

单细胞转录组测序技术在药物研究中的应用

汪茂林, 杨洪军*

(中国中医科学院中药研究所, 北京 100700)

摘要: 单细胞转录组测序 (single cell RNA sequencing, scRNA-seq) 是在单个细胞水平上研究其转录组信息的一门新兴技术。应用该技术有助于精细化解析疾病发展过程中异质性细胞图谱, 以及药物处理后, 精准发现对药物响应的特异性细胞亚群。目前, scRNA-seq 技术在药物研究领域得到广泛应用, 其在药物及其靶点筛选、药物所致不良反应、耐药性和疫苗研究中发挥重要作用。本文综述了 scRNA-seq 技术在药物研究中的应用, 为临床药物的个性化、精准化治疗提供一定的科学依据。

关键词: 单细胞转录组测序; 药物研究; 异质性; 靶点筛选; 不良反应; 疫苗开发

中图分类号: R966 文献标识码: A 文章编号: 0513-4870(2023)09-2551-09

Single cell RNA sequencing technology applied for drug discovery

WANG Mao-lin, YANG Hong-jun*

(Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China)

Abstract: Single cell RNA sequencing (scRNA-seq) is an advanced technology to study the transcriptome information at the single cell level. The application of this technology can attribute to analyze the heterogeneous map of cells in the process of disease development, and precisely identify the specific cell subsets that are responsive to pharmacological therapy. Currently, scRNA-seq technology has been widely applied in the field of drug research, including studies on therapeutic targets, drug-induced adverse reactions, drug resistance and vaccine. This work reviews the application of scRNA-seq technology in drug discovery, which offers a scientific basis for personalized and accurate medication therapy.

Key words: single cell RNA sequencing; drug discovery; heterogeneity; target discovery; adverse reaction; vaccine development

人类基因组计划 (Human Genome Project, HGP) 的完成, 开创了 21 世纪对生命科学研究的新纪元, 构建了人类对疾病个性化、精准化治疗的新蓝图, 也为生物学、生物技术、药物开发带来新的变革^[1]。人类细胞图谱 (Human Cell Atlas, HCA) 计划^[2]的实施, 进一步为全方位解析机体细胞种类、数量、空间关系和分子表达信息, 准确定义健康和疾病的细胞组成状态提供新思路。以人类基因组为蓝本, 从单个细胞水平解析疾

病发生和药物治疗的分子机制成为当前研究热点和难点。单细胞转录组测序 (single cell RNA sequencing, scRNA-seq) 技术从单个细胞水平分析其转录组信息, 从而揭示细胞群体的异质性, 弥补了普通测序 (bulk RNA-seq) 方法从整体组织、细胞水平分析的平均化效应^[3]。2009 年, 北京大学汤富酬团队^[4]首次运用 scRNA-seq 技术分析了单个小鼠卵裂球的全转录组信息, 随后该技术迅速被推广和更新, 相继出现 CEL-seq^[5]、Quartz-seq^[6]、Drop-seq^[7]等新的测序技术。随着各种测序技术的不断更新, 以 10×Genomics、BD Rhapsody 等为代表的商业化平台的出现, 极大程度地提高了测序的效率和准确性, 降低了测序成本, 从而为临床药物的研发指引

收稿日期: 2023-02-07; 修回日期: 2023-04-07.

基金项目: 国家重点研发计划 (2019YFC1708900); 国家自然科学基金资助项目 (82204797); 中国博士后科学基金 (2022M713494).

*通讯作者 Tel: 86-10-84015130, E-mail: hongjun0420@vip.sina.com

DOI: 10.16438/j.0513-4870.2023-0123

了新方向^[8]。传统的药物研究方法主要包括药代动力学 (pharmacokinetics, PK) 和药效学 (pharmacodynamic, PD) 研究。由于个体间差异、遗传背景、环境暴露的不同以及由特定细胞类型引起的基因相互作用差异所驱动的疾病,使得仅小部分患者对疾病治疗药物有效,而大部分患者可能因无效药物而遭受不必要的不良反应风险^[9]。药物基因组学 (pharmacogenomics)^[10]通过给药前预测患者的药物反应,揭示药物反应的生物学基础,架起了基因-药效-安全性的桥梁,为临床药物治疗的安全性提供了有力保障。在药物研究方面,运用 scRNA-seq 技术从单个细胞分辨率角度全方位揭示药物在疾病治疗中的作用机制,丰富了其临床安全性、有效性、稳定性的科学内涵,也为药物开发提供新方向。

本文主要介绍 scRNA-seq 技术的研究概况,重点围绕 scRNA-seq 技术在药物研究中的应用,包括药物靶点筛选、毒性、耐药性以及基因药物和疫苗研究方面进行概述,为临床药物的合理开发提供科学依据,为后续研究提供参考。

1 单细胞转录组测序技术概况

scRNA-seq 通常需要经过以下步骤:单细胞分离、cDNA 文库构建、单细胞转录组测序和数据分析,见图 1。

1.1 单细胞分离技术

scRNA-seq 的关键是如何快速、高效和稳定地获取高质量单细胞。目前,常用的单细胞分离方法主要包括显微微管法 (micro-pipetting)、激光捕获显微切割法 (laser capture microdissection, LCM)、流式细胞分选 (fluorescence activated cell sorting, FACS) 和微流控技术 (microfluidics)。显微微管法只需在显微镜下操作分离单个细胞,分离操作简单、成本

