

单细胞测序技术在肝脏疾病的应用与展望

过冬冬^{1, 3} 孙芬¹ 贺轩昂² 羊东晔² 黄来强¹

(1. 清华大学深圳国际研究生院, 深圳 518055; 2. 香港大学深圳医院消化及肝脏科, 深圳 518053; 3. 清华大学化学系, 北京 100084)

摘要: 新兴的单细胞测序技术能够从单细胞水平揭示基因组、转录组和表观遗传学等分子水平发生的基因变异与表观修饰状态, 也可用于鉴定新的细胞类型和表面标记物。这将帮助人们探明疾病发生时细胞基因、转录或表观修饰方面的变化, 了解细胞之间的联系, 以及深入理解肿瘤异质性。目前, 单细胞测序技术已用于多种疾病的研究, 其在肝脏疾病, 包括肝硬化、肝癌中已有相关成果。于此, 综述了单细胞测序技术在肝脏发育及肝病中的应用, 讨论了肝脏疾病发生的内在机制以及该技术仍存在的问题, 提出可能的解决方案, 如发展三维单细胞测序技术将更能帮助人们深刻了解肝脏疾病发生机制。

关键词: 单细胞测序技术; 组学测序; 肝脏; 肝病

DOI: 10.13560/j.cnki.biotech.bull.1985.2020-0463

Application and Prospects of Single-Cell Sequencing in Liver Disease

GUO Dong-dong^{1, 3} SUN Fen¹ HE Xuan-ang² YANG Dong-ye² HUANG Lai-qiang¹

(1. Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055; 2. Division of Gastroenterology & Hepatology, the University of Hong Kong-Shenzhen Hospital, Shenzhen 518053; 3. Department of Chemistry, Tsinghua University, Beijing 100084)

Abstract: Emerging single-cell sequencing is a potent technique, which can not only reveal mutations on genomics, transcriptomics and epigenetic modification at the molecular level but also be used to identify new cell types and surface markers. It can also help to detect changes in cell genes, transcriptomics or epigenetic modifications when a disease occurs, to understand the interactions among cells as well as to comprehensively understand tumor heterogeneity. Currently, single-cell sequencing has been applied to the studies on many diseases, and achievements on liver diseases including cirrhosis and liver cancer to some accomplishments are reported. Herein, the applications of single-cell sequencing in liver development and liver diseases are reviewed, underlying mechanisms of liver diseases and the issues of this technology are discussed, and the feasible resolutions are proposed, for instance, developing 3D single-cell sequencing technology may help us to deeply understand the mechanisms of liver diseases.

Key words: single-cell sequencing; omics sequencing; liver; liver diseases

在过去的几十年里人们对疾病的研究已逐步聚焦于分子水平的改变。自从1975年英国科学家Sanger发明第一代基因测序^[1]以来, 基因测序已发展成人们探索细胞遗传物质变化和疾病发展的有利工具。21世纪初期, 人们宣布了人类基因组计划^[2]完成, 这成为探索人类遗传物质的一个重要里程碑, 而后人们发现器官水平的测序不能完全揭示机体变

化的完整历程和潜在原因。2009年, Tang等^[3]首次在*Nature Methods*杂志上发表从单个细胞水平上的转录组测序文章, 自此打开了单细胞测序的大门。传统的“块样”大面积测序的方法得到的结果可能是由于多个细胞“平均”后的宏观结果, 掩盖了单个细胞之间存在的差异, 致使人们在研究相关疾病的时候往往聚焦在“面”上, 而单细胞测序技术的

收稿日期: 2020-04-21

基金项目: 深圳市科技计划基础研究(学科布局)项目(JCYJ20180508153013853), 深圳市科技计划项目(JCYJ20180508152130899)

作者简介: 过冬冬, 男, 硕士研究生, 研究方向: 肿瘤免疫治疗; E-mail: gdd18@mails.tsinghua.edu.cn

通讯作者: 羊东晔, 女, 博士, 主任医师, 研究方向: 肝肿瘤分子发病机制及免疫基因治疗; E-mail: yangdy@hku-szh.org

出现使得人们能够看到具体的“点”。

肝脏作为人体中心免疫器官、血流丰富器官和再生能力超强的器官,具有复杂的细胞组成与分化程度不同的细胞阶段。近年来研究表明,肝脏疾病的发生很大程度上是肝实质细胞自身遗传物质、所处的微环境与外来影响因子(如病毒)共同作用的结果^[4-5]。然而有关肝脏疾病的透彻认识与精准治疗却十分有限,原因之一在于传统疾病分析方法往往不够精准且具有一定局限性。单细胞测序技术的出现对揭示肝脏疾病内在发生机制有重要意义,它使得人们能够从分子水平了解肝脏功能和肝脏有关疾病的生理与病理意义。早期诊断、精准治疗已成为肝脏疾病治疗的终极目标。本文简要介绍单细胞测序技术在肝脏疾病中的应用发展,期望能为诠释肝脏疾病的发病机制、解析疾病发生的细胞互作网络提供参考。

