

# Effective TCR Diversity as a Superior Quantitative Representation in Tumor Immunity

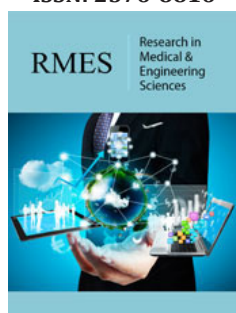
Wei Ping<sup>1,2</sup>, Li Hui<sup>2</sup>, Zhang Xiao-Min<sup>3</sup>, LI Jie-Wei<sup>2</sup>, Yang Siye<sup>1</sup> and Li Junli<sup>2\*</sup>

<sup>1</sup>Sichuan Academy of Traditional Chinese Medicine/SiChuan Institute for Translational Chinese Medicine/Translational Chinese Medicine Key Laboratory of Sichuan Province, China

<sup>2</sup>College of Computer Science, Sichuan Normal University, China

<sup>3</sup>School of Life Science and Engineering, Southwest University of Science and Technology, China

ISSN: 2576-8816



**\*Corresponding author:** Li Junli, College of Computer Science, Sichuan Normal University, China

**Submission:**  July 15, 2024

**Published:**  July 26, 2024

Volume 11 - Issue 2

**How to cite this article:** Wei Ping, Li Hui, Zhang Xiao-Min, LI Jie-Wei, Yang Siye and Li Junli\*. Effective TCR Diversity as a Superior Quantitative Representation in Tumor Immunity. Res Med Eng Sci. 11(2). RMES.000756. 2024. DOI: [10.31031/RMES.2024.11.000756](https://doi.org/10.31031/RMES.2024.11.000756)

**Copyright@** Li Junli. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

## Abstract

The occurrence and development of tumors are closely related to the immune system's response. Identifying indicators that can effectively quantify and characterize the immune status of tumors is crucial for cancer diagnosis, staging, treatment, and prognosis. T-cell receptor (TCR) diversity and specificity have been identified as potential tumor biomarkers. However, merely analyzing TCR diversity or specificity independently does not always correlate positively with effective immune presentation of tumors. We define TCR effective diversity as the diversity of specific TCRs that can recognize and bind to specific antigenic peptides, thereby eliciting a positive immune response. By leveraging a specificity clustering model, we constructed an effective diversity index and evaluated its quantitative analysis in cancer classification across two datasets. The study demonstrates that the TCR effective diversity index, which integrates both diversity and specificity, can more accurately quantify the role of TCR in tumor immunity.

**Keywords:** TCR; Effective diversity; Tumor immunity

## Introduction

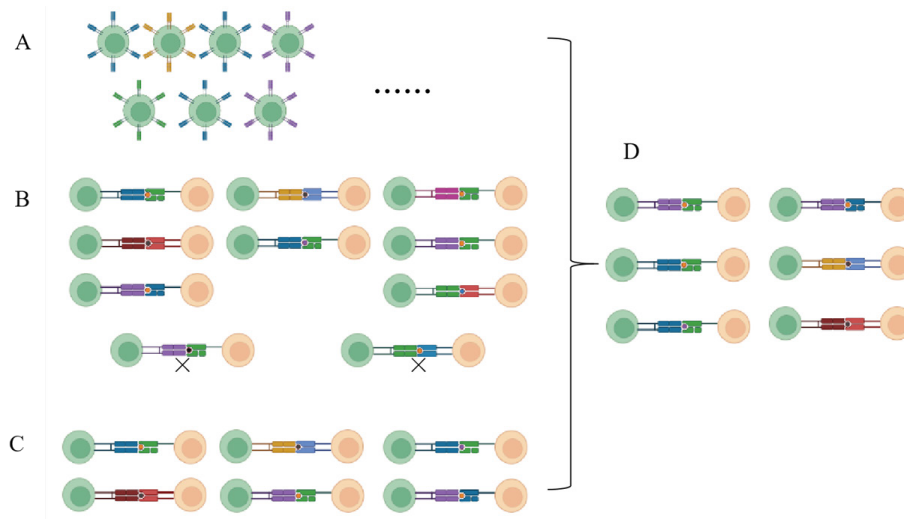
Early diagnosis of cancer is crucial, but tumors are difficult to prevent and treat. As a key immune molecule, TCR can recognize tumor neoantigens and trigger immune responses [1,2], making it an ideal tumor biomarker. However, the human TCR repertoire exhibits significant individual variability and extreme diversity, with a vast number of unique sequences, low inter-individual TCR overlap, and specificity for HLA and antigens. These characteristics complicate the quantitative characterization of tumor immunity.