较低,但分离效率低且对细胞有一定程度的机械损伤。LCM 利用低能红外激光脉冲融化转运薄膜,从而捕获与之紧密结合的靶细胞,该方法可实现细胞在组织样本中的原位分离,很大程度上保留了细胞空间位置信息,但成本较高,对组织样本有较高要求。FACS 利用流式细胞分选结合光散射特性将特定荧光标记的细胞从悬液中分离,该方法通量高、高特异性和高精度,可实现对特定类型细胞的富集,但需要较大的起始细胞浓度,对于低丰度和无特定标签的细胞捕获效率较低。这几种方法受细胞分离技术和成本所限,单细胞分离效率及可检测细胞数量通常较低。随着测序技术的不断更新,基于分子条形码 (barcode) 标签的单细胞识别的应用,以 Drop-seq^[7]、MULTI-seq^[11]和 10×Genomics^[12]等为代表的微流控技术得到迅速发展。微流控技术通过操控微尺度流体,准确地将细胞封闭在微液滴中从而实现单细胞的分离。该方法所需样本量小,具有自动化程度高、成本低、高灵敏性和高通量的优势。但受限于微孔直径大小,某些特定细胞,如成熟心肌细胞、神经元细胞和巨噬细胞等往往无法被完整的捕获,为解决这一问题,近年来又开发出提取细胞核进行单细胞核转录组测序 (single nucleus RNA sequencing, snRNA-seq)^[13]的方法。

1.2 scRNA-seq 测序技术

单细胞捕获后,通常需要经过裂解、反转录、cDNA 扩增建立测序文库等程序。根据扩增方式、测序原理的差异,目前已经发展了数十种 scRNA-seq 技术,较常用的 scRNA-seq 测序技术对比见表 1^[5-7,11,12,14-25]。

在反转录过程中,多采用 poly (T) 引物启动来执行,以降低 rRNA 或 tRNA 的捕获量,但会导致非多聚腺苷酸化编码蛋白的 RNA 信息丢失。针对这一问题,

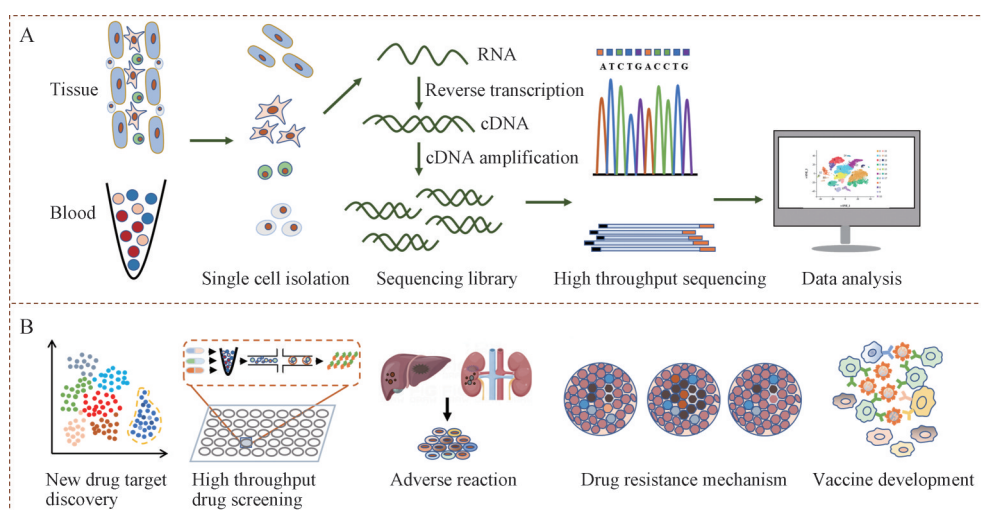


Figure 1 The process of single cell RNA sequencing (scRNA-seq) and its application in drug discovery. A: The process of scRNA-seq; B: The applications of scRNA-seq in drug discovery

Table 1 Comparison of several scRNA-seq methods

Sequencing method	Single cell separation	Capture cell count	Barcode	UMI	CDNA coverage	Amplification method	Advantage	Defect	Application
SCRB-seq ^[14]	FACS	1 000–10 000	Yes	Yes	3' end sequence	Template conversion PCR	High throughput, high sensitivity, and low cost	Highly dependent on full length mRNA; proficient manual operation required	Heterogeneous population transcriptome studies
CEL-seq ^[5] /CEL-seq ^[15]	Microtubule	100–1 000	Yes	Yes	3' end sequence	IVT	Linear amplification reduces the accumulation of non-specific fragments with high accuracy	3' end bias	Embryonic development
MARS-seq ^[16] /MARS-seq ^[17]	FACS	1 000–5 000	Yes	Yes	3' end sequence	IVT	High throughput and good stability	Proficient manual operation is required	Identification of specific cell types
Quartz-seq ^[6]	FACS	1 000–10 000	Yes	Yes	3' end biased full length	Polymeric tailings PCR	Accurate quantification	Proficient manual operation is required	Analysis of cell transcriptome differences in the same or different cycles
SUPeR-seq ^[8]	Microtubule	About 10	Yes	No	Full length	Polymeric tailings PCR	Sequence both polyadenosine and non adenosine RNA simultaneously	Low throughput and time consumption	Research on circRNA in early embryonic development of mammals
MATQ-seq ^[9]	Microtubule	10–100	Yes	Yes	Full length	Polymeric tailings PCR	High sensitivity and accurate quantification	Low throughput and time consumption	Quantitative study on single cell transcription variation
SMART-seq ^[20]	Microtubule	100–1 000	No	No	Full length	Template conversion PCR	Improve the average length and yield of cDNA libraries, enhance detection coverage and accuracy	Low detectable cell count and long separation process	Applicable to rare cell transcriptome research and analysis of alternative splicing allele expression
Drop-seq ^[7]	Microfluidics (barcode microspheres)	1 000–10 000	Yes	Yes	3' end sequence	Template conversion PCR	High throughput, fast speed, low cost	3' end bias	Construction of cell atlas
inDrops ^[21]	Microfluidic (barcode hydrogel)	1 000–10 000	Yes	Yes	3' end sequence	IVT	High throughput, high repeatability, fast cell capture speed, and low library construction cost	3' end bias	Identification of rare cell types
10×Genomics ^[12]	Microfluidics (oil droplets)	1 000–10 000	Yes	Yes	3' end sequence	Template conversion PCR	High throughput, high sensitivity, high cell capture rate, and low technical noise	3' end bias	Cell heterogeneity, immune cell population detection, and construction of cell atlas
MULTI-seq ^[11]	Microfluidics (lipid labeled emulsion)	10 000–100 000	Yes	Yes	3' end sequence	PCR	Rapid identification and removal of cell bimodal patterns	3' end bias	Immune cell response to tumor metastasis
Microwell-seq ^[22]	Nanopores (agarose micropores)	100–10 000	Yes	No	Full length	PCR	High throughput, high sequencing quality, short cell loading cycle, small reagent and sample size, and low cost	3' end bias	Construction of mammalian cell atlas
Seq-well ^[23]	Nanopores	100–10 000	Yes	Yes	3' end sequence	PCR	Portable, high-throughput, fast analysis, low cost, and minimal cross contamination	Only 10%–15% of transcripts from each cell can be analyzed	Identifying specific phenotypic cells from complex biological samples; immunocyte sequencing analysis
SCAN-seq ^[24]	Microtube	10–48	No	No	Full length	/	Long reading, high sensitivity and high robustness	High cost and high sequencing error rate	Identification of novel transcripts and analysis of allele variable splicing transcripts without annotation
HIT-scISOseq ^[25]	Microfluidic	> 5 000	Yes	Yes	Full length	/	High throughput, long read, high accuracy, and good repeatability	High cost and high sequencing error rate	Analyzing complex transcriptional isomers, selective splicing, and sequence diversity