1 单细胞测序技术简介

与传统的测序方法不同,单细胞测序大体由3个步骤组成:单细胞制备与捕获、单细胞文库构建与测序及数据处理分析。

1.1 单细胞的制备与捕获

高质量的样本是单细胞研究的关键所在。准确的测序结果与细胞的状态密不可分,有研究表明在不同的温度下细胞内的基因表达水平不尽相同,有时相差高达上千倍^[6]。待测组织先经过机械剪切成小块再经酶消化使细胞相互分离。根据样本选择使用不同的酶配方,以肝组织为例,可选用终浓度为0.16 mg/mL 胶原蛋白酶 IV 在 37°C 消化 10 min;或用 40 μg/mL 的 Liberase Blendzyme3 消化 5-8 min。消化后的悬液要经过滤网筛以除去较大的组织团块和细胞碎片,然后将细胞重悬在含有胎牛血清(Fetal bovine serum, FBS)的培养基中,并及时进行单细胞捕获以防细胞沉降聚集或发生死亡^[7],因为聚集的细胞会使得单细胞捕获过程中发生非单一细胞测序,死亡的细胞会由于自身组学发生改变造成测序结果的偏差。

现在手工分选细胞方法已不再使用,较为常见的单细胞捕获方法为荧光激活细胞分选(Fluorescent-activated cell sorting, FACS)和微流控(Microfluidics)

技术。FACS 利用荧光标记细胞表面标志物将目的细胞分选至微量滴定板中,其效率高、捕获准且可去除可能对测序产生影响的受损细胞和死亡细胞^[8]。但 FACS 由于快速的流体分选过程会对细胞造成一定的损伤,所以为保证所得的数据量需要扩大起始分选量,然而这对某些稀少组织来说并不容易实现。而微流控技术^[9]是依赖集成的微流控电路捕获含细胞的液滴至纳米孔中,可以同时处理上千个细胞,大大减少了试剂和材料的使用,在控制了成本的同时也保证了效率。激光捕获显微切割法(Laser capture microdissection, LCM)^[10]不需要事先制备细胞悬液而是利用激光对组织进行切割,其特点是可以从空间上反映出每个细胞所处的位置但是对细胞损伤较大,一般在特定情况下才会选用。

1.2 单细胞不同组学测序

1.2.1 单细胞基因组测序 细胞内极少的 DNA 含量不能达到直接测序需求,所以在进行单细胞基因组测序之前需先进行全基因组扩增(Whole genome amplification, WGA)。其中基于 PCR 扩增技术的有:简并寡核苷酸 PCR(Degenerate oligonucleotide primed-polymerase chain reaction, DOP-PCR)与扩增前引物延伸 PCR(Primer extension preamplification-PCR, PEP-PCR)^[11]。这两种方法都是在 Taq 聚合酶的作用下通过引物结合、延伸然后退火的过程执行扩增行为。因为在扩增时,即便是小的扩增条件改变(如引物、Taq 聚合酶浓度或退火温度等)也会因指数级扩增造成结果出现巨大差异,以致基因组的某些区域扩增过度而另一些区域扩增不足,产生较大的扩增后偏倚^[12]。因此基于 PCR 扩增的测序技术有很大的基因组覆盖差异,其范围可达到 10%-90% 之广^[12-13],如单核测序(Single-nucleus sequencing, SNS)基因组覆盖广度只有约 10%,而多次退火循环扩增技术(Multiple annealing and looping-based amplification cycles, MALBAC)^[12, 14]利用特异引物实现仅扩增原始基因组的效果,使得基因组覆盖广度达到 90% 以上,从而减少因为 DNA 被循环扩增而产生的扩增偏倚。多重置换扩增(Multiple displacement amplification, MDA)^[15]利用高保真聚合酶在恒温下进行链置换合成,由于

该聚合酶具有 DNA 链 3'-5' 校正作用，所以相比于上述 PCR 型扩增技术大约可将保真度提高一千倍。现在单细胞基因组扩增还可采用 PicoPLEX^[16] 扩增技术，它具有相比于 MALBAC 更短的细胞裂解时间，重复性更高，对细胞拷贝数变异（Copy number variations, CNV）更加敏感等优势，因此逐渐受到人们青睐。扩增后的基因组在传统二代测序的基础上进行测序分析。

