

DOI: 10.13376/j.cbbs/2023158

文章编号: 1004-0374(2023)11-1450-12

## 肿瘤免疫中TCR的有效多样性研究

张小敏<sup>1</sup>, 赵海洋<sup>1</sup>, 李杰威<sup>2</sup>, 刘晓艳<sup>1</sup>, 侯大斌<sup>1</sup>, 张翼冠<sup>3</sup>, 魏平<sup>3\*</sup>

(1 西南科技大学生命科学与工程学院, 绵阳 621010; 2 四川师范大学计算机科学学院, 成都610101;

3 四川省中医药科学院/四川省中医药转化医学中心/中医药转化医学四川省重点实验室, 成都 610041)

**摘要:** T细胞是抗癌免疫系统的主要组成部分, 其表面表达的特定T细胞受体(T cell receptor, TCR)负责识别抗原肽并引发免疫反应。在肿瘤免疫研究中, 分析TCR的多样性和特异性对于理解适应性免疫应答和免疫治疗至关重要。目前的研究侧重于单独分析多样性或特异性, 尽管这两者都与肿瘤的有效免疫相关, 但并非总是呈现正相关。本文探讨了肿瘤免疫中TCR的多样性和特异性研究进展及面临的挑战, 指出综合分析多样性和特异性的必要性, 给出综合性概念——TCR有效多样性的定义, 介绍并讨论了已有的少量相关研究成果。揭示有效多样性的科学意义需要综合分析TCR的多样性、特异性和有效性, 结合充分的实验来进行建模和量化表征。有效多样性对于深入了解TCR在肿瘤中的作用至关重要, 对于未来的癌症早筛、治疗和预后具有潜在的重要意义。

**摘要:** 肿瘤免疫; TCR; 有效多样性; 生物标志物

中图分类号: R392 文献标志码: A

## Effective diversity research of TCR in tumor immunity

ZHANG Xiao-Min<sup>1</sup>, ZHAO Hai-Yang<sup>1</sup>, LI Jie-Wei<sup>2</sup>, LIU Xiao-Yan<sup>1</sup>,

HOU Da-Bin<sup>1</sup>, ZHANG Yi-Guan<sup>3</sup>, WEI Ping<sup>3\*</sup>

(1 School of Life Science and Engineering, Southwest University of Science and Technology, Mianyang 621010, China;

2 College of Computer Science, Sichuan Normal University, Chengdu 610101, China; 3 Sichuan Academy of Traditional

Chinese Medicine/Sichuan Center of Translational Medicine/Sichuan Key Laboratory of Translational Medicine of Traditional Chinese Medicine, Chengdu 610041, China)

**Abstract:** T cells are a major component of the anti-cancer immune system, and their surface-expressed specific T cell receptors (TCRs) are responsible for recognizing antigen peptides and initiating immune responses. In the field of tumor immunology, understanding the diversity and specificity of TCRs is crucial for comprehending adaptive immune responses and immunotherapy. Current research primarily focuses on analyzing diversity or specificity separately, and although both are associated with effective immunity to tumors, they do not always show a positive correlation. This article discusses the research progress and challenges of TCR diversity and specificity in tumor immunity, points out the necessity of comprehensive analysis of diversity and specificity, gives the definition of comprehensive concept “TCR effective diversity”, and introduces and analyzes a few relevant research results. Revealing the scientific significance of effective diversity requires a comprehensive analysis of the diversity, specificity, and effectiveness of TCRs, combined with adequate experiments for modeling and quantitative characterization. Effective diversity is essential for in-depth understanding of the role of TCR in tumors and has potentially important implications for early screening, treatment and prognosis of cancer in the future.

**Key words:** tumor immunity; TCR; effective diversity; biomarker

收稿日期: 2023-05-19; 修回日期: 2023-07-14

基金项目: 四川省科技厅项目(2018RZ0026)

\*通信作者: E-mail: w.ping@foxmail.com

T 细胞受体 (T cell receptor, TCR) 是在 T 细胞表面表达的抗原受体, 它是由  $\alpha$  和  $\beta$  链 ( $\alpha\beta$ TCR, 约 95% 的 T 细胞 TCR 由  $\alpha$  和  $\beta$  链组成) 或  $\gamma$  和  $\delta$  链 ( $\gamma\delta$ TCR) 组合形成的跨膜糖蛋白异二聚体<sup>[1]</sup>。 $\alpha\beta$ TCR 的互补决定区 3 (complementary determining region 3, CDR3) 识别并结合主要组织相容性复合体 (major histocompatibility complex, MHC) 呈递的抗原肽<sup>[2-3]</sup>, 而  $\gamma\delta$ TCR 可以不依赖 MHC 识别抗原肽。TCR 作为体内免疫系统中与肿瘤杀伤直接相关的指标, 可以特异性识别肿瘤新抗原, 进而产生肿瘤免疫应答<sup>[4]</sup>, 其多样性可以反映个体的免疫能力。TCR 监测肿瘤具有适用性广泛、非侵入性、副作用小等优点, 且能实现个体化动态早期检测, 是更理想的潜在肿瘤生物标志物。

TCR 组库 (TCR immune repertoire) 是指机体在特定时间内免疫循环中 T 细胞克隆的总和。随着测序技术的不断发展, 高通量测序的 TCR 免疫组库正成为临床研究适应性免疫反应不可缺少的工具, 广泛应用于肿瘤研究领域<sup>[5-6]</sup>。TCR 的多样性是指由于 TCR 基因重排的组合多样性 (V、D、J 重排)、连接多样性 (N 序列的插入或缺失)、阴性选择和阳性选择产生不同克隆型的 TCR<sup>[1]</sup>, 其中 CDR3 具有高度的多样性 (图 1)。TCR 特异性是指 TCR 识别结合带有特定基序的抗原, 同时它还具有对 MHC 分子的特异性:  $CD8^+$  T 细胞的 TCR 只能识别 MHC-I 类分子呈递的抗原肽;  $CD4^+$  T 细胞的 TCR 只能识别 MHC-II 类分子呈递的抗原肽。

当前研究通常独立地分析 TCR 的多样性和特异性, 这两者对于肿瘤免疫都十分重要: 特异性 TCR 可识别特定抗原, 而多样性确保其能识别广泛的抗原, 有必要将它们综合考虑。虽然高度多样的 TCR 和特异性 TCR 都具有潜在的益处, 但它们必须在免疫监视和免疫清除方面发挥作用, 才能真正对肿瘤产生积极影响。需要通过有效性来验证多样性和特异性的综合作用, 有效性指的是 TCR 在特

定免疫响应中的有效性或功能性表现。

TCR 有效多样性是一个综合性概念, 文献<sup>[7]</sup>提到它是指有足够的 TCR 可以识别肿瘤新生抗原 (复合物)。这个定义不够准确, 因为能识别抗原的 TCR 并不一定能够产生有效的免疫反应。相关研究<sup>[8]</sup>提出了一个综合体现多样性和特异性并表现有效性的计算指标, 即特异性 TCR 克隆型数量与靶向突变数量之比。这个指标一定程度上体现了有效多样性的概念, 但适用性有限, 如  $CD8^+$  T 细胞上程序性死亡受体 1 (programmed death 1, PD-1) 的过表达可能会使细胞免疫功能丧失<sup>[9]</sup>,  $CD4^+$  T 细胞则会面临更加复杂的情况。TCR 有效多样性还没有明确的定义, 其概念需要进一步完善。

本文将 TCR 有效多样性定义为“能成功识别特定抗原、引起正免疫应答的特异性 TCR 的多样性”。这个定义表达了 TCR 的多样性、特异性和有效性的关系 (图 2)。

本文分析了 TCR 多样性指标及其在肿瘤免疫中的应用, 总结了 TCR- 表位特异性预测工具, 探讨了当前研究面临的挑战, 为更好地理解 TCR 有效多样性这一概念提供了基础和支撑。关于 TCR 的有效多样性研究还不够成熟, 后续需要结合多样性指标、特异性预测工具, 通过数学建模和实验验证来分析 TCR 的有效多样性, 以增进对 TCR 在肿瘤免疫中的作用的理