后来开发了 SUPeR-seq^[18]、MATQ-seq^[19]和 SMART-seq2^[20] 技术, 可实现对聚腺苷和非聚腺苷化 RNA 的同时测序, 保留了整个基因及其不同转录异构体的结构信息, 显著增加了 cDNA 文库测序的平均长度、产量和低表达基因的检出量, 同时还避免基因组 DNA 和 rRNA 的污染。cDNA 扩增阶段, 大部分的测序技术都是基于 PCR 的方法, 虽然扩增速度快, 但会导致高丰度转录产物的指数扩增而引起非特异性序列的积累。为此, 开发了基于体外转录线性化扩增技术, 如 CEL-seq^[5]、CEL-seq2^[15]和 MARS-seq^[16]等, 可有效避免低丰度转录物的过度扩增, 但该方法具有强烈的 3'端偏好。此外, 大多数单细胞测序都是基于 Illumina 测序平台, 为了降低 PCR 偏好性, 增加定量结果的准确性, 开发了基于分子条形码 (barcode) 和唯一分子标识符 (UMI) 标记的单细胞测序技术, 如 inDrops^[21]、10×Genomics^[12]、MULTI-seq^[26]等。其中, 10×Genomics 是目前应用较为广泛的商业化测序平台, 其通过液滴微流控系统, 将单细胞快速分配到油包水微反应系统, 给每个分子带上特定的标签, 在 PCR 和建库过程中随着核酸片段一起复制, 极大地提高细胞捕获效率、缩短建库周期和降低测序成本。

目前的单细胞测序技术主要还是基于二代测序 (next generation sequencing, NGS) 原理, 通过利用物理或化学方法将 DNA 随机打断成小片段 (100~250 bp), PCR 扩增后一次就能完成所有序列的测序, 显著地提高测序通量, 节约测序成本。但与此同时, 这种短读长测序也造成一些低丰度序列的丢失, PCR 过程也增加了碱基错配概率^[24]。为此, 以单分子长读长的第三代测序 (third-generation sequencing, TGS) 技术应运而生, 主要包括 Pacific Biosciences 公司的单分子实时测序 (single molecule real-time, SMRT) 和 Oxford Nanopore 公司的单分子纳米孔测序 (Nanopore)。PacBio^[27]技术利用在纳米孔中加入 4 种荧光标记的 dNTP, 其在 DNA 聚合酶作用下与 DNA 模板进行碱基配对, 根据发出的荧光信号波长及时间的差异识别碱基序列, 达到边合成边测序的目的。Nanopore^[28]技术核心是纳米孔, 根据不同碱基通过纳米孔时的电流变化特征来识别碱基类型, 从而实现边解链边测序。在此基础上又发展了 ScNaUmi-seq^[29]、RAGE-seq^[30]和 scCOLOR-seq^[31]的 TGS 技术, 但这些技术依然存在准确率低和灵敏度低的缺点, 主要通过结合 NGS 平台降低错误率, 或者通过牺牲通量提高检测率。最近, 汤富酬团队^[32]又开发了 SCAN-seq2 技术, 相比其他全长 RNA-seq 方法, 该技术同时表现出高通量和高灵敏度, 而且具有很高的重现性。随着测序技术的不断更新和成熟, TGS 既保

证了测序的高通量, 又可实现超长测序读长和均匀覆盖范围, 且不需要 PCR 富集就可直接测序, 极大地节省内存和计算时间, 避免了信息丢失和碱基错配的问题。但该方法单读长的错误率较高 (15%~40%), 需要修复测序纠错, 而且依赖于 DNA 聚合酶活性, 相比 NGS 需要更高的测序成本, 同时目前针对 TGS 的数据分析开发也远不如二代测序^[33]。

