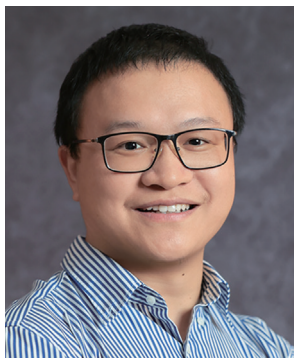


实验室介绍



王开乐博士, 2024年起任中国科学院分子细胞科学卓越创新中心/生物化学与细胞生物研究所研究员、博士生导师, 研究组长。该实验室主要通过开发和应用高通量单细胞DNA测序、单细胞多组学测序及空间多组学测序等前沿技术, 融合实验与计算生物学手段, 一方面致力于解析肿瘤细胞的起源与演化过程, 揭示驱动癌症发生、发展与转移的关键基因组事件, 进而为肿瘤的预防和精准治疗提供新的潜在靶点; 另一方面致力于将细胞数字化, 探究细胞行为与命运决定机制。

<https://wanglab.sibcb.ac.cn>

单细胞DNA测序及研究进展

熊惠¹ 王开乐^{1,2*}

(¹核糖核酸功能与应用全国重点实验室, 中国科学院分子细胞科学卓越创新中心/生物化学与细胞生物研究所, 上海 200031; ²中国科学院大学, 北京 101408)

摘要 高通量DNA测序极大推动了碱基突变、基因拷贝数和结构变异等遗传变异事件的发现和检测。传统的高通量DNA测序(bulk DNA sequencing)所需的DNA来源于大量细胞, 能检测到细胞群体共有的基因组变异事件, 但会掩盖群体内单个细胞间的差异。为了精准检测每个细胞间的差异, 单细胞DNA测序(single cell DNA sequencing, scDNA-seq)技术被开发出来, 极大推动了肿瘤基因组学、遗传学及疾病发生发展等学科领域的进展。该文简要概括了scDNA-seq技术的发展历程、面临的挑战、最新的进展及应用, 并展望了scDNA-seq的未来发展方向。

关键词 单细胞DNA测序; 单细胞DNA测序应用; 单细胞DNA测序挑战; 单细胞DNA测序未来发展方向; 单细胞测序方法

Advances in Single Cell DNA Sequencing

XIONG Hui¹, WANG Kaile^{1,2*}

(¹State Key Laboratory of RNA Innovation, Science and Engineering, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Shanghai 200031, China;

²University of Chinese Academy of Sciences, Beijing 101408, China)

收稿日期: 2025-06-17

接受日期: 2025-07-25

中国科学院战略重点研究项目(批准号: XDB0570000)和上海市浦江人才计划(批准号: 24PJA142)资助的课题

*通信作者。Tel: 021-34273040, E-mail: kailewang@sibcb.ac.cn

Received: June 17, 2025

Accepted: July 25, 2025

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Science (Grant No.XDB0570000), and Shanghai Pujiang Talent Plan (Grant No.24PJA142)

*Corresponding author. Tel: +86-21-34273040, E-mail: kailewang@sibcb.ac.cn

Abstract Bulk DNA sequencing has significantly advanced the discovery and detection of genetic variations, including point mutations, copy number alterations, and structural variants. However, because it relies on DNA extracted from large populations of cells, bulk sequencing primarily captures high-frequency events and fails to resolve rare variations present in small subpopulations and single cells. To address this limitation and enable precise characterization of cellular heterogeneity, many innovative scDNA-seq (single-cell DNA sequencing) technologies have been developed. These methods have profoundly enhanced our understanding of tumor genomics, human genetics, and the underlying mechanisms of disease initiation and progression, such as cancer. This review provides an overview of the recent advances, current challenges and applications, also discusses future directions of scDNA-seq technologies.

Keywords scDNA-seq; applications of scDNA-seq; challenges of scDNA-seq; further directions of scDNA-seq; single-cell sequencing method

细胞是生命活动的基本单位,其DNA序列携带着主要的遗传信息,具有可遗传性。DNA决定着一切生命活动特征,其序列编码的蛋白质赋予了细胞功能^[1-2]。细胞在分裂时,其基因组往往会由于细胞内外的不同因素造成多种形式的变异。这些变异往往是随机的,但当变异为有义变异时,则可能会对个体产生不同程度的影响,甚至造成疾病的发生^[3]。因此,如何有效解析细胞内基因组的特征及其变化是生物学和医学界的一大挑战。细胞中基因组的变异形式主要包括点突变、小的插入和缺失、拷贝数变异和基因组重排等。为了更深入地探究基因组的变异及其对细胞和疾病产生的影响,DNA测序技术被开发出来并逐渐成为研究基因组的必备工具。借助高通量DNA测序技术,人类基因组计划成功完成了首次对人类染色体全序列的解析,构建了完整的人类基因组图谱,极大地推动了人类对自身遗传信息的认知以及医学领域的进步^[4]。针对大量细胞的传统的高通量DNA测序揭示了大量细胞的基因组特征,包括点突变和拷贝数变异等^[5]。然而这些特征表征的是这一细胞群体的平均水平,来自特征明显的细胞信息往往会掩盖或抹去某些特征不明显或少量细胞群的基因组变化,从而造成观察结果的偏好性。同时,由于高通量测序本身的错误率(如Illumina的测序错误率约为0.1%),传统的高通量DNA测序技术很难精确检测突变频率小于1%的位点^[6-8]。为了克服传统的高通量DNA测序的缺陷,scDNA-seq技术被设计开发出来。scDNA-seq技术可以捕获到单个细胞间的基因组变异,同时可以用来检测一些罕见的克隆群体,如对肿瘤发生起关键作用的肿瘤起始克隆群体和耐药性克隆群体等。scDNA-seq技术在产前诊断、体细胞突变检测、DNA复制、发育和生殖生物学等领域也正

在被广泛应用^[9]。

人类正常细胞为二倍体,其染色体上的DNA分子在正常状态下由核小体紧密缠绕,并包含大量基因组复杂区域,如端粒、着丝粒等高度重复序列^[1]。在进行单细胞DNA测序时,首先需要去除缠绕在DNA链上的核小体,以保证后续对单个细胞的完整DNA序列进行无偏扩增;其次,正常人的单细胞仅有2套模板序列,极低的分子起始量对单细胞扩增带来了极大的挑战。再次,由于DNA分子极长,在构建文库时需要进一步将文库片段化,增加了反应步骤,也增加了添加细胞标签的难度。相比之下,进行单细胞RNA测序时,仅需要裂解细胞膜,且RNA分子的拷贝多,序列短,可在反转录时一步完成细胞标签添加。因此,相较于单细胞RNA测序(single cell RNA sequencing, scRNA-seq),scDNA-seq在技术上更为复杂,这也是其发展相对缓慢的主要原因。为了实现单细胞DNA测序,过去几十年中,研究人员相继开发出多种不同的scDNA-seq方法(表1)。

scDNA-seq技术大致可分为四个步骤:单细胞(核)解离及分离、单细胞(核)裂解、基因组序列扩增、扩增序列的检测及分析(图1)。每一个步骤都凝聚了前人大量的探索和尝试。本文将围绕这四个步骤所涉及的不同技术方案、优缺点以及面临的挑战等逐一展开讨论,并对scDNA-seq技术开发面临的挑战和未来发展方向等方面进行展望。

1 细胞(核)解离及分离

制备scDNA-seq样本的首要任务是将复杂的固体组织样本解离成单细胞或者单细胞核悬液(图1)。由于组织样本处理的方法不同,单细胞(核)解离的方法也各有不同。首先是新鲜固体组织的单细胞

