       | 17 (41.5%)       |
| Chinese other ethnic | 59 (65.6%)       | 24 (58.5%)       |
| Smoking status:      |                  |                  |
| No                   | 60 (66.7%)       | 26 (63.4%)       |
| Yes                  | 30 (33.3%)       | 15 (36.6%)       |
| Drinking status:     |                  |                  |
| No                   | 72 (80.0%)       | 33 (80.5%)       |
| Yes                  | 18 (20.0%)       | 8 (19.5%)        |
| Stage:               |                  |                  |
| Stage I-II           | 32 (35.6%)       | 6 (14.6%)        |
| Stage III-IV         | 58 (64.4%)       | 35 (85.4%)       |
| T stage:             |                  |                  |
| T1-T2                | 14 (15.6%)       | 3 (7.32%)        |
| T3-T4                | 76 (84.4%)       | 38 (92.7%)       |
| N stage:             |                  |                  |
| N0                   | 32 (35.6%)       | 7 (17.1%)        |
| N1-N3                | 58 (64.4%)       | 34 (82.9%)       |
| Tumor grade:         |                  |                  |
| G1                   | 13 (14.4%)       | 5 (12.2%)        |
| G2                   | 60 (66.7%)       | 28 (68.3%)       |
| G3                   | 17 (18.9%)       | 8 (19.5%)        |

ImMunoGeneTics information system (IMGT) database (<https://github.com/repseqio/library-imgt/releases/tag/v5>) to accurately identify V and J genes and determine CDR3 coordinates. This open-source software is specifically designed for repertoire analysis based on RNA-Seq data and can be accessed at <https://github.com/liulab-dfci/TRUST4>. The analysis was conducted with the aim of ensuring its robustness, resulting in a final dataset comprising 180 samples.

Cohort 2: The generated raw TCR sequencing data in FASTQ format undergoes quality control processing, including the removal of reads contaminated with adapter sequences, poor-quality reads, and reads with high unknown base content. Then, further analysis is performed on the remaining high-quality clean data. The IgBLAST tool provided by NCBI is used to annotate and align the clean data with the IMGT database, followed by PCR amplification and correction of sequencing errors based on clone frequency. Based on the downloaded TCR reference genome, information about TRB V, J, and D genes in each sequence is obtained. Sequences containing corresponding V, J, and D gene segments are

converted into TRB CDR3 nucleotide (nt) and amino acid (aa) sequences, while excluding samples identified as out-of-frame or stop codon sequences. The raw TCR sequencing data generated in this study in FASTQ format has been uploaded to the China National GeneBank Data Center (CNGDB) under accession code HRA006108 and can be accessed publicly through <https://ngdc.cncb.ac.cn/gsa-human> [27, 28].

#### TCR repertoire analysis

The Shannon Index and D50 are used to define the diversity and evenness of the TCR repertoire in ESCC patients. The D50 metric, a novel immunological diversity measure, quantifies the minimum number of unique clonotypes that constitute at least 50% of all sequencing reads. The richness of TCR clones in a sample is evaluated using the Shannon Index, where higher values indicate greater variety. The top 100 TCR clone types, referred to as TOP100 clones, were determined based on their proportion. The clone distribution, diversity, and overlap of TCRs were examined using the ‘vegan’ and ‘immunarch’ packages in R, and VDJtools (<https://github.com/mikessh/vdjtools>). Kaplan-Meier (K-M) survival curves were created using the ‘survival’ and ‘survminer’ R packages, and univariate Cox analysis was performed to evaluate the relationship between the diversity index and OS. Additionally, the R packages ‘pROC’ and ‘ggplot2’ were used to plot the Receiver Operating Characteristic Curve (ROC) of diversity indices at ESCC levels and to calculate the Area Under the Curve (AUC) values. When the AUC value is greater than 0.6, it is considered reliable.

#### Immunohistochemical staining

The paraffin-embedded specimens were cut into continuous slices with a thickness of 5  $\mu$ m. After dewaxing and rehydration, the slices were immersed in 3% H<sub>2</sub>O<sub>2</sub> for 20 min, followed by rinsing. Antigen retrieval was performed by adding citrate buffer and heating at high temperature for 5 min, followed by cooling to room temperature and repeating the same steps once more. After blocking with serum at room temperature, the tissue sections were covered separately with primary antibodies CD3 (ZM-0417, ZSGB-Bio, Beijing), CD4 (ZA-0519, ZSGB-Bio, Beijing), CD8 (ZA-0508, ZSGB-Bio, Beijing), PD-L1 (ab213524, Abcam, UK), and FoxP3 (ab20034, Abcam, UK) overnight at 4 °C. Subsequently, the sections were incubated with rhodamine-labeled secondary antibody (PV-6000, ZSGB-Bio, BJ) for 20 min at 37 °C. Add 100 $\mu$ L of DAB chromogen solution (ZLI-9019, ZSGB-Bio, Beijing) and treat at room temperature for 5 min. After the washing steps were completed, counterstaining was performed with hematoxylin-eosin staining solution, followed by differentiation using tap water containing hydrochloric acid-alcohol solution. Finally, the

dehydration process was carried out using xylene, and the slides were mounted with neutral gum. The expression of CD3, CD4, CD8, PD-L1, and FoxP3 was blindly evaluated by two independent pathologists with senior professional titles. The average value was calculated based on the percentage of positive cells as the immunohistochemical staining result.

### Statistical analysis

The Student t-test or Mann-Whitney U test was employed to evaluate disparities in the diversity index, V(D)J gene, and immunohistochemical scores for T cell markers (CD3, CD4, CD8, PD-L1 and FoxP3) among efficacy groups. Fisher's exact test was utilized for comparisons of non-ordinal categorical variables, while Pearson's chi-square test was employed for comparisons of ordinal categorical variables. Survival analysis of OS data was conducted using the K-M method, accompanied by a log-rank test. GraphPad Prism 10.0 and R 4.3.1 software packages were utilized for data visualization and statistical calculations. The differences are considered statistically significant at  $P < 0.05$  (\*),  $P < 0.01$  (\*\*), and  $P < 0.001$  (\*\*\*)

## Results

### Patient characteristics

To investigate the differences in TCR repertoire features among ESCC patients, we collected RNA sequencing data from 90 tumor samples and their corresponding adjacent normal tissues, extracting TCR repertoire characteristics (cohort 1). Additionally, TCR sequencing was performed on 41 ESCC tumor samples from patients receiving immunotherapy for the first time (cohort 2). Table 1 displays the clinical features of all 131 ESCC patients (Table 1). The median age of the patients was 61 years, with more than half being male (64.88%) or from ethnic minorities (63.35%). According to the eighth edition staging system of AJCC guidelines, there were 38 patients (29.77%) in stage II, 83 patients (63.35%) in stage III, and 10 patients (7.63%) in stage IV. Among the 41 patients who received immunotherapy combined with chemotherapy, Camrelizumab was administered to 13 cases (31.71%), Tislelizumab to 12 cases (29.27%), Sintilimab to 10 cases (24.39%) and Toripalimab to eight cases (19.51%). Cohort 2 treatment strategies are summarized in supplementary Table 1 (Table S1).

