

特约评述

DOI: 10.12211/2096-8280.2022-058

合成生物学在疾病信息记录与实时监测中的应用潜力

马孟丹^{1,2,3}, 刘宇辰^{1,2}

(¹ 深圳大学第一附属医院, 深圳市第二人民医院, 深圳转化医学研究院, 广东 深圳 518035; ² 广东省泌尿生殖肿瘤系统生物学与合成生物学重点实验室, 广东 深圳 518035; ³ 汕头大学医学院, 广东 汕头 515041)

摘要: 实时、高效和动态地改变储存在基因组中信息的能力是研究细胞生物学、控制细胞表型、监测疾病发展进程、研究原位生物学的一大技术进步。CRISPR-Cas系统的最新研究进展推动了体内DNA和RNA的精确编辑, 复杂多变的基因线路使细胞工程化改造成为可能。DNA具有强大的储存信息的能力, 可稳定保存数千年, 在体内利用DNA记录分子事件是监测细胞信号变化和协调细胞行为的关键技术, 能将细胞的瞬时信号转化为可持续反应, 并永久保存下来。利用该技术, 研究人员能更深入地了解在健康和疾病状态下从基因型到表型转变、临床中患者的疾病发生和用药反应情况、检测生产生活环境的变化。本文概述了合成生物学在DNA存储和细胞实时监测中的技术和应用, 以及CRISPR-Cas系统在活细胞中处理和记录各种信息的优势, 最后展望了它们在疾病研究和治疗方面的前景和挑战。

关键词: 合成生物学; DNA存储; 细胞实时监测记录系统; 疾病信息记录; CRISPR-Cas系统

中图分类号: Q81; R31 **文献标志码:** A

Potential application of synthetic biology in disease information recording and real-time monitoring

MA Mengdan^{1,2,3}, LIU Yuchen^{1,2}

(¹The first Affiliated Hospital of Shenzhen University, Shenzhen Second People's Hospital, Shenzhen Institute of Translational Medicine, Shenzhen 518035, Guangdong, China; ²Guangdong Provincial Key Laboratory of Systems Biology and Synthetic Biology for Urogenital Tumors, Shenzhen 518035, Guangdong, China; ³Shantou University Medical College, Shantou 515041, Guangdong, China)

Abstract: The ability to change information stored with genome in real time, efficiently, and dynamically is a major technological advance for *in situ* studies of cell biology and biology as well to control cell phenotype, and monitor disease progression. Since it was first used for mammalian gene editing, CRISPR-Cas technology has been widely used in research, medicine development and industrial production. In addition to indel mutations induced by Cas9 activity, recent advances in CRISPR-Cas have enabled DNA or RNA base to be edited more efficiently, and synthetic biologists

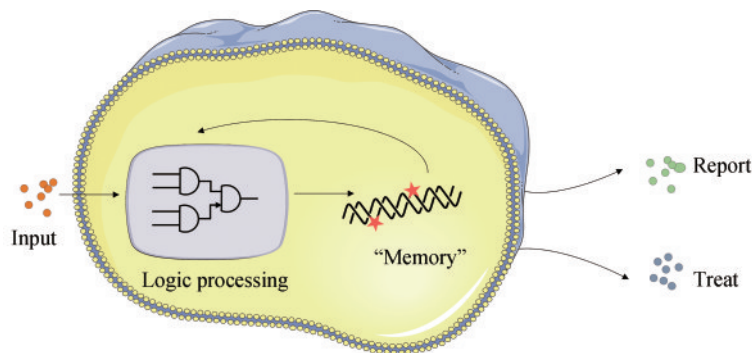
收稿日期: 2022-10-21 修回日期: 2022-12-29

基金项目: 国家重点研发计划 (2021YFA0911600)