表1 单细胞的DNA测序技术总结

Table 1 The summary of single cell DNA sequencing technologies

方法 Method	年份 Year	作者 Author	原理 Working principle	反应室 Reaction unit	细胞通量 Throughput	参考文献 References
DOP-PCR	1992, 1999	TELENIUS H, WELLS D	Employ degenerate primers to amplify the whole genome by PCR	Tube/plate	Low	[10-11]
PEP-PCR	1992	ZHANG L	Employ multiple rounds of extension with 15-base random primers to amplify the whole genome by PCR	Tube/plate	Low	[12]
MDA	2002	DEAN F	Use random primers and strand-displacing Phi29 DNA polymerase to perform isothermal WGA (whole-genome amplification)	Tube/plate	Low	[13]
MALBAC	2012	ZONG C	Perform WGA using random primers with a common sequence for multiple rounds of quasi-linear pre-amplification, where newly synthesized short DNA fragment loops to prevent amplification	Tube/plate	Low	[14]
Strand-seq	2012	FALCONER E	Employ BrdU-labeled new synthesized strands during DNA replication, followed by selective degradation of the newly synthesized strand during amplification to distinguish different DNA templates	Tube/plate	Low	[15]
MIDAS	2013	GOLE J	Employ microwell arrays to miniaturize MDA-based WGA reaction volume to the nanoliter volumes	Microwell	Low	[16]
TruePrime WGA	2015	4basebio	Use Tth PrimPol and Phi29 DNA polymerase for primer-free MDA-based WGA	Tube/plate	Low	[17]
eMDA	2015	FU Y	Perform MDA-based WGA in microdroplets	Microdroplet	Low	[18]
ddMDA	2016	RHEE M	Perform MDA-based WGA in microdroplets	Microdroplet	Low	[19]
Droplet MDA	2016	KASTON L	Perform MDA-based WGA in microdroplets	Microdroplet	Low	[20]
STT-seq	2016	WANG K	Use Tn5 transposase to insert T7 promoters into the genome, followed by T7-driven linear amplification for WGA	Tube/plate	Low	[21]
Tapestri	2017	Mission Bio	Combine microfluidics with targeted amplicon amplification for targeted scDNA-seq	Microdroplet	High	[22]
iDOP-PCR	2017	BLAGO-DATSKIKH K	Use SD polymerase with strand-displacement activity and optimized primer design to improve DOP-PCR based WGA	Tube/plate	Low	[23]
DOPlify WGA	2017	DELEYE L	Use high-fidelity enzyme to improve the accuracy of DOP-PCR based WGA	Tube/plate	Low	[24]
sdMDA	2017	HOSOKAWA M	Perform MDA-based WGA in microdroplets	Microdroplet	Low	[25]
LIANTI	2017	CHEN C	Use Tn5 to insert looped T7 promoters, followed by T7-driven linear amplification for WGA	Tube/plate	Low	[26]
DLP	2017	ZAHN H	Use microfluidics chips combined with Tn5 tagmentation for WGA	Microwell	Low	[27]
SCI-seq	2017	VITAK S	Combine tagmentation with two rounds of combinatorial indexing for scDNA-seq	Plate	High	[28]
ScRepli-seq	2018	DILEEP V	Sort S-phase cells and perform DOP-PCR based WGA	Plate	Low	[29]
10× Genomics	2018	BELHOCINE Z	Combine microfluidics and MDA for WGA	Microdroplet	High	[30]
DLP+	2019	LAKS E	Combine CellenOne single cell dispensing system with Tn5 tagmentation for scDNA-seq	Nanowell	High	[31]
MiCA-eMDA	2019	FU Y	Perform MDA-based WGA in microdroplets	Microdroplet	Low	[32]
sci-L3-WGS	2019	YIN Y	Use three rounds of combinatorial indexing for high-throughput scDNA-seq	Plate	High	[33]
ACT	2021	MINUSSI D	Combine acoustic liquid transfer with Tn5 tagmentation for scDNA-seq in nanoliter volumes	Tube/plate	Low	[34]

续表1

方法 Method	年份 Year	作者 Author	原理 Working principle	反应室 Reaction unit	细胞通量 Throughput	参考文献 References
PTA	2021	GONZALEZ-PEVA V	Spike in ddNTPs during MDA reaction for WGA	Tube/plate	Low	[35]
PiPolB MDA	2023	ORDÓÑEZ C	Use piPolB to replace Phi29 DNA polymerase for primer-independent MDA-based WGA	Tube/plate	Low	[36]
PiMDA	2023	ORDÓÑEZ C	Combine piPolB and Phi29 DNA polymerase for primer-independent MDA-based WGA	Tube/plate	Low	[36]
dd-scCNV Seq	2023	YU X	Combine microfluidics with Tn5 tagmentation for WGA	Microwell	Low	[37]
Arc-well	2023	WANG K	Use nanowell chips with Tn5 tagmentation for WGA	Nanowell	High	[38]
msCNVs	2024	LIN G	Use Tn5 transposase with distinct adapters for WGA	Tube/plate	Low	[39]

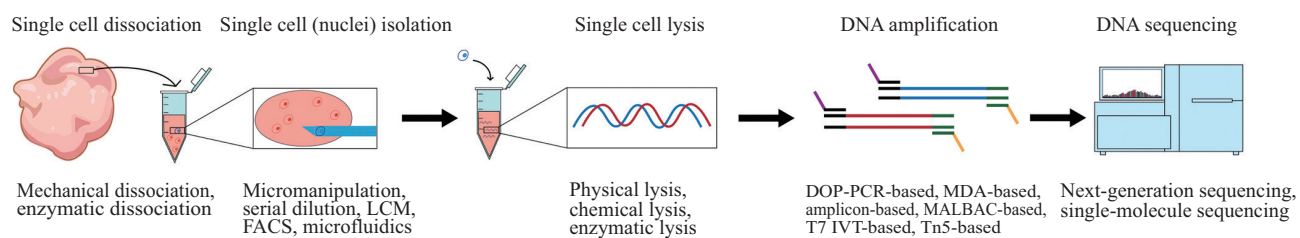


图1 scDNA-seq技术流程

Fig.1 The diagram of single cell DNA sequencing procedure

解离法,早在20世纪研究者们就利用机械研磨和酶解的方式,将新鲜的固体组织解离为单个细胞。机械研磨是指利用物理作用力将新鲜固体组织进行研磨,以拿到单细胞悬液,但机械解离法相对剧烈,容易对细胞的完整性和生物活性造成影响;酶类解离则通过利用不同特性酶的组来解离不同的组织,比如胰蛋白酶可以水解蛋白质,胶原蛋白酶可以针对性地水解结缔组织,从而使组织松散,解离成单细胞。但酶类解离法制备单细胞需要控制好酶解的时间、温度和浓度,若条件过度会使单细胞的形态和活性受到损伤^[40-41]。其次是经冰冻包埋处理后的固体组织样本的单细胞核悬液制备。这类样本由于长时间的低温储存,解冻后由于细胞内冰晶的存在,往往不能拿到完整的细胞。但是由于DNA主要存在于细胞核内,因此可以采用去污剂(detergent)的方式分离细胞核以达到对冰冻组织进行解离的目的。最后是福尔马林固定石蜡包埋(FFPE)样本的解离。由于石蜡样本在固定和包埋过程中对整个组织进行了固定、脱水和浸蜡处理,因此在单细胞(核)制备前,需要对石蜡样本进行脱蜡、复水等处理,再进行后续的机械研磨或酶解处理,以获得单细胞(核)悬液^[38]。

完成单细胞(核)悬液的制备后,需要再进一步将单个细胞(核)分离出来进行DNA扩增和测序。早期单细胞的获得是通过在显微镜下用毛细血管手动挑选或者梯度稀释从而得到单个细胞^[41]。为了拿到特定类型的细胞,激光捕获显微切割技术(laser capture microdissection, LCM)^[42]也是一个被广泛使用的技术。该技术利用激光将细胞从组织上的特定部位切割下来,实现了对特定类型的细胞的分离。虽然这两种方法能够获得单细胞,但分选效率较低,只适用于分离较少单细胞的情况。随着流式细胞荧光分选技术的飞速发展,流式分选的方法被广泛用来分离单个细胞(核),这极大地提高了单细胞(核)分离的准确性和效率^[43]。近些年微流控技术由于其精准的可控性和可操作性,微流控技术也被广泛用于单细胞(核)的捕获和分离^[44]。

综上所述,近年来针对单细胞(核)解离和单细胞(核)分离技术,研究人员开发出了包括机械分离法、酶消化法以及二者结合的混合法等多种单细胞(核)解离方法。这些方法通过优化解离条件、改进解离试剂配方等手段,有效实现了对组织样本的单细胞(核)进行解离。单细胞(核)分离则通过引入

LCM、荧光流式分选及微流控等技术提高了特定类型单细胞(核)的获得率及效率。这些方法体系的建立和完善为不同组织类型(如新鲜、冰冻、石蜡等)的单细胞研究提供了多样化的技术选择方案,极大地推动了单细胞组学技术的发展。

2 单细胞(核)裂解

细胞(核)被分离出来后,需要对单细胞(核)进行裂解以暴露出单细胞内的DNA分子。根据下游实验目的的不同,可采用不同的细胞裂解技术,从而保证从细胞中获得尽可能完整的基因组的同时,又不引入由裂解方法导致的DNA突变。单细胞(核)裂解方法按照裂解原理主要分为三大类:物理裂解法、化学裂解法和酶裂解法。

单细胞物理裂解法主要利用机械力、温度变化或电场作用等物理手段破坏细胞膜结构,实现细胞内容物的释放。其中,机械破碎法通过超声处理、高压均质或珠磨等方式产生剪切力直接破坏细胞膜^[45-46];冻融裂解法则是利用反复冻融过程中冰晶形成导致细胞膜机械性破裂^[47];电穿孔法通过施加高压电脉冲在细胞膜上形成纳米级微孔,从而实现温和可控的裂解效果^[48-49]。这些物理方法具有操作简便、无需添加化学试剂、可避免样本污染等优势,但可能存在裂解效果不充分、DNA释放不完全等局限。

单细胞化学裂解法主要利用化学试剂与细胞膜成分的相互作用实现细胞裂解。该方法主要包括表面活性剂法、有机溶剂法和碱裂解法等。表面活性剂(如SDS、Triton X-100、NP-40等)通过破坏细胞膜脂质双分子层的稳定性实现裂解^[50-51];有机溶剂(如甲醇、丙酮等)可溶解膜脂质导致细胞破裂^[52-54];碱裂解法则利用强碱(如KOH等)破坏细胞膜和细胞壁结构^[13]。化学裂解法相对操作简便,但也存在一些问题,比如有些表面活性剂,如Triton X-100、NP-40等非离子型表面活性剂不能有效去除核小体结构导致裂解不充分;强碱会导致DNA双链变成单链等。近年来,研究者通过优化试剂浓度、作用时间和温度等参数,开发出了多种温和型化学裂解方案,以更好地释放单细胞DNA^[55-56]。