## 1 TCR的多样性

### 1.1 多样性指标

在 TCR 免疫组库的分析研究中, 可以在不同时间不同组织中追踪 TCR 的多样性, 以此来捕获免疫适应度。TCR 的多样性包括丰富度 (richness) 和均匀度 (evenness), 前者测量样品中 TCR 克隆型的数量, 后者测量不同克隆型的相对丰度<sup>[10]</sup>。常见衡量 TCR 多样性的指标有 Chao1 指数 (Chao1 index)<sup>[11]</sup>、 $D_{50}$  (diversity 50)<sup>[12]</sup>、Pielou 指数 (Pielou's index)<sup>[13]</sup>、香农熵 (Shannon entropy)<sup>[14-15]</sup> 和辛普森指数 (Simpson index)<sup>[16]</sup> (表 1)。较高的 TCR 多样性表明功能性免疫系统具有良好的抗肿瘤反应的能力, 多样性的丧失可能是侵袭性肿瘤导致免疫系统失效的结果<sup>[1]</sup>。

### 1.2 多样性指标的应用

在肿瘤治疗中, TCR 多样性可以作为一个重要的生物标志物, 用于预测患者的疗效和预后, 且在肿瘤筛查中具有潜在价值<sup>[17-32]</sup> (表 2)。有研究分

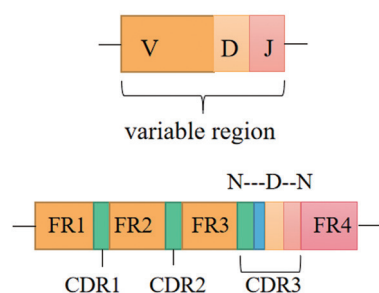
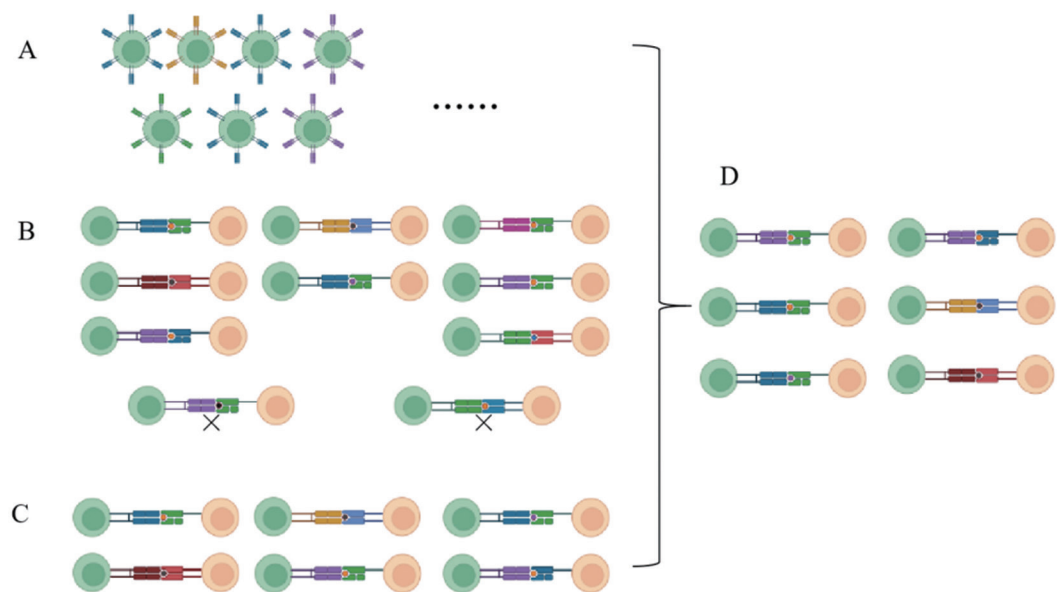


图1 TCR可变区结构示意图



A表示TCR的多样性；B表示TCR的特异性；C表示TCR的有效性；D表示TCR的有效多样性。绿色细胞：T细胞；黄色细胞：抗原递呈细胞；TCR：不同的颜色代表不同的TCR；抗原肽：不同的颜色代表不同的抗原肽；MHC：不同的颜色代表不同的MHC；×代表抗原不被TCR识别。

图2 TCR的多样性、特异性、有效性及有效多样性示意图

| 表1 TCR多样性指标 |                                                     |                                                    |
|-------------|-----------------------------------------------------|----------------------------------------------------|
| 多样性指标       | 公式                                                  | 特征                                                 |
| 丰富度         |                                                     | 衡量不同克隆型的总数                                         |
| Chao1指数     | $chao1 = S_{obs} + \frac{F_1(F_1 - 1)}{2(F_2 + 1)}$ | 反映物种丰富度，不能衡量相对丰度                                   |
| $D_{50}$    | $D_{50} = \frac{C}{X}$                              | 衡量多样性均匀度，越大越均匀                                     |
| Pielou指数    | $J = \frac{H'}{\ln(S)}$                             | 衡量物种的分布均匀性，越大越均匀                                   |
| 香农熵         | $H' = - \sum_{i=1}^{S_{obs}} p_i \ln p_i$           | 衡量物种均匀度和丰富度的综合指标，侧重于物种丰富度，既不偏爱稀有克隆，也不偏爱优势克隆；越多样性越高 |
| 辛普森指数       | $D = \sum_{i=1}^{S_{obs}} p_i^2$                    | 衡量物种均匀度和丰富度的综合指标，侧重于相对丰度，给予优势克隆型更多的权重；越多样性越低       |

注：S<sub>obs</sub>是观测到的克隆型数；F<sub>1</sub>为只有一条序列的克隆型数；F<sub>2</sub>为只有两条序列的克隆型数；C为序列数达到总数50%时的优势克隆型总数；X为总克隆型数；S为物种丰富度；p<sub>i</sub>是TCR第i种克隆型的相对频率。

表2 TCR多样性指标在肿瘤免疫中的应用

| 多样性指标    | 疗效                                                                                         | 预后                                                                                                     |
|----------|--------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| Chao1指数  | 结直肠癌 <sup>[21]</sup>                                                                       | 结直肠癌 <sup>[21]</sup>                                                                                   |
| $D_{50}$ | 黑色素瘤 <sup>[36]</sup>                                                                       | 黑色素瘤 <sup>[10]</sup>                                                                                   |
| 丰富度      | 肺癌 <sup>[20]</sup> 、黑色素瘤 <sup>[28, 35]</sup>                                               | 黑色素瘤 <sup>[10, 25]</sup>                                                                               |
| 香农熵      | 肺癌 <sup>[19-20]</sup> 、黑色素瘤 <sup>[28, 36]</sup> 、结直肠癌 <sup>[21]</sup> 、宫颈癌 <sup>[23]</sup> | 宫颈癌 <sup>[23]</sup> 、肺癌 <sup>[20]</sup> 、胃癌 <sup>[24]</sup> 、结直肠癌 <sup>[21]</sup> 、鼻咽癌 <sup>[26]</sup> |
| 辛普森指数    | 结直肠癌 <sup>[30]</sup>                                                                       | 胰腺癌 <sup>[18]</sup> 、肺癌 <sup>[27]</sup>                                                                |

析了来自 346 名健康人和癌症患者生成的大约 880 万个 TCR 读数的数据集, 发现 TCR 组库的多样性与病理之间存在高度关联性<sup>[4]</sup>。TCR 多样性指数  $D_{50}$  能用来区分正常人和癌症患者<sup>[17]</sup>, 癌症患者 TCR  $\beta$  基因库的  $D_{50}$  值显著低于健康人。正常情况下, 机体 TCR 克隆是均匀分布的, CDR3 长度呈高斯分布, 与 TCR 多样性密切相关; 而疾病状态下, 由于肿瘤抗原持续刺激导致某些 T 细胞克隆群高度扩增, TCR 组库多样性显著下降<sup>[33-34]</sup>。但在一项胰腺癌研究中, 患者和健康个体的外周多样性没有差异<sup>[18]</sup>。