1.2.2 单细胞转录组测序 单细胞转录组的测序信息能直观地反映出细胞间基因表达差异。现有的单细胞转录组测序过程都是利用含有 poly-T 的寡核苷酸来捕获含有 poly-A 尾部的 RNA 分子，之后转录成稳定的 cDNA 分子。现为了方便后续建库^[17]，用十几个碱基组成的特殊 barcodes 对同一个细胞的转录本进行标记，使得在之后的细胞池中可以准确区分每一个细胞的单细胞信息，避免不同细胞信息混合在一起难以区分彼此的情况，如 STRT-seq、CEL-seq 便采用此种方法。特殊分子标记（Unique Molecular Identifier, UMI）的引入可以在很大程度上纠正扩增造成的结果偏倚和降低背景噪音，如 CEL-seq2、Drop-seq 则是应用 UMI 特殊标记的建库方法^[18]。得到 cDNA 后再通过体外转录技术（*In vitro* transcription, IVT）进行扩增以满足后续测序要求。测序过程可根据目的和条件选择不同的测序手段，一般来说有两种转录本的测序方法：全长测序和 3' 或 5' 端测序。全长测序能够完整的表达转录组序列信息并能识别出基因变异以及转录部分的基因改变，如单核苷酸变异和融合转录等。对于 3' 或 5' 端 50-100 bp 信息已经能满足实验的要求，考虑到成本和全长转录的实际价值，末端测序的方法则会大幅降低成本。

1.2.3 单细胞表观遗传学测序 近年来，表观遗传学的研究已成为人们从分子水平上了解细胞行为的重要生物学分支。表观修饰为人们理解由基因组和转录组控制的细胞行为之外的行为提供了途径。DNA 甲基化在表观遗传中起着重要作用，而哺乳动物中 5-胞嘧啶甲基化（5mC）则在 DNA 甲基化中占主导地位，因此利用单细胞测序技术测定 DNA 甲基化程度十分必要。当前亚硫酸氢盐测序的原理是利用亚硫酸氢钠处理 DNA 后将未修饰的胞嘧啶转化为

尿嘧啶，但不影响 5mC^[19]。基于还原替代亚硫酸氢盐测序（Reduced representation bisulfite sequencing, RRBS）和全基因组重亚硫酸盐处理后的接头和引物扩增（Post-bisulfite adaptor tagging, PBAT）测序策略，建立了单细胞 DNA 甲基化组谱分析方法，避免了 DNA 丢失，确保下一代基因测序（Next generation sequencing, NGS）中覆盖整个基因组。亚硫酸氢盐处理后的 DNA 再经过 PCR 扩增，至满足后续测序要求再测序。如已普及的 ATAC-seq 技术^[20]，通过识别开放染色质并引入引物在染色质开放区域进行高通量测序达到对单细胞表观遗传修饰进行测序的目的，这是一种集细胞捕获、文库建立与测序一体的技术，大大简化了测序的过程。

1.2.4 单细胞多组学测序 对于某些单细胞样本获取并不容易，并且在制备单细胞悬液的过程中不可避免地造成细胞损失，这使得期望利用某一样本进行复杂的多组学分析变得异常困难。然而对多组学的平行测序是帮助我们全面综合分析细胞行为的重要一步。现已有多个课题组开发了两组学甚至多组学同时测序的方法，如由 Dey 等^[21]报道的 DR-seq 技术，在裂解单细胞后同时扩增细胞内的基因组与转录组，然后将裂解物分成两个部分，分别用于基因组测序和转录组测序。该课题组利用 DR-seq 分别对小鼠胚胎干细胞系（E14）和乳腺癌细胞系（SK-BR-3）进行了测序发现，细胞到细胞之间基因的变化程度与拷贝数变异呈负相关趋势，表明拷贝数变异可能驱动个体细胞的基因表达。但这个方法存在不能完全隔离 DNA 和 RNA 的问题，无法避免相互干扰的风险。另一种基因组与转录组同时测序的 G&T-seq 技术由 Macaulay 等^[16]报道，该课题组利用 G&T-seq 技术对已经基因组测序的乳腺癌细胞和 B 淋巴瘤细胞系再次测序发现，B 淋巴瘤细胞中存在 11 号染色体三体的细胞亚群。同时对分裂期的小鼠胚胎细胞和诱导多能干细胞进行测序，明确了细胞分裂中染色体的错配与染色体表达量的关系，表明该测序方法可获得的信息远远超过之前的测序方法。随着技术的不断成熟，单细胞三组学测序（scTrio-seq）也逐渐为人们所用，它是将细胞裂解液的上清液进行转录组测序，而对裂解物沉淀进行基因组和 DNA 甲基化测序的方法，Hou 等^[22]利