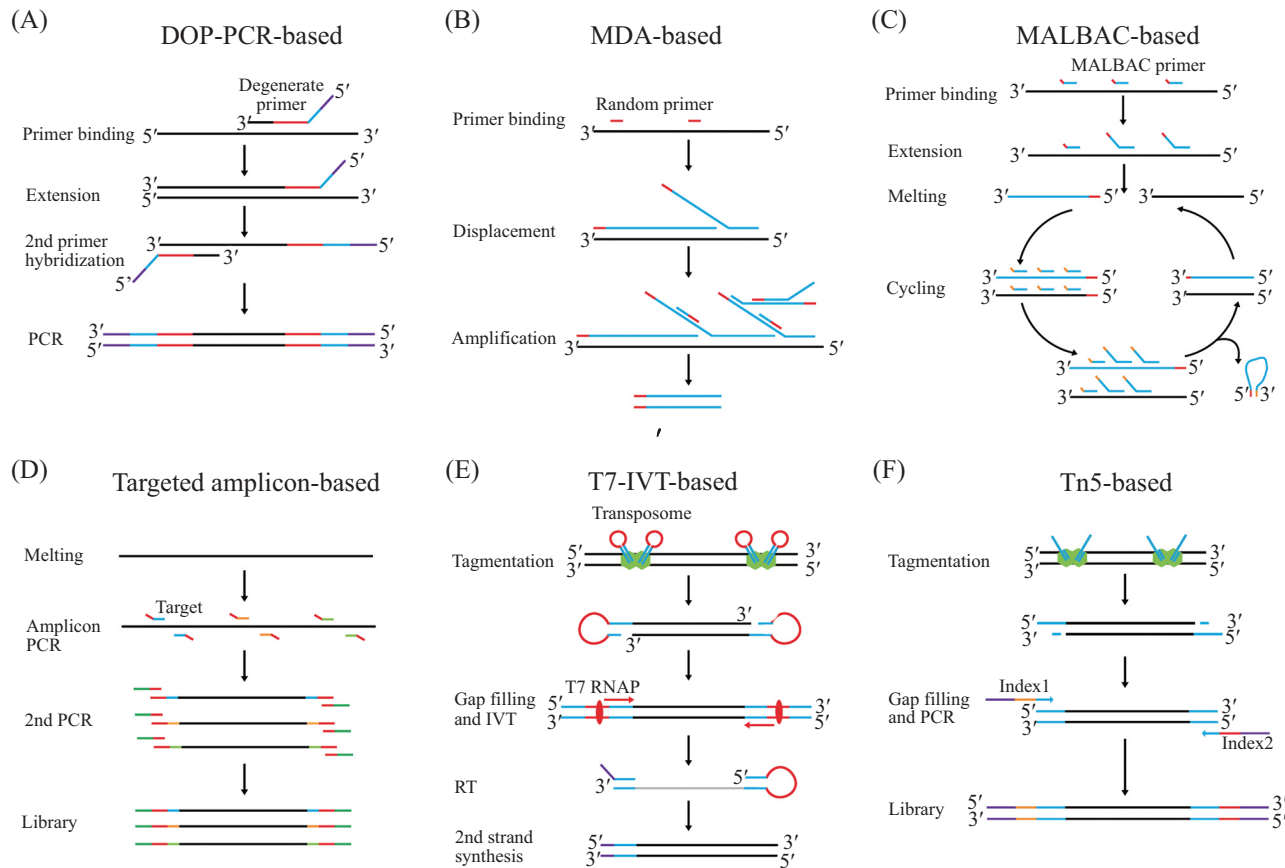
单细胞酶裂解法利用生物酶的特异性催化作用实现细胞膜和细胞壁的温和裂解。该方法主要包括蛋白酶法(如胰蛋白酶、蛋白酶K)和复合酶法等

多种形式。其中,蛋白酶通过特异性降解细胞膜蛋白和核小体蛋白等组分实现裂解,而复合酶法则通过多种酶的协同作用提高裂解效率,比如裂解植物或微生物等单细胞时,需要联合果胶酶、溶菌酶或几丁质酶等^[57]。酶裂解法具有条件温和、特异性强、对DNA损伤小等优势。

总之,单细胞裂解技术主要包含物理裂解法、化学裂解法以及酶裂解法等。其中,物理裂解法具有快速、直接的特点;化学裂解法中的表面活性剂法成本低,但存在裂解不充分的问题。碱裂解法和酶解法因其高效性和对细胞内容物的低损伤特性,在scDNA-seq中被广泛应用。除了单一裂解法外,组合裂解法也被应用。比如通过组合表面活性剂法和酶解法,利用表面活性剂破坏细胞膜结构,利用蛋白酶消化掉细胞核和核小体等蛋白,可在保持基因组完整性的同时实现温和、彻底的单细胞DNA释放。该组合方法已成为单细胞基因组学研究中的首选方法^[14,26,31,34,38,58]。随着单细胞多组学技术的快速发展,裂解方法的选择需要综合考虑样本类型、目标分子特性以及下游分析需求等因素,以达到理想的裂解效果。

3 单细胞DNA扩增

哺乳动物二倍体基因组DNA仅有两个拷贝,在进行测序前需要将基因组扩增到足够的模板量才能进行测序,因此怎么才能把极少量的基因组DNA进行扩增从而用于测序是scDNA-seq的核心难题。1986年,MULLIS等^[59]发明了具有时代意义的聚合酶链式反应(polymerase chain reaction, PCR),只需要在反应体系中提供合适的反应条件,就可以在体外进行核酸扩增。由于单细胞基因组模板量少和长度过长的原因,PCR难以对其进行有效扩增,因此后来的研究者们基于PCR技术进行了改良,开发出了多种适配DNA测序的扩增方法,按照DNA扩增的原理不同大致可分为六类,包含基于DOP-PCR(degenerate oligonucleotide-primed polymerase chain reaction)^[10]、基于MDA(multiple displacement amplification)^[13]、基于MALBAC(multiple annealing and looping-based amplification cycles)^[14]、基于特异序列的扩增子(targeted amplicon)扩增^[22]以及基于T7体外转录^[21,26]和基于Tn5转座酶^[26-27,34,38]的方法(图2)。



IVT: 体外转录; RT: 反转录; RNAP: RNA聚合酶。基于DOP-PCR(A)、MDA(B)、MALBAC(C)、特异序列的扩增子(D)、T7体外转录(E)、Tn5(F)原理的scDNA-seq技术。

IVT: *in vitro* transcription; RT: reverse transcription; RNAP: RNA polymerase. DOP-PCR-based (A), MDA-based (B), MALBAC-based (C), targeted amplicon-based (D), T7-IVT-based (E) and Tn5-based (F) based scDNA-seq technologies.

图2 基于六大不同原理的单细胞DNA扩增技术示意图

Fig.2 Principles of the six major classes of single cell DNA amplification technologies

3.1 基于DOP-PCR的全基因组扩增

DOP-PCR^[10]是一种基于部分简并碱基随机引物的全基因组扩增方法(图2A)。该方法所使用的引物是经过优化设计的部分简并随机引物,其降低了PEP-PCR(primer extension preamplification PCR)中由全随机引物扩增导致的扩增产物的不均一性。简并随机引物在低温退火温度中结合到基因组DNA的不同位置,然后升温进行延伸,第一轮延伸后变性,再进行第二轮延伸,第二轮延伸时形成了两端均带接头分子的片段,随后加入跟接头分子匹配的引物,进行PCR扩增,再后续构建高通量测序文库,进行单细胞测序。

DOP-PCR的方法在scDNA-seq早期尤其是低通量scDNA-seq中被广泛应用(如检测拷贝数变异、点突变等)。为了进一步改善DOP-PCR,后续又相继开发了多种scDNA-seq方法,包含: iDOP-PCR(improved

DOP-PCR)^[23]、DOPlify WGA^[24]和scRepli-seq(single cell replication timing sequencing)^[60]等,并且通过使用具有强链置换活性的新型热稳定DNA聚合酶(SD polymerase)和调整引物设计改进了扩增DNA文库的质量^[24]。但总体来讲,DOP-PCR的扩增产物存在扩增偏好性和扩增错误率相对较高、覆盖度相对较低的问题。

3.2 基于MDA的全基因组扩增

在2002年,DEAN等^[13]开发出了基于多重链置换的单细胞DNA扩增方法MDA(图2B),这种方法主要是利用具有链置换活性的DNA聚合酶(如Phi29 DNA聚合酶、Bst DNA聚合酶等)对单链DNA分子进行恒温扩增。由于DNA聚合酶的链置换活性,在扩增过程中会形成大量DNA序列的分支,这些DNA分支序列进一步作为新的随机引物结合的模板进行扩增,从而在恒温下短时间内即可扩增出大量的基因组