患者的免疫状态可能会影响治疗反应, 治疗前较高的 TCR  $\beta$  多样性增加了存在免疫相关抗肿瘤 T 细胞群的可能性, 可能有助于患者针对特定肿瘤抗原的免疫激活<sup>[19, 35]</sup>, 预示了对治疗的良好反应。在大部分研究中, 高多样性和更好的预后相关, 较低的多样性水平与较差的预后相关<sup>[10, 20-25]</sup>。但也有例外: 在鼻咽癌患者中, 外周血多样性越高, 预后越差<sup>[26]</sup>; 肺癌患者肿瘤中较高的 TCR 多样性与临床预后差显著相关<sup>[27]</sup>。因此, 需要更多的研究或者扩大研究队列来确认 TCR 多样性与预后之间的联系是否具有肿瘤特异性。经免疫治疗后 TCR 多样性增加的患者临床疗效好<sup>[20, 28-29, 36]</sup>, 治疗后 TCR 多样性的增加可能更多地与患者免疫系统的一般激活能力有关; 治疗后 TCR 多样性下降的患者具有更好的治疗反应<sup>[21, 30]</sup>, 表明免疫系统正在增强对肿瘤的有效应答, 这说明光凭 TCR 多样性的动态变化不足以评估疗效。

此外, 肿瘤组织和正常组织之间 TCR 序列的多样性和相似性水平也已被证明可以提供患者临床结局的相关信息, 如肿瘤和相邻正常组织之间 TCR 组库的相似性与肝细胞癌的预后相关<sup>[31]</sup>; 根据 TCR 多样性的动态变化也能预测癌症的复发, 部分晚期直肠癌患者接受放化疗前后 TCR 组库的较大变化与更好的无复发生存期相关<sup>[32]</sup>。

TCR 多样性已被用来评估中药的作用, 如研究含何首乌制剂对 TCR 多样性的影响<sup>[37]</sup>; 经过中药恢复治疗的新冠康复者的 TCR 多样性指数高于未经过恢复治疗的康复者<sup>[38]</sup>。部分中药材或成分通过靶向免疫系统显示出免疫调节功能和抗肿瘤作用<sup>[39-41]</sup>。与未使用中药的患者相比, 使用中药的患者死亡风险降低了 32%<sup>[42]</sup>, 表明癌症患者可以从中药的免疫调节作用中受益, 中药辅助治疗可以提高癌症患者的总体生存率, 但最好在癌症早期阶段进

行干预, 以发挥中药的最大效力。目前已有将 TCR 多样性用于中药治疗的肿瘤免疫研究<sup>[43]</sup>: 含有 12 种中草药的金复康口服液可以上调外周血  $CD4^+$  T、 $CD8^+$  T 和 NKT 细胞的比例, 通过调节 TCR 基因的表达, 逆转肿瘤转移引起的 TCR 变化, 提高 T 细胞对肿瘤抗原的识别, 表明该口服液可以通过调节 TCR 来抑制肺癌的转移, 科学客观地评估了中药免疫调节作用; 但金复康对 TCR 多样性没有显著影响, 我们推测它可能引起了特异性 TCR 的变化, 有待进一步研究验证。

## 2 TCR的特异性

被 TCR 所识别的抗原肽称之为 T 细胞表位, TCR-表位结合特异性反映了机体的抗肿瘤能力, 同时这种特异性可以为免疫治疗提供思路, 如 TCR-T 疗法<sup>[2]</sup>。因此, TCR 与抗原肽-MHC 复合物的结合是一个研究热点, 对于监测免疫系统与肿瘤之间的相互作用至关重要。大量的生化实验被用来测定它们之间的结合, 如四聚体测定和抗原肽刺激实验(图 3、4), 但是这样可能错过重要的抗原特异性 TCR 或错误地报告非结合性 TCR 是抗原特异的<sup>[44-45]</sup>, 并且耗时长、成本高。同时, 由于一个抗原肽可以由多个 MHC 分子呈现<sup>[46]</sup>, 一个 TCR 可能与不同的抗原肽-MHC 复合物相互作用<sup>[47]</sup>, 这种复杂度导致实验操作难度大, 技术上具有挑战性。因此, 人们开发了大量的生物信息学算法来预测这种相互作用, 节省了繁琐的实验时间和昂贵的实验成本。

计算模型已被成功训练以预测抗原肽-MHC 结合, 但不是所有 MHC 呈递的抗原肽都具有免疫原性。为了预测哪些 MHC 限制性抗原肽确实能成为 T 细胞表位, 需要更好地理解 TCR 与其同源靶标之间的相互作用<sup>[48]</sup>。IEDB 数据库、VDJdb、McPAS-TCR 和 PIRD 的 TBAdb 储存了大量的 TCR-表位配对信息<sup>[49-52]</sup>, 为 TCR-抗原肽识别预测模型的构建提供了良好的数据<sup>[48, 53-72]</sup>。与同一抗原肽相互作用的 TCR 序列之间可以推断出共同的模式, 从而可以基于序列相似性将 TCR 特异性聚类<sup>[53]</sup>; 随机森林分类器的创建进一步证明了开发高性能 TCR-抗原肽结合预测模型的可行性<sup>[54]</sup>。特异性模型主要分为聚类模型和 TCR-抗原肽亲和力预测模型这两类, 部分模型总结如下(表 3、4)。

除了以上模型, 还有 CDRdist<sup>[68]</sup>、TCRMatch<sup>[69]</sup>等也可以用来分析 TCR 的抗原特异性。部分模



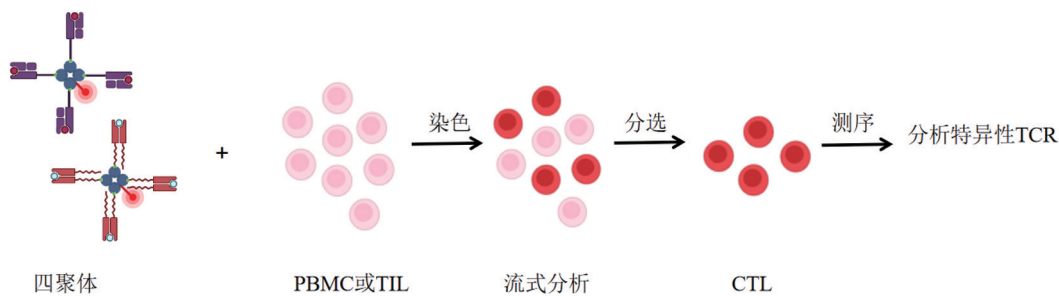


图3 四聚体染色分析T细胞特异性

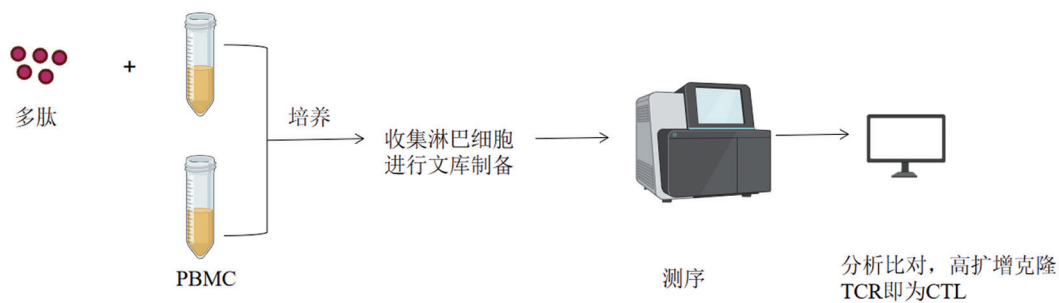


图4 多肽刺激PBMC分析T细胞特异性

表3 聚类模型

| 工具                       | 特征                                                      | 优缺点                                                                                      |
|--------------------------|---------------------------------------------------------|------------------------------------------------------------------------------------------|
| GLIPH <sup>[53]</sup>    | 基于全局和局部CDR3序列相似性，将TCR聚类成具有相同特异性的群组                      | 考虑CD4 <sup>+</sup> T和CD8 <sup>+</sup> T细胞 $\alpha/\beta$ 链CDR3区域；可利用MHC信息；存在误差和假阳性，特异性较低 |
| TCRdist <sup>[55]</sup>  | 根据TCRs空间上的距离度量鉴定出识别相同抗原决定簇的TCR序列                        | 考虑了 $\alpha$ 链和 $\beta$ 链的8个CDR(包括CDR2.5)区域；不适用于大型TCR队列                                  |
| iSMART <sup>[56]</sup>   | 基于CDR3 $\beta$ 比对，将相似的TCR CDR3序列分组到抗原特异性簇中              | 可应用于肿瘤单细胞测序和RNA-seq数据；仅考虑 $\beta$ 链CDR3区域                                                |
| SwarmTCR <sup>[57]</sup> | 使用CDR $\beta$ 或 $\alpha$ 序列之间的距离预测TCR的特异性，优化了各个CDR区域的权重 | 考虑 $\alpha$ 链和 $\beta$ 链的8个CDR(包括CDR2.5)区域；不适用于分析大型数据集                                   |
| DeepTCR <sup>[58]</sup>  | 通过学习TCR的CDR3序列和V/D/J基因使用来探索TCR的特征                       | 改进了TCR序列的特征化，可分析配对的 $\alpha/\beta$ 链，利用了V/D/J基因；在嘈杂的数据集中存在局限性                            |
| GIANA <sup>[59]</sup>    | 基于几何等距，将序列比对和聚类问题转化为高维欧几里得空间中的经典近邻搜索                    | 可以快速处理大型TCR队列，高聚类特异性；仅适用于具有相同长度的序列，未考虑MHC                                                |