2 单细胞转录组测序技术在药物研究中的应用

2.1 药物靶点发现

药物靶点的筛选一直是新药研发过程中重要的环节, 也是融合多学科的系统性工程。scRNA-seq 技术在药物靶点研究方面, 通过分析正常组织细胞和疾病组织细胞转录组信息, 为进一步精准化发现与疾病相关生物标志物和药物靶标提高效率。Cheng 等^[34]通过 Hydro-seq 分析 21 例乳腺癌患者中 666 个循环肿瘤细胞 (circulating tumor cells, CTCs) 的转录组信息, 确定了激素和靶向治疗乳腺癌的药物靶点, 并发现了特异性表达肿瘤干细胞以及上皮/间充质细胞状态转变标记物的细胞亚群, 对这些细胞的转录组分析为监控靶向治疗和肿瘤转移的过程提供科学依据。Gladka 等^[35]利用 scRNA-seq 测序技术分析了心肌梗死小鼠的心肌梗死区域 509 个细胞及假手术组小鼠左室游离壁 426 个细胞的 mRNA 差异, 通过 K-medoids 相关性分析、t-SNE 降维可视化处理等生信分析, 找到了在激活的成纤维细胞中特异性表达 Ckap4 的细胞亚群, 提示该基因可作为心肌梗死后心肌纤维化的生物标志物, 为靶向该基因的药物开发奠定基础。Johnston 等^[36]利用 scRNA-seq 技术研究赖氨酸特异性去甲基化酶 1 (lysine specific demethylase 1, LSD1) 抑制剂 INCB059872 在急性髓系白血病 (acute myeloid leukemia, AML) 中的作用机制, 在模型小鼠骨髓中, 分别给药 4、6 天后, 观察到在非巨核细胞群中的浆细胞样树突状细胞 (plasmacytoid dendritic cell) 和 Csflr⁺/Cd36⁺ 前单核细胞群消失, 在 INCB059872 处理的白血病患者骨髓中, 发现大量新的细胞亚群, 并且独立生长因子-1 (growth factor independent 1, GF11) 和 GF11B 基因显著上调, 进一步研究表明, INCB059872 通过中断 LSD1 与 GF11 的结合, 导致 GF11 结合位点 LSD1/CoREST 活性丧失, 从而诱导白血病细胞向髓系细胞分化。

2.2 高通量药物筛选

在新药研发过程中, 准确、高效和特异性地筛选药物至关重要, 利用单细胞基因表达的高通量测序研究手段可以显著提高药物筛选效率, 加速药物研发进程。Shin 等^[26]开发了一种利用短条形码核苷酸瞬时转染, 可以同时分析不同条件下多个样本的测序方法, 并将其应用于抗慢性髓系白血病药物研究, 通过对 45 种药物处理和 5 个药物处理时间

下的单细胞测序研究,从单个细胞水平上揭示每种药物的特异性转录应答和靶向基因,为临床药物选择提供科学依据。Ye等^[37]将scRNA-seq应用于药物化合物分析,开发了一种大规模并行化、自动化、低成本的高通量药物筛选平台DRUG-seq,通过对433种具有潜在靶向骨肉瘤U20S的化合物8个剂量作用下的转录谱分析,从DRUG-seq生成的转录谱中成功地将化合物根据预期作用机制(MoAs)分组,发现了88个有效化合物具有50个以上基因显著改变,进一步对这些差异基因t-SNE聚类分析找到了具有相似药理作用机制(MoA)的功能簇。利用该平台可以快速鉴定具有相似MoA的化合物,促进新化合物的机制研究,极大地缩短了药物筛选周期和节约成本。

2.3 揭示药物所致不良反应机制 药物所致不良反应是临床药物研究所面临的另一重大问题,也是药物导致发病甚至死亡的重要因素,在一定程度上制约了药物的发现和应用。将scRNA-seq技术运用在药物不良反应研究中,将加强研究者理解机体对药物反应的病理过程和药物的毒性机制。Zhang等^[38]基于单细胞测序技术提出了精准毒理学(precision toxicology)概念,即通过测序一群细胞或一种甚至单个细胞研究药物诱导毒性的作用机制。对乙酰氨基酚(acetaminophen, APAP)过量服用会导致急性肝衰竭。Umbaugh等^[39]利用scRNA-seq技术对服用APAP后小鼠的肝脏进行分析,结果发现中央周围(pericentral, PC)肝细胞和门周(peripoortal, PP)肝细胞对APAP的感知性存在较大差异,PP肝细胞会沿着APAP应激轴向氧化应激发展,然后由于先天和适应性机制而消除损伤。但是PC肝细胞中的促分裂原活性蛋白激酶(mitogen-activated protein kinase, MAPK)基因被激活从而沿着应激轴继续发展,其中一部分高表达谷氨酰胺合成酶的PC肝细胞亚群会因为应激程序的发展最终导致细胞死亡,这些结果为揭示APAP诱导肝脏毒性机制提供新的科学依据。顺铂在治疗肿瘤过程中可能会诱导急性或慢性肾毒性,Gupta等^[40]在一项顺铂诱导肾脏类器官损伤研究中,利用snRNA-seq技术发现同源定向修复基因FANCD2和RAD51在早期损伤修复过程中上调,而在晚期不完全修复过程中下调,随后结合梗阻性肾损伤小鼠模型和免疫介导损伤人类肾脏样本中scRNA-seq结果分析验证了这两个基因在肾小管修复过程中上调,而在肾纤维化样本中FANCD2丢失,最后利用DNA连接酶IV抑制剂SCR7可以挽救FANCD2/RAD51介导的肾损伤修复,从而抑制顺铂诱导的肾脏类器官损伤进程。

