TISSUE AND PERIPHERAL T-CELL REPERTOIRE PREDICTS IMMUNOTHERAPY RESPONSE AND PROGRESSION-FREE SURVIVAL IN NSCLC PATIENTS

Manuel Pino-González1, Martín Lázaro-Quintela1,2#, Irene Alonso-Álvarez1, María Gallardo-Gómez1, Laura Juaneda-Magdalena1, Alejandro Francisco-Fernández3, Silvia Calabuig-Fariñas5,6, Eloisa Jantus-Lewintre5,6,7, Mónica Martínez-Fernández1*

6. Mixed Unit TRIAL. General University Hospital of Valencia Research Foundation and Príncipe Felipe Research Center. Hospital General Universitario de Valencia, 46014 (Valencia), Spain.
7. Department of Biotechnology. Universitat Politècnica de València, 46022 (Valencia), Spain.

# Equal contribution

Email: monica.martinez@iisgaliciasur.es

Running title

TCR Predicts Immunotherapy Response and Progression in NSCLC Patients

Keywords

non-small cell lung cancer; immunotherapy; predictive biomarker; immunotherapy response; T cell receptor; TCR repertoire.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Immunotherapy has opened new avenues of treatment for patients with advanced non-small cell lung cancer (NSCLC) without previous hope of survival. Unfortunately, only a small percentage of patients benefit from it, and it is still not well understood which tumor characteristics can be used to predict immunotherapy response. As the key cellular effectors of antitumor immunity, T cells are endowed with specialized T cell receptors (TCRs) to recognize and eliminate cancer cells. Here, we evaluated the potential of TCR repertoire as a predictive biomarker in patients treated with immunotherapy. With this aim, advanced NSCLC patients treated with immunotherapy at first-line were included. After obtaining peripheral blood and tissue samples at baseline, next-generation sequencing targeting TCRβ/γ was performed. Beyond TCR metrics, clonal space of the most frequent clones was determined. We found a positive association between uneven tumor-infiltrating TCRβ repertoire and the immunotherapy response. Moreover, the use of various tumor-infiltrating and circulating TRBV/J genes predicted the immunotherapy response. Our results indicate the importance of evaluating tissue and circulating TCRβ repertoire prior immunotherapy, showing it as a promising immunotherapy response biomarker in NSCLC patients.
Introduction

According to the latest estimates from the International Agency for Research on Cancer (IARC), lung cancer is the most common and deadliest cancer worldwide [1]. Regarding its histology, it can be classified into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC accounts for 85% of cases and is further subdivided into adenocarcinoma, squamous cell carcinoma, and large cell carcinoma [2–4]. Although environmental exposures (such as fuels, arsenic, radon, and air pollution) and medical conditions (as chronic obstructive pulmonary disease, HIV infection, or presence of driver mutations), contribute to its incidence and mortality, smoking is the main risk factor associated with lung cancer [2,3,5]. NSCLC is a highly heterogeneous disease with nonspecific symptoms, leading to a common late diagnosed at advanced stages (III-IV), where surgery is not feasible [2,4,5]. This greatly compromise the patient's prognosis, which is clearly reflected in the 5-year survival rate, decreasing from 60% in early stages to 4-6% in metastatic settings [2,4,5]. In fact, 90% of cancer deaths are due to metastasis [6]. Nowadays, targeted therapy against many actionable mutations (EGFR, ALK, RET, BRAF, ROS1, NTRK, MET, HER2 and KRAS) represent the standard treatment for molecularly defined populations [2,7]. Unfortunately, only 25% of patients carry any of these mutations [2,8]. In 2015, the combination of immunotherapy based on immune checkpoint inhibitors (ICIs) and chemotherapy (CTX) as first-line treatment significantly improved overall survival (OS) in NSCLC patients compared to those treated with CTX alone [2,9,10]. Thus, ICIs has been established as the standard treatment for patients with advanced NSCLC who do not have actionable mutations, either as monotherapy or in combination with CTX, leading to a 5-year survival rate up to 20% [2,9,10]. Unfortunately, only a small percentage of patients (20-40%) benefit from it [2,9,10]. Its high cost, the limited time for therapy selection, and the possibility of developing autoimmune-related side effects (such as inflammatory arthritis, colitis and pneumonitis) underscore the importance of identifying those patients who can truly benefit from the treatment [2,9,10]. Currently, PD1/PD-L1 immunostaining is the only approved predictive biomarker for immunotherapy response, so patients with a tumor proportion score (TPS) ≥ 50% are classified as PD-L1 positive and thus, expected to respond [2,9–11]. However, its reproducibility has been limited by the heterogeneity of results obtained in subsequent studies, largely due to lack of standardization (calculation systems, thresholds, and use of different antibodies) [2,9–11]. Over time, other alternatives such as tumor mutational burden (TMB), tumor-infiltrating

CC-BY-NC-ND 4.0 International license
It is made available under a CC-BY-NC-ND 4.0 International license.
lymphocytes (TILs) and other molecular characteristics have been proposed, but none have been able to reliably predict patient response in the real clinical practice [2,9–11].