The diversity of the TCR repertoire is highly correlated with pathological states, and studies have extensively explored tumor associated TCR diversity. However, sheer diversity alone does not consistently correlate with effective tumor immunity: while higher TCR diversity generally indicates a robust anti-tumor immune capacity, predicting better therapeutic responses and outcomes [3-10], exceptions exist [11-13]. The specific binding of TCRs to antigen-MHC complexes is crucial for monitoring immune-tumor interactions. Recent studies have begun to explore experimental techniques and predictive models for TCR specificity, such as GLIPH [14], Panpep [15], and PMTnet [16]. Current research efforts have focused on analyzing TCR diversity or specificity alone.

It is hypothesized that an effective diversity metric combining diversity and specificity can more accurately quantify the role of TCRs in tumor immunity (Figure 1). Effective diversity of TCRs is defined as the diversity of specific TCRs that can successfully recognize a specific antigen and elicit a positive immune response [17]. Based on the hypothesis that structural

similarity determines functional similarity, a TCR structural similarity clustering algorithm was utilized to construct the TCR effective diversity index. The predictive performance of this index

in distinguishing between cancer patients and healthy individuals was analyzed. The results validated that the effective diversity index can more accurately measure the body's anti-tumor immune levels.



**Figure 1:** Schematic Diagram of TCR Diversity, Specificity, Efficacy, and Effective Diversity.

**Note:** A represents TCR diversity; B represents TCR specificity; C represents TCR efficacy; D represents TCR effective diversity. Green cells - T cells, yellow cells - antigen-presenting cells; TCR - different colors represent different TCRs; antigen peptide - different colors represent different antigen peptides; MHC - different colors represent different MHCs; × indicates antigen not recognized by TCR.

### Methods and Result

The TCR diversity index is calculated at the amino acid sequence level using commonly employed metrics such as Shannon entropy, Gini-Simpson index, D50, and Pielou, as shown in formulas 1-4.

#### Shannon Entropy:

$$H = -\sum_{i=1}^N p_i \ln p_i \quad (1)$$

In these formulas, H represents Shannon entropy, N denotes the total observed number of clonotypes, also referred to as richness, and  $P_i$  represents the relative frequency of the i-th TCR clonotype.

#### Gini-Simpson Index:

$$D = 1 - \sum_{i=1}^N P_i^2 \quad (2)$$

In the formula, D represents the Gini-Simpson diversity index, N denotes the total observed number of clonotypes, also referred to as richness, and  $p_i$  represents the relative frequency of the i-th TCR clonotype.

#### D50:

$$D_{50} = \frac{C}{N} \quad (3)$$

In the formula, C represents the number of dominant clonotypes that account for 50% of the total sequences, and N denotes the total observed number of clonotypes.

#### Pielou:

$$J = \frac{H}{\ln(n)} \quad (4)$$

In the formula, H represents Shannon entropy, and N denotes the total observed number of clonotypes.

Based on the reasonable hypothesis that TCR structural similarity largely determines functional similarity, we constructed the TCR effective diversity index (EDI) using a specificity clustering model:

$$EDI = n^2 \times \left( -\sum_{i=1}^n (S_i \times T_i) \ln(S_i \times T_i) \right) \quad (5)$$

where n represents the number of clusters;  $S_i$  represents the proportion of TCR sequences in the i-th cluster relative to the total number of TCR sequences;  $T_i$  represents the proportion of observed TCR clonotypes in the i-th cluster relative to the total number of clonotypes.