药物所致不良反应机制研究同时还面临着缺乏

对应的疾病模型问题,如药物诱导的超敏反应综合征伴有嗜酸性粒细胞增多和全身症状的药物反应(drug-induced hypersensitivity syndrome/drug reaction with eosinophilia and systemic symptoms, DiHS/DRESS)。scRNA-seq从单细胞分辨率水平为解决这些问题提供了一种全新的策略。Kim等^[41]利用scRNA-seq技术对1例难治性DiHS/DRESS患者的皮肤和血液样本分析,从中发现大量富含人类疱疹病毒6b基因的记忆性CD4⁺T细胞亚群,进一步研究表明,该不良反应的发生与JAK-STAT信号通路激活所诱导的炎症密切相关,随后在体外使用JAK抑制剂tofacitinib和抗病毒药可显著抑制T细胞增殖,从而证明诱导DiHS/DRESS的发生机制与JAK-STAT通路相关。单细胞测序技术在药物不良反应中的应用为在临床动物模型缺乏、病理生理学难以解释、治疗效果不佳的情况下,开展药物不良反应实验研究开辟新思路,加深了研究者对药物所致机体不良反应的病理生理学认识,也为个性化、精准化药物治疗提供了一种新途径。

2.4 揭示耐药机制 耐药性是临床药物治疗面临的一大难题,尤其在恶性肿瘤治疗中,耐药性是导致化学药物治疗失败的重要原因。肿瘤细胞群的异质性(heterogeneity)介导了肿瘤的发生、耐药和转移^[42]。常规的检测方法多数是通过整体组织水平分析,从而忽视瘤内异质性问题,单细胞测序技术在揭示肿瘤异质性和检测稀有细胞亚群方面具有独特优势。三阴性乳腺癌(triple-negative breast cancer, TNBC)患者在接受新辅助化疗(neoadjuvant chemotherapy, NAC)后,约有30%~50%患者会产生耐药,最终导致整体生存率降低。目前,对TNBC患者接受化疗后耐药的分子机制研究尚不明确,其中一个关键问题是化学耐药到底是通过预先存在的稀有亚克隆扩增产生(适应性耐药),还是通过诱导形成耐药表型新突变(获得性耐药)。Kim等^[43]利用scDNA-seq、scRNA-seq及外显子测序(bulk exome sequencing)技术对20例接受NAC的TNBC患者样本进行分析,结果发现TNBC耐药基因在治疗前后一直存在,NAC产生的耐药机制是通过诱导这些基因组进行适应性选择而引起转录重编程。表明肿瘤耐药性并不是通过获得新的基因突变,而是由特异性的克隆通过适应性耐药产生。Sharma等^[44]利用scRNA-seq研究原发性口腔鳞状细胞癌(oral squamous cell carcinoma, OSCC)在顺铂选择压力下的演化和转移扩散机制,发现表型均一的肿瘤细胞群体在化疗药物诱导下通过表观遗传机制诱导细胞重编程获得新的表型状态,从而产生适应性耐药。进一步研究发现,顺铂诱导肿瘤干细胞表型转变是由于Y染

色体性别决定域 (sex-determining region of Y chromosome, SRY)-盒转录因子 9 (SRY-box transcription factor 9, SOX9) 的表达增加, 促进了染色质开放, 使得含溴结构域蛋白 4 (bromodomain containing 4, BRD4) 促进超强启动子的 H3K27ac 位点乙酰化, 最终激活细胞重编程。短期内使用 JQ1 抑制 BRD4 可以逆转顺铂诱导的耐药性, 进一步证明顺铂耐药机制与肿瘤细胞表观遗传调控密切相关, 也为抑制细胞表观遗传靶向肿瘤药物开发提供新的策略。Miyamoto 等^[45]利用微流控富集技术对来源于 13 例前列腺癌患者的 77 个完整 CTCs 进行 scRNA-seq 分析, 发现每个个体的单个 CTC 表现出很大的异质性, 包括雄激素受体 (androgen receptor, AR) 基因突变和剪接变异体的表达, 接受 AR 抑制剂治疗患者的 CTCs 中表现出非常规 Wnt 信号通路的激活, 进一步研究发现 Wnt5a 在前列腺癌细胞中的异位表达减弱了 AR 抑制剂的抗增殖作用, 而在耐药细胞中通过抑制 Wnt5a 则可恢复其对药物的部分敏感性。该研究利用 scRNA-seq 技术对前列腺癌 CTCs 分析, 从单个细胞水平上揭示前列腺癌耐药的信号通路异质性, 为后续针对该疾病的治疗方案优化提供科学参考。Wang 等^[46]利用 scRNA-seq 技术对 HER2⁺ 乳腺癌靶向治疗药物的潜在耐药性进行建模, 发现了一种可导致耐药性的肿瘤浸润免疫抑制细胞 (immunosuppressive immature myeloid cell, IMC), 通过靶向 IMC, 发现新的药物组合 CDK4/6 和曲妥珠单抗的联合用药成功逆转肿瘤耐药性, 为肿瘤免疫疗法开辟了新途径。

炎症性肠病 (inflammatory bowel disease, IBD) 包括克罗恩病 (Crohn's disease, CD) 和溃疡性结肠炎, 其特征是胃肠道间歇性慢性炎症, 能够导致肠道损伤和残疾, 临床针对该疾病主要通过免疫治疗, 然而部分患者治疗过程中并未获得持久的免疫应答反应, 导致治疗失败。为了研究与免疫治疗耐药性相关因素, Martin 等^[47]通过对 11 例 CD 患者手术切除的回肠组织中分离的固有层细胞 scRNA-seq 分析, 构建发炎性回肠病变和邻近非发炎回肠的细胞图谱, 发现了一种与炎症相关的特殊细胞模块 GIMATS, 该模块包括 IgG 浆细胞、炎性单核吞噬细胞、活化的 T 细胞和间质细胞, 其分泌的肿瘤坏死因子 α (tumor necrosis factor α , TNF α)、白细胞介素 (interleukin, IL)-1 β 、IL-1 α 、C-C 基序趋化因子配体 (C-C motif chemokine ligand, CCL)-2、CCL7、抑瘤素 M (oncostatin M, OSM) 等形成特定的细胞因子-趋化因子网络, 进一步队列分析发现在治疗前 GIMATS 模块就已经在炎症性回肠组织中富集, 从而导致机体对抗肿瘤坏死因子治疗的免疫不应答。

