

doi:10.3971/j.issn.1000-8578.2023.22.1027

• 专家论坛 •

基于批量转录组测序分析肿瘤免疫微环境及其应用

陈肖，刘泉，任欢

Analysis of Tumor Immune Microenvironment Based on Bulk RNA-sequencing Data and Its Application

CHEN Xiao, LIU Quan, REN Huan

School of Medicine, Southern University of Science and Technology, Shenzhen 518055, China

Corresponding Author: REN Huan, E-mail: renh@sustech.edu.cn



任欢 南方科技大学医学院教授，博士生导师。2001—2005年在英国利物浦大学医学院工作并获得博士学位，2009年哈尔滨医科大学药学博士后出站。2006—2018年在哈尔滨医科大学担任免疫学教研室主任、教授、博士生导师，省级领军人才梯队学术带头人。目前担任中国免疫学会理事，中华医学核心期刊《国际免疫学》副主编，国家自然基金、波兰国家自然基金、国家教育部学位点论文评审专家。作为主编、副主编或参编医学免疫学教材和参考书10部；主编或参编中英文学术专著9部。主持完成多项国家自然基金面上项目以及重大研究计划培育项目等科研课题，获得教育部提名科技进步一等奖、黑龙江省科技进步二等奖、世界卫生组织奖学金等。课题组采用生物医学实验、生物信息学、临床医学、生物物理及数学模型等多学科研究方法，以肿瘤生物免疫与靶向治疗基础与应用研究为主要方向，重点关注恶性肿瘤多分子靶

标诊断、肿瘤免疫微环境在诊疗中的作用，以及建立体外药物筛选检测平台和肿瘤疫苗研发等。近年来在国际学术期刊*Nuclear Acid Research, Neuro-oncology, Cell Death & Differentiation, Mol Cancer Ther, Oncotarget, BMC Immunology* 等发表论文50多篇。

Abstract: Cancer immunotherapy is one of the most promising biological therapies. The dynamic changes of the immune microenvironment of heterogeneous tumors are critical factors in determining the interaction and therapeutic efficacy between tumor and immune microenvironment. Therefore, quantitative analysis of its constituent cells, related genes, and phenotypes in real time is of great importance. The methods used to analyze tumor immune microenvironment include immunohistochemistry, flow cytometry and so on. With the development of next-generation sequencing technology, transcriptome RNA sequencing data analysis has become one of the important methods to determine the composition of tumor immune microenvironment. This article focuses on the common methods of bulk RNA sequencing data analysis and related research progress.

Key words: Tumor; Immune microenvironment; Immunotherapy; Bulk transcriptome sequencing; Deconvolution; Algorithmic tools

Funding: National Natural Science Foundation of China (No. 81871993, 81472367)

Competing interests: The authors declare that they have no competing interests.

摘要：肿瘤免疫治疗是目前最有希望的生物学疗法之一。异质性肿瘤免疫微环境的动态变化是判定肿瘤与免疫微环境相互作用和治疗疗效的关键因素，因此，实时对其组成细胞、相关基因和表型的量化分析至关重要。用于肿瘤免疫微环境的分析方法包括免疫组织化学、流式细胞术等。随着新一代测序技术的发

展，转录组RNA测序数据分析成为确定肿瘤免疫微环境组成的重要方法之一。本文重点总结了常用的批量RNA测序数据分析方法及相关研究进展。

关键词：肿瘤；免疫微环境；免疫治疗；批量转录组测序；反卷积；算法工具

中图分类号：R730.51

开放科学(资源服务)

标识码(OSID):

收稿日期：2022-09-02；修回日期：2023-01-02

基金项目：国家自然基金面上项目（NSFC-81871993；NSFC-81472367）

作者单位：518055 深圳，南方科技大学医学院

通信作者：任欢（1967-），女，博士，教授，主要从事肿瘤生物免疫与靶向治疗基础与应用研究，E-mail: renh@sustech.edu.cn, ORCID: 0000-0001-7114-2012