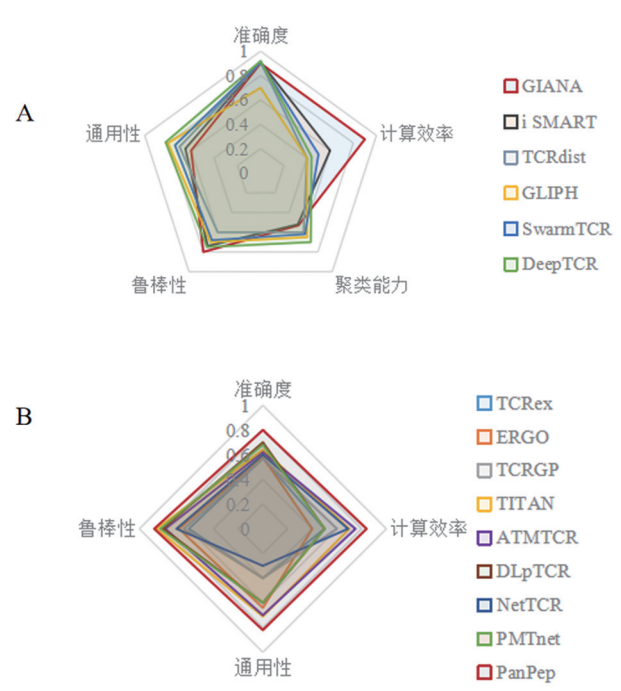
型 如 NetTCR-2.0<sup>[70]</sup>、ERGO-II<sup>[71]</sup>、GLIPH2<sup>[72]</sup> 等 经过改进，预测性能和范围均得到了提升。模型的性能受训练数据集的类型、大小、质量和所用方法学的影响。我们对聚类模型和 TCR- 抗原结合预测模型的性能各进行了几个方面的比较，其中聚类模型的特有评估指标“聚类能力”指的是可聚类 TCR 的数量。综合对比得出，聚类模型中 GIANA 的总体性能最好，TCR- 抗原结合预测模型中 PanPep 总体性能最好 ( 图 5)。

在使用这些工具时，应当根据具体的研究问题和可用数据来评估和选择合适的模型。(1) 在仅已

知 TCR 序列或包含 V、J 基因的情况下可以使用聚类模型 ( 使用 MHC 准确性更高 )，可将 TCR 序列进行初步特异性分析，分为不同的特异性组，但不能预测可能识别的表位；TCR 序列的多样性和复杂性使得聚类结果可能存在误差和偏差，幼稚 TCR 群体也会具有一定相似度的 TCR<sup>[53]</sup>，但是它们可能并不具有抗原特异性。(2) 在抗原肽信息已知时，则可以使用 TCR- 表位识别预测工具预测 TCR 可能识别哪些抗原肽，有助于发现新抗原；使用时注意某些模型的适用范围仅限于特定的抗原肽，如 TCRex<sup>[60]</sup>、TCRGP<sup>[62]</sup>；这类模型忽略了 MHC 的多

表4 TCR-抗原肽结合预测模型

| 工具                      | 特征                                               | 优缺点                                                             |
|-------------------------|--------------------------------------------------|-----------------------------------------------------------------|
| NetTCR <sup>[48]</sup>  | 基于卷积神经网络预测TCR CDR3β序列和HLA-A*02:01所呈现的抗原肽序列之间的亲和力 | 只适用于HLA-A*02:01限制肽和β链CDR区域; 表位/CDR3s均短于10个氨基酸                   |
| TCRex <sup>[60]</sup>   | 基于随机森林算法, 利用TCR CDR3β序列和V/J基因来学习识别特定抗原肽所共有的TCR模式 | 具有较高的识别率, 可能发现其他技术遗漏的表位特异性TCRs; 仅限于β链和有限表位, 未考虑MHC              |
| ERGO <sup>[61]</sup>    | 基于LSTM和自动编码器预测给定TCR CDR3β序列和抗原肽序列是否结合            | 可以预测更广泛的肽, CD4 <sup>+</sup> 和CD8 <sup>+</sup> T细胞均可预测; 性能因肽类型而异 |
| TCRGP <sup>[62]</sup>   | 基于高斯过程预测TCRs是否识别特定的表位                            | 可以了解CDR类型在识别不同表位中的权重, 预测精度较高; 仅限于有限表位, 未考虑MHC                   |
| TITAN <sup>[63]</sup>   | 使用双峰神经网络对TCR序列和表位进行编码, 用于预测TCR-表位结合概率            | 可以预测未见表位; 仅限于β链, 泛化能力受限                                         |
| ATMTTCR <sup>[64]</sup> | 基于多头自注意机制预测给定的一对TCR和表位序列之间的结合亲和力                 | 对未见TCR的预测精确度较高; 对罕见或未见表位的结合亲和力预测表现不佳                            |
| DLpTCR <sup>[65]</sup>  | 多模态集成深度学习框架, 用于预测TCR的单链/配对链与MHC分子呈现的肽之间的相互作用     | 同时考虑了α链和β链; 不能推广到未见肽或外源性肽                                       |
| PMTnet <sup>[66]</sup>  | 基于迁移学习预测TCR CDR3β序列、I类MHC和其呈递的抗原肽之间的亲和力          | 预测精度较高, 考虑了MHC; 仅限于MHC I类限制肽和CD8 <sup>+</sup> T细胞β链CDR3区域       |
| PanPep <sup>[67]</sup>  | 基于元学习和神经图模型的TCR-抗原亲和力预测模型, 可预测任何抗原肽的TCR结合特异性     | 可以预测TCR与任何肽的结合特异性, 准确度较高; 未考虑MHC和α链                             |



A为聚类模型性能比较雷达图; B为TCR-抗原结合预测模型性能比较雷达图。图中的数值代表评分。

图5 模型性能比较雷达图

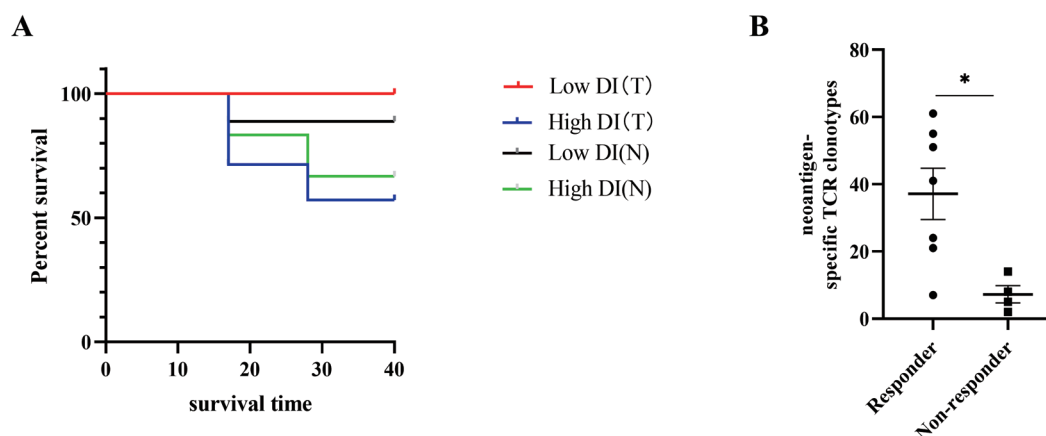
多样性对 TCR 识别抗原的影响。(3) 如果 MHC 也已知, 则可以使用三合一模型 (TCR-P-MHC), 能加速有效靶表位和特异性 TCR 的初步选择。总之, 这些