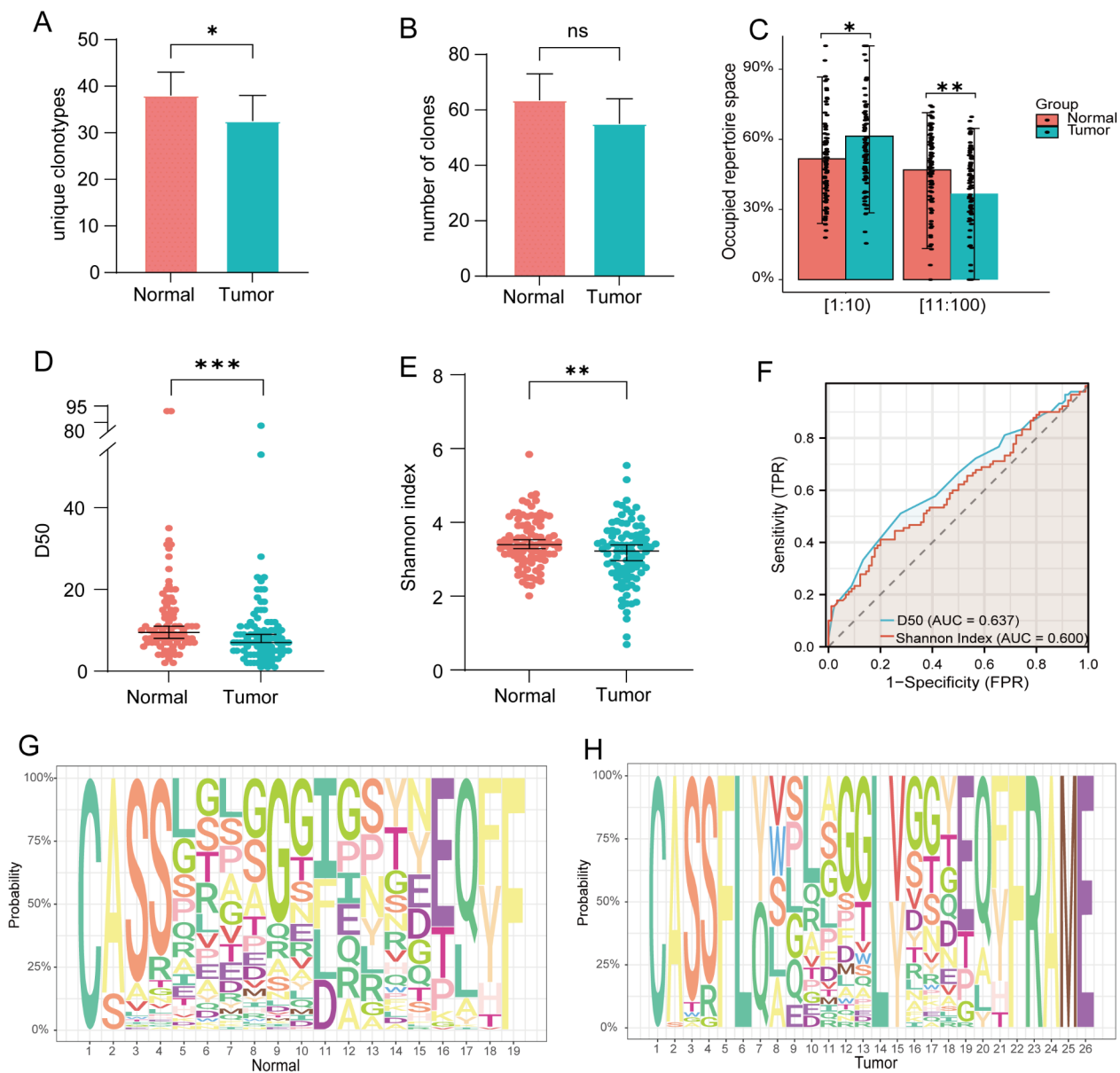
### Compared to the adjacent normal tissues, TCR repertoire diversity within the tumor tissue is reduced

We measured clone counts, unique clone counts, and utilization of top clones in both tissue types to examine the differences in the TCR repertoire between ESCC tumor tissues and adjacent normal tissues. The results revealed significant disparities in unique clone species

and the frequency of top clones between ESCC tumor tissues and adjacent normal tissues. However, there was no discernible distinction in overall clone numbers (Fig. 2A-C). Using the richness and evenness indicators D50 and Shannon Index, we compared the tumor tissue to the adjacent normal tissue. The findings showed that the tumor tissue had lower D50 and Shannon Index values than the adjacent normal tissue (Fig. 2D-E). Furthermore, we assessed the diagnostic potential of D50 and Shannon Index for ESCC and its adjacent tissues using ROC curves. The findings demonstrated that both D50 and Shannon Index exhibited a high diagnostic value for ESCC ( $AUC > 0.60$ ) (Fig. 2F). The reduced diversity of tumor tissue may be ascribed to the preferential selection of specific clones during tumor infiltration. Additionally, in the process of gene rearrangement, multiple codons can encode the same amino acid, and different V, J and D rearrangements can ultimately yield identical TCR amino acid sequences. This phenomenon is referred to as TCR convergence or convergent recombination (CR). The richness of CDR3 nt sequences also serves as an indicator of TCR clone characteristics; therefore, we further investigated the disparities in CDR3 aa motifs between the two tissue types and observed a higher abundance of these motifs in adjacent normal tissues (Fig. 2G-H).