引用本文: 马孟丹, 刘宇辰. 合成生物学在疾病信息记录与实时监测中的应用潜力[J]. 合成生物学, 2023, 4(2): 301-317

Citation: MA Mengdan, LIU Yuchen. Potential application of synthetic biology in disease information recording and real-time monitoring[J]. Synthetic Biology Journal, 2023, 4(2): 301-317

are developing devices for information storage by harnessing the versatility of CRISPR-Cas-based tools and gene circuits to engineer cells with modifications. DNA has a strong ability to store information that is stable for thousands of years. Key technology for monitoring cell signal change and behavior coordination is to use DNA as the recorder of molecular events in the body, which can transfer transient signals in cells into sustainable reactions, and store them permanently. With this key technology, researchers could get an in-depth understanding on transformation from genotype to phenotype in health and disease states, drug reactions with diseases in clinical trials for patients, and environmental changes associated with activities of human being. The goal of synthetic biology is to engineer cells with genetic circuits for new biological functions. CRISPR-Cas based tools are useful in developing genetic circuits because they can be easily repurposed by designing complementary gRNAs that interfere or act on any arbitrary nucleic acid sequence of interest. Although they are still in their infancy, CRISPR-Cas based tools are gaining popularity in encoding biological memory and tracing and forwarding genetic screening for each formed lineage. In this article, we summarize the progress and application of synthetic biology in DNA storage and real-time monitoring in cell, as well as the advantages of the CRISPR-Cas system processes and records of information in living cells. Finally, we highlight their prospects and challenges in research on diseases and treatments.



Keywords: synthetic biology; DNA store; real-time monitoring and recording systems in cell; disease information records; CRISPR-Cas systems

合成生物学 (synthetic biology) 是生物学和工程学交叉的领域, 与主要集中在特定、单独基因上的基因工程不同, 合成生物学协调了生物体一系列分子成分的变化, 如 DNA、RNA 和蛋白质, 针对细胞不同的状态形成基因线路 (gene circuits) 和网络, 经过工程方式设想、革新甚至重新分解有特定功用的生物系统^[1]。合成生物学的一个基本目标是可预测地和有效地对细胞进行重新编程, 以执行计算和执行特定的生物任务^[2], 在工程化策略指导下有目的地设计合成标准化生物元件, 具有不同功能的生物元件按照一定逻辑构建基因线路, 多种基因线路组装集成系统, 并且与宿主细胞融合, 使合成生物学研究能够以程序化的方式改变细胞的行为。研究人员已经建造出许多用

于细胞计算机的遗传元件, 包括转录型逻辑门^[3]、计时器^[4]、计数器^[5]、记忆元件、可协调的传感器, 甚至还可进行复杂计算的 DNA 系统。利用基因线路进行改造的细胞在改进治疗^[6-8]、诊断^[9-10]、动物模型^[11] 和工业生物技术^[12] 中有巨大的应用前景, 适用于所利用的任何生产平台, 如植物^[13]、细菌^[14-15]、人类细胞^[16] 或酵母^[17]。

重新设计生物系统并对疾病治疗和监测进行工程改造, 需要全面了解疾病如何在没有干预的情况下发生和发展, 但在生物系统中的许多分子事件是短暂的, 很难在自然条件下对其进行监测和研究。在活细胞中处理和存储信息的能力对于开发新一代疗法和生物学机制研究至关重要, 也使研究细胞内部的精细状态成为可能, 包括基因

组的序列、表观遗传修饰的状态以及共同决定细胞状态的RNA、蛋白质和代谢物的特性和丰度^[18-19]。研究人员已经在活细胞中实现可靠且可扩展的分子记录和计算的平台等一系列研究,从研究发育和癌症中的信号动力学和细胞谱系,从构建活细胞的生物传感器和适应性疗法,到编码逻辑和编程细胞表型^[20]。了解这些生物信号和细胞状态的时空表达,有助于我们研究疾病发生和发展,探索生命奥秘。但这些数据的获得难以实现,需要侵入性小、无损害、可持续时间长的技术。