Although the antitumor response is mediated by different cellular populations, cytotoxic T cells play a key role by recognizing neoantigens and killing cancer cells [12,13]. Neoantigens are recognized by the T cell receptors (TCRs). This transmembrane glycoprotein is composed of an α and a β chain in Tαβ cells (95%) or by a γ and a δ chain in Tγδ cells (5%). Its synthesis occurs through a somatic, stochastic, and imprecise recombination of various non-contiguous homologous gene sets during T cell differentiation in the thymus [11,13–17]. Each chain consists of a variable and a constant region. The variable region of the α and γ chains results from recombination between the V and J genes, while β and δ chains result from recombination between the V, D, and J genes. The result of this V(D)J recombination creates a variable domain that is transcribed and joins one of the constant genes (C), resulting in a functional chain. Each chain contains three hypervariable loops (CDR1, CDR2, and CDR3) in its structure, with CDR3 mediating peptide recognition and thus determining TCR specificity [11,13–17]. CDR3 is encoded by V(D)J genes and undergoes a series of random nucleotide insertions and deletions (N) during recombination at gene junctions (VN(D)NJ). This entire process ensures the creation of a vast array of unique TCRs capable of recognizing a wide range of antigens [11,13–17].

Considering its key role in neoantigen recognition, TCR repertoire analysis provides valuable information about the antitumor response [11,12,14,15,18,19]. Furthermore, advancements in sequencing technologies and bioinformatics tools have enabled sensitive and precise characterization of TCR repertoire in various types of tumors, including NSCLC, making it feasible to study as a potential biomarker for immunotherapy response [11,14,15,18,19]. In recent years, various studies have suggested the potential use of the TCR repertoire as a biomarker for immunotherapy response, both for anti-CTLA4 and anti-PD-L1/PD1 therapies in different types of cancer (such as melanoma, breast and liver), including NSCLC [20–39]. These studies have yielded contradictory results but indicate a potential capacity for the TCR repertoire to be used as an immunotherapy response biomarker. Consequently, the main objective of this study is to evaluate the potential of TCR repertoire analysis as a biomarker for response to immunotherapy in patients with advanced NSCLC.

**Materials and Methods**

**Study design and sample collection**
This study included twelve patients diagnosed with advanced (IIIA-IVB) NSCLC treated with pembrolizumab either as first-line (N = 11) or second-line treatment (N = 1) at Álvaro Cunqueiro Hospital in Vigo. The study was conducted with appropriate authorization from the Galician Regional Research Ethics Committee (2019/046) following the Helsinki Declaration of 1975 and all patients signed informed consent approving their participation. Patients were classified as responders (R: those showing complete/partial response or stable disease) and non-responders (NR: those with tumor progression or unable to be evaluated due to death) regarding the Response Evaluation Criteria for Solid Tumours (RECIST) [40] at 3 months. To determine both tumor and peripheral TCR repertoire, formalin-fixed paraffin-embedded (FFPE) tissue samples were obtained from all patients prior to immunotherapy and blood samples were drawn from 10 out of 12 patients before the first ICIs dose. Peripheral blood mononuclear cells (PBMCs) were isolated from blood samples by gradient density centrifugation using Ficoll® (Sigma-Aldrich) and cryopreserved until use. DNA was chosen over RNA for analysis due to its superior stability in FFPE samples.

**DNA Extraction**

DNA extraction from PBMCs was performed using the QIAamp® DNA Blood Mini Kit (QIAGEN) following the manufacturer's instructions. DNA from FFPE samples was extracted using the AllPrep® DNA/RNA FFPE Kit (QIAGEN) following the manufacturer's instructions. DNA quantification was carried out using NanoDropTM 2000c Spectrophotometer (Thermo Fisher Scientific).

**Library preparation and TCR sequencing**

A total of 250 ng of genomic DNA was used to prepare the libraries for TCR CDR3β/γ region sequencing following the Oncomine™ TCR Pan-Clonality Assay (Thermo Fisher Scientific). This targeted next-generation sequencing (NGS) assay has been specifically designed to sequence the FR3-J regions of the TCR beta and TCR gamma chains. Libraries were pooled on an Ion 540 chip at 25 pmol/L and sequenced using an Ion GeneStudio S5 Plus Series (Thermo Fisher Scientific).