2.5 助推疫苗的开发 在肿瘤、感染性疾病中, 如感

染人类免疫缺陷病毒 (human immunodeficiency virus, HIV)、登革热病毒 (dengue virus, DENV)、人乳头瘤病毒 (human papilloma virus, HPV)、流感 (influenza)、新型冠状病毒 (corona virus disease 2019, COVID-19) 等, 免疫治疗发挥重要作用。机体免疫系统是个复杂的网络, 且由于个体的遗传背景、年龄、环境暴露因素的不同, 导致不同的人群对疫苗的应答反应有所差异, 部分患者可能会出现免疫不应答或者过表达产生不良反应现象。运用 scRNA-seq 技术分析患者在接种疫苗后体内免疫细胞应答异质性, 可以评估疫苗在治疗过程中的疗效, 揭示不同人群抗原识别谱多样性的科学内涵, 从而优化疫苗治疗方案指导个性化疫苗接种。靶向肿瘤新生抗原的个性化肿瘤疫苗是当前疫苗研究领域的前沿热点, 但个性化疫苗长期使用的安全性和临床疗效尚存在争议。Hu 等^[48]利用 10 \times Genomics 测序平台对接受个性化新抗原疫苗 NeoVax 治疗的黑色素瘤复发患者外周血细胞测序, 评估疫苗治疗的长期临床疗效以及引起的免疫反应。结果发现, 记忆表型和细胞毒性的新抗原特异性 CD4⁺ T 细胞在体外长期持久性存在, 且随着时间的推移, 这些细胞会分化出多种具有不同功能亲和性的 T 细胞受体 (T-cell receptor, TCR) 克隆亚型, 起到持久抗肿瘤效果。此外, 还发现 NeoVax 诱导 T 细胞反应的表位扩散及 CD4⁺ T 细胞克隆肿瘤浸润, 表明新抗原疫苗诱导肿瘤杀伤。该研究利用单细胞测序技术对外周血免疫细胞分析, 从单个细胞层面揭示了个性化新抗原疫苗治疗过程中 T 细胞在不同时间段的分化状态, 记忆性新抗原特异性 T 细胞的持久性存在反映了新抗原疫苗治疗的临床效果, 为肿瘤个性化疫苗的研究提供科学依据。

免疫系统的异质性以及宿主控制与疫苗诱导的细胞介导免疫大小、多样性和持久性机制是疫苗学的知识空白, 为研究该科学问题, Waickman 等^[49]采用 scRNA-seq 分析登革热疫苗 TAK-003 对患者血液样本中 CD8⁺ T 细胞免疫激活的效应机制, 结果发现暴露于 TAK-003 的 CD8⁺ T 细胞具有高度多克隆性, 且找到了可用于准确识别 CD8⁺ T 细胞对抗原刺激或体内对 TAK-003 反应的代谢标记物 Tfr1/CD71, 其能调控 DENV-2 非结构蛋白 NS-1 和 NS3 反应性记忆前体 CD8⁺ T 细胞对 Fe³⁺ 的摄取, 使得 CD8⁺ T 细胞具有高增殖和效应/记忆潜能。

基于树突细胞 (dendritic cell, DC) 呈递的免疫疗法在肿瘤治疗中具有较好的应用前景, 但由于生产方式差异从外周血白细胞中大规模分离 DCs, 会产生细胞异质性监管难题, 有可能引发疫苗安全性的担忧。Li 等^[50]利用 10 \times Genomics 单细胞转录组测序平台系统

研究了CUD-002疫苗的细胞组成,并经FACS验证,结果发现,其中80%以上都是DCs,剩下的主要为骨髓细胞和淋巴细胞,且这些非树突细胞并不会影响DCs的抗肿瘤活性、也无明显的毒副作用,该分析方法可以指导GMP级别树突细胞疫苗CDU-002的生产,确保疫苗质量和安全性。

3 总结与展望

随着测序技术的不断更新和成熟,可检测的细胞、分子实现高通量化、精准化、高效化。scRNA-seq技术在药物研究领域发挥推波助澜作用。运用该技术可以构建机体完整的细胞图谱,追踪与疾病相关的特异性细胞亚群,指导药物靶点筛选,揭示机体对药物治疗后异质性应答反应机制,为临床个性化、精准化药物治疗提供一个全新的视角。

然而,任何技术都存在自身的缺陷,scRNA-seq技术也不例外。首先,在药物靶点筛选过程中,scRNA-seq相比bulk-seq具有去平均化效应,可以提高靶点筛选精度,但操作步骤较多、测序成本高、测序深度较低依然是制约该技术推广的一大因素,将scRNA-seq和bulk-seq联合应用可以进一步验证scRNA-seq数据准确性和结果的可靠性。其次,大部分药物的主要靶点是蛋白质而非mRNA,并且mRNA和蛋白质的表达水平存在差异,scRNA-seq只能反映单细胞的转录组信息,因此将scRNA-seq和蛋白质组学联合分析可以揭示调控蛋白的细胞来源,阐释疾病发生发展规律以及药物调控机制。此外,在肿瘤化疗过程中,细胞对药物的刺激会做出一系列复杂的反应,包括DNA、RNA和蛋白质等的变化,scRNA-seq显然不能全面反映这种变化规律,将单细胞转录组、基因组、蛋白质组和代谢组等多组学技术联合应用更加有利于理解基因与表型之间的关联性,有助于从多维度 and 更深层次地解析机体的奥秘以及发现新的药物靶标。单细胞测序样本处理是保证测序结果好坏的关键步骤,scRNA-seq需要制备单细胞悬液,因此丢失了细胞在特定空间位置信息,导致细胞间通讯网络被打破使得转录本水平发生改变。如何在空间分辨率水平研究单细胞的转录组信息也是一大难题。基于质谱成像MALDI-MSI的空间蛋白技术与scRNA-seq联合应用可以在空间上揭示细胞异质性。空间转录组(spatial transcriptomics, ST)技术的发展也为在空间水平解析细胞转录特征提供有力依据,如基于10×Genomics Visium平台整合scRNA-seq数据和ST数据建立的MIA分析方法,能够准确注释不同组织区域的精确细胞组成,从而真正实现在空间维度检测单细胞转录本分布信息。与此同时,基于干细胞和3D细胞培养技术发展的类器官模型和器官芯