In Formula 5, the clustering algorithm uses GLIPH [14], an algorithm for identifying groups of T-cell receptors (TCRs) with similar functional properties. Compared to diversity indices, the effective diversity index comprehensively reflects both specificity and diversity.

We analyzed and compared the classification performance of individual diversity indices and the effective diversity index in distinguishing between healthy individuals and cancer patients using data from 83 healthy individuals and 87 cancer patients in Data 1 [1], and 20 healthy individuals and 16 colorectal cancer patients in Data 2 [4]. Among the diversity indices, Shannon entropy exhibited the best classification performance, while D50 showed the weakest classification performance. The AUC values of the effective diversity index were higher than those of the diversity indices (Table 1). This indicates that the single effective diversity index is effective for cancer classification and performs better than the diversity indices.

**Table 1:** AUC Values for TCR Diversity Indices and Effective Diversity Index Classification in Data 1 and Data 2.

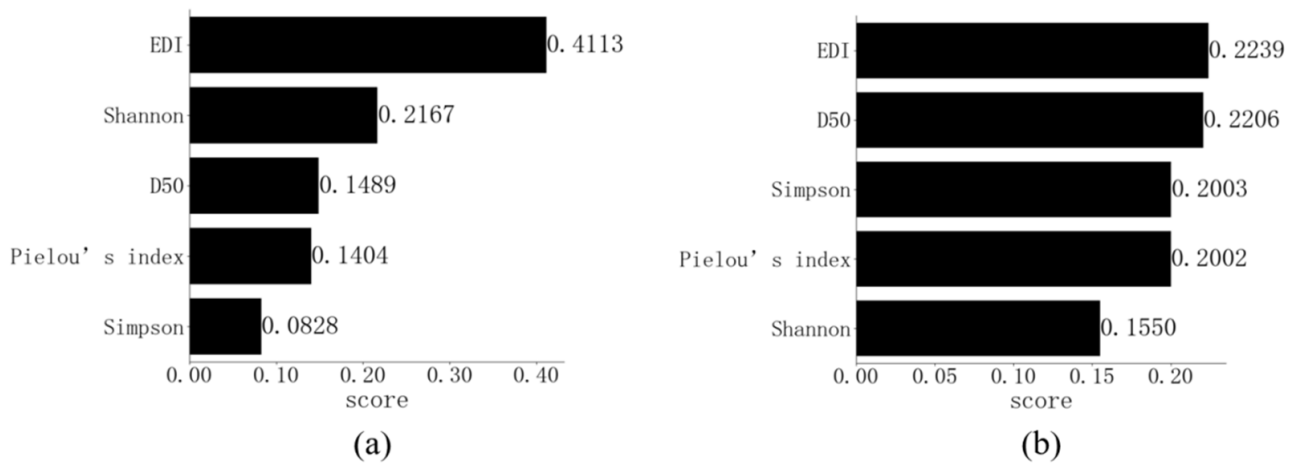
Index Types		AUC Values in Data 1	AUC Values in Data 2
TCR Diversity Indices	Shannon	0.8061	0.5844
	Simpson	0.6572	0.5344
	D50	0.6034	0.5125
	Pielou's index	0.6164	0.5531
Effective Diversity Index	EDI	<b>0.915</b>	<b>0.7281</b>

We constructed Random Forest (RF) classifiers based on diversity indices, and a combination of both types of indices.

We divided the dataset into a training set and a test set, which accounted for 80% and 20% of the total data, respectively. The RF classifier is trained on the training set and its performance is evaluated on the test set. The results show that the RF classifier incorporating the combined indices demonstrated the best classification performance, indicating that the inclusion of effective diversity indices improved classification performance (Table 2). This suggests that TCR effective diversity indices are effective for cancer classification, enhancing classification performance. Furthermore, feature importance rankings (Figure 2) show that effective diversity indices are the most important classification factors in the constructed RF model, significantly improving the classifier's performance.