作者简介：陈肖（1999-），女，硕士，主要从事肿瘤生物免疫与靶向治疗的基础与应用研究，ORCID: 0009-0008-9274-6147



0 引言

肿瘤免疫微环境 (tumor immune microenvironment, TIME) 主要由肿瘤微环境中各种免疫细胞及其相关因子构成。近年来, 免疫治疗在癌症治疗和预防方面取得了重大进展, 已发展成为癌症治疗的支柱产业^[1], 其目的是激发机体的免疫细胞对癌细胞进行特异性识别和攻击^[2], 主要包括癌症疫苗^[3]、免疫检查点抑制剂^[4]和过继细胞疗法^[5-6]。免疫治疗在治疗某些恶性肿瘤中取得了显著疗效^[7], 但多数肿瘤患者的临床获益还很有限。其中TIME组成差异及其在肿瘤进展不同阶段的动态变化是不同肿瘤患者疗效差异的重要因素^[8]。因此, TIME组成的量化分析对于判定患者免疫状态、定义预测性生物标志物和确定新的个体化治疗方案至关重要。

1 肿瘤免疫微环境分析方法及其发展

免疫组织化学和免疫荧光染色法是量化肿瘤组织TIME中细胞组成的常用方法, 通过以标志物如荧光素、酶、某种金属等显色物质标记抗体, 利用抗体和抗原的特异性结合来检测免疫细胞或因子的定位和定量方法^[9]。近年来多重免疫组织化学/免疫荧光技术逐渐被应用, 可实现单张组织切片上的多个标志物同时检测分析^[10-11]。例如一种基于酪胺信号放大的多重免疫荧光技术可实现在同一组织切片样本上复染7种以上抗原^[12]。此外, 通过提取组织细胞并对其进行荧光标记, 流式细胞术可依据细胞大小和荧光标记进行细胞分类和计数。相比于传统的流式细胞仪和分选仪, 目前该技术也进行了升级换代, 可用于检测更多种荧光标记和复杂细胞构成等^[13]; 然而, 流式细胞术与免疫组织化学法相比, 丢失了样本的空间信息。因此, 将CyTOF (Cytometry by Time-Of-Flight) 大规模细胞术与免疫细胞化学和免疫组织化学技术、高分辨率激光消融系统和低色散激光消融室相结合的成像质谱流式细胞术应运而生, 该方法可在石蜡包埋的肿瘤组织切片上读取三十多种标志物^[14]。

近年来随着新一代测序技术快速发展, 高通量测序方法使生物医学领域相关研究发生了重要转变, 例如获取转录组测序 (RNA sequencing, RNA-Seq) 数据及其分析技术的发展已产生了许多重要发现^[15]。批量转录组测序 (bulk RNA-seq) 技术的出现推进了对细胞群体的特性认识, 单细胞转录组测序 (scRNA-seq) 技术将细胞基因表达

的认识推进到单细胞的层面, 目前已在新细胞类型的发现、揭示细胞异质性等方面展现出优势。近几年出现的空间转录组测序技术可以在获得基因表达数据的同时获取细胞的空间位置信息, 进一步推进了对组织原位细胞真实基因表达的研究^[16]。空间转录组测序及scRNA-seq等技术目前过于昂贵, 对样本质量要求较高, 无法常规使用。此外, 单细胞解离效率的细胞特异性差异可能会使细胞比例产生偏差^[17]。虽然bulk RNA-seq在揭示细胞异质性等方面还有所欠缺, 但考虑到scRNA-seq及空间转录组测序应用到大量样本的可行性, bulk RNA-seq依旧在肿瘤免疫微环境分析等方面发挥着重要作用。目前很多研究采用体现组织原态的系列时间点bulk RNA-seq数据分析样本细胞构成, 或者基于表达模式相关性分析辅助证明scRNA-seq或空间转录组测序结果。肿瘤浸润性免疫细胞的组成可以使用基于一组免疫特异性标记基因或表达特征的计算方法从大量肿瘤RNA-seq数据中表征。本文主要对肿瘤组织批量RNA-Seq数据的免疫浸润分析方法及其研究进展进行综述和分析。