用其对 25 个来源于人肝细胞癌组织的单细胞进行基因组、甲基化组和转录组同时测序, 跟据 CNVs, DNA 甲基化和单个细胞转录组确定了肝细胞癌的两个细胞亚群, 揭开了每个细胞亚群之间存在的异质性差异。另外, Bian 等^[23]在结直肠癌肿瘤病人原发瘤、淋巴结和远处转移瘤的多部位取样, 运用优化的单细胞多组测序方法 (scTrio-seq2), 进一步了解结直肠癌的内部异质性。因此多组学同时测序的方法能为人们提供更加全面综合的视野。

1.3 数据处理与分析

单细胞测序后的原始数据需要经过一系列步骤转换为可读的基因表达矩阵。在生成 FASTQ 读数后要通过质量控制步骤, 筛选掉不符合要求的数据然后用 barcodes 进行复读, 再由计算机对映射读数进行量化, 以创建一个表达矩阵^[24]。标准化的原始数据处理流程和计算机处理所用的数据处理包都可在公开数据库下载。通过对数据标准化处理^[25], 以去除非生物技术效应相关的高水平的噪音和差异性, 包括在样品制备过程中由于随机 RNA 丢失、偏向扩增和文库测序不完整而导致的非准确情形。其他不稳定性也可能来自于对加工单元 (如板或阵列)、时间点、设备和其他来源的批量影响。因此, 数据标准化成为单细胞数据分析的重要步骤。

单细胞的数据分析要根据自己的实验要求设定相应参数。以 scRNA-seq 为例, 主要是对测试样本进行异质性分析, 评估新的细胞类型或分析细胞发育过程中发生的基因变化。通过已知的特定基因对细胞进行分群, 而后常见的可视化分析是降维处理并把各个群的细胞投影到二维或者三维坐标空间中去。普遍被人们采用的数据处理方法是主成分分析 (Principal components analysis, PCA) 和 t-分布随机邻域嵌入 (t-distributed stochastic neighbor embedding, t-SNE) 算法。尽管种群标记时允许监督聚类, 但在大多数情况下, 无假设的非监督聚类是首选的^[26]。不同实验间的测试数据已经被上传至公开数据库, 允许科研人员免费访问。

2 单细胞测序在肝脏疾病中的研究

单细胞测序现已应用于肝脏疾病研究中, 并已取得一定成果。现有研究主要集中在对肝组织中所

含有的各类细胞进行分群探索, 期望揭示各细胞间的相互联系。在肝病方面, 肝癌和肝硬化则是人们探索的重点, 通过单细胞测序技术对疾病的起因、发展和治疗奠定理论基础和提供治疗依据。

2.1 单细胞测序描绘肝细胞类型图谱

肝脏内含有丰富细胞群、细胞亚群、并且各群细胞之间相互作用, 共同构建平衡的肝内环境, 细胞的异质性和强大再生能力是其主要特点。在探索肝内异质性方面, Aizarani 等^[27]通过对 9 个人的近 10 000 个细胞进行了单细胞 RNA 测序, 构建了一幅人类肝脏图谱。该课题组基于 mCEL-seq2 技术对肝脏中所有细胞类型进行确定, 并发现一种具有祖细胞特征的 EPCAM⁺ TROP2^{int} 细胞群, 它对肝脏内平衡的维持, 肝再生和疾病发生发展十分关键。这次测序结果帮助人们寻找到此前从未发现的内皮细胞亚群, 并对找到正常肝脏和癌变肝脏之间的变化起到重要作用。这个肝脏全细胞群图谱的描绘, 有助于人们从宏观逐步走向微观了解肝脏。同时, 新发现的细胞群也表明, 以往分析组织病变的方法可能忽略了某些数量稀少的细胞群, 因为这些细胞被其他数量庞大的细胞所掩盖, 而这些少数细胞很可能是决定组织正常或病变的执牛耳者。肝脏作为人体免疫过程中的重要器官, 也需要揭示其所包含的免疫细胞特征。MacParland 等^[28]首次利用单细胞测序确定了肝细胞内独特的巨噬细胞群并明确了其功能通路。而此前由于肝内巨噬细胞的分离难度和复杂的个体基因导致人们对其知之甚少。作者发现了两类不同的 CD68⁺ 巨噬细胞群, 其中表达丰富的 LYZ, CSTA, CD74 的一群被定义为肝内的一种炎性巨噬细胞, 然而以前仅凭借细胞表面标志物区分调节性巨噬细胞或炎性巨噬细胞的方法并不能准确区分同种标志物下的不同亚群细胞。对肝内免疫细胞群的精确划分将有助于人们构建肝内免疫系统图谱, 帮助人们更为深刻的认识正常与异常肝脏之间的差异。