将细胞转化为分子记录器可以解决这一问题:对细胞DNA进行“编程”,打上一个永久性的“标签”,记录细胞周围发生的变化,有利于人们认识疾病的原因,进而达到精准医疗。DNA编辑器可以对活细胞内所含DNA进行插入、删除、倒位或碱基替换突变等形式,并可用于区分不同DNA记忆状态;短暂的细胞事件,例如蛋白质-蛋白质相互作用,可以作为转录输出被DNA信息储存^[21]。利用DNA的四种单体作为存储介质,把化学信号转变成数字信号成为二进制的数据存储材料,其较高的稳定性、耐用性、存储量、广泛适用性以及与生物功能的兼容性而成为人工生物信息存储的理想介质^[22],基于细胞DNA的记忆是长期存储的一种有用的实现方式,因为它在细胞分裂时自然繁殖,即使在细胞死亡之后也能保持稳定^[5, 23]。文章主要介绍了随着基因编辑技术的进步,利用合成生物学的思维,在生物体内构建可针对细胞状态进行DNA编辑的基因线路,动态地改变遗传信息,作为活细胞中信息处理和存储媒介的存储器结构,并通过测序获取信息的研究方法和进展,及其在疾病信息记录与实时监测中的应用潜力^[22, 24]。

1 DNA读写和分子记录仪的应用

1.1 DNA存储

利用DNA创建分子记录器的原理将活细胞转化为记录设备,将其自身信号动力学的历史存储为永久的DNA信息,可以提供对其原始环境中的生物过程的纵向研究。条件诱导型启动子启动

DNA编写,例如细胞因子(如代谢物、蛋白质)或环境因素诱导(如光、污染物、温度),该程序可以记录诱发条件的存在与否、出现时间、强度、时长、顺序等^[25]。通过捕捉自然环境中的瞬时时空分子事件,分子记录技术可以在不同的学科中具有广泛的实用价值。例如生物学家可以研究肿瘤的发展,并更深入地了解与癌症异质性有关的肿瘤微环境中的细胞和环境线索;免疫学家可以研究免疫细胞成熟、记忆形成和免疫反应中的信号;微生物学家可以研究细菌群落和生物膜中的信号动力学和分子相互作用^[20]。

1.2 活细胞生物传感器

DNA写入技术可以用来制造生物传感器,用于健康和环境监测。与基础研究目的类似,生物传感应用需要具有最小体积和较大记录容量的存储体系结构,以实现连续和稳定的记录^[20]。Oishi等^[26]证明了在细胞中建立有限状态机(finite state machine, FSM)与基因调控网络(gene regulatory network, GRN)的等价性,FSM是计算的基本模型,可以描述对象所经历的状态,具有存储器,并且可以根据其当前状态对相同的输入做出不同的响应。在活细胞中建立GRN,包括启动子(结构型启动子和/或诱导型启动子)、可扩散的小信号分子和具有可编程的DNA结合域的转录抑制因子。前三个组分代表通过抑制转录因子进行基因调控的基序,而第四个组分将用于检测小的输入信号分子,其中状态是表观遗传编码,只需使用容易获得的抑制转录因子的基因线路就可以构建任何细胞的状态记录仪。作者通过合成细菌微菌落边缘检测,探索了一个设计更复杂系统的框架,微菌落中的细胞都可以发出可扩散的信号分子的脉冲,被邻近的细胞感知和传递,通过在传递波后的短时间内感知信号分子的局部浓度,可以确定细胞是否处于微菌落的边缘,利用这一技术可以发现扩散性强的癌细胞亚群。

内源性生物标记物仍然处于早期疾病检测工作的前沿,但许多生物标记物缺乏影响疾病所需的敏感性和特异性。Aalipour等^[27]利用发生在肿瘤浸润免疫细胞中的代谢改变,开发了一种基于

细胞的体内传感器,用于高灵敏度的早期癌症检测。通过将荧光素酶表达与精氨酸酶-1启动子的激活耦合,使巨噬细胞可以感知M2型肿瘤相关代谢谱产生荧光素酶。在结直肠和乳腺小鼠肿瘤模型中过继转移后,工程巨噬细胞迁移到肿瘤中并激活精氨酸酶-1,从而通过生物发光成像和血液荧光素酶测定来检测。巨噬细胞传感器通过血液荧光素酶检测到 $25\sim 50\text{ mm}^3$ 的肿瘤,即使存在伴随炎症,也比临床使用的蛋白质和核酸癌症生物标志物更敏感。巨噬细胞传感器还有效地跟踪肌肉和肺部炎症模型的免疫反应,表明这种方法在癌症以外的疾病状态中具有潜在的效用。Jiang等^[28]设计了一种由两个RNA组成的可编程的RNA生物传感器平台,一条是引导区,可以与目标RNA结合,另一个可以编码目的蛋白,如荧光信号或细胞毒性蛋白,其中引导区含有终止信号会阻止目的蛋白的翻译,因此目的蛋白并不表达。当引导RNA与靶向RNA结合时会产生一个短的双链RNA序列,其中包含序列中处于终止信号范围内的两个A-C之间的错配,这种错配吸引了一个天然存在的RNA的腺苷脱氨酶(ADAR),ADAR修复错配并使终止信号失效,允许目的蛋白的翻译。该系统可以特异性地针对并杀死具有某些序列的细胞,或进行细胞特异性基因组编辑,在帮助检测和选择性地杀死肿瘤细胞以及疾病监测和治疗方面有巨大潜力。