DNA片段。Phi29 DNA聚合酶由于其延伸性强、错误率低和链置换活性高的特性被广泛用于MDA反应。采用Phi29 DNA聚合酶的MDA方法扩增的基因组能覆盖到甚至超过90%的单细胞的全基因组。同时,由于该酶的极低错误率,MDA的方法也拥有极低的扩增错误率,适用于点突变的检测^[60]。随着MDA方法的报道,衍生出了一系列基于MDA原理的商业化试剂。例如: GenomiPhi V2 DNA amplification kit和Qiagen发布的REPLI-g single cell kit等,都是基于MDA技术实现了从单细胞或微量样本中进行高效的全基因组扩增,为单细胞基因组学研究提供了有力的工具。但MDA方法也并非完美,由于Phi29 DNA聚合酶对不同碱基组成序列复制的速度不同,造成了模板的不均一扩增。为了进一步解决MDA扩增过程中复制不均匀的问题,又开发出了不依赖引物的MDA(TruePrime WGA^[17]、piPolB MDA^[36]、piMDA^[36]);基于油包水的MDA[(dMDA(digital MDA)^[61]、MIDAS(microwell displacement amplification system)^[16]、eMDA(emulsion MDA)^[18]、ddMDA(digital droplet MDA)^[19]、droplet MDA^[20]、sd-MDA(single droplet multiple displacement amplification)^[25]和MiCA-eMDA^[32]]等。其中无引物的MDA法,如TruePrime利用TthPrimPol聚合酶先合成一段引物,然后再用Phi29 DNA聚合酶完成后续的链置换扩增,该方法减少了由于随机引物带来的假阳性扩增。为了降低MDA扩增方法的偏好性,2013年GOLE等^[16]通过将MDA的反应体系从微升水平降低至纳升水平开发了MIDAS,该方法极大降低了MDA扩增的倍数,从而显著改善了MDA扩增的偏好性。后续随着微流控技术的发展,多种结合油包水和MDA扩增的方法被开发出来,这些方法通过将MDA扩增反应包进皮升或纳升大小的液滴里面,进一步降低了MDA扩增的倍数,改善了MDA扩增的偏好性。10× Genomics基于油包水的原理开发出了高通量单细胞全基因组测序平台Chromium Single Cell CNV^[30]。除了通过减小MDA反应体积的方式来限制MDA扩增的偏好性外,2021年GAWAD团队^[35]报道了PTA-seq(primary template-directed amplification),该技术通过在MDA反应液内添加ddNTP的方式,极大缩短了扩增子的长度,从而提高了每个模板被下一轮随机引物结合的概率,进而降低了扩增偏好性。基于MDA的全基因组扩增由于使用随机六聚体引物和

具有强置换活性与高保真性的Phi29聚合酶,MDA的方法大大增加了单细胞扩增的覆盖度,并极大降低了扩增的错误率,使其被广泛用于单细胞点突变的检测。

3.3 基于MALBAC的全基因组扩增

第三类是在2012年由谢晓亮团队^[14]报道的MALBAC,该方法是一种拟线性扩增方法。其原理是使用带有相同接头的半随机引物(5'端为固定序列,3'端为随机序列)在全基因组多个位置退火,并通过Bst DNA聚合酶进行延伸后形成一端带有接头的长扩增子(图2C)。后续经过解链后,这些一端带有接头的长扩增子和原本的模板基因组DNA均作为下一轮引物结合的模板链。第二轮循环时,带有固定接头的半随机引物再与上述长扩增子结合,扩增出带有两端互补接头的DNA片段。短的这些片段由于会自我形成发卡结构,在后续扩增中不易被再次扩增。而长扩增子不能有效自我形成发夹结构而会被继续扩增,进而实现相对均匀的全基因组扩增。该方法有效降低了扩增的偏好性。但是由于Bst DNA聚合酶的复制错误率相对较高,因此该方法更适合于检测拷贝数变异。

3.4 基于特异性扩增子的DNA扩增

基于特异性扩增子的DNA扩增是通过设计跟目标片段匹配的特异的引物进行PCR反应来实现对目的片段特异性扩增(图2D)。2017年,由Mission Bio公司推出的单细胞DNA特异性扩增子测序平台(Tapestry),它将微流控技术与基于特异性扩增子的DNA扩增技术结合,实现了高通量单细胞DNA特异性扩增子测序^[22]。该平台通过第一次油包水,将细胞和细胞裂解液一起包裹到液滴中进行裂解,再通过第二次油包水的形式,将细胞标签微珠包裹到液滴中实现了对单细胞特异性扩增子的标记,最后再通过PCR的方式将扩增出来含细胞标签的文库进行再次富集后完成文库构建^[62]。由于该方法利用特异性引物序列对目标基因进行靶向扩增,因此主要适用于已知目标位点的DNA点突变检测。为实现染色体层面的拷贝数变异分析,Mission Bio公司针对不同染色体设计了多对染色体特异性扩增引物。通过计算各条染色体上特异扩增产物的覆盖度,进而推断细胞内的染色体拷贝数变化。因此,该方法在一定程度上也可用于估算染色体层面的拷贝数变异。

3.5 基于T7体外转录的全基因组扩增

下一类是基于T7体外转录的DNA扩增技术。该类技术通过利用Tn5转座酶将T7启动子序列插入到基因组中,然后通过体外转录的方式将基因组DNA转化为RNA并进行线性扩增,然后再将RNA进行反转录和二链合成后构建DNA文库(图2E)。2016年,阮珏和王开乐等^[21]开发了STT-seq(single cell tagmentation and *in vitro* transcription)。该方法将T7启动子序列与Tn5转座酶组装成转座子,利用Tn5的转座酶活性将T7启动子序列插入基因组DNA中,通过T7启动子体外转录后得到的RNA可进一步用于测序文库的构建。类似地,2017年由谢晓亮团队^[26]开发的LIANTI(linear amplification via transposon insertion),利用Tn5转座酶结合发卡结构的T7转录启动子序列,随机识别并插入基因组DNA,使DNA片段通过线性扩增得到扩增产物。由于体外转录过程是一个线性扩增过程,且T7体外转录时可以同时对正负链进行转录,因此该类方法不但解决了DNA测序中扩增偏好性的问题,而且极大地降低了单细胞DNA扩增时的错误率,实现了低偏好性以及高保真度的扩增优势,适用于单细胞点突变的检测。

3.6 基于Tn5转座酶的方案

最后一类是基于Tn5转座酶的单细胞DNA扩增技术。这类技术通过利用Tn5转座酶将单细胞DNA片段化后,直接进行PCR,从而制备scDNA-seq文库(图2F)。该方案将单细胞DNA扩增与文库构建集合为一步,省去了额外的单细胞全基因组预扩增步骤。相比其他扩增方法(如基于DOP-PCR、MDA的方法等),该方法不但极大节省了单细胞DNA文库构建的时间,而且也降低了由于全基因组预扩增引入的扩增偏好性,可以用于检测拷贝数变异和点突变等。基于此原理,2016年,ZAHN等^[27]开发了DLP(direct library preparation)的方法。作者通过结合微流控技术和Tn5转座酶技术,实现了同时对几十个单细胞进行全基因组DNA文库构建。但由于DLP本身的通量较低,APARICIO团队^[31]后续又开发了DLP+(direct library preparation plus)。该方法通过结合cellenONE自动化加样系统和纳升小孔芯片,实现了在皮升体积下同时对几千个单细胞进行scDNA-seq。2017年,ADEY团队^[28]开发了SCI-seq(single-cell combinatorial indexed sequencing),该方法利用带有

不同序列标签的Tn5转座酶对单细胞悬液内的DNA进行第一轮标记,然后将来自不同孔的标记好的细胞核进行混合后,再分到其他不同的孔里,然后利用不同的PCR引物进行扩增,即通过组合索引的方式添加两轮标签实现对单细胞的唯一标记,该方案极大地提高了测序通量^[28]。然而该方法由于无法完全去除DNA上结合的组蛋白,导致其获得的数据非常稀疏。2021年,NAVIN团队^[34]报道了另一种scDNA-seq技术ACT-seq(acoustic cell tagmentation),该方法利用超声移液系统将纳升体系的反应液加入到384孔板中。该方法使用Tn5转座酶片段化DNA,并利用PCR两端不同的引物组合进行单细胞标记。由于该方法利用表面活性剂和蛋白酶的组合物彻底去除了结合在DNA上的各种蛋白质,保证了后续扩增反应的高效进行,因此该方法获得了均一的、高质量的单细胞基因组扩增数据。2023年,同样由来自NAVIN团队的王开乐等^[38]开发出了首个兼容福尔马林固定FFPE样本的高通量scDNA-seq技术Arc-well(archival nanowell sequencing),这一技术通过进一步优化基于Tn5转座酶的方法,并同时结合了纳升小孔芯片,既实现了高质量的单细胞DNA扩增,又实现了高通量。同时该方法还特别针对FFPE样本进行了优化,使得第一次可以对存储在医院样本库的、具有详细临床跟踪数据的FFPE样本进行高通量的scDNA-seq。基于Tn5转座酶的方法,由于省去了对单细胞DNA进行预扩增的步骤,只通过一轮PCR就实现了单细胞DNA扩增、测序文库构建和细胞标签添加三个功能,同时由于Tn5转座酶插入的位点是唯一的,可以通过分子断点来区分不同的DNA分子,从而达到单分子检测DNA的拷贝数变异和点突变的目的,因此该技术是目前最简洁高效、扩增偏好性最低的单细胞DNA文库制备方法,被广泛用来检测DNA的拷贝数变异及DNA点突变。但是该方案也并非完美,其所采用的Illumina Tn5转座酶(如:TDE1、ATM等)对文库片段化时会产生50%的分子前后接头相同,而这些分子在扩增时会自我形成发卡结构,无法在测序时被有效检出。因此,在基因组覆盖度水平上,该方案最多覆盖约50%的单细胞基因组(对一套模板而言)。为了克服这个问题,后续可通过改造Tn5的接头序列的组成来获得更高的覆盖度,如MULQUEEN等^[64](2021年)报道的s3-WGS的方法以及XING等^[65](2021年)报道的META-