片技术也在一定程度上保留了细胞空间信息,弥补了某些疾病因为缺乏对应的模型而限制药物开发的不足,将scRNA-seq与类器官模型相结合用于药物筛选,将为新药研发提供新的视角,也为精准化、个性化药物开发提供新的策略。最后,单细胞测序技术能够深入揭示化学药物、天然药物及生物药在疾病治疗过程中的作用机制,但是对于中药及其复方这种复杂体系的作用解析,依然面临巨大挑战,将scRNA-seq数据、中药驱动网络大数据和人工智能相结合,开发符合中药整合作用特点的数据分析软件或平台将有助于揭示中药治疗疾病的科学内涵。

scRNA-seq技术在药物研究领域具有广阔的应用前景,随着测序技术的发展及多学科的融合,降低测序成本、实现高质量样本制备、3D立体重建保存细胞空间信息、整合生物信息学数据库搭建单细胞测序数据分析平台等将有助于该技术在实验室及临床研究的推广,基于单细胞水平的高通量药物筛选平台也将为新药研发提供新的方向。

作者贡献:汪茂林负责文献检索和文章撰写;杨洪军负责文章内容核对和修改。

利益冲突:所有作者声明本文无任何利益冲突。

References

- [1] Gibbs RA. The human genome project changed everything [J]. *Nat Rev Genet*, 2020, 21: 575-576.
- [2] Regev A, Teichmann SA, Lander ES, et al. The human cell atlas [J]. *Elife*, 2017, 6: e27041.
- [3] Van den Berge K, Perraudeau F, Sonesson C, et al. Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications [J]. *Genome Biol*, 2018, 19: 24.
- [4] Tang F, Barbacioru C, Wang Y, et al. mRNA-seq whole-transcriptome analysis of a single cell [J]. *Nat Methods*, 2009, 6: 377-382.
- [5] Hashimshony T, Wagner F, Sher N, et al. Cel-seq: single-cell RNA-seq by multiplexed linear amplification [J]. *Cell Rep*, 2012, 2: 666-673.
- [6] Sasagawa Y, Nikaido I, Hayashi T, et al. Quartz-seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity [J]. *Genome Biol*, 2013, 14: R31.
- [7] Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets [J]. *Cell*, 2015, 161: 1202-1214.
- [8] Gao C, Zhang M, Chen L. The comparison of two single-cell sequencing platforms: BD rhapsody and 10× genomics chromium [J]. *Curr Genomics*, 2020, 21: 602-609.
- [9] van der Wijst M, de Vries DH, Brugge H, et al. An integrative approach for building personalized gene regulatory networks for