现阶段基因线路型全细胞微生物传感器的设计和应用已取得了重大突破,为代谢途径改造、定向进化等提供了高效工具,在生物制造、环境监测、食品安全以及疾病诊断等领域得到了广泛的应用^[29]。未来生物传感器可进行更大范围的应用,例如:被赋予疾病生物标记物传感器的细菌细胞与DNA记录器相结合,可以被患者服用,通过胃肠道记录疾病生物标记物,当它们离开体内时报告这一信息;携带分子记录器的工程化人类细胞可以被输入到人体内,以报告疾病的早期迹象,如癌症或神经退化;具有记录能力的工程细胞和动物可以用来连续监测和记录生物和环境线索(如毒素、重金属、代谢物和光)的水平和活动,而不需要人工电源,也可以在非生物传感器不易接触到的环境下进行。

1.3 谱系追踪

确定控制细胞在某一状态下如何分化的分子机制是干细胞和发育生物学中的一个长期关注的焦点,随着组织和器官的发育,对细胞状态变化的全面记录可以深入了解细胞在胚胎发生或再生过程中选择其终端身份的分子机制和事件顺序,细胞在组织动态平衡、修复和疾病过程中的行为^[30],它可以为如何在体内操纵细胞命运、预测发育病理和癌症的起源以及在体外重建细胞分化过程提供线索^[31]。由于遗传变化、环境差异和细胞性质的可逆变化,同一肿瘤内的癌细胞之间出现了表型和功能的异质性。一些癌症还包含致癌干细胞分化为非致癌后代的层次结构,使用谱系追踪和深度测序的研究可能会对癌症干细胞模型产生影响,并可能有助于确定它在多大程度上解释了治疗抵抗和癌症进展^[32]。

使用DNA条形码(DNA barcoding)来构建谱系系统发育图^[33],通常分为三种类型:①外源DNA序列的转基因整合^[34-35];②转基因DNA的体内重组^[36-37];③CRISPR-Cas9对转基因DNA靶标的体内编辑^[35]。基于CRISPR-Cas的分子记忆装置已在哺乳动物细胞中建立了多种谱系追踪线路,基于三个成分:Cas内切酶、DNA靶点序列和一组sgRNA。利用CRISPR-Cas9活性导致的靶点的累积可变性,在细胞内产生高度多样性的“进化”DNA条形码,并促进细胞谱系树的系统发育重建。任何谱系示踪剂的关键在于它不会改变标记细胞、其后代和它相邻细胞的属性,标签必须传递给起始细胞的所有后代,随着时间的推移长期保留,且不会转移到无关的相邻细胞。

2 基于细胞的记录装置系统技术

动态地提取细胞信息需要将细胞内的信息(如代谢物的存在或基因的表达)转换成与记录系统兼容的格式(如诱导记录成分表达的生物信号),这些信息必须借助各种DNA修饰酶(如核酸酶、整合酶或重组酶),通过碱基的改变、缺失或添加直接写入DNA(表1)。最后,使用测序或成像等多种技术从DNA中读出存储的数据。所存储

表1 基于细胞的记录装置系统技术汇总

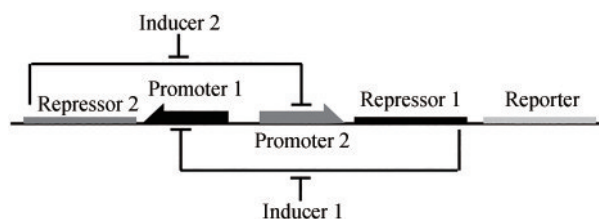
Table 1 Summary for real-time monitoring and recording systems in cell