**Data analysis**

After sequencing, data analysis was conducted using Ion Reporter version 5.20.2.0 (Thermo Fisher Scientific). Regarding quality control, off target and unproductive reads were removed. Read classification is shown in Supp. File S1. Afterwards, the software reported the VDJ rearrangements and the main repertoire metrics, such as richness, convergence, diversity, and
evenness. TCR richness represented the total number of clones, defined as unique TCR beta/gamma nucleotide sequences. TCR convergence was determined as the aggregate frequency of clones which shared a variable gene and CDR3 amino acid sequence. Diversity (Shannon’s diversity) and evenness (normalized Shannon’s diversity) were calculated using the formulas below, in which \( p_i \) is the frequency of clone \( i \) for the sample with \( n \) unique clones. Evenness describes how evenly distributed is the TCR repertoire. It ranges from 0, meaning the repertoire is unbalanced by a reduced number of predominant clones, to 1, meaning the repertoire is balanced with similar frequencies of all the clones.

\[
\text{Shannon’s diversity (H)} = - \sum_{i=1}^{n} p_i \log(p_i)
\]

\[
\text{Evenness (E)} = \frac{H}{\log n}
\]

In addition, clones were classified according to their relative abundance in the repertoire, named total clonal space, for each sample. We divided it into top 1% clonal space, top 3% clonal space and top 5% clonal space, so the top 1% clonal space is defined as the aggregate frequencies of the top 1% most frequent clones. All the variables were categorized as “high” or “low” based on their median value.

**Statistical analysis**

Statistical analyses were performed using R-Studio version 2023.12.0 and Bioconductor version 3.19 environment. Statistical significance was considered by an overall p-value < 0.05. TCR repertoire variables were correlated with clinical-pathological data (age, sex, tumor histology, PD1/PD-L1 expression, and smoking status), clinical response and survival data (OS and PFS). OS was defined as the time from treatment initiation to death or last follow-up, and PFS was defined as time from treatment initiation to progression or death (whichever is earlier), or last follow-up. Mann-Whitney U test and Student’s t-tests were used to study the relationship between TCR repertoire variables and clinical-pathological data and clinical response. Choice of appropriate test (Mann-Whitney or Student’s t) was based on normality (Shapiro-Wilk test) and variance homogeneity (Levene test). Spearman rank non-parametric test was used for variable correlations. Contingency table analyses when comparing clinical-pathological data with clinical response or clinical-pathological data with TCR repertoire variables were done using Fisher exact test. Kaplan-Meier survival analyses and Log-Rank test were used for time-dependent variables such as OS and PFS. Receiver operating characteristics (ROC) curve analyses were performed to evaluate response prediction, and area under the curve (AUC),
sensitivity and specificity values were obtained. The Youden Index method was used to select the best cut-off values [41]. Random Forest Analysis (number of trees = 1000) was used to evaluate the combined predictive signatures performances.

Results

Age is associated with immunotherapy response and tumor-infiltrating evenness

This retrospective observational study included a total of twelve patients diagnosed with locally advanced or advanced NSCLC (IIIA-IVB) treated with pembrolizumab at first-line (N = 11) or second-line (N = 1). The most relevant clinico-pathological characteristics are represented in Table 1. None of the patients had actionable mutations. Both tissue (N = 12) and PBMCs (N = 10) samples were collected before treatment and were subjected to TCR sequencing. We obtained valid TCR data in all cases (100%).

First, we evaluated the effect of immunotherapy response on patient prognosis. As expected, R had statistically significant longer PFS (p = 0.015) than NR (Fig. 1A). We found the same trend regarding OS (p = 0.057) (Supp. File S2A). Interestingly, we detected an association between response and age, where R were statistically significant older than NR (p = 0.008; R age range = 71-81 years; NR age range = 43-71 years) (Fig. 1B). All R were above the median age of the cohort (70.5 years) (Supp. File S2B). No other association was detected between clinico-pathological characteristics and response or prognosis.

Second, we sought for possible associations between the clinico-pathological characteristics and TCR repertoire characterization. We detected that squamous tumors presented statistically significant higher tumor-infiltrating TCRβ convergence (p = 0.026) than adenocarcinomas (Fig. 1C). Furthermore, we found an association between age and tumor-infiltrating evenness in both TCRβ and TCRγ, where younger patients (<71 years old) had statistically significant higher tumor-infiltrating evenness than older patients (≥71 years old) (p = 0.022 and p = 0.041, respectively) (Fig. 1D-E).

Tumor-infiltrating TCRβ evenness is associated with immunotherapy response and Progression-Free Survival (PFS)