**Table 2:** Comparison of RF model classification performance based on different indices.

Data	RF Model	Accuracy	Recall	F1 Score	AUC
Data 1	TCR Diversity Indices	0.8529	0.7857	0.8148	0.9179
	Composite Indices	<b>0.8824</b>	<b>0.9286</b>	<b>0.8667</b>	<b>0.9286</b>
Data2	TCR Diversity Indices	0.75	1	0.6667	0.8333
	Composite Indices	0.75	1	0.6667	<b>0.9167</b>



**Figure 2:** RF Feature Importance Analysis.

- a) Ranking of RF feature importance scores for Dataset 1.
- b) Ranking of RF feature importance scores for Dataset 2.

**Conclusion**

Numerous studies on TCR diversity and specificity have demonstrated their significant roles in tumor immunity research. However, most of these studies analyze diversity or specificity in isolation. In tumor immunity, diversity and specificity are often interrelated, and their combined effect needs to be validated through efficacy. TCR diversity does not distinguish between specific and non-specific TCRs, and the role of non-specific TCRs in diversity is minimal. While some tumor therapy studies show a correlation between high diversity and better prognosis, and low diversity levels with poorer prognosis, there are exceptions. Therefore, the diversity that exerts positive immune effects must be effective diversity; if diversity lacks efficacy, it still fails to achieve its purpose.

The comprehensive analysis of TCR diversity, specificity, and effectiveness can provide a more accurate understanding of the immune system's function and improve strategies for the prevention, diagnosis, and treatment of immune-related diseases. The TCR effective diversity index, constructed by combining structural similarity expression of specificity with diversity quantification, to some extent reflects effective diversity. If the TCR effective diversity is sufficiently high, T cells can recognize a sufficient number of tumor neoantigens, thereby increasing the coverage and efficiency of the immune system and inhibiting tumor occurrence and development. The TCR effective diversity index is effective for cancer classification and outperforms diversity indices. Preliminary validation has shown that the effective diversity index is a better quantitative measure of the body's anti-tumor immune level.

However, inferring TCR antigen specificity based solely on TCR sequences remains challenging. It may be feasible to construct a TCR effective diversity index based on receptor-ligand binding by incorporating TCR-antigen affinity prediction models. Nevertheless, the high false positive rate of current TCR-antigen affinity models is a significant issue that needs to be addressed. Additionally, a more refined effective diversity calculation model must consider the impact of non-linear factors such as immune evasion and the synergistic effects of various cell types. These issues require further research and exploration.

With the ongoing development and refinement of single-cell TCR sequencing and TCR-pMHC structure prediction technologies [18,19], TCR effective diversity is expected to become an important tool for cancer immune monitoring and prognosis prediction in the future.