工具在肿瘤诊断、新生抗原预测、过继细胞疗法等诸多领域具有广泛的应用价值。

3 TCR的有效多样性

关于 TCR 多样性、特异性的多项研究表明它们在肿瘤免疫研究中均具有重要作用, 但是大多都是单独分析多样性或特异性。多样性和特异性在肿瘤免疫中通常是相互关联的, 需要通过有效性来验证其综合作用。TCR 多样性不关注特异性 TCR 和非特异性 TCR 的区别, 且非特异性 TCR 在多样性中的作用很小: 如在一项肺癌研究中, 公开的临床数据表明高 TCR 多样性的患者预后更差<sup>[27]</sup>, 可能是因为多样的 TCR 中不具有或仅具有少量特异性 TCR (图 6A)。单独研究特异性则忽略了 TCR 多样性的潜在益处, 足够多样的特异性 TCR 会产生更好的有效免疫应答: 如在一项黑色素瘤研究中, 每个患者体内都具有特异性 TCR, 分析对比数据后发现特异性 TCR 足够多样的患者疗效更好<sup>[8]</sup>(图 6B)。综合分析 TCR 多样性、特异性和有效性能更全面准确地理解免疫系统的功能, 改善免疫相关疾病的预防、诊断、治疗策略。

本文将 TCR 的有效多样性定义为能识别和结合特定抗原肽、引起正免疫应答的特异性 TCR 的多样性, 即有效 TCR 的多样性, 能很好地将 TCR 的多样性、特异性和有效性结合起来, 具有更好的



A图描述了TCR低/高多样性的肺癌患者预后情况, Low DI表示患者TCR多样性低于中位数, High DI表示患者TCR多样性高于中位数。T表示肿瘤组织, N表示正常组织。数据来源于参考文献<sup>[27]</sup>。B图是对治疗有/无反应的黑色素瘤患者体内新抗原特异性TCR数量的比较, X轴为对抗PD-1治疗有反应和无反应的患者, Y轴为新抗原特异性TCR克隆型的数量。数据来源于参考文献<sup>[8]</sup>。

图6 TCR多样性和特异性与疗效、预后的关系

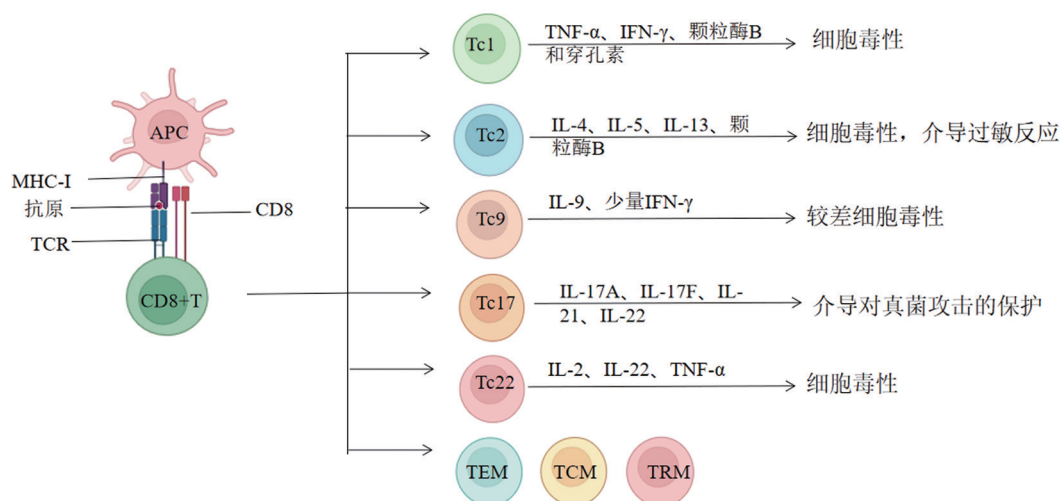
研究价值。TCR的有效多样性如果足够高, T细胞就可以识别足够多的肿瘤新生抗原, 从而提高免疫系统的覆盖范围和效率, 抑制肿瘤的发生和发展, 故TCR不仅要多, 关键还要有效<sup>[7]</sup>。部分TCR-T细胞临床试验显示出抗肿瘤疗效, 表现出了特异性TCR的有效性<sup>[73-75]</sup>。还可通过四聚体实验或抗原肽刺激实验来获取特异性的TCR, 流式细胞仪分选出特异性T细胞或克隆这种特异性TCR用来修饰自体T细胞, 再通过其与肿瘤细胞共培养来验证其有效性<sup>[7, 76-78]</sup>。

Beshnova等<sup>[79]</sup>基于与癌症相关的TCR可能存在共同的生化特征这一假设开发了DeepCAT, 可以用其来预测抗原特异性TCR, 引入的癌症评分可以作为临床早期癌症的判断指标。对于输入的每个TCR都会有一个癌症相关概率, 癌症评分则是这些癌症相关概率的平均值, 这在一定程度上反映了涉及不同患者和肿瘤类型的癌症相关TCR的多样性。但是在这些与癌症相关的TCR中, 可能会存在非特异性的TCR, 是否有效尚未可知, 并且慢性炎症的存在也会影响癌症评分, 在实际应用中有待进一步的实验验证和临床研究。该工具和癌症评分为后续TCR有效多样性的研究提供了有价值的参考。

最新发表的一篇文章阐述了肿瘤和血液中效应CD8<sup>+</sup>T细胞多样性是免疫治疗的关键<sup>[8]</sup>, 文中提出的TCR多克隆指标(特异性TCR克隆型数量与其靶向突变的数量之比)类似于有效多样性指标, 可

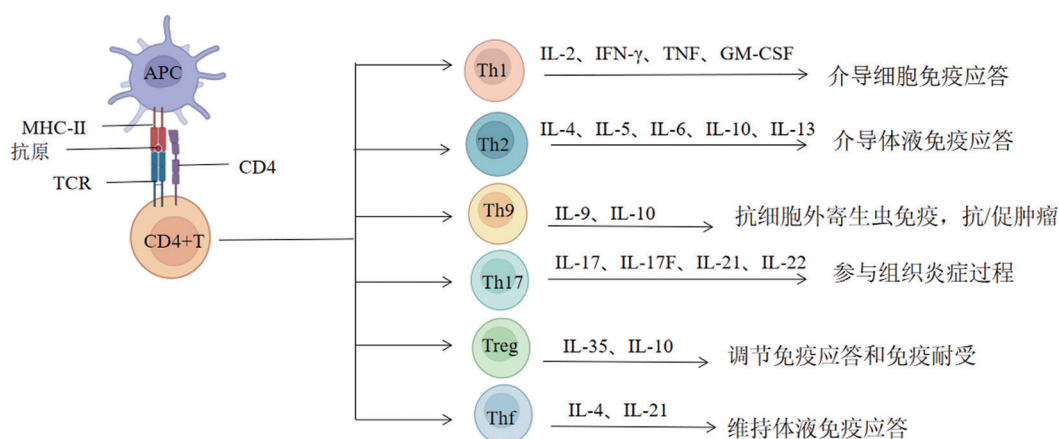
作为抗原特异性T细胞的肿瘤反应性和治疗反应的生物标志物。通过四聚体染色筛选的特异性TCR经实验鉴定为具有识别和结合抗原的能力, 从中选择部分TCR测试了其对自体肿瘤细胞系的杀伤功能, 证明了其有效性。通过研究结果可以发现, 在对黑色素瘤抗PD-1治疗有反应的患者中, 单个新抗原对应多样的特异性TCR; 而对治疗无反应的患者, 单个新抗原仅对应少数或单一的特异性TCR。该TCR多克隆指标结合了TCR特异性和有效性, 但是仅考虑了特异性TCR的丰富度, 不能完全表现出其多样性, 且该指标的构建依赖于实验, 所需成本较高, 难以大规模推广。

在大部分研究中, CD8<sup>+</sup>T细胞是肿瘤免疫的主要研究目标, 它具有确定的细胞毒性反应(图7), 负责识别并杀伤感染细胞或肿瘤细胞, 结合强度较高的TCR-抗原肽复合物通常意味着更强的抗原识别和更强的免疫反应。随着研究的深入, 人们逐渐意识到CD4<sup>+</sup>T细胞在肿瘤免疫中的重要性, 其TCR识别抗原并接受激活信号, 受特定的细胞因子或TCR传递的信号强度(P:MHC-II复合物的数量、TCR与P:MHC-II复合物的结合强度)的影响分化为不同的Th亚群<sup>[80-81]</sup>(图8), 产生不同的免疫应答。理解这些亚型的发展、功能和相互关系仍然是一个复杂的问题。CD4<sup>+</sup>T细胞和CD8<sup>+</sup>T细胞的协同作用对于发挥强大的抗肿瘤免疫至关重要, 综合评估两者的有效多样性, 有助于更好地理解免疫系统的功能。