2 基于批量RNA-seq数据的分析方法

2.1 基于标记基因的分析方法

常用的标记基因分析方法主要基于基因富集分析实现。基因富集分析是根据基因在样本中的相对表达量对基因进行排序, 通过判断一组细胞特异性标记基因倾向于出现在列表的顶部(或底部)来判断细胞表型, 并计算富集分数^[18-19]。然而, 富集分数通常不能代表微环境中不同细胞类型的相对数量, 也无法比较不同细胞类型的相对丰度比例, 这是该分析方法的弱点。目前, 基于标记基因的分析方法已开发了包括MCP-counter、TIminer、xCell、ImmuCellAI等一系列工具, 使批量RNA-seq数据分析简单易行。

由Becht等开发的MCP-counter能够对异质性肿瘤组织中多种免疫细胞和基质细胞群的丰度进行量化分析, 已被用于量化19 000多个样本中32种非造血系肿瘤的免疫细胞和其他基质细胞^[20], 由于其输出结果用任意单位表示, 它们不能直接解释为细胞分数, 也不能在细胞类型之间进行比较。TIminer是首个通过RNA-Seq数据执行综合免疫基因组分析的工具, 在分析免疫浸润特征之外, 还可以进行人白细胞抗原分型、新抗原预测、肿瘤免疫原性的量化等分析^[21]。与MCP-counter、

TIminer直接得到细胞富集分数不同，xCell将富集分数转换为线性比例，并使用溢出补偿技术减少密切相关细胞类型之间的依赖性，从而克服了基因富集分析不能跨细胞类型比较的缺点^[22]。当然，xCell一个重要限制是估计的结果是浓缩分数，不能解释为细胞比例。与xCell一样基于单样本基因富集开发的ImmucellAI将量化目标放在了T细胞亚群，可以较精确地估算24种免疫细胞类型（包括18个T细胞亚群）相对丰度^[23]，对T细胞亚群的分析具有一定的优势。

小鼠TME免疫浸润分析可以通过与人的基因进行同源转换后使用以上工具进行分析，但可能会因为部分基因的无法转换使结果出现误差，针对这一问题，MCP-counter、ImmucellAI开发了适用于小鼠组织样本转录组数据的分析方法，mMCP-counter可以量化小鼠的12种免疫细胞和4种基质细胞^[24]，ImmucellAI-mouse以三层的分层策略从bulk RNA-seq数据中估计36种免疫细胞（亚）类型的丰度^[25]。

2.2 反卷积算法

反卷积算法的原理是将混合物中的基因表达建模为该基因在不同类型细胞中表达的线性组合，其基因表达谱汇总在特征矩阵中，由混合物中细胞类型的相对分数加权^[19]。该算法可直接从异质性样本数据中提取特定细胞信息，并对其进行量化和计算比例^[26]。然而，组织TIME中细胞基因表达谱可能与该算法所参考的基因表达谱存在差异，因而产生误差。此外，由于参考细胞谱系特异性基因相对较少，也可能出现多重共线性问题^[27]。基于反卷积算法开发的在线工具有CIBERSORT、TIMER、EPIC、quanTIseq等。

CIBERSORT是目前常用的基于反卷积原理开发的免疫浸润分析工具，以微阵列方式构建了一种称为LM22的白细胞基因特征矩阵，该矩阵可区分包含7种T细胞，以及记忆B细胞和自然杀伤细胞在内的22种人类血细胞表型，使用nu-持向量回归（v-SVR）估算细胞分数^[28]。CIBERSORT可以以单细胞测序数据为参考计算bulk RNA-seq数据的免疫浸润，在噪声、未知混合物含量和密切相关的细胞类型方面优于其他方法，但由于受到参考资料保真度的限制，可能会在经历异型相互作用、表型可塑性或疾病引起的失调的细胞中结果发生偏差。2020年该团队为了提高数字化细胞计数的性能和多功能性将CIBERSORT扩展到一个名为CIBERSORTx的新计算框架中，与其他方