CS方法里面所提出的克服 Tn5 50%有效率的设计策略。

4 单细胞DNA检测

扩增后的单细胞基因组需要进一步解析其DNA所含有的变异信息。为此,研究者开发了一系列DNA检测方法,包括单细胞实时荧光定量PCR检测(quantitative real-time PCR, qPCR)、比较基因组杂交[CGH(comparative genomic hybridization)和aCGH(array CGH)]和高通量测序等。其中,qPCR是一种基于荧光信号的核酸定量技术,通过设计特异性引物或荧光探针,在PCR扩增过程中实时监测荧光信号强度,从而精确量化目标序列的拷贝数。在2010年TREFF等^[65]通过设计Taqman探针的qPCR方法验证WGA多位点的扩增均一性,以支持后续微阵列分析的准确性以准确筛查单细胞24染色体中的非整倍体。该方法主要是作为一种辅助手段,具有快速、高效和低成本的优势。CGH是一种在全基因组范围内进行拷贝数变异检测的技术,其利用不同荧光染料将待测DNA和参考DNA分别标记后杂交到正常的中期染色体涂片上,根据荧光信号的差异区分待测DNA上的拷贝数变异。aCGH是一种基于探针杂交的比较基因组杂交方法,该方法将寡核苷酸探针有序固定在玻片上,然后将带有不同颜色荧光染料(如Cy3和Cy5)标记的单细胞DNA和参考DNA与上述探针进行杂交,从而检测染色体全基因组范围内的拷贝数变异^[66]。CGH和aCGH技术被广泛应用于检测单细胞扩增产物中的染色体非整倍体变异,如联合DOP-PCR和CGH检测单个卵裂球中的染色体非整倍体(如21三体)^[67],联合DOP-PCR和aCGH检测体外受精(*in vitro* fertilization, IVF)时囊胚阶段的染色体重排情况^[68],检测不同肿瘤细胞系之间的微小染色体拷贝数变异等^[66]。尽管CGH和aCGH在单细胞基因组检测中发挥了重要作用,但是这类基于探针杂交的方法由于其低灵敏度和低分辨率等特点,限制了其广泛应用。

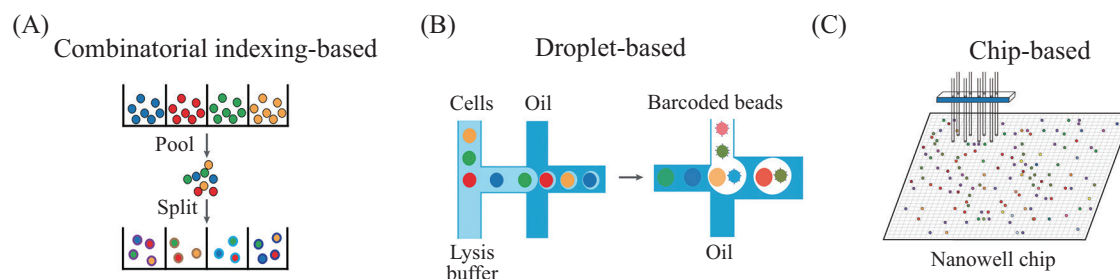
随着二代测序技术的兴起,单细胞扩增产物的检测也逐渐由芯片杂交变为测序^[69-70]。2011年,NAVIN等^[43]首次将基于MDA的单细胞全基因组扩增技术与二代高通量测序技术相结合,成功实现了单细胞水平上的全基因组测序。通过这一方法,作者在乳腺癌样本中揭示了肿瘤的异质性,并提出

了肿瘤的演化模型。这一工作不仅证明了scDNA-seq的可行性,还为肿瘤基因组学、发育生物学和遗传学等领域的研究提供了重要的技术框架。随着测序技术的不断发展,近些年涌现出了多种单细胞DNA扩增和二代高通量测序技术相结合的方法。如MALBAC、Strand-seq、MIDAS、eMDA、ddMDA、iDOP-PCR、DOPlify WGA、sd-MDA、LIANTI、DLP、SCI-seq、ScRepli-seq、DLP+、MICA-eMDA、sci-L3-WGS、PiPolBMDA、PiMDA、PTA、dd-scCNV Seq、Arc-well、msCNVs等单细胞全基因组扩增技术均结合了二代高通量测序平台^[14-16,18-19,23-29,31-32,38-39]。这些单细胞DNA扩增技术与二代测序技术的结合已成为单细胞基因组学研究中的主流,极大地推动了scDNA-seq在各领域中的应用。

除了二代测序外,随着近些年三代单分子测序技术的发展,单细胞DNA的扩增产物开始利用三代测序的方法进行检测,以克服二代测序短读长不能解决的问题。例如:2021年,汤富酬团队^[58]开发了SMOOTH-seq(single-molecule real-time sequencing of long fragments amplified through transposon insertion)测序方法,SMOOTH-seq通过结合基于Tn5的基因组扩增方法和PacBio SMRT三代测序来实现单细胞DNA的测序,该方法通过长读长测序技术实现了高精度的单倍型解析和基因组变异检测。2023年,HÅRD等^[71]报道了基于MDA全基因组扩增方法并结合三代测序技术PacBio,完成了对人类单个T细胞的DNA检测。相信随着三代测序的成本不断降低和测序错误率不断降低,未来将有更多的单细胞DNA扩增方法结合三代测序来实现scDNA-seq。

5 高通量单细胞DNA测序

随着单细胞测序技术的快速发展,通量问题逐渐成为制约scDNA-seq广泛应用的瓶颈。传统的PCR管,96孔板或384孔板方法虽然在一定程度上实现了单细胞的DNA测序,但由于操作繁琐、成本高昂,难以满足大规模研究的需求。为了解决通量问题,近年来涌现出多种创新性的解决方案,显著提升了scDNA-seq的通量和效率。这些创新方案大体可分为基于组合索引、基于微液滴和基于微孔芯片的三种方案(图3)。



A: 基于组合索引的方案; B: 基于两步油包水的方案; C: 基于纳升微孔芯片的方案。

A: combinatorial indexing-based strategy; B: droplet-based strategy; C: nanowell chip-based strategy.

图3 三种高通量scDNA-seq技术的设计策略

Fig.3 Three design strategies for high throughput single cell DNA sequencing

5.1 组合索引

组合索引的方案被广泛用于提高单细胞测序的吞吐量,尤其是在单细胞RNA-seq和单细胞ATAC-seq领域。组合索引的方案首先将单细胞悬液分配到多个不同的小孔内(每个小孔内含多个细胞),然后将各个小孔分别加入不同的标签序列,当每个小孔内所有细胞的所有目标分子标上相同的第一轮小孔标签后,将所有细胞混合,然后再分配到多个不同的小孔内,进行第二轮标签加入,该过程可多次进行,进而无限量增加标记单细胞的吞吐量。基于此原理,2017年ADEY团队^[28]开发了SCI-seq,通过两轮组合索引,实现了对几千个细胞的scDNA-seq。2019年SHENDURE团队^[33]开发了sci-L3-WGS方法则通过3轮组合索引进一步提高了细胞的吞吐量。然而不同于ATAC和RNA,细胞内的DNA被组蛋白层层缠绕。而组合索引实现需要两个前提条件,第一是一个细胞内的所有分子必须在每一轮加标签的时候保持在同一个细胞内,第二是每一轮都需要对目标分子的每个分子添加标签。因此,为了实现对DNA进行组合索引,既需要去除缠绕在DNA上的组蛋白,又不能破坏细胞膜或者核膜的完整性。这种苛刻的、相互矛盾的要求导致了DNA组合索引技术一直无法摆脱数据低质量的困境,限制了该类方法的广泛应用。

5.2 微液滴

第二类是基于微液滴的方法。基于微液滴的方法开创了高通量单细胞RNA测序领域,其在DNA领域也开始被使用。Mission Bio公司2017年推出了基于微液滴的Tapestri。该方案首先将单细胞包裹进第一个微液滴里面实现单细胞的裂解,然后将含有

裂解细胞的液滴与细胞标签微珠和PCR反应试剂及引物融合进第二个更大的微液滴,在第二个微液滴里面完成扩增子的PCR扩增和细胞标签添加,实现了对几千个单细胞的DNA扩增子进行测序。10× Genomics公司则在2018年推出了基于微流控油包水(droplet-based)的scDNA-seq方案^[30]。该技术首先将细胞包裹进一个水凝胶球内,以完成裂解和DNA单链化,然后再将上述水凝胶球包裹进第二个油滴内进行单细胞标记和扩增反应,从而实现高通量的scDNA-seq。基于微液滴的方案初步实现了高通量的scDNA-seq,但是由于需要进行多步微液滴制备和融合的过程,该类方案有着较高的细胞损失率。