Tc, 细胞毒性T细胞; TRM, 组织驻留记忆性T细胞; TCM, 中枢记忆性T细胞; TEM, 效应记忆性T细胞。

图7 CD8<sup>+</sup> T细胞识别抗原及分化过程示意图



Treg, 调节性T细胞; Thf, 滤泡辅助T细胞。

图8 CD4<sup>+</sup> T细胞识别抗原及分化过程示意图

## 4 讨论

T 细胞在抗肿瘤反应中发挥着重要作用, 研究癌症患者的 TCR 组库有助于了解抗肿瘤反应。随着高通量测序技术的进步, TCR 组库分析变得更加容易。较长的测序读长可以分析更加完整、更加复杂的免疫组库, 通常更有利于全面、准确地分析 TCR 库, 选择读长时需要综合考虑研究目的、资源和成本等因素。评估 TCR 真实多样性的关键是获得准确的 T 细胞含量和种类, 样品选择和扩增方法会造成 TCR 多样性评估偏差; 目前的 TCR 多样性指标都受采样不足的影响<sup>[82]</sup>, TCR 多样性测量指标有待优化, 也可以开发能准确计算多样性指标的模型, 从而解决仅仅取样少量细胞而低估 TCR 多

样性的问题。肿瘤浸润淋巴细胞 (tumor infiltrating lymphocytes, TIL) 的 TCR 多样性与相应外周血 TCR 多样性差异较大<sup>[83]</sup>, 且 VJ 区域也具有差异, 但外周血中某些特异性 TCR 可以区分健康人和肿瘤患者, 且外周血无创易获取, 大多数研究选取的是外周血而非 TIL。

虽然大量研究表明了 TCR 多样性的动态变化能作为预测疗效、预后的生物标志物, 但是当前研究样本量有限, 仍然需要研究队列来证明 TCR 多样性的确具有预测价值。根据 TCR 的多样性评估患者治疗反应和预后不够准确<sup>[17, 25-26]</sup>, 重点要关注肿瘤特异性 TCR, 其与治疗响应密切相关<sup>[84]</sup>。由于  $\beta$  链具有更大的组合和连接多样性<sup>[85]</sup>, 大多数研究通常局限于  $\beta$  链, 而 TCR 的完整功能是  $\alpha$  链和  $\beta$



链协同作用的结果,加入对 $\alpha$ 链的分析会更准确地反映免疫状态。CDR3区域具有高度多样性,是与抗原肽直接接触的区域,通常只对CDR3区域测序分析,忽略了CDR1和CDR2在抗原识别和免疫应答中的作用。在研究和理解免疫状态时,综合分析 $\alpha$ 链和 $\beta$ 链的CDR区域至关重要。

确定哪些TCRs与患者特异性新抗原结合是治疗设计的一个重要问题。T细胞库分析的目标之一是仅使用序列信息预测个体T细胞的特异性<sup>[54]</sup>,人工智能预测T细胞特异性的发展得益于生化实验提供的大规模高质量TCR-抗原肽结合配对信息及机器学习的发展。TCR的交叉反应性使T细胞能够识别广泛的病原体特异性表位<sup>[86-89]</sup>,且TCR序列长度不均一,预测TCR-抗原结合十分具有挑战性。由于预测模型的准确性和适用范围受限,结合实验验证和优化机器学习算法仍然是提高预测模型性能的关键。准确预测模型的开发取决于成对的 $\alpha/\beta$ TCR序列数据<sup>[90]</sup>,CDR1和CDR2也携带与TCR特异性相关的重要信息,整合 $\alpha$ 、 $\beta$ 链所有六个CDR区域的信息会获得更高的总体性能<sup>[53,55]</sup>。也应考虑MHC的有效多样性,以确保能呈递多样的抗原供TCR识别,MHC等位基因的加入对于预测性能的提升也有帮助。目前的TCR-表位预测结合模型主要是关于CD8<sup>+</sup>TCR和I类MHC呈递表位之间的相互作用,MHC II类抗原肽结合的特征使MHC II免疫原性表位的鉴定复杂化,CD4<sup>+</sup>T细胞和MHC-II呈递表位之间相互作用的模型可能是未来研究的一个难点及重点。

关于TCR多样性在中药治疗肿瘤方面的研究还处于初级阶段,需要更多的实验室研究和临床研究来探讨中药在肿瘤治疗中的潜在应用以及与TCR多样性、特异性的关系。同时,由于TCR有效多样性与肿瘤的发生发展密切相关,中药具有免疫调节作用,TCR有效多样性在中药治疗肿瘤方面极具应用前景,有待深入研究。对TCR多样性的分析没有考虑到其特异性,多样的TCR中可能包括不针对患者新抗原的TCR,不能够准确反映患者抗肿瘤能力;部分T细胞特异性预测工具可以分析TCR的特异性和相似性,但不能反映出其是否有效,还需结合实验。少量研究展现了TCR有效多样性显著的预测价值,综合考虑CD4<sup>+</sup>T细胞和CD8<sup>+</sup>T细胞,结合实验、TCR多样性指标和特异性预测工具,有助于构建高质量的有效多样性指标用于肿瘤免疫

治疗。

## 5 总结和展望

综上所述,TCR可以作为一种病原体感染、癌症和自身免疫疾病的个性化诊断工具,分析其多样性和特异性对于发展生物标志物至关重要,还有助于指导患者的后续治疗。如前所述,实验研究中存在的不足有待改进,尤其需要对综合性指标——TCR的有效多样性进行深入探究,融合TCR的多样性、特异性和有效性,有助于更好地理解 and 利用TCR在免疫应答和免疫治疗中的作用。

TCR有效多样性的研究还处于起步阶段,未来随着TCR免疫组库测序成本的下降,预测工具的更加可靠,可以通过获得人体外周血或TIL,搜集肿瘤新抗原库或对患者癌组织进行外显子测序,进行多肽刺激细胞实验、四聚体染色实验、T细胞与肿瘤细胞或肿瘤类器官共培养实验,结合高通量测序技术和生物信息学工具,从以下几个方向实现对TCR有效多样性的综合评估。

(1) 免疫监测:评估免疫系统状态,有助于开发更有效的疾病诊断工具,为诊断、预防、治疗疾病提供指导性的方案。

(2) 临床评估:疗效、预后评估,了解TCR有效多样性与治疗效果、疾病进展和患者生存率之间的关系,分析疗效差异原因及与MHC的相关性,指导用药以提升疗效。同时,可以结合中药疗效来探究其对免疫系统的调节作用及中药在肿瘤预防、预后中的潜在价值,以对患者实施最佳的中医药治疗方案。

(3) 药物研发:了解TCR与不同疾病发病机制和免疫应答之间的关联,研发通过调节免疫力或TCR活性对抗疾病的新型药物,如中药新药的研发,特别是药食同源中药,可用于增强免疫力并抵抗不同类型的疾病。结合传统中医药深入研究TCR有效多样性,有望客观量化、揭示中医特色治未病技术对重大疾病的预防作用。

## [参 考 文 献]

- [1] Aran A, Garrigós L, Curigliano G, et al. Evaluation of the TCR repertoire as a predictive and prognostic biomarker in cancer: diversity or clonality? *Cancers (Basel)*, 2022, 14: 1771
- [2] Zhang J, Wang L. The emerging world of TCR-T cell trials against cancer: a systematic review. *Technol Cancer Res Treat*, 2019, 18: 1533033819831068