法相比，CIBERSORTx具有可以在没有抗体、组织解离或活细胞的情况下从复杂组织中剖析细胞异质性的优势^[29]。癌症免疫治疗研究的数量和复杂性不断增加，对分析和可视化快速积累的临床和癌症基因组学数据提出了新的挑战，癌症研究人员缺乏一种综合的计算工具来方便地探索和可视化肿瘤免疫学及基因组学数据，TIMER是第一种允许用户对肿瘤免疫学、临床和基因组学数据进行综合分析的方法，为每种癌症类型选择与肿瘤纯度（肿瘤组织中恶性细胞的百分比）负相关的信息基因，并对所选基因的表达进行线性最小二乘回归法进行反卷积求解，应用免疫细胞表达特征分析了32种癌症类型中6种免疫细胞浸润情况，允许用户交互式探索免疫浸润与基因表达、临床结果、体细胞突变和体细胞拷贝数改变等广泛因素之间的关联^[30-31]，但其结果不能直接解释为细胞分数，也不能在不同的免疫细胞类型和数据集之间进行比较。EPIC是Racle等使用约束性最小二乘回归将非负性约束条件引入反卷积运算，从而开发出可以从大量的肿瘤基因表达数据中同时估算癌细胞和免疫细胞比例的工具^[32]，但没有考虑与癌症免疫学相关的免疫细胞，如调节性T细胞（Treg）细胞、树突状细胞等。quanTIseq使用了约束最小二乘回归执行反卷积运算，以强制细胞分数为非负且其总和不超过1，来量化与肿瘤免疫学相关的十种免疫细胞类型^[33]。EPIC和quanTIseq具有直接生成细胞分数的优势。TIMER 2.0是在TIMER基础上的更新版本，该版本集成了TIMER、CIBERSORT、EPIC、quanTIseq、MCP-counter和xCell这6种免疫浸润细胞的算法，为癌症基因组图谱或肿瘤表达谱的免疫浸润分析提供了可靠的工具^[34]。

3 批量RNA-seq分析方法的应用

目前，RNA-seq数据的分析方法广泛应用于生物医学领域研究。对TIME组成的量化分析有助于鉴定肿瘤预后标志物，筛选肿瘤免疫治疗新靶点。xCell、CIBERSORT、TIMER等批量RNA-seq数据分析方法中较为常用的算法工具常与空间转录组测序和scRNA-seq数据进行联合分析^[35-36]。Zhou等^[37]应用CIBERSORT评估了白介素-1野生型组（IL-1-WT）和突变组（IL-1-MT）结肠腺癌患者TME中22种免疫细胞的比例，发现肿瘤浸润淋巴细胞在IL-1-MT结肠腺癌患者组织TIME中显著富集，免疫反应相关评分更高。Bao等^[38]通

过scRNA-seq鉴定了三阴性乳腺癌的肿瘤间和肿瘤内的异质性，使用CIBERSORT估计bulk RNA-seq数据中M2样TAM的丰度，发现M2样TAM与三阴性乳腺癌患者的不良预后相关。Xiong等^[39]利用xCell分析bulk RNA-seq数据结合scRNA-seq及空间转录组数据综合分析发现非转移性糖蛋白B（GPNMB）在胶质母细胞瘤细胞从原神经细胞型到间质型转变过程中发挥重要作用，并通过细胞共培养进一步揭示这些源自单核细胞的高GPNMB巨噬细胞亚群也可以有效地阻止T细胞不被树突状细胞激活。

毋庸置疑，应用上述工具估算的相关结果还需要独立的相关实验进行验证，例如流式细胞术、免疫组织化学及免疫荧光等。例如，Newman等^[28]为比较CIBERSORT估算结果与组织中白细胞真实含量，利用流式细胞术对早期非小细胞肺癌肺组织标本和来自滤泡性淋巴瘤中的免疫细胞亚群进行比较分析，结果显示CIBERSORT结果与流式细胞术测量值显著相关。xCell相关结果也通过应用流式细胞术等验证了xCell输出的富集分数等^[22]。