When comparing tumor tissues, we found that R had statistically significant lower tumor-infiltrating TCRβ evenness than NR (p = 0.044) (Fig. 2A). Thus, its ability of potentially predicting the response and survival of these patients was analysed. As represented in Fig. 2B, tumor-infiltrating TCRβ evenness predicted response with an AUC of 0.86, where a tumor-infiltrating TCRβ evenness lower than 0.795 had a 60% sensitivity and 100% specificity of
predicting response. Moreover, patients with low tumor-infiltrating TCRβ evenness (< 0.8441) had statistically significant longer PFS than patients with high (≥ 0.8441) tumor-infiltrating TCRβ evenness (p = 0.013) (Fig. 2C). Similarly, we found the same trend regarding circulating TCRβ evenness (p = 0.139) (Fig. 2D). Despite they did not reach the statistical significance the same trends were found regarding TCRγ in both tissue and blood, respectively (p = 0.149 and p = 0.114), showing a lower evenness in R patients (Fig. 2E-F). It is worth to mention that R patients showed a tendency to a higher tumor-infiltrating and circulating TCRβ top 3% and top 5% clonal space compared with NR patients (Supp. File S3A-D). Accordingly, TCRβ evenness and top 3% and 5% clonal space were always strongly negatively correlated both in tissue (R = -0.81, p = 0.001; R = -0.91, p < 0.001, respectively) and blood (R = -0.99, p < 0.001; R = -0.97, p < 0.001, respectively) (Supp. File S4). No other association was detected between clinicopathological characteristics and TCR repertoire variables.

**Tumor-infiltrating and circulating TCRβ TRBV and TRBJ genes usage is associated with immunotherapy response and Progression-Free Survival**

Next, we sought for any relation between TRBV/TRGV and TRBJ/TRGJ genes frequencies and the immunotherapy response. Regarding TCRγ, there was no association between TRGV and TRGJ genes frequency and clinical response, neither in tissue nor in blood. Interestingly, tumor-infiltrating TCRβ TRBV6.5, TRBV11.3 and TRBJ2.1 genes frequencies were statistically significant lower in R than in NR (Table 2; p = 0.011, p = 0.048 and p = 0.002, respectively). Moreover, circulating TCRβ TRBV5.3 gene frequency was statistically significant higher in R than in NR (p = 0.019), while TRBV27, TRBV28, TRBJ2.1 and TRBJ2.6 genes frequencies were statistically significant lower in R than in NR (p = 0.039, p= 0.029, p= 0.017, and p = 0.008, respectively), as shown in Table 3.

Then, we analysed their ability of potentially predict the PFS and the immunotherapy response in these patients. As represented in Fig. 3A, patients with low tumor-infiltrating TCRβ TRBV11.3 gene frequency presented statistically significant longer PFS than patients with high frequency (p = 0.002). Tumor-infiltrating TCRβ TRBV11.3 gene frequency predicted response with an AUC of 0.86, where a lower frequency than 0.0057 had an 80% sensitivity and 85.7% specificity predicting response (Fig. 3B). In the same line, but not reaching the statistical significance (p = 0.051), patients with low tumor-infiltrating TCRβ TRBJ2.1 gene frequency had higher PFS than patients with high frequency (Fig. 3C). Tumor-infiltrating TCRβ TRBJ2.1 gene frequency predicted response with an AUC of 0.94, where a lower frequency than 0.0460 had an 100% sensitivity and 85.7% specificity predicting response (Fig. 3D). Moreover, patients...
with low circulating TCRβ TRBJ2.6 gene frequency had statistically significant higher PFS than
patients with high frequency (p = 0.003) (Fig. 3E). Circulating TCRβ TRBJ2.6 gene frequency
predicted response with an AUC of 1.00, where a lower frequency than 0.0055 had an 100%
sensitivity and 100% specificity predicting response (Fig. 3F). Regarding the other genes,
despite most of them showed a clear tendency predicting PFS, they did not reach the statistical
significance, as represented in Supp. File S5. In fact, most of them predicted response with high
AUC (0.88-0.96) (Supp. File S6).

**Combined TCRβ signatures to predict response**

Although some of these TRBV/TRBJ genes did not reach the statistical significance in predicting
PFS, they were still able to accurately predict the clinical response. Based on this observation,
we decided to develop two predictive signatures that combine the previously described tumor-
infiltrating or circulating TCRβ variables associated with clinical response.

For the tissue predictive signature, we included variables such as evenness, TRBV6.5,
TRBV11.3, and TRBJ2.1. By performing Random Forest analyses, this signature was able to
predict clinical response with an AUC of 0.83. A positive result from this signature had a
sensitivity of 100% and a specificity of 71.4% in predicting response (Fig. 4A). In contrast, the
predictive signature for blood samples included the variables TRBV5.3, TRBV27, TRBV28,
TRBJ2.1, and TRBJ2.6. After conducting Random Forest analyses, this signature demonstrated a
predictive ability with an AUC of 0.92. In this case, a positive result had a sensitivity of 75% and
a specificity of 100% in predicting response (Fig. 4B).

These findings highlight the potential of using combined TCRβ variables to create effective
predictive signatures for both tissue and blood, aiding in the accurate prediction of clinical
response.