### The TCR repertoire is associated with immune therapy response

Immunotherapy has been extensively utilized in the treatment of ESCC, and there is a growing clinical need for predicting its efficacy. As the TCR repertoire evolves concomitantly with novel antigens over time, it implies that specific TCR repertoires can reflect immune response characteristics and disease progression [29]. Therefore, we collected tumor tissue samples from 41 ESCC patients who were undergoing immunotherapy combined with chemotherapy for the first time at baseline status and conducted TCR sequencing analysis to investigate the association between immunotherapy and TCR repertoire features. Our study unveiled significant variations in the effectiveness of immunotherapy across different treatment modalities, independent of factors such as age, gender, and ethnicity (Table S2). Next, individual patients were classified as either responders or non-responders based on how well their immunotherapy worked. We found that respondents had a higher percentage of the top 100 clones in terms of frequency as well as a greater diversity in unique clone species. However, although there were more clones overall in the non-responders' display, there was no statistically significant difference (Fig. 3A-C). Moreover, using the Jaccard method to evaluate TCR repertoire overlap, we found low overlap between responders and non-responders, suggesting significant differences in TCR repertoire clone

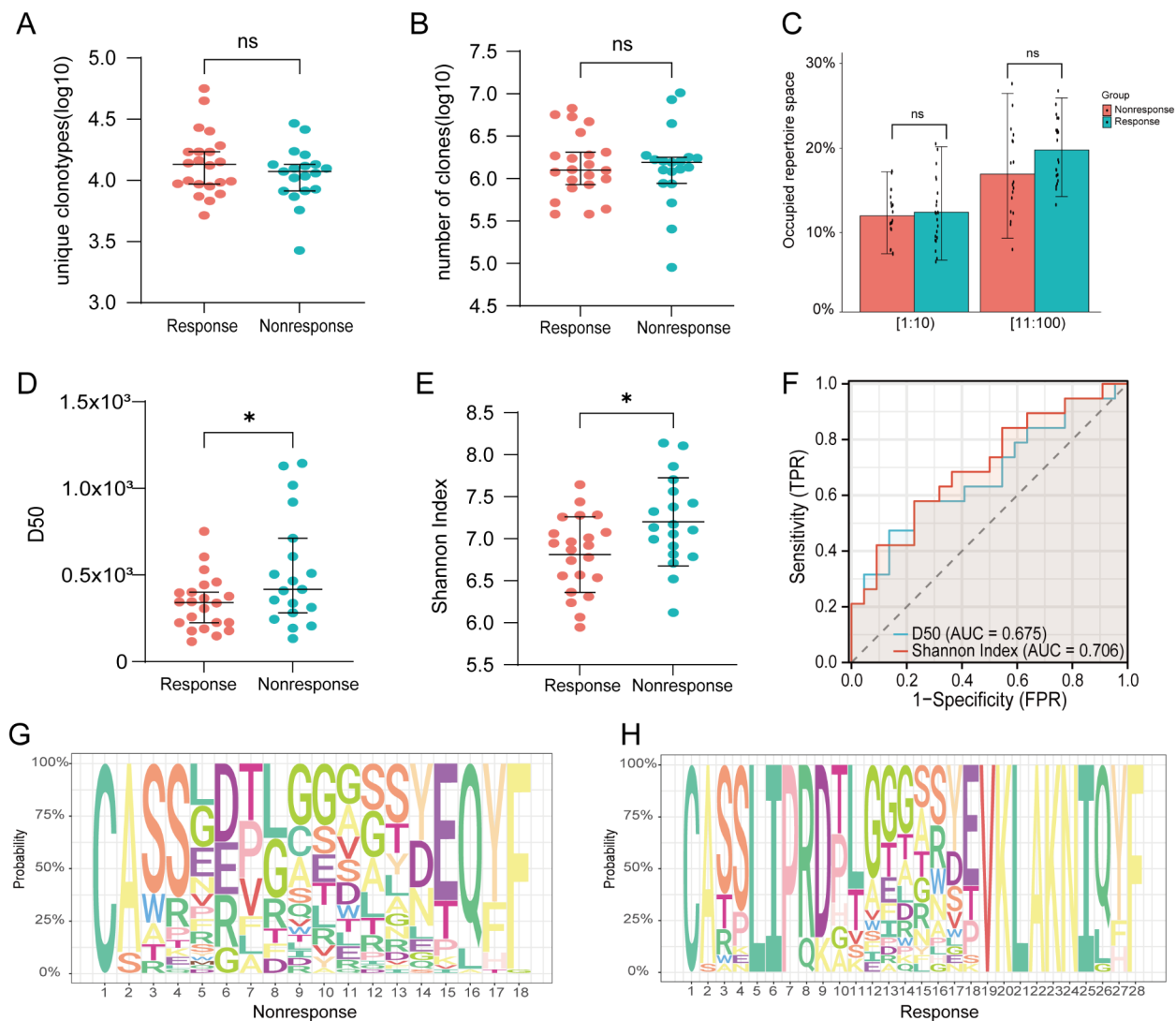


**Fig. 2** TCR repertoire characteristics of ESCC tumor tissue and adjacent normal tissue. **(A-C)** Differences in unique clone types, total clone numbers, and top clone frequencies between tumor tissue and adjacent normal tissue. **(D-E)** Comparison of diversity indices (D50 and Shannon Index) between tumor tissue and adjacent normal tissue. **(F)** Diagnostic value of D50 and Shannon Index for tumor tissue and adjacent normal tissue. **(G-H)** CDR3 aa motifs in tumor tissue and adjacent normal tissue

composition (Figure S2E). Furthermore, we conducted a comparative analysis of D50 and Shannon Index between non-responsive and responsive patients. Our findings indicated considerably higher values for both indices in non-responsive patients, indicating a greater diversity that was statistically significant (Fig. 3D-E). Using ROC curves, we assessed the diagnostic capacity of D50 and Shannon’s Index in predicting the response to ESCC immunotherapy. We found that they had a strong diagnostic value for ESCC (AUC>0.60) (Fig. 3F).

Additionally, we noticed that the responders had a greater number of CDR3 aa motifs (Fig. 3G-H).

Additionally, we conducted an analysis of the forms of TRB V-J gene combinations and V-D-J gene combinations. The bar graph illustrates disparities in the utilization rates of V and J genes between the two groups, with TRBV20-1 and TRBV28 being the most commonly used V genes, while TRBJ2-7 and TRBJ2-1 were found to be the most commonly used J genes in both cases (Figure S2A-B). Previous investigations have also documented that diseases such as hepatocellular carcinoma,



**Fig. 3** TCR repertoire characteristics of endoscopic biopsy paraffin-embedded samples obtained from ESCC patients with different immunotherapy responses. **(A-C)** Differences in unique clone types, total clone numbers, and top clone frequencies between responsive and non-responsive patients. **(D-E)** Comparison of diversity indices (D50 and Shannon Index) between responsive and non-responsive patients. **(F)** Diagnostic value of D50 and Shannon Index for predicting immune therapy response. **(G-H)** CDR3 aa motifs in responsive and non-responsive patients. Responsive patients: PR+CR. Non-responsive patients: SD+PD

nasopharyngeal carcinoma, and pancreatic cancer exhibit enrichment of gene segments including TRBV20-1, TRBJ2-7, and TRBJ2-1 [22, 30, 31]. The Venn diagram showed that the V and J genes were paired similarly. The volcano plot revealed that, in comparison to responsive patients, non-responsive patients exhibited a higher abundance of upregulated combinations of VJ and VDJ genes (Figure S2C-D, G-H). Subsequently, we utilized VDJtools to acquire information regarding the existing TCR antigens. We generated a list of potential antigens, which included antigen species, antigen epitopes, antigen genes, and MHC class, that were found in both responsive and non-responsive patients (>0.01%) (Table S3). While non-responsive patients exhibited a wider range