记录装置系统		种群分布 vs 单细胞记录	写入周期	记录能力	发生顺序	持续时间	灵敏度	保真度
双稳态开关		种群	短	一般	否	是	低	低
DNA 重组酶技术		种群	长	一般	否	否	低	低
ssDNA 编辑技术	HiSCRIBE ^[25]	种群	长	强	否	是	一般	高
CRISPR 系统	Record-seq ^[38]	种群	长	强	是	是	高	高
	CAMERA ^[39]	单细胞	长	强	是	是	高	较高
DNA typewriter ^[40]		种群	长	强	是	是	高	高
LINNAEUS ^[41]		种群	短	弱	否	否	一般	低
mSCRIBE ^[42]		单细胞	长	强	是	是	高	高
iTracer ^[43]		单细胞	长	强	是	是	高	高
DOMINO ^[44]		种群	长	强	否	否	一般	一般

的信息还可用于直接启动或引发一组特定的易于检测的生物反应，例如基因表达^[45]。可靠稳健的基因线路的设计面临着许多挑战，包括：需要一个特异性良好的基因开关文库，可以扩大到复杂的调节线路；稳定、长期的多线路结合，以确保细胞中正确和动态的功能；合成存储设备，以保证由遗传程序计算的信息的持续存储^[46]。

2.1 双稳态开关

第一批合成记忆设备是双稳态开关，又称“拨动开关”，经典的双稳态开关由两个抑制器和两个启动子组成，每个启动子都受到由相反启动子转录的抑制物的抑制，通过调控实现基因线路在两种不同稳定状态间的切换（图1）^[47-49]。这种设计作为触发开关只需要最少的基因和顺式调节元件来实现鲁棒性强的双稳态行为，并且所选择的触发设计不需要任何专门的启动子，任何一组启动子和阻遏子都可以通过合理的设计实现双稳性^[48]。Bhomkar等^[50]在大肠杆菌中将细胞分裂抑制基因 *minC* 的表达与双稳态开关耦合，以设计表型反应（即丝状反应），可用于计算初始信号暴露的时间，它可以通过引起细胞伸长来长期记录细胞短暂暴露于化学物质后的表型反应，根据长度来估计暴露于化学物质后的时间。这种细菌时间记录装置的目的包括：对短暂的刺激作出反应，保持对该刺激的记忆，标记刺激的时间。虽然这些生物纳米记录仪展示了许多设计所需的功能，但它们的灵敏度、保真度和曝光时间方面仍需改

图1 双稳态开关设计^[48]

抑制物1抑制启动子1的转录，由诱导物1诱导；
抑制物2抑制启动子2的转录，由诱导物2诱导

Fig. 1 Design for toggle switch^[48]

Repressors 1 and 2 inhibit transcription driven by Promoters 1 and 2, respectively, which is induced by Inducers 1 and 2 correspondingly

善。在生产和临床应用中构建纳米记录仪，可用于感知污染物、毒素（如重金属、酚类）和临床有用的代谢物（硝酸盐、尿素）。

2.2 DNA 重组酶技术

位点特异性重组酶（site-specific recombinases, SSR）可以根据DNA识别位点，翻转或切除位于其同源位点之间的一段DNA，利用DNA重组酶在基因组中建立新型的记录装置，可以更加持久地记录信号^[5]。已有报道将正交化的重组酶植入更加复杂的环境（如哺乳动物肠道中）^[51-53]，可以实现复杂微生物群的实时监测，促进活体诊断与医疗的发展。Siuti等^[54]使用了两种丝氨酸重组酶 Bxb1 和 phiC31，可以根据周围一对识别位点的方向不可逆转地颠倒或切除DNA，通过在不同方向上直接组装重组的目标启动子、终止子和输出基因模块，研究人员建立了16种二输入逻辑功能，从而能够通过记忆对细胞线路进行简单的定义编