**Discussion**

ICI-based immunotherapy is the standard treatment for advanced NSCLC patients who lack of
actionable mutations and cannot undergo surgery. Despite ICIs have significantly improved
both life expectancy and quality of life for patients who benefit from it, the response rate is
still low (20-40%) [2,9,10]. To date, the lack of a reliable predictive response biomarkers leads
to the fact that most of patients are treated with ICIs, highlighting the huge need of identifying
those patients who can truly benefit from the treatment and offer alternative treatments to
the others [2,9,10]. Recently, the study of the TCR repertoire has emerged as a potential
predictive biomarker for immunotherapy response in various types of cancer (such as lung,
melanoma, liver or breast) [20–39]. It is noteworthy and important to consider the lack of methodological standardization. Although most of the starting materials, technologies, library preparation approaches, statistical methods and bioinformatics tools used are valid and approved, the combination of all these factors generates variability that greatly complicates the comparison of results [15,18,19,42,43]. Therefore, in this study we employed a reliable and efficient commercial kit for library preparation, along with user-friendly and practical software designed to work with FFPE samples, which are commonly used by clinicians. Additionally, we opted to use DNA instead of RNA due to its superior stability, a particularly important consideration when working with FFPE samples [18,42,43].

To the best of our knowledge this is the first study sequencing the tumor-infiltrating and circulating TCRβ/γ repertoire at baseline in advanced NSCLC patients treated with immunotherapy. Most studies regarding the TCR repertoire have focused solely on Tαβ cells [20–26,29–32,34–39], with very few studying the Tγδ cells, especially in solid tumors [27,28]. Although the mechanism of action of Tγδ cells is still not precisely known and they represent only a small portion of the T cell population, it has been shown that they can play an important role in antitumor response in both systemic and local immunosurveillance [14,17,27,28,33]. For this reason, we decided to include also the γ chain in this work. Moreover, many studies have analysed the TCR repertoire dynamics before and after treatment [20,23,25,26,30–32,34,37,39], providing valuable insights into the mechanisms behind the antitumor response reactivation. However, most advanced NSCLC patients do not survive long enough to start a second-line therapy. Thus, the study of the TCR repertoire as a biomarker for ICIs response should be applicable before the first-line treatment. Therefore, this study focused on the TCR repertoire at baseline.

In this study, we included 12 advanced NSCLC patients treated with ICIs at first and second line. Our results confirmed that R achieved greater PFS, with the same trend observed for OS. Interestingly, we found that R were older than NR. On the contrary, a comprehensive comparative analysis using a cohort of anti-PD1-treated patients with renal cell carcinoma identified no statistically significant difference in PFS or OS with age [44]. However, due the limited numbers of older patients treated with ICIs available for their analyses there is a need for further investigation of this subject in other cancer types, including NSCLC [44]. We also found older patients had a lower TCRβ/γ evenness in tissue but not in blood. In concordance, Erbe et al. reported a decrease in tissue TCR evenness with age in patients with adenocarcinoma [44]. Moreover, Dong et al. did not find such an association regarding blood
either [32]. On the contrary, following a different methodology, Liu et al. found that circulating TCRβ diversity was lower in older patients (>60 years) [26].

Regarding the histological classification and the TCR repertoire, previous studies showed contradictory associations. For instance, some studies described how patients with squamous cell carcinomas had lower tumor-infiltrating TCRβ evenness [24,38], clonal richness [38] or diversity [24] compared to patients with adenocarcinoma. In contrast, and as far as we know for the first time, we observed squamous had higher tumor-infiltrating TCRβ convergence than adenocarcinomas, while others did not find any association between histology and any TCR repertoire metric [31,32]. Although it should be verified in larger cohorts, this disparity might be due to our cohort containing only three squamous cell carcinomas.

Previous studies showed smoking can inhibit T cell proliferation and that smoking duration negatively correlates with TCR evenness [32,38]. While some studies found lower circulating TCRβ clonal richness in smokers [32], others noted higher tumor-infiltrating TCRβ evenness in never-smokers [24]. However, other studies found no association between TCR metrics and smoking [21,29]. Discrepancies may arise from how former smokers are classified. Given that smoking effects persist long-term, we included former smokers with smokers. Due to having only one non-smoker in our cohort, we lacked the statistical power to assess smoking’s impact on TCR repertoire metrics.

Our results showed that R exhibit low tumor-infiltrating TCRβ evenness at baseline, as well as been able of predicting a longer PFS. This could be due to the clonal expansion of the clones mediating the response. Concordant findings were also found in clear cell renal cell carcinoma, where R were observed to have lower tumor-infiltrating TCRβ evenness at baseline [39]. Moreover, Casarrubios et al. reported a high tumor-infiltrating TCRβ top 1% clonal space and a low evenness at baseline were involved in a pro-inflammatory tumor microenvironment associated with the complete pathological response of partially advanced stage NSCLC patients treated with neoadjuvant chemo-immunotherapy [34]. Additionally, they described how these same top 1% clones expanded in peripheral blood after treatment, suggesting a possible role of systemic immunosurveillance in preventing relapse. Despite us and Zhang et al., found this same trend regarding circulating TCRβ there is no study reporting such association [20,30,32,34,35,39].