单细胞测序技术除了帮助人们识别不同的细胞群, 还可以帮助了解各类细胞在发育过程中的变化, 评估微环境与组织生长发育关系。肝脏作为在发育成熟过程中生理功能变化的一个典型代表器官, 已

经发现在发育中出现造血功能的改变，而其内在的原因并不清楚。因此，Popescu 等^[29]通过对大约 140 000 个肝脏和 74 000 个皮肤、肾脏和卵黄囊细胞的单细胞转录组分析，确定了人类血液和免疫细胞在发育过程中的全部功能。并从造血干细胞/多能祖细胞（Haematopoietic stem cells and multipotent progenitors, HSC/MPPs）中推断出分化轨迹，以及评估了组织微环境对血液和免疫细胞发育的影响。实验证实，在妊娠期间胎儿肝脏的造血成分发生了变化，不再以红细胞为主，而是伴有造血干细胞和多能干细胞平行分化。胎龄对 HSC/MPPs 分化潜能的调节提示这可能是在妊娠第一和第二阶段调节胎儿肝脏造血功能输出的一种额外的功能机制。而机体也是处于不断变化过程中的，揭示各个器官以及血液细胞发育时的动态信息对了解和实时监测人体动态变化十分关键，同时也对解密肝脏造血机制和勾画儿科肝脏、血液与免疫疾病的关系蓝图起重要作用。

2.2 单细胞测序在肝硬化中的研究

肝硬化是肝脏疾病中致死率较高的一类疾病，其主要特征是肝脏的大范围纤维化。目前，人们对肝纤维化的治疗并无十分有效的手段，对其背后的机理也缺乏深入的理解。然而近期一篇发表在 *Nature* 上的单细胞测序工作^[30]帮助人们了解肝纤维化背后的细胞互作情况。Ramachandran 等^[30]通过对超过 100 000 个人类细胞进行 scRNA-seq 后发现一个与疤痕形成有关的 TREM2⁺ CD9⁺ 巨噬细胞亚群，这是一类在肝纤维化过程中显著促进纤维化的细胞亚群。同时定义了一种在纤维化微环境下才存在的 ACKR1⁺ 和 PLVAP⁺ 内皮细胞，研究表明它能够扩大肝纤维化的面积并促进白细胞迁移。另外，该文章还揭示了细胞背后的促纤维化信号通路，如 TNFRSF12A, PDGFR 和 NOTCH 信号。据此我们可以深入了解到参与肝纤维化细胞和背后的分子机理，为寻找合适的治疗靶点提供参考。以上两篇文献表明，肝脏内巨噬细胞与多种生理状态相关，是一个不可忽略的细胞群体。Krenkel 等^[31]和 Dobie 等^[32]两个课题组的工作表明肝纤维化还与肝星状细胞有关。前者单细胞转录组测序结果表明，在肝损伤之后，

肝星状细胞向胶原分泌型肌成纤维细胞转分化会促进肝纤维化过程。数据显示了肝星状细胞和肌成纤维细胞的异质性，表明肝纤维化中存在功能相关的亚群。而 Dobie 课题组发现中心静脉相关的肝星状细胞（Central vein-associated HSCs, CaHSCs）作为胶原生成细胞是肝小叶中心纤维化的主要控制因素，并且鉴定出 LPAR1 是其治疗靶点，为肝纤维化的治疗提供了新方向。肝纤维化的研究已经因单细胞测序技术而加快了脚步，但是内在机制和治疗靶点的寻找仍需要不断深入。

2.3 单细胞测序在肝癌中的研究

肝癌是全球第六大癌症，也是导致癌症死亡的第二大原因，仅中国就占新发病例和死亡病例的一半以上。在中国所有的癌症中，肝癌是存活率最低的肿瘤之一，5 年相对生存率仅为 10.1%^[33]。为揭示免疫细胞在肝细胞癌中的动态变化，Zhang 等^[34]用商业成熟的 SMART-seq2 和 10× Genomics Chromium3 技术对肝癌患者身体多个部位的细胞进行单细胞测序，通过描绘不同组织中细胞的动态变化构建出更加广泛的细胞联系。他们发现不同组织的免疫细胞组成差别巨大，来自于患者腹水的细胞有很强的组织特异性，并通过生物学分析得出存在巨噬细胞从肿瘤迁移到腹水的过程。除此之外，与患者不良预后有关的基因 SLC40A1 和 GPNMB 也在肝脏肿瘤相关巨噬细胞中显著表达，基因敲除验证表明这两个基因与肿瘤内炎症反应息息相关。这是人们首次对人体不同组织之间细胞状态关联性探究。人体是一个有机的整体，割裂某一部分必不能观察到全貌，因此这也为研究肿瘤发生时对整体产生的影响提供了范例。值得注意的是，肝脏作为一个功能复杂的器官，其组成成分的复杂造成了发生肿瘤病变时肝脏内并非只有肝细胞癌的发生。Xue 等^[35]利用单细胞基因组和转录组测序手段分析鉴别了 133 例肝癌合并肝内胆管癌（Combined hepatocellular and intrahepatic cholangiocarcinoma, HCC-ICC）患者的组织样本发现，包括分离、合并和混合亚型。将 cHCC-ICC 与肝细胞癌、肝内胆管癌进行综合比较发现，合并型和混合型 cHCC-ICCs 是不同的亚型，具有不同的临床和分子特征，并从分析数据中得出可