## References

1. Simnica D, Akyüz N, Schliffke S, Mohme M, Wenserski LV, et al. (2019) T cell receptor next-generation sequencing reveals cancer-associated repertoire metrics and reconstitution after chemotherapy in patients with hematological and solid tumors. *Oncoimmunology* 8(11): e1644110.
2. Klebanoff CA, Chandran SS, Baker BM, Quezada SA, Ribas A (2023) T cell receptor therapeutics: immunological targeting of the intracellular cancer proteome. *Nat Rev Drug Discov* 22(12): 996-1017.
3. Charles J, Mouret S, Challende I, Leccia MT, Fraipont FD, et al. (2020) T-cell receptor diversity as a prognostic biomarker in melanoma patients. *Pigment Cell Melanoma Res* 33(4): 612-624.
4. Chen YT, Hsu HC, Lee YS, Liu H, Tan BCM, et al. (2021) Longitudinal high-throughput sequencing of the t-cell receptor repertoire reveals dynamic change and prognostic significance of peripheral blood TCR diversity in metastatic colorectal cancer during chemotherapy. *Front Immunol* 12: 743448.
5. Manuel M, Tredan O, Bachelot T, Clapisson G, Courtier A, et al. (2012) Lymphopenia combined with low TCR diversity (divpenia) predicts poor overall survival in metastatic breast cancer patients. *Oncoimmunology* 1(4): 432-440.
6. Cui JH, Lin KR, Yuan SH, Jin YB, Chen XP, et al. (2018) TCR repertoire as a novel indicator for immune monitoring and prognosis assessment of patients with cervical cancer. *Front Immunol* 9: 2729.
7. Jia Q, Zhou J, Chen G, Shi Y, Yu H, et al. (2015) Diversity index of mucosal resident T lymphocyte repertoire predicts clinical prognosis in gastric cancer. *Oncoimmunology* 4(4): e1001230.
8. Arakawa A, Vollmer S, Tietze J, Galinski A, Heppt MV, et al. (2019) Clonality of CD<sub>4</sub><sup>+</sup> blood T cells predicts longer survival with CTLA<sub>4</sub> or PD<sub>1</sub> checkpoint inhibition in advanced melanoma. *Front Immunol* 10: 1336.
9. Han J, Duan J, Bai H, Wang Y, Wan R, et al. (2020) TCR repertoire diversity of peripheral PD<sub>1</sub><sup>+</sup> CD8<sup>+</sup> T cells predicts clinical outcomes after immunotherapy in patients with non-small cell lung cancer. *Cancer Immunol Res* 8(1): 146-154.
10. Postow MA, Manuel M, Wong P, Yuan J, Dong Z, et al. (2015) Peripheral T cell receptor diversity is associated with clinical outcomes following ipilimumab treatment in metastatic melanoma. *J Immunother Cancer* 3: 23.
11. Jin YB, Luo W, Zhang GY, Lin KR, Cui JH, et al. (2018) TCR repertoire profiling of tumors, adjacent normal tissues, and peripheral blood predicts survival in nasopharyngeal carcinoma. *Cancer Immunol Immunother* 67(11): 1719-1730.
12. Wang X, Zhang B, Yang Y, Zhu J, Cheng S, et al. (2019) Characterization of distinct T cell receptor repertoires in tumor and distant non-tumor tissues from lung cancer patients. *Genomics Proteomics Bioinformatics* 17(3): 287-296.
13. Bai X, Zhang Q, Wu S, Zhang X, Wang M, et al. (2015) Characteristics of tumor infiltrating lymphocyte and circulating lymphocyte repertoires in pancreatic cancer by the sequencing of T cell receptors. *Sci Rep* 5: 13664.
14. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, et al. (2017) Identifying specificity groups in the T cell receptor repertoire. *Nature* 547(7661): 94-98.
15. Gao Y, Gao Y, Fan Y, Zhu C, Wei Z, et al. (2023) Pan-peptide meta learning for T-cell receptor-antigen binding recognition. *Nat Mach Intell* 5(3): 236-249.
16. Lu T, Zhang Z, Zhu J, Wang Y, Jiang P, et al. (2021) Deep learning-based prediction of the T cell receptor-antigen binding specificity. *Nat Mach Intell* 3(10): 864-875.
17. Zhang Xiao-Min, Zhao Hai-Yang, Li Jie-Wei, et al. (2023) Effective diversity research of TCR in tumor immunity. *Chinese Bulletin of Life Sciences* 35(11): 1450-1461.
18. Oliveira G, Wu CJ (2023) Dynamics and specificities of T cells in cancer immunotherapy. *Nat Rev Cancer* 23(5): 295-316.
19. Tan CL, Lindner K, Boschert T, Meng Z, Ehrenfried AR, et al. (2024) Prediction of tumor-reactive T cell receptors from scRNA-seq data for personalized T cell therapy. *Nat Biotechnol*.