In our case, as well as in other studies [26,30,32,34], we did not find any association between TCRβ diversity at baseline and the clinical response. On the contrary, other studies observed patients with greater circulating TCRβ diversity at baseline showed a better response and
longer PFS to ICIs + CTX or ICIs alone, highlighting the role of cytotoxic T cells in the antitumor response [25,31,35]. A possible explanation for this disparity could be that in peripheral blood, the abundance of tumor-specific T cell may be diluted by other factors (age, antigen exposure, or immunosuppression), affecting especially the diversity.

After studying the TCRβ gene usage regarding the response and the PFS to immunotherapy, we found some of them (TRBV5.3, TRBV6.5, TRBV11.3, TRBV27, TRBV28, TRBJ2.1 and TRBJ2.6) were associated with the immunotherapy response either regarding tissue or blood. Besides, TRBV11.3 and TRBJ2.6 also predicted PFS. Interestingly, R had a higher circulating TCRβ TRBV5.3 than NR. This could be due some of the neoantigens commonly expressed in the tumor presented high affinity for the sequence encode by this gene. As consequence, it would promote the recognition and eliminations of the tumor cells, mediating the antitumoral response. Similarly, Dong et al. reported a circulating TCRβ low TRBV6.5 and high TRBV20.1 at baseline showed a longer PFS and OS [32]. Plus, TRBV20.1 has been reported previously as one of the most used TCRβ V-gene in tumor-infiltrating T cells in several solid tumors [45]. Therefore, it seems the TRBV/J gene usage is also, being able predict the response both in tissue and blood. However, it is not clear which genes are more decisive, if there are ones.

Consequently, we decided to combine all the variables associated with response either in a tumor-infiltrating or circulating predictive signature using a random forest classifier. In both cases, they predicted the response accurately (AUC of 0.83 with 100% sensitivity and 71.4% specificity for tissue, and AUC of 0.92 with 75% sensitivity and 100% specificity for blood), suggesting that a combined signature could be more reliable than a specific variable to predict immunotherapy response.

Regarding TCRγ, we detected the same trend found in TCRβ repertoire without reaching the statistical significance, where R had a lower tumor-infiltrating and circulating TCRγ evenness at baseline than NR. However, we did not find association between TCR variables or V/J genes usage and immunotherapy response. Although a high number of distinct TCRγδ can theoretically be generated, certain rearrangements and chain pairings are substantially over-represented, biasing V(D)J recombination and the selection of functional clones [14]. This leads to various oligoclonal populations predominating in the circulation and in different tissues. Furthermore, there is minimal knowledge reported to date regarding the TCRγ repertoire as an immunotherapy response biomarker [27,28], so further validation will be necessary to verify these findings and draw more definitive conclusions.
It is noteworthy that in our study, we described both tumor-infiltrating and circulating TCRβ gene usage is associated with response and PFS, although different genes are present in each compartment. One possible explanation is that a group of Tαβ cells clones had recognized a series of neoantigens, expanding in peripheral blood and started infiltrating into the tumor, which would explain the low TCRβ evenness found in the tissue. However, some of the clones were remaining in circulation due to the expression of immune checkpoints by the tumor cells. This would change with treatment, allowing all the clones to identify these cancer cells as malignant, thereby promoting their complete infiltration and the elimination of the target. The fact of finding potential biomarkers in blood is particularly important because, although the tumor-infiltrating TCR repertoire can provide more detailed information about the immune response, it can be biased by biopsy due to intratumoral heterogeneity [7,8]. Additionally, finding minimally invasive circulating biomarkers that are easily measurable throughout therapy would facilitate their implementation in daily clinical practice, as the number of recent studies has demonstrated [23,25,26,31,32,35].

Conclusions

In summary, although much remains to be understood about the antitumor response, this study highlights the potential of TCRβ repertoire analysis as a potential biomarker of response to immunotherapy in NSCLC patients, as we found various tumor-infiltrating and circulating TCRβ variables at baseline able to predict response and progression-free survival. However, the present retrospective observational study is primarily exploratory, aiming to generate hypotheses rather than confirm them definitively. Consequently, further validation in larger cohorts will be necessary to substantiate these findings and for better understanding of the dynamics of the antitumor response.

Data accessibility

The datasets supporting the findings of this study are available from the corresponding authors upon reasonable request.

Author contributions

MPG performed formal analytics and statistical analysis and wrote the manuscript. MLQ enrolled patients, reviewed the clinical data annotations and wrote the manuscript. IAA performed and optimized DNA isolation and libraries preparation. MGG gave analytics and statistical analysis support. LJM evaluated tissue availability for FFPE sample selection. AFF
reviewed the clinical data annotations. SCF and EJL performed TCR sequencing. MMF conceived and designed the study, provided supervision and oversight for the management and execution of this study, acquired financial support for the project, and correct the manuscript. All authors provided critical review, edited the manuscript, and approved the final version.