- [3] Rossjohn J, Gras S, Miles JJ, et al. T cell antigen receptor recognition of antigen-presenting molecules. *Annu Rev Immunol*, 2015, 33: 169-200
- [4] Simnica D, Akyüz N, Schliffke S, et al. T cell receptor next-generation sequencing reveals cancer-associated repertoire metrics and reconstitution after chemotherapy in patients with hematological and solid tumors. *Oncoimmunology*, 2019, 8: e1644110
- [5] 李静. 弥漫大B细胞淋巴瘤患者化疗前后TCR免疫组库分析[D]. 郑州: 郑州大学, 2021
- [6] 杨桂淇, 薛雯, 戴勇. T细胞抗原受体 $\beta$ 链互补决定区3受体库高通量测序分析及其在重要疾病中的研究应用. *国际免疫学杂志*, 2016, 39: 522-6
- [7] 张翼冠, 赵军宁. 基于肿瘤超早期免疫应答信息放大指标naCTL, TCR/BCR/HLA有效多样性等免疫信息学多维指标的中医(肿瘤)治未病测知、评价模型(体系)构建. *中国中药杂志*, 2019, 44: 3129-34
- [8] Puig-Saus C, Sennino B, Peng S, et al. Neoantigen-targeted CD8<sup>+</sup> T cell responses with PD-1 blockade therapy. *Nature*, 2023, 615: 697-704
- [9] Liu Y, Liang X, Dong W, et al. Tumor-repopulating cells induce PD-1 expression in CD8<sup>+</sup> T cells by transferring kynurenine and AhR activation. *Cancer Cell*, 2018, 33: 480-94
- [10] Charles J, Mouret S, Challende I, et al. T-cell receptor diversity as a prognostic biomarker in melanoma patients. *Pigment Cell Melanoma Res*, 2020, 33: 612-24
- [11] Chao A. Nonparametric estimation of the number of classes in a population. *Scand J Statist*, 1984, 11: 265-70
- [12] Han J. Immunodiversity assessment method and its use. U.S. 0183969A1[P]. 2012-7-19
- [13] Pielou EC. The measurement of diversity in different types of biological collections. *J Theor Biol*, 1966, 13: 131-44
- [14] Chiffelle J, Genolet R, Perez MA, et al. T-cell repertoire analysis and metrics of diversity and clonality. *Curr Opin Biotechnol*, 2020, 65: 284-95
- [15] Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*, 1948, 27: 379-423
- [16] Simpson EH. Measurement of diversity. *Nature*, 1949, 163: 688
- [17] Zhuo Y, Yang X, Shuai P, et al. Evaluation and comparison of adaptive immunity through analyzing the diversities and clonalities of T-cell receptor repertoires in the peripheral blood. *Front Immunol*, 2022, 13: 916430
- [18] 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. *Sci Rep*, 2015, 5: 13664
- [19] Han J, Duan J, Bai H, et al. TCR repertoire diversity of peripheral PD-1<sup>+</sup>CD8<sup>+</sup> T cells predicts clinical outcomes after immunotherapy in patients with non-small cell lung cancer. *Cancer Immunol Res*, 2020, 8: 146-54
- [20] Liu YY, Yang QF, Yang JS, et al. Characteristics and prognostic significance of profiling the peripheral blood T-cell receptor repertoire in patients with advanced lung cancer. *Int J Cancer*, 2019, 145: 1423-31
- [21] Chen YT, Hsu HC, Lee YS, et al. 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*, 2022, 12: 743448
- [22] Manuel M, Trédan O, Bachelot T, et al. Lymphopenia combined with low TCR diversity (divpenia) predicts poor overall survival in metastatic breast cancer patients. *Oncoimmunology*, 2012, 1: 432-40
- [23] Cui JH, Lin KR, Yuan SH, et al. TCR repertoire as a novel indicator for immune monitoring and prognosis assessment of patients with cervical cancer. *Front Immunol*, 2018, 9: 2729
- [24] Jia Q, Zhou J, Chen G, et al. Diversity index of mucosal resident T lymphocyte repertoire predicts clinical prognosis in gastric cancer. *Oncoimmunology*, 2015, 4: e1001230
- [25] Arakawa A, Vollmer S, Tietze J, et al. Clonality of CD4<sup>+</sup> blood T cells predicts longer survival with CTLA4 or PD-1 checkpoint inhibition in advanced melanoma. *Front Immunol*, 2019, 10: 1336
- [26] Jin Y, Luo W, Zhang G, et al. TCR repertoire profiling of tumors, adjacent normal tissues, and peripheral blood predicts survival in nasopharyngeal carcinoma. *Cancer Immunol Immunother*, 2018, 67: 1719-30
- [27] 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. *Genomics Proteomics Bioinformatics*, 2019, 17: 287-96
- [28] Robert L, Tsoi J, Wang X, et al. CTLA4 blockade broadens the peripheral T-cell receptor repertoire. *Clin Cancer Res*, 2014, 20: 2424-32
- [29] Kansy BA, Shayan G, Jie HB, et al. T cell receptor richness in peripheral blood increases after cetuximab therapy and correlates with therapeutic response. *Oncoimmunology*, 2018, 7: e1494112
- [30] Tamura K, Hazama S, Yamaguchi R, et al. Characterization of the T cell repertoire by deep T cell receptor sequencing in tissues and blood from patients with advanced colorectal cancer. *Oncol Lett*, 2016, 11: 3643-9
- [31] Lin KR, Deng FW, Jin YB, et al. T cell receptor repertoire profiling predicts the prognosis of hbv-associated hepatocellular carcinoma. *Cancer Med*, 2018, 7: 3755-62
- [32] Akiyoshi T, Gotoh O, Tanaka N, et al. T-cell complexity and density are associated with sensitivity to neoadjuvant chemoradiotherapy in patients with rectal cancer. *Cancer Immunol Immunother*, 2021, 70: 509-18
- [33] Hou X, Wang M, Lu C, et al. Analysis of the repertoire features of TCR  $\beta$  chain CDR3 in human by high-throughput sequencing. *Cell Physiol Biochem*, 2016, 39: 651-67
- [34] 赵静静, 简星星, 王广志, 等. TCR组库及其在肿瘤免疫中的应用前景与挑战. *生命科学*, 2021, 33: 466-71
- [35] Postow MA, Manuel M, Wong P, et al. Peripheral T cell receptor diversity is associated with clinical outcomes following ipilimumab treatment in metastatic melanoma. *J Immunother Cancer*, 2015, 3: 23