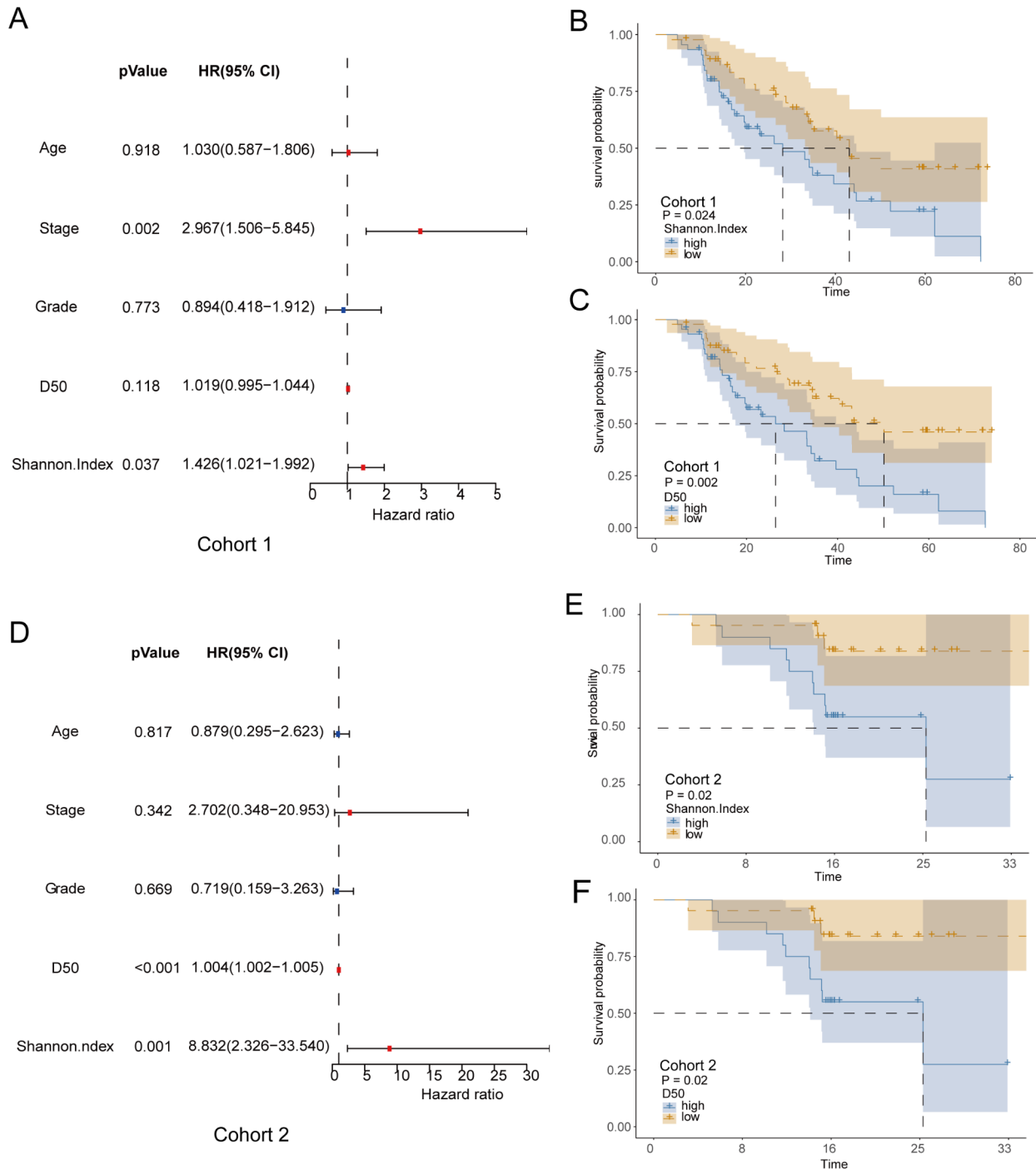
of potential antigens, MHC class I, CMV, and Influenza A were shared among all potential antigens. However, MHC class I was found to be more predominant in responsive patients. Of all the antigen epitopes, GILG-FVFTL, KLGGALQAK, and NLVPMVATV were shown to be the three most commonly detected epitopes.

#### The diversity of TCR repertoire can predict overall survival

To investigate whether the TCR repertoire diversity index can serve as an effective biomarker for predicting the survival of patients with ESCC, we conducted a univariate Cox analysis including variables such as age, tumor stage, pathological grade, and TCR repertoire diversity indices (D50 and Shannon Index) in cohort 1 of ESCC patients.

The results showed that higher D50 and Shannon Index were positively correlated with poor prognosis in ESCC patients, while age and histological grade had no significant correlation with prognosis (Fig. 4A). Subsequently, the association between D50, Shannon Index, and OS was further confirmed using K-M survival analysis, which

yielded results consistent with those of univariate Cox regression: patients with higher D50 and Shannon Index exhibited shorter OS (Fig. 4B-C). The results obtained from patients who received immunotherapy in cohort 2 also demonstrated similar findings, providing further support for the hypothesis that the TCR repertoire may



**Fig. 4** Survival prognosis of ESCC patients in cohort 1 (n=90) and cohort 2 (n=41). **(A)** Univariate COX model for patients in cohort 1. **(B-C)** The impact of D50 and Shannon Index on overall survival for patients in cohort (1) **(D)** Univariate Cox model for patients in cohort (2) **(E-F)** The impact of D50 and Shannon Index on overall survival for patients in cohort 2



serve as a prognostic risk factor influencing the survival outcome of ESCC patients (Fig. 3D-F). This finding also corroborates previous research indicating a relationship between the TCR repertoire diversity index and survival prognosis [32].