5.3 微孔芯片

第三个方案是基于微孔芯片的方案。此类方案克服了微液滴方案不能进行多步加样的限制,实现了更高效率、更高质量的scDNA-seq。基于此类的方法如Arc-well和DLP+。这些技术通过结合精密的点样仪和纳升小孔芯片,实现了高通量的scDNA-seq的同时极大提高了单细胞测序的质量。尤其是Arc-well,其通过对整个反应体系的大量优化,不但能对新鲜和冰冻组织进行极高质量的scDNA-seq,同时还可以对保存几十年的FFPE组织进行高通量scDNA-seq。

scDNA-seq的吞吐量逐渐从最初的单个细胞,几十上百个细胞,到现在的几千个细胞,样本类型从新鲜组织扩展到FFPE组织。随着技术的持续创新,我们坚信scDNA-seq的吞吐量将会进一步增加,成本进一步的降低,其应用也会变得越来越广,进一步推进生命科学和生物医学的发展。

6 单细胞DNA测序的应用

scDNA-seq技术正在深刻改变着多个生命科学领域。该技术能够揭示传统高通量测序所无法检测到的细胞间基因组异质性,为理解多种生命过程的发生和发展提供了全新的研究视角。目前,该技术在肿瘤研究、耐药靶点发现、精准医疗和发育生物学等多个领域展现出广阔的应用前景。

在肿瘤研究领域scDNA-seq方法主要应用于肿瘤演化、获得性耐药、癌前病变、循环肿瘤细胞检测等过程中。传统的高通量测序技术由于其数据来源于细胞群体,无法精确解析肿瘤克隆谱系结构^[43,72]。scDNA-seq技术的开发,可以从单个肿瘤细胞水平出发,获得单个肿瘤细胞的基因组拷贝数和突变图谱。与高通量测序技术相比,scDNA-seq技术可以精确解析肿瘤亚克隆,通过基因组突变图谱推测肿瘤的起源及演化进程。2011年,NAVIN等^[43]首次应用scDNA-seq技术,探究了乳腺癌中存在的肿瘤异质性和肿瘤演化模型,开辟了用scDNA-seq技术研究肿瘤的异质性和演化的领域。2015年,APARICIO团队^[73]将乳腺癌病人肿瘤异体移植到免疫缺陷小鼠中形成肿瘤,利用scDNA-seq技术构建了克隆演化图谱,发现了相似克隆扩增模式的肿瘤异体移植到不同的免疫缺陷小鼠上,所形成的乳腺肿瘤中的克隆也有一定的相似性。2016年,GAO等^[74]利用scDNA-seq技术在三阴性乳腺癌(triple-negative breast cancer, TNBC)中揭示了肿瘤拷贝数图谱和肿瘤的演化模式。同时在另一项研究中利用scDNA-seq技术对TNBC和高浆卵巢癌(high-grade serous carcinoma, HGSC)进行探究,识别出细胞或克隆特异性的高水平扩增、平行发生的单倍型特异性拷贝数改变,以及拷贝数片段长度变异三种由细胞间结构变异定义的不同“前景”突变模式,且发现这些变异具有可测量的表型影响和演化后果^[75]。2023年,由王开乐等^[38]开发Arc-well对原发的原位乳腺导管癌(ductal carcinoma *in situ*, DCIS)和复发的乳腺癌样本进行了系统性研究。他们发现许多原发性DCIS已经经历了拷贝数倍增和克隆多样化,并且与复发乳腺癌样本中的克隆存在直接的演化谱系关系。进一步分析表明,大多数DCIS在复发过程中经历了演化瓶颈,该研究还发现了导致复发的肿瘤亚克隆,识别了与复发相关的基因组变异事件。

scDNA-seq技术也被用于研究肿瘤耐药克隆。

通过分析药物处理前后肿瘤细胞的基因组变化,能够更准确地评估药物作用机制和耐药性产生的真正原因。2018, KIM等^[76]报道了另一项关于TNBC新辅助化疗耐药性的研究,通过使用scDNA-seq技术和单细胞转录组测序技术的联合分析表明, TNBC中的耐药基因型是预先存在于肿瘤块中并经过适应后被选择的。在2019年, MCMAHON等^[77]通过scDNA-seq技术探究急性髓性白血病(acute myeloid leukemia, AML), 确定了FLT3抑制剂治疗后出现了新的基因突变从而为临床治疗提供了可能的治疗靶点。

scDNA-seq可以绘制癌前病变发展到浸润性恶性肿瘤过程中的基因组图谱,从而回答癌前病变如何逐步发展为浸润性恶性肿瘤的问题。2018年,由CASASENT等^[78]开发的TCSC(topographic single cell sequencing)方法,利用此方法通过研究共发的原位乳腺导管癌和浸润型乳腺癌之间的克隆演化关系,揭示了原位肿瘤亚克隆与侵袭性肿瘤亚克隆之间直接的克隆谱系关系,并进一步表明了大部分突变和拷贝数异常都发生在侵袭之前。研究进一步发现从原位乳腺癌到侵袭性乳腺癌遵循了多克隆演化模型。2024年, WILLIAMS等^[79]应用scDNA-seq技术,针对BRCA1和BRCA2突变的携带者和非携带者的管腔乳腺上皮以及正常乳腺上皮组织,发现在肿瘤的演化过程中拷贝数的变化是导致恶性肿瘤的早期事件,有助于预防和监测乳腺癌的发展。

scDNA-seq也可以应用于稀有肿瘤细胞拷贝数变异或基因组突变的检测,例如循环肿瘤细胞(circulating tumor cell, CTC)。CTC细胞是指肿瘤细胞从发生处混入循环系统中的细胞^[80-81]。对CTC的基因检测对肿瘤早期检测和临床治疗具有重要的指导意义。例如,在一项对小细胞肺癌(small cell lung cancer, SCLC)患者的研究中,利用CellRearch系统富集化疗前后血样中的CTC细胞并进行scDNA-seq分析,其数据被用来预测分类对化疗敏感和耐药的群体^[82]。2024年, SHEN等^[83]开发了一种基于scDNA-seq的方法,命名为scMet-Seq(single-cell metabolic assay and sequencing)。这一方法通过结合代谢特异性标记物和拷贝数特征,对CTC进行富集后测序,发现CTC衍生的CNA与肿瘤组织的CNA一致,从而作为一种临床诊断工具,用于多种癌症的液体活检诊断^[83]。

在发育生物学研究领域, scDNA-seq技术可以用于分析胚胎发育早期的染色体异常, 帮助识别潜在的非整倍性或嵌合现象。通过检测胚胎中的染色体异常, 筛选出更健康的胚胎用于辅助生殖技术, 从而提高胚胎移植的成功率。2019年, 有团队利用灵长类动物(恒河猴)胚胎, 结合scDNA-seq技术, 探究了胚胎从受精卵到囊胚阶段的染色体变化^[84]。他们揭示了胚胎从受精卵到囊胚阶段如何通过细胞碎片化和卵裂球排除来应对染色体错误, 从而维持胚胎的发育稳定性, 为理解胚胎发育过程中的染色体动态变化提供了新的视角, 并对辅助生殖技术(如胚胎筛选和植入前遗传学诊断)产生了重要影响。scDNA-seq也被广泛应用于植入前遗传学诊断(pre-implantation genetic diagnosis, PGD), 帮助识别携带遗传疾病的胚胎, 从而降低遗传病的传递风险^[85]。2013年, 乔杰、谢晓亮和汤富酬团队^[86]首次使用MALBAC技术对单个人类卵母细胞进行全基因组扩增和测序, 全面研究了人类卵母细胞中的交叉和遗传干扰, 使用极体活检检测到了从母体遗传到的非整倍体和疾病的等位基因, 从而排除了遗传疾病的发生。

7 单细胞DNA测序的未来发展方向

单细胞DNA检测技术在过去十年中取得了显著性进展, 但其在扩增偏好性、精确度、覆盖度、成本效益、单倍体解析、多组学整合以及空间基因组学等方面仍然存在挑战和优化的空间。当前的单细胞DNA扩增技术(如DOP-PCR、MDA等)仍存在扩增偏好性的问题, 导致某些基因组区域覆盖不足。未来则需要通过改进扩增酶、引物设计和反应体系等条件, 进一步提高扩增的偏好性, 确保更均匀的全基因组范围。scDNA-seq还存在测序错误率较高的问题, 未来可以通过开发其他单细胞DNA扩增方法或者通过组合分子标签矫正的方法来提高测序的准确性。另外, 当前的绝大多数scDNA-seq技术依托于二代测序技术, 导致其读长短, 无法准确解析单倍型信息, 未来可以通过长读长测序技术(如PacBio和Oxford Nanopore)以及单细胞多组学整合方法, 实现高精度的单倍型解析, 揭示单细胞中的等位基因连锁现象。随着单细胞基因组、转录组和表观基因组测序技术的发展, 未来也可以通过整合多种组学数据, 全面地解析单细胞的基因型如何调控细胞的表

型和功能异质性。近年来空间组学已成为研究的热点, 因此可以通过结合scDNA-seq和空间组学技术, 以实现高分辨率的空间基因组分析, 进而揭示肿瘤细胞在组织微环境中的相互作用和功能, 为肿瘤微环境和克隆演化研究提供新的视角。随着大语言模型人工智能的快速发展, 将scDNA-seq、多组学测序及空间组学测序等数据通过人工智能技术进行训练来实现对细胞状态的预测也是单细胞测序技术未来重要的发展方向。