程, 并至少在 90 代细胞保持记忆。Roquet 等^[55] 使用基于重组酶的状态机 (recombinase-based state machines, RSMs), 通过利用化学控制的 DNA 切除和反转操作来编码 DNA 序列中的状态, 从而在活细胞中建立记录系统 (图 2)。这一策略能够方便地读出状态 (通过测序和/或聚合酶链反应) 以及对基因表达的复杂调控, 研究人员在大肠杆菌中设计记录系统来验证这一框架时, 该系统能够记录所有输入的时间顺序, 并执行多输入、多输出的基因表达控制。RSM 可以记录生物事件的组合和它们的发生顺序, 并且对这些事件作出反应。可以利用这一技术追踪干细胞或其他未成熟细胞分化为成熟细胞的轨迹, 以及追踪癌症这类疾病的发展。将 RSM 系统用于患者肠道中, 通过获取记录细菌或细胞得到患者使用药物后情况, 进一步指导临床用药。

2.3 单链 DNA 编辑技术

由于正交化的重组酶数量有限, 每个开关只能记录一项简单的信息, 难以在复杂信息系统中实现多层次的应用。Lu 团队^[25] 没有使用传统的依赖于位点特异性识别的重组酶, 而是依靠反转录子组成一些细菌的遗传系统 (图 3), 这种方法被称为 SCRIBE (synthetic cellular recorders integrating biological events), 包含基于编码逆转录酶的细菌来源的“逆转录”盒、RNA 模板和引物, 它们共同合成单链、共价连接的 RNA-DNA 分子^[56-57], 将其插入细菌的细胞中, 只有当细菌响应特定的刺激 (如化学物质、光) 时, 该反转录子才会开启, 以响应一系列调节信号, 产生特定的单链 DNA (single stranded DNA, ssDNA)。这些 ssDNA 根据序列同源性定位到特定的靶基因座, 并将精确的

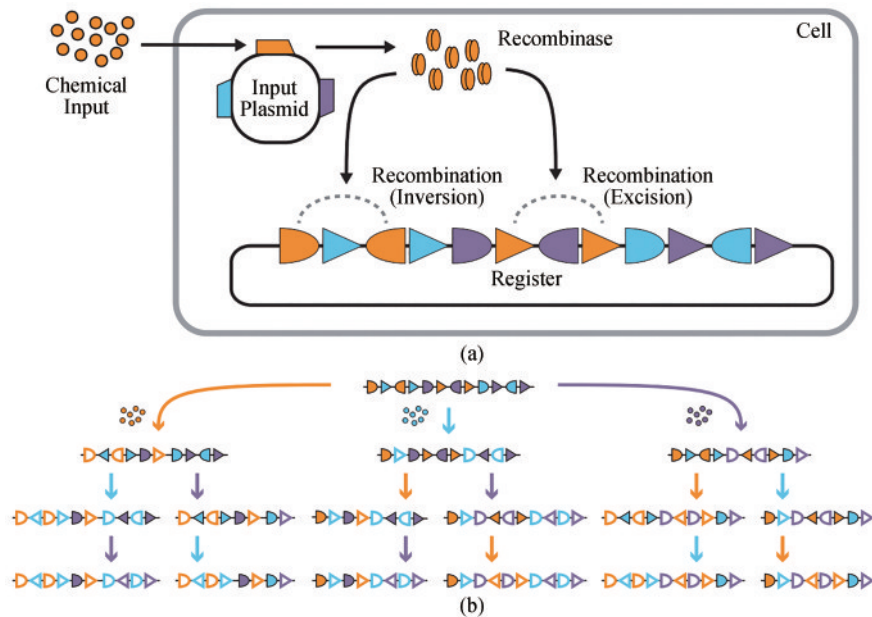


图 2 三输入、16 状态的 RSM 图示^[55]

(a) RSM 机制。化学输入诱导重组酶的表达 (来自输入质粒上的基因), 由重叠和正交的重组酶识别位点组成的 DNA 登记表。不同的重组酶可以由不同的输入来控制。这些重组靶向多个正交对同源识别位点 (显示为三角形和半椭圆形), 以催化反转 (当这些位点是反对齐的) 或切除 (当这些位点对齐的)。(b) 登记表的设计是为每种身份和输入顺序采用不同的 DNA 状态。三种不同的输入 (橙色、蓝色和紫色) 由彩色箭头表示, 每个箭头表示一个不同的重组酶。未重组的识别位点有阴影; 重组的识别位点有轮廓