4 总结与展望

免疫疗法为恶性肿瘤患者治疗带来了治愈新希望，TIME组成和动态变化是评估疗效和转换实施方案的重要参考指标。新一代RNA测序技术的广泛应用可逐步实现对TIME的动态检测和分析。除了估算TIME中免疫细胞组成和相对丰度比例，还可以通过估算RNA-seq数据中特定免疫表型的大规模变化，评估治疗后免疫应答类型的动力变化，以分析并预测患者对免疫治疗的反应。此外，数据中免疫组库包含TCR/BCR互补决定区的多样性，可用于评估免疫系统与肿瘤进展的关联，联合TMB相关数据实时检测TIME，预测肿瘤进展并及时调整治疗策略。总之，虽然目前对于TIME组成和量化分析方法还有诸多限制，相信随着技术发展和多种研究方法的联合应用，可深入揭示TIME构成和演化机制，从而为临床治疗提供更可靠的指标参考。

利益冲突声明：

所有作者均声明不存在利益冲突。

参考文献：

- [1] Abbott M, Ustoyev Y. Cancer and the Immune System: The History and Background of Immunotherapy[J]. Semin Oncol Nurs, 2019, 35(5): 150923.
- [2] Lei X, Lei Y, Li JK, et al. Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy[J]. Cancer Lett, 2020, 470: 126-133.
- [3] Igarashi Y, Sasada T. Cancer Vaccines: Toward the Next Breakthrough in Cancer Immunotherapy[J]. J Immunol Res, 2020, 2020: 5825401.
- [4] Abdou Y, Pandey M, Sarma M, et al. Mechanism-based treatment of cancer with immune checkpoint inhibitor therapies[J]. Br J Clin Pharmacol, 2020, 86(9): 1690-1702.
- [5] Rath JA, Arber C. Engineering Strategies to Enhance TCR-Based Adoptive T Cell Therapy[J]. Cells, 2020, 9(6): 1485.
- [6] Lesch S, Gill S. The promise and perils of immunotherapy[J]. Blood Adv, 2021, 5(18): 3709-3725.
- [7] Kruger S, Ilmer M, Koboldt S, et al. Advances in cancer immunotherapy 2019 - latest trends[J]. J Exp Clin Cancer Res, 2019, 38(1): 268.
- [8] Binnewies M, Roberts EW, Kersten K, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy[J]. Nat Med, 2018, 24(5): 541-550.
- [9] Magaki S, Hojat SA, Wei B, et al. An Introduction to the Performance of Immunohistochemistry[J]. Methods Mol Biol, 2019, 1897: 289-298.
- [10] Tan WCC, Nerurkar SN, Cai HY, et al. Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy[J]. Cancer Commun (Lond), 2020, 40(4): 135-153.
- [11] Widodo SS, Hutchinson RA, Fang Y, et al. Toward precision immunotherapy using multiplex immunohistochemistry and in silico methods to define the tumor immune microenvironment[J]. Cancer Immunol Immunother, 2021, 70(7): 1811-1820.
- [12] Lim JCT, Yeong JPS, Lim CJ, et al. An automated staining protocol for seven-colour immunofluorescence of human tissue sections for diagnostic and prognostic use[J]. Pathology, 2018, 50(3): 333-341.
- [13] Mckinnon KM. Flow Cytometry: An Overview[J]. Curr Protoc Immunol, 2018, 120: 5.1.1-5.1.11.
- [14] Giesen C, Wang HA, Schapiro D, et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry[J]. Nat Methods, 2014, 11(4): 417-422.
- [15] Hedlund E, Deng Q. Single-cell RNA sequencing: Technical advancements and biological applications[J]. Mol Aspects Med, 2018, 59: 36-46.
- [16] Li X, Wang CY. From bulk, single-cell to spatial RNA sequencing [J]. Int J Oral Sci, 2021, 13(1): 36.
- [17] Plattner C, Finotello F, Rieder D. Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq[J]. Methods Enzymol, 2020, 636: 261-285.
- [18] Fan R, Cui Q. Toward comprehensive functional analysis of gene lists weighted by gene essentiality scores[J]. Bioinformatics, 2021: btat475.
- [19] Finotello F, Trajanoski Z. Quantifying tumor-infiltrating immune cells from transcriptomics data[J]. Cancer Immunol Immunother, 2021: 1-10.