以 Nestin 可作为 cHCC-ICCs 生物标志物的结论,为临床上治疗 cHCC-ICCs 提供了潜在靶点。近年来肿瘤干细胞愈发受到人们关注,作为肿瘤顽固且难以治愈的罪魁祸首,人们期望对肿瘤干细胞加以分析,以找到精确有效的肿瘤干细胞靶点。Zheng 等^[36]通过对肿瘤干细胞转录组的分析发现肝肿瘤干细胞在肝内是独特的存在,其异质性高于普通肝癌细胞。研究表明不同标志物的肝癌干细胞可能受到不同的肿瘤驱动因子驱动,同时单细胞水平的肝癌干细胞具有表型、功能和转录组异质性。不同的肿瘤干细胞可能给肿瘤带来不同的异质性,这就解释肝癌治疗效果因个体而异的原因。这个工作为研究肿瘤细胞亚群和肿瘤发生发展提供了独特的视角,提示对肿瘤的治疗应该由普遍化走向个体化,找到适合每一个人的精准治疗方法才是肿瘤治疗的未来之路。

3 总结与展望

自从单细胞测序技术面世以来,已逐渐揭开多种疾病的神秘面纱,尤其在胚胎发育和肿瘤机制及治疗领域有着广阔的应用前景。目前在其肿瘤领域已有很多优秀的工作,如人们利用单细胞测序技术深入了解肺癌细胞和结直肠癌微环境特征^[37-38]、乳腺癌细胞微环境的免疫细胞表型^[39]、头颈部鳞状细胞癌^[40]、肾癌细胞特征^[41]及卵巢癌^[42]等。此外,单细胞测序技术在其他疾病上的应用也体现出巨大的价值,如对脑细胞图谱的绘制帮助人们理解神经系统疾病的发生原因以及奥尔默兹海默症相关的疾病线索^[43-44],高分辨率小肠上皮细胞表达图谱的绘制为肠道如何防止病原体入侵提供了线索^[45]以及通过分析细胞组成和基因表达动态变化剖析了糖尿病的发生与发展^[46]。但是目前单细胞测序技术还有待进一步提升,如单细胞悬液的制备。由于某些样本十分稀少且难以获得,细胞悬液制备过程中操作复杂、条件苛刻,稍有不慎会造成样本的细胞死亡率过高,或被污染致使在后期数据处理过程中达不到质控要求,造成数据失真或损失。再者由于单个细胞的 DNA 含量极少,需要扩增至可测序的数量,尽管人们在各个阶段都采取措施(如采用高保真酶、特殊引物、数据处理标准化等),但全基因组扩增产生的扩增偏倚仍难以避免,且较为微弱的拷贝

数变异仍然可能在背景噪音中掩盖。这就不单单需要单细胞测序技术的创新发展,也需要辅助测序手段,如捕获、扩增技术的共同进步。另外,现有的单细胞测序技术是针对含有 poly-A 尾的 mRNA,对于其他可能有重要功能却没有该特点的 microRNA, lncRNA 等尚未建立完善的测序手段,这可能是单细胞测序技术未来需要克服的困难之一。

现有文献表明肝脏是个庞杂的细胞群体,多种肝疾病(如肝硬化、肝癌等)发生发展也是细胞与细胞、细胞与微环境共同作用的结果。虽然已有相关工作^[28, 34]表明肝脏中的免疫细胞之间联系和功能复杂,但更有意义的是进一步描绘其在疾病发展全过程的动态变化。利用单细胞测序技术对肝脏中的免疫细胞进行详细的分群也能帮人们进一步了解肝脏疾病的发病机制。当下成熟的单细胞测序在数据处理上多用 t-SNE 降维方法将细胞群投影到二维空间中,但实际上这会导致细胞的某个方向上的空间信息的丢失,因此将细胞一一映射到三维空间是单细胞测序未来需要完善的。保存完整的空间信息对分析很多器官细胞的动态变化至关重要,近期 Yu 等^[47]通过分析来自 13 位中国人的神经胶质瘤表达图谱,生成了胶质瘤的时空景观,揭示了胶质瘤不同子区域之间的侵袭模式,这是单细胞测序在肿瘤空间信息方面的又一次探索,虽然已有部分科研人员对肝脏细胞的空间信息作出了初步尝试^[48-49],但仍然不够深入,也许对肝内多种细胞群的空间解析会带来更加有趣的结果。