#### **The diversity of TCR repertoire remains unaffected by the tumor characteristics**

In order to explore the correlation between clinical characteristics and TCR repertoire, we examined alterations in TCR repertoire diversity index across different pathological grades, N stages, and tumor stages. Initially, within cohort 1, we observed a progressive increase in TCR repertoire diversity with worsening pathological grade, displaying significant differential expression between G1 and G3 as well as G2 and G3 (Figure S3A). Unfortunately, no correlation was found between pathological grade and patient survival prognosis; furthermore, the observed trend in cohort 2 did not reach statistical significance (Figure S3B-D). We observed differential expression of TCR repertoire diversity between N0 and N1-N3 stages in cohort 1 in terms of N staging, which was also associated with an unfavorable prognosis (Figure S3E-G). This suggests that the state of lymph nodes may impact the TCR repertoire. However, the TCR repertoire diversity index observed in cohort 2 did not show any statistically significant differences (Figure S3F-H). Subsequently, we examined the variations in TCR repertoire diversity and prognosis across tumor stages. Consistent with previous findings, the differential expression of TCR repertoire diversity was evident in cohort 1 and correlated with an adverse prognosis. However, there were no statistically significant differences observed in cohort 2, possibly due to its smaller sample size (Figure S3I-L).

#### **The diversity of TCR repertoire is associated with the TIME**

The tumor microenvironment is a complex and dynamic system consisting of immune cells, stromal cells, extracellular matrix components, cancer cells, and intricate cytokines [33]. T cell subsets are widely acknowledged as pivotal immune cells that orchestrate robust anti-tumor responses and confer primary immunity within the tumor microenvironment. To investigate the correlation between TIME and diversity index of TCR repertoire, we employed immunohistochemical staining to assess the expression levels of molecular markers associated with various subsets of T cells (including CD3, CD4, CD8, FoxP3, and PD-L1) within TIME. The patients in cohort 1 with higher TCR repertoire diversity indices (Shannon Index and D50) exhibited increased expression of CD8 and PD-L1. Unfortunately, no significant differences were observed in D50 expression (Fig. 5A-D, Figure S4A-B). Similarly, in cohort 2, we found that patients with higher Shannon Index and D50 showed increased expression

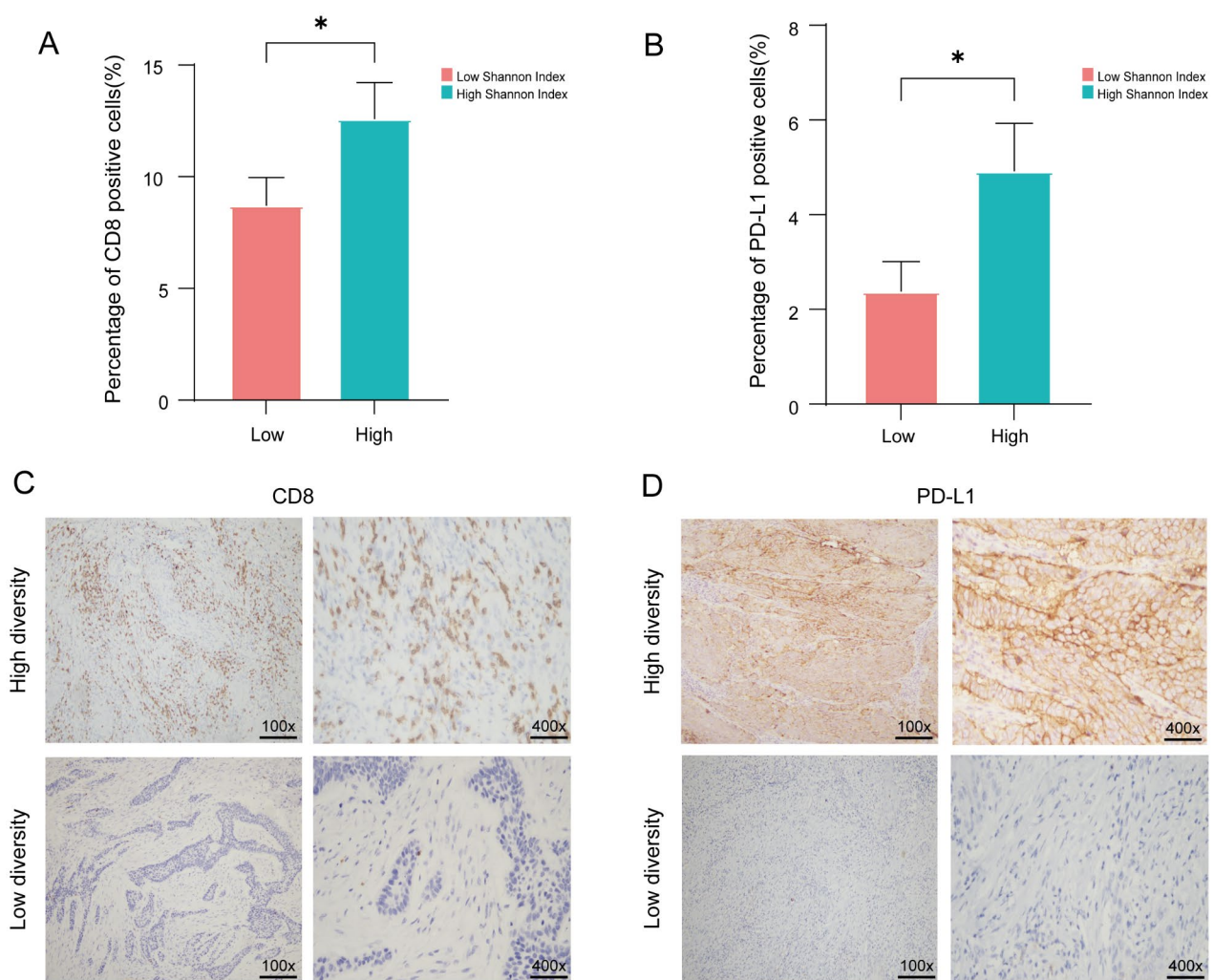
of CD3, CD4, CD8, FoxP3, and PD-L1 (Fig. 6A-H, Figure S4C-E, Figure S5A-C). Therefore, we speculate that there is a correlation between TIME and the diversity of tumor-infiltrating TCRs. Higher levels of T lymphocyte infiltration may be accompanied by a more diverse repertoire of TCRs.

#### **Discussion**

In this study, we collected data from two cohorts of ESCC patients, constituting the largest TCR repertoire cohort to date in ESCC, and characterized their TCR repertoire landscape. The TCR repertoire landscape was primarily delineated by assessing the diversity of TCR Diversity index of TCR repertoire within tumor and adjacent normal tissues, evaluating immune therapy response, and predicting patient survival prognosis. Additionally, we investigated the correlation between the TCR repertoire and the TIME. Our findings revealed that the diversity of TCR repertoire within tumor tissues was comparatively restricted compared to surrounding tissues. Non-responsive patients exhibited higher TCR repertoire diversity indices than responsive patients following immune therapy. Moreover, the TCR repertoire diversity index demonstrated potential as a prognostic indicator for survival outcomes in ESCC patients. Furthermore, immunohistochemical staining provided evidence for an association between TIME and TCR repertoire.