- 2018, 67(7): 1031-1040.
- [20] Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression[J]. *Genome Biol*, 2016, 17(1): 218.
- [21] Tappeiner E, Finotello F, Charoentong P, et al. TIminer: NGS data mining pipeline for cancer immunology and immunotherapy[J]. *Bioinformatics*, 2017, 33(19): 3140-3141.
- [22] Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape[J]. *Genome Biol*, 2017, 18(1): 220.
- [23] Miao YR, Zhang Q, Lei Q, et al. ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy[J]. *Adv Sci (Weinh)*, 2020, 7(7): 1902880.
- [24] Petitprez F, Levy S, Sun CM, et al. The murine Microenvironment Cell Population counter method to estimate abundance of tissue-infiltrating immune and stromal cell populations in murine samples using gene expression[J]. *Genome Med*, 2020, 12(1): 86.
- [25] Miao YR, Xia M, Luo M, et al. ImmuCellAI-mouse: a tool for comprehensive prediction of mouse immune cell abundance and immune microenvironment depiction[J]. *Bioinformatics*, 2022, 38(3): 785-791.
- [26] Shen-Orr SS, Gaujoux R. Computational deconvolution: extracting cell type-specific information from heterogeneous samples[J]. *Curr Opin Immunol*, 2013, 25(5): 571-578.
- [27] Lau D, Bobe AM, Khan AA. RNA Sequencing of the Tumor Microenvironment in Precision Cancer Immunotherapy[J]. *Trends Cancer*, 2019, 5(3): 149-156.
- [28] Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles[J]. *Nat Methods*, 2015, 12(5): 453-457.
- [29] Steen CB, Liu CL, Alizadeh AA, et al. Profiling Cell Type Abundance and Expression in Bulk Tissues with CIBERSORTx[J]. *Methods Mol Biol*, 2020, 2117: 135-157.
- [30] Li B, Severson E, Pignon JC, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy[J]. *Genome Biol*, 2016, 17(1): 174.
- [31] Li T, Fan J, Wang B, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells[J]. *Cancer Res*, 2017, 77(21): e108-e110.
- [32] Racle J, De Jonge K, Baumgaertner P, et al. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data[J]. *Elife*, 2017, 6: e26476.
- [33] Finotello F, Mayer C, Plattner C, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data[J]. *Genome Med*, 2019, 11(1): 34.
- [34] Li T, Fu J, Zeng Z, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells[J]. *Nucleic Acids Res*, 2020, 48(W1): W509-W514.
- [35] Yang H, Zhao L, Zhang Y, et al. A comprehensive analysis of immune infiltration in the tumor microenvironment of osteosarcoma[J]. *Cancer Med*, 2021, 10(16): 5696-5711.
- [36] Lv L, Zhang Y, Zhao Y, et al. Effects of 1p/19q Codeletion on Immune Phenotype in Low Grade Glioma[J]. *Front Cell Neurosci*, 2021, 15: 704344.
- [37] Zhou X, Liu Y, Xiang J, et al. Analysis of Interleukin-1 Signaling Alterations of Colon Adenocarcinoma Identified Implications for Immunotherapy[J]. *Front Immunol*, 2021, 12: 665002.
- [38] Bao X, Shi R, Zhao T, et al. Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels tumour heterogeneity plus M2-like tumour-associated macrophage infiltration and aggressiveness in TNBC[J]. *Cancer Immunol Immunother*, 2021, 70(1): 189-202.
- [39] Xiong A, Zhang J, Chen Y, et al. Integrated single-cell transcriptomic analyses reveal that GPNMB-high macrophages promote PN-MES transition and impede T cell activation in GBM[J]. *EBioMedicine*, 2022, 83: 104239.

[编辑校对：邱颖慧]

作者贡献：

陈 肖：选题构思、文献检索及论文撰写

刘 泉：论文审校

任 欢：论文设计指导及审校