参考文献

- [1] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors [J]. *Proceedings of the National Academy of Sciences*, 1977, 74 (12): 5463-5467.
- [2] Collins FS, Morgan M, Patrinos A. The human genome project: lessons from large-scale biology [J]. *Science*, 2003, 300 (5617): 286-290.
- [3] Tang F, Barbacioru C, Wang Y, et al. mRNA-Seq whole-transcriptome analysis of a single cell [J]. *Nature Methods*, 2009, 6 (5): 377-382.
- [4] Lee JW, Stone ML, Porrett PM, et al. Hepatocytes direct the formation of a pro-metastatic niche in the liver [J]. *Nature*, 2019,

- 567 (7747) : 249-252.
- [5] Duncan AW, Dorrell C, Grompe M. Stem cells and liver regeneration [J] . *Gastroenterology*, 2009, 137 (2) : 466-481.
- [6] Adam M, Potter AS, Potter SS. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts : a molecular atlas of kidney development [J] . *Development*, 2017, 144 (19) : 3625-3632.
- [7] Lafzi A, Moutinho C, Picelli S, et al. Tutorial : guidelines for the experimental design of single-cell RNA sequencing studies [J] . *Nature Protocols*, 2018, 13 (12) : 2742-2757.
- [8] Rinke C, Lee J, Nath N, et al. Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics [J] . *Nature Protocols*, 2014, 9 (5) : 1038-1048.
- [9] Brouzes E, Medkova M, Savenelli N, et al. Droplet microfluidic technology for single-cell high-throughput screening [J] . *Proceedings of the National Academy of Sciences*, 2009, 106 (34) : 14195-14200.
- [10] Picelli S, Björklund ÅK, Faridani OR, et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells [J] . *Nature Methods*, 2013, 10 (11) : 1096-1098.
- [11] Gawad C, Koh W, Quake SR. Single-cell genome sequencing : current state of the science [J] . *Nature Reviews Genetics*, 2016, 17 (3) : 175-188.
- [12] Huang L, Ma F, Chapman A, et al. Single-cell whole-genome amplification and sequencing : methodology and applications [J] . *Annual Review of Genomics and Human Genetics*, 2015, 16 : 79-102.
- [13] Navin NE. Cancer genomics : one cell at a time [J] . *Genome Biology*, 2014, 15 (8) : 452-465.
- [14] Lasken RS. Single-cell sequencing in its prime [J] . *Nature Biotechnology*, 2013, 31 (3) : 211-212.
- [15] Lasken RS. Single-cell genomic sequencing using multiple displacement amplification [J] . *Current Opinion in Microbiology*, 2007, 10 (5) : 510-516.
- [16] Macaulay IC, Haerty W, Kumar P, et al. G&T-seq : parallel sequencing of single-cell genomes and transcriptomes [J] . *Nature Methods*, 2015, 12 (6) : 519-522.
- [17] Cui P, Lin Q, Ding F, et al. A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing [J] . *Genomics*, 2010, 96 (5) : 259-265.
- [18] Ziegenhain C, Vieth B, Parekh S, et al. Comparative analysis of single-cell RNA sequencing methods [J] . *Molecular Cell*, 2017, 65 (4) : 631-643.
- [19] Lorthongpanich C, Cheow LF, Balu S, et al. Single-cell DNA-methylation analysis reveals epigenetic chimerism in preimplantation embryos [J] . *Science*, 2013, 341 (6150) : 1110-1112.
- [20] Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data [J] . *Cell*, 2019, 177 (7) : 1888-1902.
- [21] Dey SS, Kester L, Spanjaard B, et al. Integrated genome and transcriptome sequencing of the same cell [J] . *Nature Biotechnology*, 2015, 33 (3) : 285-289.
- [22] Hou Y, Guo H, Cao C, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas [J] . *Cell Research*, 2016, 26 (3) : 304-319.
- [23] Bian S, Hou Y, Zhou X, et al. Single-cell multiomics sequencing and analyses of human colorectal cancer [J] . *Science*, 2018, 362 (6418) : 1060-1063.
- [24] Van Der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high-confidence variant calls : the Genome Analysis Toolkit best practices pipeline [J] . *Current Protocols in Bioinformatics*, 2013, 43 (1110) : 11.10.1-11.10.33.
- [25] Ding B, Zheng L, Zhu Y, et al. Normalization and noise reduction for single cell RNA-seq experiments [J] . *Bioinformatics*, 2015, 31 (13) : 2225-2227.
- [26] Kobak D, Berens P. The art of using t-SNE for single-cell transcriptomics [J] . *Nature Communications*, 2019, 10 (1) : 5416.
- [27] Aizarani N, Saviano A, Sagar, et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors [J] . *Nature*, 2019, 572 (7768) : 199-204.
- [28] MacParland SA, Liu JC, Ma XZ, et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations [J] . *Nature Communications*, 2018, 9 (1) : 4383.
- [29] Popescu DM, Botting RA, Stephenson E, et al. Decoding human fetal liver haematopoiesis [J] . *Nature*, 2019, 574 (7778) : 365-371.
- [30] Ramachandran P, Dobie R, Wilson-Kanamori J, et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level [J] . *Nature*, 2019, 575 (7783) : 512-518.