Fig. 2 Summary for three-input and 16-state RSM^[55]

(a) RSM mechanism. A chemical input induces the expression of a recombinase encoded by a gene on the input plasmid, which modifies a DNA register with overlapping and orthogonal recombinase recognition sites. Specific recombinases can be controlled by corresponding inputs. Each of these recombinases can target multiple orthogonal pairs of their cognate recognition sites (shown as triangles and half-ovals) to catalyze inversion (when the sites are anti-aligned) or excision (when the sites are aligned). (b) The register is designed to adopt a specific DNA state for every identity and order of inputs. Three different inputs are represented by colored arrows (orange, blue, and purple), each of which expresses a specific recombinase. Unrecombined recognition sites are shown by solid symbols, and symbols without filling highlight recombined recognition sites

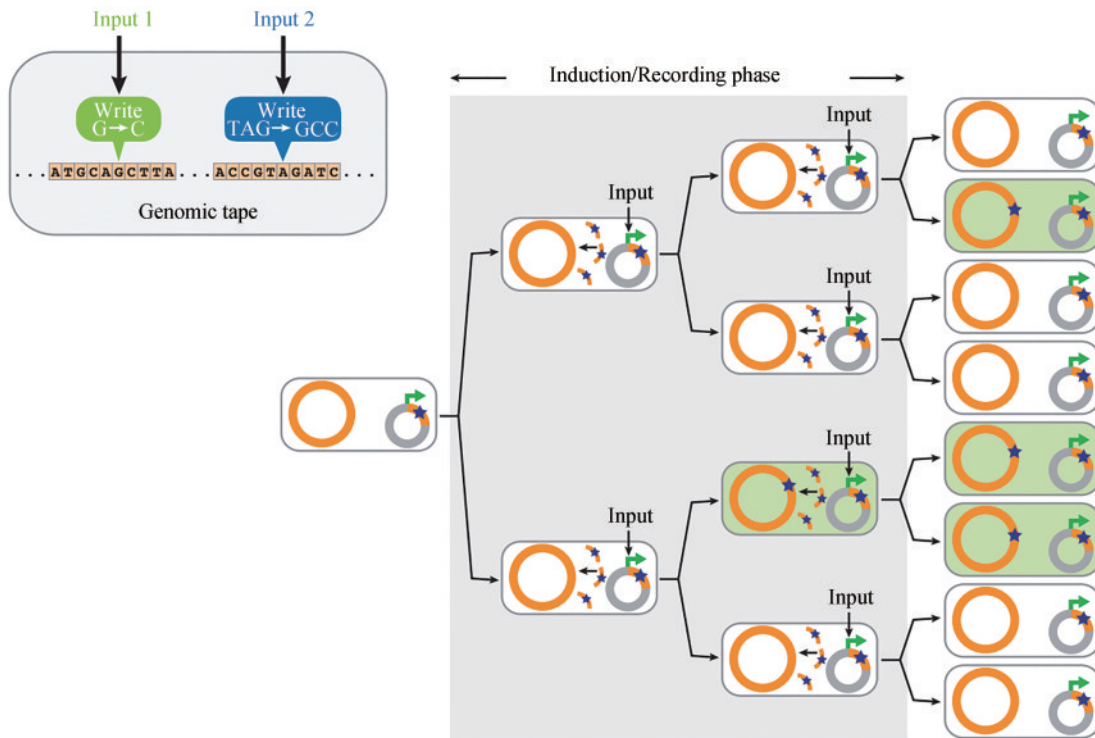


图3 SCRIBE实现了分布式基因组编码记忆^[25]

在输入存在的情况下，ssDNA（橙色曲线）从质粒携带的盒（灰色圆圈）中产生，并重组基于序列同源性的特定基因组位点（橙色圆圈）。这就导致了精确突变的积累（绿色细胞中的星星），产生输入信号的大小和持续时间的相关函数