综上所述, 单细胞DNA测序技术为我们深入理解细胞基因组异质性、揭示罕见克隆变异以及解析肿瘤演化和耐药机制等提供了前所未有的分辨率。尽管当前的scDNA-seq技术仍面临诸多挑战, 但随着新技术的不断开发, 这些挑战将有望逐步被克服。与此同时, 随着人工智能和机器学习在生命科学中的深度融合, 未来可望通过整合scDNA-seq、多组学数据和空间信息, 实现对肿瘤演化模式和单细胞命运轨迹等精准建模与预测。scDNA-seq作为一项具有变革性的前沿技术, 正在推动从群体平均水平向细胞个体层面的基因组研究转变。未来, 随着通量的提升、成本的降低以及跨学科的深度融合, scDNA-seq将在基础生物学、临床医学和精准治疗等多个领域发挥更加重要的作用, 为深入理解细胞层面的基因组变化及其功能提供更加有力的技术支撑。

参考文献 (References)

- [1] TURNER W. The cell theory, past and present [J]. J Anat Physiol, 1890, 24(Pt 2): 253-87.
- [2] AVERY O T, MACLEOD C M, MCCARTY M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III [J]. J Exp Med, 1944, 79(2): 137-58.
- [3] MCCONNELL M J, LINDBERG M R, BRENNAND K J, et al. Mosaic copy number variation in human neurons [J]. Science, 2013, 342(6158): 632-7.
- [4] SANGER F, NICKLEN S, COULSON A R. DNA sequencing with chain-terminating inhibitors [J]. Proc Natl Acad Sci USA, 1977, 74(12): 5463-7.
- [5] TOURDOT R W, BRUNETTE G J, PINTO R A, et al. Determination of complete chromosomal haplotypes by bulk DNA sequencing [J]. Genome Biol, 2021, 22(1): 139.
- [6] SCHMITT M W, KENNEDY S R, SALK J J, et al. Detection of ultra-rare mutations by next-generation sequencing [J]. Proc Natl Acad Sci USA, 2012, 109(36): 14508-13.
- [7] WANG K, LAI S, YANG X, et al. Ultrasensitive and high-efficiency screen of *de novo* low-frequency mutations by o2n-seq [J]. Nat

- Commun, 2017, 8: 15335.
- [8] WANG K, MA Q, JIANG L, et al. Ultra-precise detection of mutations by droplet-based amplification of circularized DNA [J]. BMC Genomics, 2016, 17: 214.
 - [9] EVRONY G D, HINCH A G, LUO C. Applications of single-cell DNA sequencing [J]. Annu Rev Genomics Hum Genet, 2021, 22: 171-97.
 - [10] TELENUS H, CARTER N P, BEBB C E, et al. Degenerate oligo-nucleotide-primed PCR: general amplification of target DNA by a single degenerate primer [J]. Genomics, 1992, 13(3): 718-25.
 - [11] WELLS D, SHERLOCK J K, HANDYSIDE A H, et al. Detailed chromosomal and molecular genetic analysis of single cells by whole genome amplification and comparative genomic hybridisation [J]. Nucleic Acids Res, 1999, 27(4): 1214-8.
 - [12] ZHANG L, CUI X, SCHMITT K, et al. Whole genome amplification from a single cell: implications for genetic analysis [J]. Proc Natl Acad Sci USA, 1992, 89(13): 5847-51.
 - [13] DEAN F B, HOSONO S, FANG L, et al. Comprehensive human genome amplification using multiple displacement amplification [J]. Proc Natl Acad Sci USA, 2002, 99(8): 5261-6.
 - [14] ZONG C, LU S, CHAPMAN A R, et al. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell [J]. Science, 2012, 338(6114): 1622-6.
 - [15] FALCONER E, HILLS M, NAUMANN U, et al. DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution [J]. Nat Methods, 2012, 9(11): 1107-12.
 - [16] GOLE J, GORE A, RICHARDS A, et al. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells [J]. Nat Biotechnol, 2013, 31(12): 1126-32.
 - [17] 4BASEBIO. 4BB™ TruePrime® Whole Genome Amplification (WGA) Kit Handbook. [Cambridge,England]:4basebio [Z]. 2015. https://www.4basebioenzymes.com/wp-content/uploads/2018/01/3x0xxx_TruePrime-WGA-Kit-1.pdf.
 - [18] FU Y, LI C, LU S, et al. Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification [J]. Proc Natl Acad Sci USA, 2015, 112(38): 11923-8.
 - [19] RHEE M, LIGHT Y K, MEAGHER R J, et al. Digital droplet multiple displacement amplification (ddMDA) for whole genome sequencing of limited DNA samples [J]. PLoS One, 2016, 11(5): e0153699.
 - [20] LEUNG K, KLAUS A, LIN B K, et al. Robust high-performance nanoliter-volume single-cell multiple displacement amplification on planar substrates [J]. Proc Natl Acad Sci USA, 2016, 113(30): 8484-9.
 - [21] 阮珏, 王开乐, 沈栩, 等. 一种DNA扩增方法. CN105463066A [P]. 2016-06-04.
 - [22] MISSIONBIO. Tapestri [Z]. <https://missionbio.com/products/platform/>. 2017.
 - [23] BLAGODATSKIKH K A, KRAMAROV V M, BARSOVA E V, et al. Improved DOP-PCR (iDOP-PCR): a robust and simple WGA method for efficient amplification of low copy number genomic DNA [J]. PLoS One, 2017, 12(9): e0184507.
 - [24] DELEYE L, TILLEMANN L, VANDER PLAETSEN A S, et al. Performance of four modern whole genome amplification methods for copy number variant detection in single cells [J]. Sci Rep, 2017, 7(1): 3422.
 - [25] HOSOKAWA M, NISHIKAWA Y, KOGAWA M, et al. Massively parallel whole genome amplification for single-cell sequencing using droplet microfluidics [J]. Sci Rep, 2017, 7(1): 5199.
 - [26] CHEN C, XING D, TAN L, et al. Single-cell whole-genome analyses by linear amplification via transposon insertion (LIANTI) [J]. Science, 2017, 356(6334): 189-94.
 - [27] ZAHN H, STEIF A, LAKS E, et al. Scalable whole-genome single-cell library preparation without preamplification [J]. Nat Methods, 2017, 14(2): 167-73.
 - [28] VITAK S A, TORKENCZY K A, ROSENKRANTZ J L, et al. Sequencing thousands of single-cell genomes with combinatorial indexing [J]. Nat Methods, 2017, 14(3): 302-8.
 - [29] DILEEP V, GILBERT D M. Single-cell replication profiling to measure stochastic variation in mammalian replication timing [J]. Nat Commun, 2018, 9(1): 427.
 - [30] KAMILA B, RAJIV B, HINDSON C. Methods and systems for droplet-based single cell barcoding. United States, US10428326B2 [P]. 2018-08-02.
 - [31] LAKS E, MCPHERSON A, ZAHN H, et al. Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing [J]. Cell, 2019, 179(5): 1207-21.e22.
 - [32] FU Y, ZHANG F, ZHANG X, et al. High-throughput single-cell whole-genome amplification through centrifugal emulsification and eMDA [J]. Commun Biol, 2019, 2(1): 147.
 - [33] YIN Y, JIANG Y, LAM K G, et al. High-throughput single-cell sequencing with linear amplification [J]. Mol Cell, 2019, 76(4): 676-90.e10.
 - [34] MINUSSI D C, NICHOLSON M D, YE H, et al. Breast tumours maintain a reservoir of subclonal diversity during expansion [J]. Nature, 2021, 592(7853): 302-8.
 - [35] GONZALEZ-PENA V, NATARAJAN S, XIA Y, et al. Accurate genomic variant detection in single cells with primary template-directed amplification [J]. Proc Natl Acad Sci USA, 2021, 118(24): e2024176118.
 - [36] ORDÓÑEZ C D, MAYORAL-CAMPOS C, EGAS C, et al. A primer-independent DNA polymerase-based method for competent whole-genome amplification of intermediate to high GC sequences [J]. NAR Genom Bioinform, 2023, 5(3): lqad073.
 - [37] YU X, RUAN W, LIN F, et al. Digital microfluidics-based digital counting of single-cell copy number variation (dd-scCNV Seq) [J]. Proc Natl Acad Sci USA, 2023, 120(20): e2221934120.
 - [38] WANG K, KUMAR T, WANG J, et al. Archival single-cell genomics reveals persistent subclones during DCIS progression [J]. Cell, 2023, 186(18): 3968-82.e15.
 - [39] LIN G, PENG B, CHEN C, et al. msCNVS: medium throughput single cell copy number variation sequencing with barcoded library construction free of preamplification toward clinical implementation [J]. bioRxiv, 2024, doi: 2024.04.01.587505.
 - [40] DUONG A, WONG A, RAMENDRA R, et al. A rapid human lung tissue dissociation protocol maximizing cell yield and minimizing cellular stress [J]. Am J Respir Cell Mol Biol, 2024, 71(5): 509-18.
 - [41] GROSS A, SCHOENDUBE J, ZIMMERMANN S, et al. Technologies for single-cell isolation [J]. Int J Mol Sci, 2015, 16(8): 16897-919.
 - [42] EMMERT-BUCK M R, BONNER R F, SMITH P D, et al. Laser capture microdissection [J]. Science, 1996, 274(5289): 998-1001.
 - [43] NAVIN N, KENDALL J, TROGE J, et al. Tumour evolution inferred by single-cell sequencing [J]. Nature, 2011, 472(7341): 90-