- [31] Krenkel O, Hundertmark J, Ritz TP, et al. Single cell RNA sequencing identifies subsets of hepatic stellate cells and myofibroblasts in liver fibrosis [J] . *Cells*, 2019, 8 (5) : 503-513.
- [32] Dobie R, Wilson-Kanamori JR, Henderson BEP, et al. Single-cell transcriptomics uncovers zonation of function in the mesenchyme during liver fibrosis [J] . *Cell Reports*, 2019, 29 (7) : 1832-1847.
- [33] Fu J, Wang H. Precision diagnosis and treatment of liver cancer in China [J] . *Cancer Letters*, 2018, 412 : 283-288.
- [34] Zhang Q, He Y, Luo N, et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma [J] . *Cell*, 2019, 179(4) : 829-845.
- [35] Xue R, Chen L, Zhang C, et al. Genomic and transcriptomic profiling of combined hepatocellular and intrahepatic cholangiocarcinoma reveals distinct molecular subtypes [J] . *Cancer Cell*, 2019, 35 (6) : 932-947.
- [36] Zheng H, Pomyen Y, Hernandez MO, et al. Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma [J] . *Hepatology*, 2018, 68 (1) : 127-140.
- [37] Guo X, Zhang Y, Zheng L, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing [J] . *Nature Medicine*, 2018, 24 (7) : 978-985.
- [38] Zhang L, Yu X, Zheng L, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer [J] . *Nature*, 2018, 564 (7735) : 268-272.
- [39] Azizi E, Carr AJ, Plitas G, et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment [J] . *Cell*, 2018, 174 (5) : 1293-1308.
- [40] Puram SV, Tirosh I, Parikh AS, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer [J] . *Cell*, 2017, 171 (7) : 1611-1624.
- [41] Young MD, Mitchell TJ, Braga FAV, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors [J] . *Science*, 2018, 361 (6402) : 594-599.
- [42] Hu Z, Artibani M, Alsaadi A, et al. The repertoire of serous ovarian cancer non-genetic heterogeneity revealed by single-cell sequencing of normal fallopian tube epithelial cells [J] . *Cancer Cell*, 2020, 37 (2) : 226-242.
- [43] Zhong S, Zhang S, Fan X, et al. A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex [J] . *Nature*, 2018, 555 (7697) : 524-528.
- [44] Mathys H, Davila-Velderrain J, Peng Z, et al. Author correction : Single-cell transcriptomic analysis of Alzheimer's disease [J] . *Nature*, 2019, 571 (7763) : E1-E1.
- [45] Haber AL, Biton M, Rogel N, et al. A single-cell survey of the small intestinal epithelium [J] . *Nature*, 2017, 551 (7680) : 333-339.
- [46] Veres A, Faust AL, Bushnell HL, et al. Charting cellular identity during human *in vitro* β -cell differentiation [J] . *Nature*, 2019, 569 (7756) : 368-373.
- [47] Yu K, Hu YQ, Wu F, et al. Surveying brain tumor heterogeneity by single-cell RNA sequencing of multi-sector biopsies [J] . *National Science Review*, 2020, 7 (8) : 1306-1318.
- [48] Halpern KB, Shenhav R, Massalha H, et al. Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells [J] . *Nature Biotechnology*, 2018, 36 (10) : 962-970.
- [49] Halpern KB, Shenhav R, Matcovitch-Natan O, et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver [J] . *Nature*, 2017, 542 (7641) : 352-356.

(责任编辑 朱琳峰)