Acknowledgements

Authors thank all the enrolled patients and their families. Samples were collected and stored by the Galicia Sur Health Research Institute (IIS Galicia Sur) Biobank (registry B.0000802).

Funding sources and disclosure of conflicts of interest

This work was supported by the Instituto de Salud Carlos III (ISCIII) and the European Social Fund ("Investing in your future") (PI21/00348, CP20/00188). MPG is currently supported by a grant for the predoctoral stage (IN606A-2024/017) from the Axencia Galega de Innovación-GAIN, Xunta de Galicia. MGG is currently supported by a grant for the postdoctoral stage (IN606B-2024/014) from the Axencia Galega de Innovación-GAIN, Xunta de Galicia. MMF is currently supported by the Miguel Servet program (CP20/00188) from the Instituto de Salud Carlos III (ISCIII).

The authors declare no competing interests.

References


### Table 1: Clinico-pathologic characteristics of patients.

<table>
<thead>
<tr>
<th>Clinico-pathologic characteristics</th>
<th>Responders (N = 5)</th>
<th>Non-responders (N = 7)</th>
<th>Cohort (N = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>74 (71-81)</td>
<td>60 (43-71)</td>
<td>70.5 (43-81)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoker</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Former smoker</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Never smoker</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squamous</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIIA</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IIIC</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IVA</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>IVB</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>PD1/PD-L1 staining</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Negative</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Response type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PR</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>SD</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>PD</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>NE</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>PFS (days)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>328 (116-952)</td>
<td>72 (3-148)</td>
<td>100 (4-952)</td>
</tr>
<tr>
<td>OS (days)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>405 (126-952)</td>
<td>113 (4-364)</td>
<td>225 (4-952)</td>
</tr>
</tbody>
</table>

Table 2: Tumor-infiltrating TCRβ TRBV and TRBJ genes frequency in responder (R) compared to non-responders (NR).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Responder (N = 5)</th>
<th>Non-responder (N = 7)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRBV6.5</td>
<td>Median</td>
<td>0.0082</td>
<td>0.0359</td>
</tr>
<tr>
<td>TRBV11.3</td>
<td>Median</td>
<td>0.0054</td>
<td>0.0077</td>
</tr>
<tr>
<td>TRBJ2.1</td>
<td>Median</td>
<td>0.0230</td>
<td>0.0581</td>
</tr>
</tbody>
</table>

P-value was obtained using Student’s t-test. N: number of patients, *: statistically significance, **: strongly statistically significance.

It is made available under a CC-BY-NC-ND 4.0 International license.
Table 3: Circulating TCRβ TRBV and TRBJ genes frequency in responder (R) compared to non-responder (NR). N, number; *, statistically significance; **, very statistically significance.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Responder (N = 4)</th>
<th>Non-responder (N = 6)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRBV5.3</td>
<td>Median 0.0003</td>
<td>0.0001</td>
<td>0.019*</td>
</tr>
<tr>
<td>TRBV27</td>
<td>Median 0.0423</td>
<td>0.0897</td>
<td>0.039*</td>
</tr>
<tr>
<td>TRBV28</td>
<td>Median 0.0176</td>
<td>0.0633</td>
<td>0.029*</td>
</tr>
<tr>
<td>TRBJ2.1</td>
<td>Median 0.0257</td>
<td>0.0662</td>
<td>0.017*</td>
</tr>
<tr>
<td>TRBJ2.6</td>
<td>Median 0.0032</td>
<td>0.0092</td>
<td>0.008**</td>
</tr>
</tbody>
</table>

P-value was obtained using Student’s t test. N: number of patients, *: statistically significance, **: strongly statistically significance.
Figure legends

Figure 1: A) Responders (R) have statistically significant longer PFS than non-responders (NR) (p = 0.015). P-value was obtained using Log-Rank test. B) There are an association between response and age, where R are statistically significant older than NR (p = 0.008). P-value was obtained using Student’s t test. C) Squamous have statistically significant higher tumor-infiltrating TCR8 convergence than adenocarcinomas (p = 0.026). P-value was obtained using Mann-Whitney test. D) Younger patients (< 71 years old) have statistically significant higher tumor-infiltrating TCR8 evenness than older patients (≥ 71 years old) (p = 0.022). P-value was obtained using Student’s t test. E) Younger patients have statistically significant higher tumor-infiltrating TCRy evenness than older patients (p = 0.041). P-value was obtained using Mann-Whitney test. N: number of patients, PFS: progression-free survival, *: statistically significance, **: strongly statistically significance.