- [36] Hosoi A, Takeda K, Nagaoka K, et al. Increased diversity with reduced “diversity evenness” of tumor infiltrating T-cells for the successful cancer immunotherapy. *Sci Rep*, 2018, 8: 1058
- [37] 刘福梅, 谢雁鸣, 王志飞, 等. 含何首乌制剂致药物性肝损伤患者TCR免疫组库特征研究. *中国中药杂志*, 2019, 44: 4397-404
- [38] Feng B, Zheng D, Yang L, et al. Post-hospitalization rehabilitation alleviates long-term immune repertoire alteration in COVID-19 convalescent patients. *Cell Prolif*, 2023: e13450
- [39] Lee JH, Han Y. Ginsenoside Rg1 helps mice resist to disseminated candidiasis by Th1 type differentiation of CD4<sup>+</sup> T cell. *Int Immunopharmacol*, 2006, 6: 1424-30
- [40] Wang Y, Liu Y, Yu H, et al. Structural characterization and immuno-enhancing activity of a highly branched water-soluble  $\beta$ -glucan from the spores of *Ganoderma lucidum*. *Carbohydr Polym*, 2017, 167: 337-44
- [41] Jiang J, Wu C, Gao H, et al. Effects of astragalus polysaccharides on immunologic function of erythrocyte in chickens infected with infectious bursa disease virus. *Vaccine*, 2010, 28: 5614-6
- [42] Liao YH, Li CI, Lin CC, et al. Traditional Chinese medicine as adjunctive therapy improves the long-term survival of lung cancer patients. *J Cancer Res Clin Oncol*, 2017, 143: 2425-35
- [43] Luo B, Wang P, Tian J, et al. Jinfukang inhibits lung cancer metastasis by regulating T cell receptors. *J Ethnopharmacol*, 2023, 318: 116885
- [44] Rius C, Attaf M, Tungatt K, et al. Peptide-MHC class I tetramers can fail to detect relevant functional T cell clonotypes and underestimate antigen-reactive T cell populations. *J Immunol*, 2018, 200: 2263-79
- [45] Dolton G, Zervoudi E, Rius C, et al. Optimized peptide-MHC multimer protocols for detection and isolation of autoimmune T-cells. *Front Immunol*, 2018, 9: 1378
- [46] Martin MD, Jensen IJ, Ishizuka AS, et al. Bystander responses impact accurate detection of murine and human antigen-specific CD8<sup>+</sup> T cells. *J Clin Invest*, 2019, 129: 3894-908
- [47] Robins HS, Srivastava SK, Campregher PV, et al. Overlap and effective size of the human CD8<sup>+</sup> T cell receptor repertoire. *Sci Transl Med*, 2010, 2: 47ra64
- [48] Jurtz VI, Jessen LE, Bentzen AK, et al. NetTCR: sequence-based prediction of TCR binding to peptide-MHC complexes using convolutional neural networks. *BioRxiv*, 2018, doi: 10.1101/433706
- [49] Vita R, Overton JA, Greenbaum JA, et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res*, 2015, 43: D405-12
- [50] Shugay M, Bagaev DV, Zvyagin IV, et al. VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res*, 2018, 46: D419-27
- [51] Tickotsky N, Sagiv T, Prilusky J, et al. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics*, 2017, 33: 2924-9
- [52] Zhang W, Wang L, Liu K, et al. PIRD: pan immune repertoire database. *Bioinformatics*, 2020, 36: 897-903
- [53] Glanville J, Huang H, Nau A, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature*, 2017, 547: 94-8
- [54] De Neuter N, Bittremieux W, Beirnaert C, et al. On the feasibility of mining CD8<sup>+</sup> T cell receptor patterns underlying immunogenic peptide recognition. *Immunogenetics*, 2018, 70: 159-68
- [55] Dash P, Fiore-Gartland AJ, Hertz T, et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature*, 2017, 547: 89-93
- [56] Zhang H, Liu L, Zhang J, et al. Investigation of antigen-specific T-cell receptor clusters in human cancers tumor. *Clin Cancer Res*, 2020, 26: 1359-71
- [57] Ehrlich R, Kamga L, Gil A, et al. SwarmTCR: a computational approach to predict the specificity of T cell receptors. *BMC Bioinformatics*, 2021, 22: 422
- [58] Sidhom JW, Larman HB, Pardoll DM, et al. DeepTCR is a deep learning framework for revealing sequence concepts within T-cell repertoires. *Nat Commun*, 2021, 12: 1605
- [59] Zhang H, Zhan X, Li B. GIANA allows computationally-efficient TCR clustering and multi-disease repertoire classification by isometric transformation. *Nat Commun*, 2021, 12: 4699
- [60] Gielis S, Moris P, Bittremieux W, et al. TCRex: detection of enriched T cell epitope specificity in full T cell receptor sequence repertoires. *Front Immunol*, 2019, 10: 2820
- [61] Springer I, Besser H, Tickotsky-Moskovitz N, et al. Prediction of specific TCR-peptide binding from large dictionaries of TCR-peptide pairs. *Front Immunol*, 2020, 11: 1803
- [62] Jokinen E, Huuhtanen J, Mustjoki S, et al. Predicting recognition between T cell receptors and epitopes with TCRGP. *PLoS Comput Biol*, 2021, 17: e1008814
- [63] Weber A, Born J, Rodriguez Martínez M. TITAN: T-cell receptor specificity prediction with bimodal attention networks. *Bioinformatics*, 2021, 37: i237-44
- [64] Cai M, Bang S, Zhang P, et al. ATM-TCR: TCR-epitope binding affinity prediction using a multi-head self-attention model. *Front Immunol*, 2022, 13: 893247
- [65] Xu Z, Luo M, Lin W, et al. DLpTCR: an ensemble deep learning framework for predicting immunogenic peptide recognized by T cell receptor. *Brief Bioinform*, 2021, 22: bbab335
- [66] Lu T, Zhang Z, Zhu J, et al. Deep learning-based prediction of the T cell receptor-antigen binding specificity. *Nat Mach Intell*, 2021, 3: 864-75
- [67] Gao Y, Gao Y, Fan Y, et al. Pan-peptide meta learning for T-cell receptor-antigen binding recognition. *Nat Mach Intell*, 2023, 5: 236-49
- [68] Thakkar N, Bailey-Kellogg C. Balancing sensitivity and specificity in distinguishing TCR groups by CDR sequence similarity. *BMC Bioinformatics*, 2019, 20: 241
- [69] Chronister WD, Crinklaw A, Mahajan S, et al. TCRMatch: predicting T-cell receptor specificity based on sequence similarity to previously characterized receptors. *Front Immunol*, 2021, 12: 640725

- [70] Montemurro A, Schuster V, Povlsen HR, et al. NetTCR-2.0 enables accurate prediction of TCR-peptide binding by using paired TCR $\alpha$  and  $\beta$  sequence data. *Commun Biol*, 2021, 4: 1060
- [71] Springer I, Tickotsky N, Louzoun Y. Contribution of T cell receptor  $\alpha$  and  $\beta$  CDR3, MHC typing, V and J genes to peptide binding prediction. *Front Immunol*, 2021, 12: 664514
- [72] Huang H, Wang C, Rubelt F, et al. Analyzing the *Mycobacterium tuberculosis* immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat Biotechnol*, 2020, 38: 1194-202
- [73] Johnson LA, Morgan RA, Dudley ME, et al. Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. *Blood*, 2009, 114: 535-46
- [74] Parkhurst MR, Yang JC, Langan RC, et al. T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. *Mol Ther*, 2011, 19: 620-6
- [75] Rapoport AP, Stadtmauer EA, Binder-Scholl GK, et al. NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. *Nat Med*, 2015, 21: 914-21
- [76] Roth T L, Puig-Saus C, Yu R, et al. Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature*, 2018, 559: 405-9
- [77] Oh SA, Senger K, Madireddi S, et al. High-efficiency nonviral CRISPR/Cas9-mediated gene editing of human T cells using plasmid donor DNA. *J Exp Med*, 2022, 219: e20211530
- [78] Kohyama M, Saijyo K, Hayasida M, et al. Direct activation of human CD8<sup>+</sup> cytotoxic T lymphocytes by interleukin-18. *Jpn J Cancer Res*, 1998, 89: 1041-6
- [79] Beshnova D, Ye J, Onabolu O, et al. *De novo* prediction of cancer-associated T cell receptors for noninvasive cancer detection. *Sci Transl Med*, 2020, 12: eaaz3738
- [80] Kim C, Williams MA. Nature and nurture: T-cell receptor-dependent and T-cell receptor-independent differentiation cues in the selection of the memory T-cell pool. *Immunology*, 2010, 131: 310-7
- [81] Tubo NJ, Jenkins MK. TCR signal quantity and quality in CD4<sup>+</sup> T cell differentiation. *Trends Immunol*, 2014, 35: 591-6
- [82] Bortone DS, Woodcock MG, Parker JS, et al. Improved T-cell receptor diversity estimates associate with survival and response to anti-PD-1 therapy. *Cancer Immunol Res*, 2021, 9: 103-12
- [83] Sims JS, Grinshpun B, Feng Y, et al. Diversity and divergence of the glioma-infiltrating T-cell receptor repertoire. *Proc Natl Acad Sci U S A*, 2016, 113: E3529-37
- [84] Ott PA, Hu-Lieskovan S, Chmielowski B, et al. A phase Ib trial of personalized neoantigen therapy plus anti-PD-1 in patients with advanced melanoma, non-small cell lung cancer, or bladder cancer. *Cell*, 2020, 183: 347-62.e24
- [85] Schrama D, Ritter C, Becker JC. T cell receptor repertoire usage in cancer as a surrogate marker for immune responses. *Semin Immunopathol*, 2017, 39: 255-68
- [86] Arstila TP, Casrouge A, Baron V, et al. A direct estimate of the human  $\alpha\beta$  T cell receptor diversity. *Science*, 1999, 286: 958-61
- [87] Robins HS, Campregher PV, Srivastava SK, et al. Comprehensive assessment of T-cell receptor  $\beta$ -chain diversity in  $\alpha\beta$  T cells. *Blood*, 2009, 114: 4099-107
- [88] Wooldridge L, Ekeruche-Makinde J, van den Berg HA, et al. A single autoimmune T cell receptor recognizes more than a million different peptides. *J Biol Chem*, 2012, 287: 1168-77
- [89] Sewell AK. Why must T cells be cross-reactive? *Nat Rev Immunol*, 2012, 12: 669-77
- [90] Dean J, Emerson RO, Vignali M, et al. Annotation of pseudogenic gene segments by massively parallel sequencing of rearranged lymphocyte receptor loci. *Genome Med*, 2015, 7: 123