The heterogeneity of the TCR repertoire between tumor tissue and surrounding tissue has been thoroughly studied in past studies pertaining to a variety of tumor types, such as lung cancer, bladder cancer, hepatocellular carcinoma, and pancreatic cancer [12, 14, 34, 35]. However, characterizing the TCR repertoire in ESCC has been the subject of only a few studies. By conducting a comprehensive investigation into the disparities in TCR repertoire characteristics between ESCC tumor tissue and its surrounding environment, we discovered that within the tumor tissue, there is reduced diversity in terms of unique clone types as well as top clone abundance. Additionally, decreased diversity was observed in terms of diversity index and CDR3 aa motifs. This phenomenon may be attributed to selective clonal expansion during tumor infiltration. Therefore, when observing diminished diversity within tumor tissue samples, it can be inferred that there are more constraints on selection within this subset, which could potentially be associated with an increased degree of CR. CR leads to multiple CDR3 nt sequences encoding the same CDR3 aa sequence due to TCR selection during the immune response [36].

A number of ICIs have been licensed in recent years to treat metastatic malignancies, including bladder cancer, lung cancer, and malignant melanoma [37–39]. Immunotherapy has demonstrated unprecedented therapeutic benefits in a significant proportion of cancer patients,

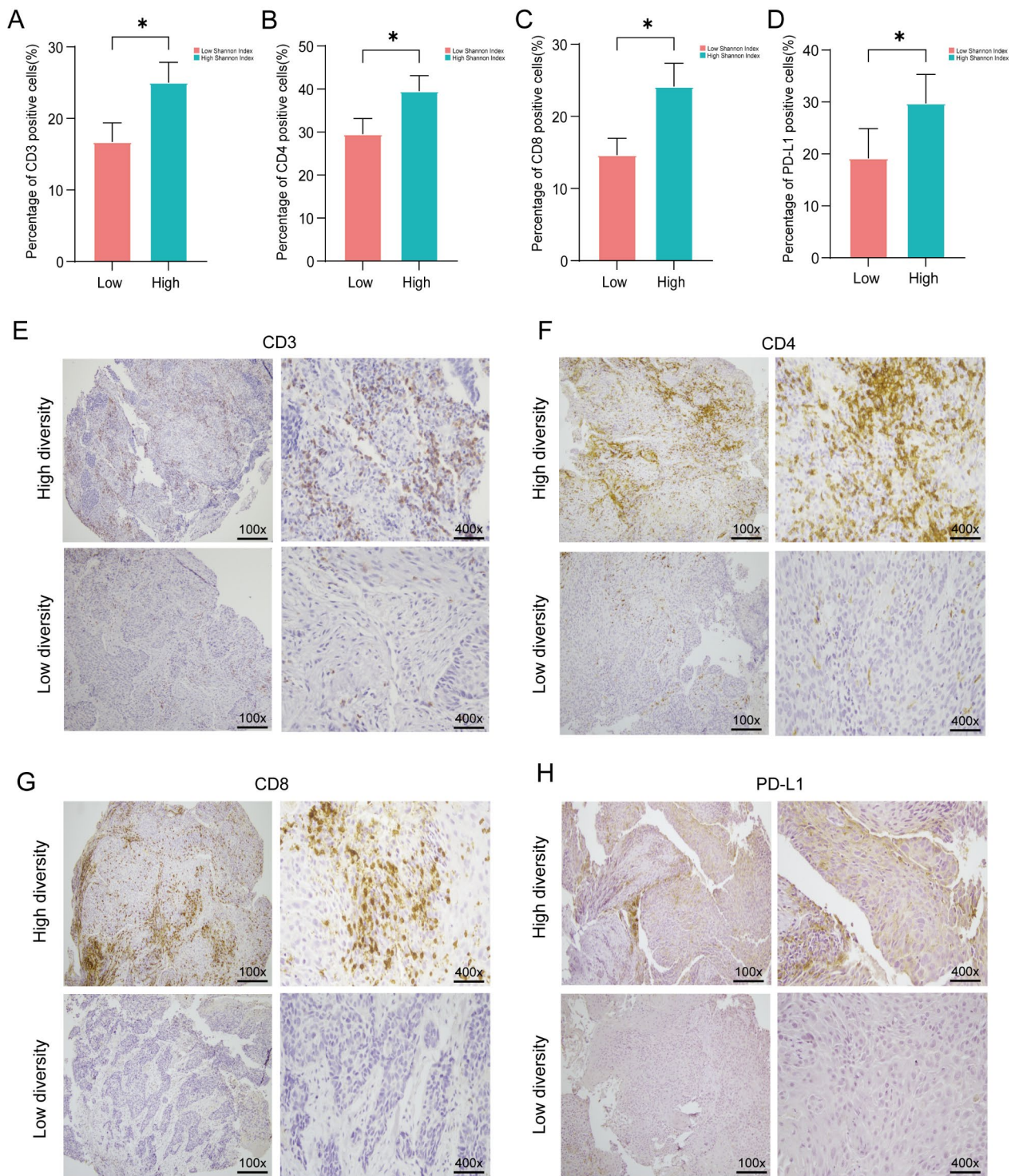


**Fig. 5** Expression of CD8 and PD-L1 between Shannon Index in cohort 1 ( $n=90$ ). **(A-B)** Statistical analysis was conducted on the Shannon Index for different groups based on immunohistochemical staining results of CD8 and PD-L1 expression. **(C-D)** Representative immunohistochemical staining images of CD8 and PD-L1 expression in patients from cohort 1

highlighting the growing clinical demand for predicting immunotherapy efficacy. Investigations into TCR repertoire characteristics may serve as effective biomarkers for prognosticating patient outcomes and response to ICI therapy [40–44]. Our study investigated the characteristics of the TCR repertoire in patients exhibiting diverse responses to immunotherapy. The findings revealed that non-responsive patients displayed higher TCR repertoire diversity index and a higher prevalence of common V and J gene usage. Moreover, non-responsive patients exhibited an increased abundance of potential antigens, while MHC class I expression was more frequently observed among all potential antigens in responsive patients. These results suggest that MHC class I expression on tumor cells plays a pivotal role in facilitating T cell recognition and cytotoxicity, particularly within tumor micro-environments. Additionally, high-frequency TCRs also

recognized other antigenic sites with potential implications for engineered T cell therapies targeting cancer cell proliferation, invasion, and metastasis.