Figure 2: A) Responders (R) have statistically significant lower tumor-infiltrating TCR8 evenness than non-responders (NR) (p = 0.044). P-value was obtained using Student’s t test. B) Tumor-infiltrating TCR8 evenness predicts response with an AUC of 0.86. A lower tumor-infiltrating TCR8 evenness than 0.795 have a 60% sensitivity and 100% specificity predicting response. C) Patients with low tumor-infiltrating TCR8 evenness have statistically significant higher PFS than patients with high tumor-infiltrating TCR8 evenness (p = 0.013). P-value was obtained using Log-Rank test. D) R have a non-statistically significant lower circulating TCR8 evenness than NR (p = 0.139). P-value was obtained using Student’s t test. E) R have a non-statistically significant lower tumor-infiltrating TCRy evenness than NR (p = 0.149). P-value was obtained using Mann-Whitney test. F) R have a non-statistically significant lower circulating TCRy evenness than NR (p = 0.114). P-value was obtained using Mann-Whitney test. AUC: Area Under the Curve, N: number of patients, PFS: progression-free survival, *: statistically significance.

Figure 3: A) Patients with low tumor-infiltrating TCR8 TRBV11.3 gene frequency have statistically significant higher PFS than patients with high tumor-infiltrating TCR8 TRBV11.3 gene frequency (p = 0.002). P-value was obtained using Log-Rank test. B) Tumor-infiltrating TCR8 TRBV11.3 gene frequency predicts response with an AUC of 0.86. A lower tumor-infiltrating TCR8 TRBV11.3 gene frequency than 0.0057 have an 80% sensitivity and 85.7% specificity predicting response. C) Patients with low tumor-infiltrating TCR8 TRBJ2.1 gene frequency have non-statistically significant higher PFS than patients with high tumor-infiltrating TCR8 TRBJ2.1 gene frequency (p = 0.051). P-value was obtained using Log-Rank test. D) Tumor-infiltrating TCR8 TRBJ2.1 gene frequency predicts response with an AUC of 0.94. A lower tumor-infiltrating TCR8 TRBJ2.1 gene frequency than 0.0460 have an 100% sensitivity and 85.7% specificity predicting response. E) Patients with low circulating TCR8 TRBJ2.6 gene frequency have statistically significant higher PFS than patients with high circulating TCR8 TRBJ2.6 gene frequency (p = 0.003). P-value was obtained using Log-Rank test. F) Circulating TCR8 TRBJ2.6 gene frequency predicts response with an AUC of 1.00. A lower circulating TCR8 TRBJ2.6 gene frequency than 0.0055 have an 100% sensitivity and 100% specificity predicting response. AUC: Area Under the Curve, N: number of patients, PFS: progression-free survival, *: statistically significance, **: strongly statistically significance.

Figure 4: A) A predictive signature including tumor-infiltrating TCR8 variables (evenness, TRBV6.5 gene frequency, TRBV11.3 gene frequency and TRBJ2.1 gene frequency) predicts response with an AUC of 0.83. A positive result has an 100% sensitivity and 71.4% specificity predicting response. B) A predictive signature including circulating TCR8 variables (TRBV5.3 gene frequency, TRBV27 gene frequency, TRBV28 gene frequency, TRBJ2.1 gene frequency and TRBJ2.6 gene frequency) predicts response with an AUC of 0.92. A positive result has an 75% sensitivity and 100% specificity predicting response. AUC: Area Under the Curve.
Figure 1

A. PFS (%)

Time (days)

- Responder (N = 5)
- Non-responder (N = 7)

p = 0.015**

B. Age (years)

Response

- Responder (N = 5)
- Non-responder (N = 7)

p = 0.003**

C. Tissue TCRβ Convergency

Tumor type

- Adenocarcinoma (N = 9)
- Squamous (N = 3)

p = 0.026**

D. Tissue TCRβ Evenness

Age (years)

- < 71 (N = 6)
- ≥ 71 (N = 6)

p = 0.022**

E. Tissue TCRγ Evenness

Age (years)

- < 71 (N = 6)
- ≥ 71 (N = 6)

p = 0.041**
Figure 2

A. Tissue TCRβ Evenness

B. Tissue TCRβ Evenness

C. PFS (%)

D. Blood TCRβ Evenness

E. Tissue TCRγ Evenness

F. Blood TCRγ Evenness
Figure 3

(A) High tissue TCRβ TRBV11.3 (N = 6) and Low tissue TCRβ TRBV11.3 (N = 6).

(B) Tissue TCRβ TRBV11.3

(C) High tissue TCRβ TRBJ2.1 (N = 6) and Low tissue TCRβ TRBJ2.1 (N = 5).

(D) Tissue TCRβ TRBJ2.1

(E) High blood TCRβ TRBJ2.6 (N = 5) and Low blood TCRβ TRBJ2.6 (N = 5).

(F) Blood TCRβ TRBJ2.1

AUC: 0.86
AUC: 0.94
AUC: 1.00

*p = 0.002**
*p = 0.051*
*p = 0.003**
Figure 4

A. Tissue Predictive Signature

B. Blood Predictive Signature

Sensitivity vs. Specificity

AUC: 0.83

AUC: 0.92
A

B

C

D

E

F

It is made available under a CC-BY-NC-ND 4.0 International license. The copyright holder for this version posted August 2, 2024. doi: medRxiv preprint