- 4.
- [44] LI P C, HARRISON D J. Transport, manipulation, and reaction of biological cells on-chip using electrokinetic effects [J]. *Anal Chem*, 1997, 69(8): 1564-8.
- [45] DI CARLO D, JEONG K H, LEE L P. Reagentless mechanical cell lysis by nanoscale barbs in microchannels for sample preparation [J]. *Lab Chip*, 2003, 3(4): 287-91.
- [46] MEISTER A. Lactic dehydrogenase activity of certain tumors and normal tissues [J]. *J Natl Cancer Inst*, 1950, 10: 1263-71.
- [47] BOTTOMLEY R H, LOCKE S J, INGRAM H C. Lactic dehydrogenase of human chronic lymphocytic leukemic leukocytes [J]. *Blood*, 1966, 27(1): 85-92.
- [48] JOHNSON K R, GAO Y, GREGUŠ M, et al. On-capillary cell lysis enables top-down proteomic analysis of single mammalian cells by CE-MS/MS [J]. *Anal Chem*, 2022, 94(41): 14358-67.
- [49] PUNJIYA M, MOCKER A, NAPIER B, et al. CMOS microcavity arrays for single-cell electroporation and lysis [J]. *Biosens Bioelectron*, 2020, 150: 111931.
- [50] SHEHADUL ISLAM M, ARYASOMAYAJULA A, SELVAGANAPATHY P R. A review on macroscale and microscale cell lysis methods [J]. *Micromachines*, 2017, 8(3): 83.
- [51] BROWN R B, AUDET J. Current techniques for single-cell lysis [J]. *J R Soc Interface*, 2008, 5(Suppl 2): S131-8.
- [52] ZHU Y, WANG W, YANG Z. Combining mass spectrometry with paterno-büchi reaction to determine double-bond positions in lipids at the single-cell level [J]. *Anal Chem*, 2020, 92(16): 11380-7.
- [53] CHEN J, CHEUNG F, SHI R, et al. PBMC fixation and processing for chromium single-cell RNA sequencing [J]. *J Transl Med*, 2018, 16(1): 198.
- [54] MARTIN B K, QIU C, NICHOLS E, et al. Optimized single-nucleus transcriptional profiling by combinatorial indexing [J]. *Nat Protoc*, 2023, 18(1): 188-207.
- [55] SVEC D, ANDERSSON D, PEKNY M, et al. Direct cell lysis for single-cell gene expression profiling [J]. *Front Oncol*, 2013, 3: 274.
- [56] HODNE K, WELTZIEN F A. Single-cell isolation and gene analysis: pitfalls and possibilities [J]. *Int J Mol Sci*, 2015, 16(11): 26832-49.
- [57] DANAEIFAR M. New horizons in developing cell lysis methods: a review [J]. *Biotechnol Bioeng*, 2022, 119(11): 3007-21.
- [58] FAN X, YANG C, LI W, et al. SMOOTH-seq: single-cell genome sequencing of human cells on a third-generation sequencing platform [J]. *Genome Biol*, 2021, 22(1): 195.
- [59] MULLIS K, FALOONA F, SCHARF S, et al. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction [J]. *Cold Spring Harb Symp Quant Biol*, 1986, 51(Pt 1): 263-73.
- [60] DE BOURCY C F, DE VLAMINCK I, KANBAR J N, et al. A quantitative comparison of single-cell whole genome amplification methods [J]. *PLoS One*, 2014, 9(8): e105585.
- [61] BLAINEY P C, QUAKE S R. Digital MDA for enumeration of total nucleic acid contamination [J]. *Nucleic Acids Res*, 2011, 39(4): e19.
- [62] MISSIONBIO. Tapestry [Z]. <https://portal.missionbio.com/>. 2017
- [63] MULQUEEN R M, POKHOLOK D, O'CONNELL B L, et al. High-content single-cell combinatorial indexing [J]. *Nat Biotechnol*, 2021, 39(12): 1574-80.
- [64] XING D, TAN L, CHANG C H, et al. Accurate SNV detection in single cells by transposon-based whole-genome amplification of complementary strands [J]. *Proc Natl Acad Sci USA*, 2021, 118(8): e2013106118.
- [65] TREFF N R, SU J, TAO X, et al. Accurate single cell 24 chromosome aneuploidy screening using whole genome amplification and single nucleotide polymorphism microarrays [J]. *Fertil Steril*, 2010, 94(6): 2017-21.
- [66] FIEGLER H, GEIGL J B, LANGER S, et al. High resolution array-CGH analysis of single cells [J]. *Nucleic Acids Res*, 2007, 35(3): e15.
- [67] VOULLAIRE L, WILTON L, SLATER H, et al. Detection of aneuploidy in single cells using comparative genomic hybridization [J]. *Prenat Diagn*, 1999, 19(9): 846-51.
- [68] CHRISTODOULOU C, DHEEDENE A, HEINDRYCKX B, et al. Preimplantation genetic diagnosis for chromosomal rearrangements with the use of array comparative genomic hybridization at the blastocyst stage [J]. *Fertil Steril*, 2017, 107(1): 212-9, e3.
- [69] RONAGHI M, KARAMOHAMED S, PETTERSSON B, et al. Real-time DNA sequencing using detection of pyrophosphate release [J]. *Anal Biochem*, 1996, 242(1): 84-9.
- [70] MARGULIES M, EGHOLM M, ALTMAN W E, et al. Genome sequencing in microfabricated high-density picolitre reactors [J]. *Nature*, 2005, 437(7057): 376-80.
- [71] HÅRD J, MOLD J E, EISFELDT J, et al. Long-read whole-genome analysis of human single cells [J]. *Nat Commun*, 2023, 14(1): 5164.
- [72] GAWAD C, KOH W, QUAKE S R. Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics [J]. *Proc Natl Acad Sci USA*, 2014, 111(50): 17947-52.
- [73] EIREW P, STEIF A, KHATTRA J, et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution [J]. *Nature*, 2015, 518(7539): 422-6.
- [74] GAO R, DAVIS A, MCDONALD T O, et al. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer [J]. *Nat Genet*, 2016, 48(10): 1119-30.
- [75] FUNNELL T, O'FLANAGAN C H, WILLIAMS M J, et al. Single-cell genomic variation induced by mutational processes in cancer [J]. *Nature*, 2022, 612(7938): 106-15.
- [76] KIM C, GAO R, SEI E, et al. Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing [J]. *Cell*, 2018, 173(4): 879-93, e13.
- [77] MCMAHON C M, FERNG T, CANAANI J, et al. Clonal selection with RAS pathway activation mediates secondary clinical resistance to selective f1t3 inhibition in acute myeloid leukemia [J]. *Cancer Discov*, 2019, 9(8): 1050-63.
- [78] CASASENT A K, SCHALCK A, GAO R, et al. Multiclonal invasion in breast tumors identified by topographic single cell sequencing [J]. *Cell*, 2018, 172(1/2): 205-17, e12.
- [79] WILLIAMS M J, OLIPHANT M U J, AU V, et al. Luminal breast epithelial cells of BRCA1 or BRCA2 mutation carriers and non-carriers harbor common breast cancer copy number alterations [J]. *Nat Genet*, 2024, 56(12): 2753-62.
- [80] RAO C, BUI T, CONNELLY M, et al. Circulating melanoma cells and survival in metastatic melanoma [J]. *Int J Oncol*, 2011, 38(3): 755-60.
- [81] VINCENT-SALOMON A, BIDARD F C, PIERGA J Y. Bone marrow micrometastasis in breast cancer: review of detection methods, prognostic impact and biological issues [J]. *J Clin Pathol*, 2008,

- 61(5): 570-6.
- [82] CARTER L, ROTHWELL D G, MESQUITA B, et al. Molecular analysis of circulating tumor cells identifies distinct copy-number profiles in patients with chemosensitive and chemorefractory small-cell lung cancer [J]. *Nat Med*, 2017, 23(1): 114-9.
- [83] SHEN X, DAI J, GUO L, et al. Single-cell low-pass whole genome sequencing accurately detects circulating tumor cells for liquid biopsy-based multi-cancer diagnosis [J]. *NPJ Precis Oncol*, 2024, 8(1): 30.
- [84] DAUGHTRY B L, ROSENKRANTZ J L, LAZAR N H, et al. Single-cell sequencing of primate preimplantation embryos reveals chromosome elimination via cellular fragmentation and blastomere exclusion [J]. *Genome Res*, 2019, 29(3): 367-82.
- [85] VAN DER AA N, ESTEKI M Z, VERMEESCH J R, et al. Preimplantation genetic diagnosis guided by single-cell genomics [J]. *Genome Med*, 2013, 5(8): 71.
- [86] HOU Y, FAN W, YAN L, et al. Genome Analyses of Single Human Oocytes [J]. *Cell*, 2013, 155(7): 1492-506.