During tumor development, the TCR repertoire undergoes continuous changes as cancer cells emerge and are cleared. When cancer cells evade immune destruction and reach a state of equilibrium or escape, tumor-associated T cells are continuously generated, shaping the TIME [45]. Numerous studies have shown that the baseline TCR repertoire in tumors is closely linked to the TIME, with TCR repertoire diversity reflecting T cell activation across various cancers within this environment [46]. Our research primarily evaluated immune cell-related biomarkers, with a main focus on T-cell subtypes, in order to elucidate the association between baseline TCR repertoire diversity and the tumor immune micro-environment. We also confirmed the significant value of



**Fig. 6** Expression of CD3, CD4, CD8 and PD-L1 between Shannon Index in cohort 2 ( $n = 41$ ). **(A-D)** Statistical analysis was performed on the immunohistochemical staining results of CD3, CD4, CD8 and PD-L1 expression in different groups based on Shannon Index. **(E-H)** Representative immunohistochemical staining images of CD3, CD4, CD8 and PD-L1 expression in patients from cohort 2 are shown

the TCR repertoire diversity index in predicting patient survival prognosis. While previous studies have investigated the impact of immunotherapy on the immune microenvironment by analyzing samples collected before and after treatment, our current data only includes baseline samples. In future investigations, we aim to further elucidate the impact of immunotherapy on the immune microenvironment.

There are several limitations to this study. Firstly, the extraction of TCR data from cohort 1 may not fully capture the functional diversity of T cells, whereas employing high-throughput TCR sequencing could yield more comprehensive insights. Secondly, the utilization of patient data solely from a single center and the absence of tumor samples from stage I patients necessitate collaboration with other medical centers to augment the sample size. Lastly, incorporating pre-immunotherapy and post-immunotherapy patient samples would provide a more comprehensive understanding of the impact of immunotherapy on the TCR repertoire.

Through high-throughput RNA and TCR sequencing, this study identified a reduction in the diversity of TCR repertoire in tumor tissues compared to adjacent normal tissues. Furthermore, it unveiled inter-patient heterogeneity in immune therapy response based on their respective TCR repertoires. Moreover, our research substantiated the prognostic significance of TCR repertoire diversity for patient survival prognosis and its association with TIME. These findings hold promise for guiding novel immunotherapy approaches, particularly personalized TCR-T therapy, in patients with ESCC.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-024-05825-0>.

Supplementary Material 1

### Acknowledgements

Not applicable.

### Author contributions

Meng Yang and Dan He drafted the manuscript, conducted in vitro experiments, and visualized the results. Yu Sun and Yu Ma collected the data and performed the data analysis. Yunquan Guo completed the immunohistochemical staining experiment. Article direction by Lin Feng and Meng Liu. All authors read and approved the final manuscript.

### Funding

This study was supported by the Key Laboratory open project of Xinjiang Uygur Autonomous Region (Grant no. 2021D04022), National Natural Science Foundation of China (Grant no.82260600), Tianshan Talents Training Program of Xinjiang Uygur Autonomous Region (Grant no.2023TSYCC0058), and Distinguished Young Project of Natural Science Foundation of Xinjiang Uygur Autonomous Region (Grant no. 2022D01E76).

### Data availability

The T cell receptors raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics &

Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2022), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA006108 and GSA-Human: HRA000178 ) that are publicly accessible at <https://ngdc.cnbc.ac.cn/gsa-human>.

### Declarations

#### Ethical approval

The clinical tissue samples of this study have obtained ethical approval from the Ethics Committee of Xinjiang Medical University Affiliated Tumor Hospital, with approval number K-2023006, and informed consent for the use of relevant clinical tissue samples. All methodologies strictly adhere to relevant standards and regulations, in accordance with the Declaration of Helsinki (revised 2013).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

Received: 10 September 2024 / Accepted: 31 October 2024

Published online: 27 November 2024

### References

- Liu C Q, Ma Y L, Qin Q, et al. Epidemiology of esophageal cancer in 2020 and projections to 2030 and 2040 [J]. *Thorac Cancer*. 2023;14(1):3–11.
- Sung H, Ferlay J, Siegel RL, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries [J]. *CA Cancer J Clin*. 2021;71(3):209–49.
- Siegel RL, Miller KD Jemala. Cancer statistics, 2018 [J]. *CA Cancer J Clin*. 2018;68(1):7–30.
- Taziki M H Rajaees, Behnampour N, et al. Esophageal cancer: 5-year survival rate at south-east of Caspian sea of northern Iran [J]. *J Cancer Res Ther*. 2011;7(2):135–7.
- Zhao Q, Jiang Y, Xiang S, et al. Engineered TCR-T cell immunotherapy in Anti-cancer Precision Medicine: pros and cons [J]. *Front Immunol*. 2021;12:658753.
- Rosenberg S A, Restifo N P. Adoptive cell transfer as personalized immunotherapy for human cancer [J]. *Science*. 2015;348(6230):62–8.
- Bagchi S, Yuan R. Immune Checkpoint inhibitors for the treatment of Cancer: clinical impact and mechanisms of response and resistance [J]. *Annu Rev Pathol*. 2021;16:223–49.
- Garon E B, Hellmann M D, Rizvi N A, et al. Five-year overall survival for patients with Advanced non-small-cell lung Cancer treated with Pembrolizumab: results from the phase I KEYNOTE-001 study [J]. *J Clin Oncol*. 2019;37(28):2518–27.
- Hamid O, Robert C. Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001 [J]. *Ann Oncol*. 2019;30(4):582–8.
- Sudo T, Nishida R, Kawahara A, et al. Clinical impact of Tumor-infiltrating lymphocytes in esophageal squamous cell carcinoma [J]. *Ann Surg Oncol*. 2017;24(12):3763–70.
- Chen K, Zhu Z, Zhang N, et al. Tumor-infiltrating CD4+ lymphocytes predict a favorable survival in patients with operable esophageal squamous cell carcinoma [J]. *Med Sci Monit*. 2017;23:4619–32.
- Wang X, Zhang B, Yang Y, et al. Characterization of distinct T cell receptor repertoires in Tumor and distant non-tumor tissues from Lung Cancer patients [J]. *Genomics Proteom Bioinf*. 2019;17(3):287–96.
- Wei M, Shen D, Mulmi Shrestha S et al. The Progress of T Cell Immunity Related to Prognosis in Gastric Cancer [J]. *Biomed Res Int*, 2018, 2018: 3201940.
- Ma J, Sun G, Zhu P, et al. Determination of the complexity and diversity of the TCR  $\beta$ -chain CDR3 repertoire in bladder cancer using high-throughput sequencing [J]. *Oncol Lett*. 2019;17(4):3808–16.
- Hopkins A C, Yarchoan M, Durham J N et al. T cell receptor repertoire features associated with survival in immunotherapy-treated pancreatic ductal adenocarcinoma [J]. *JCI Insight*, 2018, 3(13).