- precision medicine [J]. *Genome Med*, 2018, 10: 96.
- [10] De T, Park CS, Perera MA. Cardiovascular pharmacogenomics: does it matter if you're black or white?[J]. *Annu Rev Pharmacol Toxicol*, 2019, 59: 577-603.
- [11] Mcginnis CS, Patterson DM, Winkler J, et al. Multi-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices [J]. *Nat Methods*, 2019, 16: 619-626.
- [12] Zheng GX, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells [J]. *Nat Commun*, 2017, 8: 14049.
- [13] Grindberg RV, Yee-Greenbaum JL, McConnell MJ, et al. RNA-sequencing from single nuclei [J]. *Proc Natl Acad Sci U S A*, 2013, 110: 19802-19807.
- [14] Soumillon M, Cacchiarelli D, Semrau S, et al. Characterization of directed differentiation by high-throughput single-cell RNA-seq [J]. *bioRxiv*, 2014, 65: 631-643.
- [15] Hashimshony T, Senderovich N, Avital G, et al. Cel-seq2: sensitive highly-multiplexed single-cell RNA-seq [J]. *Genome Biol*, 2016, 17: 77.
- [16] Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types [J]. *Science*, 2014, 343: 776-779.
- [17] Keren-Shaul H, Kenigsberg E, Jaitin DA, et al. Mars-seq2.0: an experimental and analytical pipeline for indexed sorting combined with single-cell RNA sequencing [J]. *Nat Protoc*, 2019, 14: 1841-1862.
- [18] Fan X, Zhang X, Wu X, et al. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos [J]. *Genome Biol*, 2015, 16: 148.
- [19] Sheng K, Cao W, Niu Y, et al. Effective detection of variation in single-cell transcriptomes using matq-seq [J]. *Nat Methods*, 2017, 14: 267-270.
- [20] Picelli S, Faridani OR, Bjorklund AK, et al. Full-length RNA-seq from single cells using Smart-seq2 [J]. *Nat Protoc*, 2014, 9: 171-181.
- [21] Klein AM, Mazutis L, Akartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells [J]. *Cell*, 2015, 161: 1187-1201.
- [22] Han X, Wang R, Zhou Y, et al. Mapping the mouse cell atlas by Microwell-seq [J]. *Cell*, 2018, 172: 1091-1107.
- [23] Gierahn TM, Wadsworth MN, Hughes TK, et al. Seq-well: portable, low-cost RNA sequencing of single cells at high throughput [J]. *Nat Methods*, 2017, 14: 395-398.
- [24] Fan X, Tang D, Liao Y, et al. Single-cell RNA-seq analysis of mouse preimplantation embryos by third-generation sequencing [J]. *PLoS Biol*, 2020, 18: e3001017.
- [25] Zheng Y, Chen Z, Shi Z, et al. Hit-scisoseq: high-throughput and high-accuracy single-cell full-length isoform sequencing for corneal epithelium [J]. *bioRxiv*, 2020. DOI: 10.1101/2020.07.27.222349.
- [26] Shin D, Lee W, Lee JH, et al. Multiplexed single-cell RNA-seq *via* transient barcoding for simultaneous expression profiling of various drug perturbations [J]. *Sci Adv*, 2019, 5: v2249.
- [27] Rhoads A, Au KF. Pacbio sequencing and its applications [J]. *Genomics Proteomics Bioinformatics*, 2015, 13: 278-289.
- [28] Lin B, Hui J, Mao H. Nanopore technology and its applications in gene sequencing [J]. *Biosensors*, 2021, 11: 214.
- [29] Lebrigand K, Magnone V, Barbry P, et al. High throughput error corrected nanopore single cell transcriptome sequencing [J]. *Nat Commun*, 2020, 11: 4025.
- [30] Singh M, Al-Eryani G, Carswell S, et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes [J]. *Nat Commun*, 2019, 10: 3120.
- [31] Philpott M, Watson J, Thakurta A, et al. Nanopore sequencing of single-cell transcriptomes with scCOLOR-seq [J]. *Nat Biotechnol*, 2021, 39: 1517-1520.
- [32] Liao Y, Liu Z, Zhang Y, et al. High-throughput and high-sensitivity full-length single-cell RNA-seq analysis on third-generation sequencing platform [J]. *Cell Discov*, 2023, 9: 5.
- [33] van Dijk EL, Jaszczyszyn Y, Naquin D, et al. The third revolution in sequencing technology [J]. *Trends Genet*, 2018, 34: 666-681.
- [34] Cheng YH, Chen YC, Lin E, et al. Hydro-seq enables contamination-free high-throughput single-cell RNA-sequencing for circulating tumor cells [J]. *Nat Commun*, 2019, 10: 2163.
- [35] Gladka MM, Molenaar B, de Ruyter H, et al. Single-cell sequencing of the healthy and diseased heart reveals cytoskeleton-associated protein 4 as a new modulator of fibroblasts activation [J]. *Circulation*, 2018, 138: 166-180.
- [36] Johnston G, Ramsey HE, Liu Q, et al. Nascent transcript and single-cell RNA-seq analysis defines the mechanism of action of the LSD1 inhibitor INCB059872 in myeloid leukemia [J]. *Gene*, 2020, 752: 144758.
- [37] Ye C, Ho DJ, Neri M, et al. Drug-seq for miniaturized high-throughput transcriptome profiling in drug discovery [J]. *Nat Commun*, 2018, 9: 4307.
- [38] Zhang B, Huang K, Zhu L, et al. Precision toxicology based on single cell sequencing: an evolving trend in toxicological evaluations and mechanism exploration [J]. *Arch Toxicol*, 2017, 91: 2539-2549.
- [39] Umbaugh DS, Ramachandran A, Jaeschke H. Spatial reconstruction of the early hepatic transcriptomic landscape after an acetaminophen overdose using single-cell RNA-sequencing [J]. *Toxicol Sci*, 2021, 182: 327-345.
- [40] Gupta N, Matsumoto T, Hiratsuka K, et al. Modeling injury and repair in kidney organoids reveals that homologous recombination governs tubular intrinsic repair [J]. *Sci Transl Med*, 2022, 14: j4772.
- [41] Kim D, Kobayashi T, Voisin B, et al. Targeted therapy guided by

- single-cell transcriptomic analysis in drug-induced hypersensitivity syndrome: a case report [J]. *Nat Med*, 2020, 26: 236-243.
- [42] Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing [J]. *N Engl J Med*, 2012, 366: 883-892.
- [43] Kim C, Gao R, Sei E, et al. Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing [J]. *Cell*, 2018, 173: 879-893.
- [44] Sharma A, Cao EY, Kumar V, et al. Longitudinal single-cell RNA sequencing of patient-derived primary cells reveals drug-induced infidelity in stem cell hierarchy [J]. *Nat Commun*, 2018, 9: 4931.
- [45] Miyamoto DT, Zheng Y, Wittner BS, et al. RNA-seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance [J]. *Science*, 2015, 349: 1351-1356.
- [46] Wang Q, Guldner IH, Golomb SM, et al. Single-cell profiling guided combinatorial immunotherapy for fast-evolving CDK4/6 inhibitor-resistant HER2-positive breast cancer [J]. *Nat Commun*, 2019, 10: 3817.
- [47] Martin JC, Chang C, Boschetti G, et al. Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti-TNF therapy [J]. *Cell*, 2019, 178: 1493-1508.
- [48] Hu Z, Leet DE, Allesoe RL, et al. Personal neoantigen vaccines induce persistent memory T cell responses and epitope spreading in patients with melanoma [J]. *Nat Med*, 2021, 27: 515-525.
- [49] Waickman AT, Victor K, Li T, et al. Dissecting the heterogeneity of DENV vaccine-elicited cellular immunity using single-cell RNA sequencing and metabolic profiling [J]. *Nat Commun*, 2019, 10: 3666.
- [50] Li Q, Yang C, Tian H, et al. Development of a personalized dendritic cell vaccine and single-cell RNA sequencing-guided assessment of its cell type composition [J]. *Cytotherapy*, 2023, 25: 210-219.