16. Inoue H, Park J H, Kiyotani K, et al. Intratumoral expression levels of PD-L1, GZMA, and HLA-A along with oligoclonal T cell expansion associate with response to nivolumab in metastatic melanoma [J]. *Oncoimmunology*. 2016;5(9):e1204507.
17. Ham B, Kim S Y, Kim Y A et al. Persistence and enrichment of dominant T cell clonotypes in expanded tumor-infiltrating lymphocytes of breast cancer [J]. *Br J Cancer*, 2024.
18. Robins H S, Srivastava S K, Campregher P V, et al. Overlap and effective size of the human CD8+ T cell receptor repertoire [J]. *Sci Transl Med*. 2010;2(47):47ra64.
19. Attaf M, Huseby E, Sewell A K.  $\alpha\beta$  T cell receptors as predictors of health and disease [J]. *Cell Mol Immunol*. 2015;12(4):391–9.
20. Rosati E, Dowds C M. Overview of methodologies for T-cell receptor repertoire analysis [J]. *BMC Biotechnol*. 2017;17(1):61.
21. Rossjohn J, Gras S. T cell antigen receptor recognition of antigen-presenting molecules [J]. *Annu Rev Immunol*. 2015;33:169–200.
22. Jin Yb, Luo W, Zhang G Y, et al. TCR repertoire profiling of tumors, adjacent normal tissues, and peripheral blood predicts survival in nasopharyngeal carcinoma [J]. *Cancer Immunol Immunother*. 2018;67(11):1719–30.
23. Yan C, Ma X, Guo Z, et al. Time-spatial analysis of T cell receptor repertoire in esophageal squamous cell carcinoma patients treated with combined radiotherapy and PD-1 blockade [J]. *Oncoimmunology*. 2022;11(1):2025668.
24. Liu M, An H, Zhang Y, et al. Molecular analysis of Chinese oesophageal squamous cell carcinoma identifies novel subtypes associated with distinct clinical outcomes [J]. *EBioMedicine*. 2020;57:102831.
25. Eisenhauer E A, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1) [J]. *Eur J Cancer*. 2009;45(2):228–47.
26. Song L, Cohen D, Ouyang Z, Et Al. Trust4: immune repertoire reconstruction from bulk and single-cell RNA-seq data [J]. *Nat Methods*. 2021;18(6):627–30.
27. Database Resources of the National Genomics Data Center. China National Center for Bioinformatics in 2022 [J]. *Nucleic Acids Res*. 2022;50(D1):D27–38.
28. Chen T, Chen X, Zhang S, et al. Genomics Proteom Bioinf. 2021;19(4):578–83. The Genome Sequence Archive Family: Toward Explosive Data Growth and Diverse Data Types [J].
29. Woodsworth D J, Castellarin M, HOLT RA. Sequence analysis of T-cell repertoires in health and disease [J]. *Genome Med*. 2013;5(10):98.
30. Lin K R, Deng F W, JIN Y B, et al. T cell receptor repertoire profiling predicts the prognosis of HBV-associated hepatocellular carcinoma [J]. *Cancer Med*. 2018;7(8):3755–62.
31. Bai X, Zhang Q, WU S, et al. Characteristics of Tumor infiltrating lymphocyte and circulating lymphocyte repertoires in pancreatic Cancer by the sequencing of T cell receptors [J]. *Sci Rep*. 2015;5:13664.
32. Hogg G D Pothuriv. Intratumoral T-cell receptor repertoire composition predicts overall survival in patients with pancreatic ductal adenocarcinoma [J]. *Oncoimmunology*. 2024;13(1):2320411.
33. Gajewski T F, Schreiber H, FU Y X. Innate and adaptive immune cells in the tumor microenvironment [J]. *Nat Immunol*. 2013;14(10):1014–22.
34. Wang Z, Zhong Y, Zhang Z, et al. Characteristics and clinical significance of T-Cell receptor repertoire in Hepatocellular Carcinoma [J]. *Front Immunol*. 2022;13:847263.
35. Wang H, Yuan Y, Lu C, et al. Analysis of T-cell receptor repertoire in peripheral blood of patients with pancreatic cancer and other pancreatic diseases [J]. *J Cell Mol Med*. 2021;25(8):3991–4000.
36. Gordin M, Philip H, Zilberberg A, et al. Breast cancer is marked by specific, public T-cell receptor CDR3 regions shared by mice and humans [J]. *PLoS Comput Biol*. 2021;17(1):e1008486.
37. LOMmatszsch M, Bratke K, Neoadjuvant STOLL. PD-1 blockade in Resectable Lung Cancer [J]. *N Engl J Med*. 2018;379(9):e14.
38. Liu X, Cui Y, Zhang Y, et al. A comprehensive study of immunology repertoires in both preoperative stage and postoperative stage in patients with colorectal cancer [J]. *Mol Genet Genomic Med*. 2019;7(3):e504.
39. Lerner S P, Bajarin D F, Dinney C P, et al. Summary and recommendations from the National Cancer Institute's clinical trials planning meeting on Novel therapeutics for non-muscle invasive bladder Cancer [J]. *Bladder Cancer*. 2016;2(2):165–202.
40. Schrama D, Ritter C, Becker JC. T cell receptor repertoire usage in cancer as a surrogate marker for immune responses [J]. *Semin Immunopathol*. 2017;39(3):255–68.
41. Aran A, Garrigós L, Curigliano G et al. Evaluation of the TCR Repertoire as a predictive and prognostic biomarker in Cancer: diversity or clonality? [J]. *Cancers (Basel)*, 2022, 14(7).
42. Akyüz N, Brandt A, Stein A, et al. T-cell diversification reflects antigen selection in the blood of patients on immune checkpoint inhibition and may be exploited as liquid biopsy biomarker [J]. *Int J Cancer*. 2017;140(11):2535–44.
43. Looney T J, Topacio-Hall D, Lowman G, et al. TCR convergence in individuals treated with Immune Checkpoint Inhibition for Cancer [J]. *Front Immunol*. 2019;10:2985.
44. Postow M A, Manuel M, Wong P, et al. Peripheral T cell receptor diversity is associated with clinical outcomes following ipilimumab treatment in metastatic melanoma [J]. *J Immunother Cancer*. 2015;3:23.
45. Zhang Z, Liu S, Zhang B, et al. T cell dysfunction and exhaustion in Cancer [J]. *Front Cell Dev Biol*. 2020;8:17.
46. Schina A, Sztupinszki Z, Marie Svane I et al. Intratumoral T-cell and B-cell receptor architecture associates with distinct immune tumor microenvironment features and clinical outcomes of anti-PD-1/L1 immunotherapy [J]. *J Immunother Cancer*, 2023, 11(8).

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.