Fig. 3 SCRIBE-based distributed encoded memory at genome levels^[25]

In the presence of an input, ssDNAs (orange curved lines) are produced from a plasmid-borne cassette (gray circles) and recombined into specific genomic loci (orange circles) that are targeted on the basis of sequence homology. This results in the accumulation of precise mutations (stars in green cells) as a function of the magnitude and duration of exposure to the input

突变引入基因组DNA，且突变频率与信号输入量呈线性相关，通过改变ssDNA模板，存储设备可以针对不同的基因组位置重编程。SCRIBE可以记录任意输入的存在/不存在和一定范围内的变化，通过将信息编码到细胞群体的基因组DNA中，跟踪输入的大小和时间行为。这种模拟存储体系结构利用细菌培养中的大量细胞进行分布式信息存储，并将携带特定突变的群体中一小部分细胞的事件历史存档，适合于诸如环境和医学监测传感器等应用。随后研究者开发更安全高效的HiSCRIBE，这种方法通过重写细菌的DNA，有效地编辑细菌基因组，并将记忆程序写入细菌细胞。研究团队设计了DNA条形码，通过将这些DNA条形码整合到细菌的基因组中，随后与其他细菌进行交换，通过测序检测出哪些细菌携带了DNA条形码，由此确定哪些细菌进行了细菌接合（在此过程中细菌会交换DNA片段），这种图谱可以帮助研究人员研究细

菌是如何在生物膜等聚集物中相互交流的。类似的方法应用于哺乳动物细胞时，它可以用来绘制神经元等其他类型细胞之间的相互作用，这为绘制大脑连接体提供了一种新方法。

2.4 基于CRISPR系统的信息记录

2.4.1 利用Cas1-Cas2获取间隔序列

上述存储设备的容量和可伸缩性受到能够一同起作用的正交调控元件（例如，转录因子和重组酶）的数量限制，只能捕获一小部分确定的刺激。使用CRISPR间隔序列捕获并将细胞内RNA转化为DNA，实现基于DNA的转录信息存储介导的RNA记录，并进行深度测序，用于重建复杂的细胞行为或病理状态的演变历史^[58]。

Tanna团队^[38]根据CRISPR-Cas机制研发出一种记录肠道细菌响应肠道环境变化的方法——Record-

seq, 通过从RNA获取CRISPR间隔区 (spacer) 来实现转录记录的技术 (图4)。将细菌CRISPR序列引入肠道细菌大肠杆菌 (*Escherichia coli*) 菌株, 在Record-seq工程大肠杆菌细胞内, 转录物被融合到Cas1的RT domain 逆转录为DNA, 并作为spacer储存在CRISPR阵列中, 实现以质粒DNA形式永久存储转录信息。在实验中, 研究人员给小鼠注射了工程大肠杆菌, 收集动物粪便样本并分离出细菌DNA, 利用高通量DNA测序以及生物信息学通过大量的数据来重建mRNA片段的遗传信息, 使研究人员以非侵入性方式测定肠道细菌在机体内产生特定mRNA分子的频率来识别哪些基因处于活跃状态, 并从这些序列上推断和检测到一些特殊的微生物, 例如有些菌可分泌某些酶, 而有些则不存在这种功能。前哨细胞从时间尺度上, 在转录水平上长期记录小鼠肠道内与饮食、疾病和微生物相互作用的信息。作为一种基于种群的全局转录, Record-seq不用考虑单个细胞之间转录的变化或偶然性, 可以实现对积累的转录变化进行无偏好地记录, 能够记录长时间发生的瞬态转

录响应, 适用于研究广泛的群体转录反应记录, 这一转录记录技术使肠道的无创检测成为可能, 有望在生物医学研究和未来的生物医学诊断应用中发挥作用^[59]。

2.4.2 核酸酶产生突变

基于CRISPR-Cas的谱系追踪技术对gRNA靶向的DNA序列有几种设计, 包括“条形码”的串联序列, McKenna等^[35]开发了一种可追踪几代癌细胞的新谱系追踪方法, 使用CRISPR技术为每个细胞嵌入可遗传和可进化的DNA条形码, 当细胞分裂时, 它的条形码都会被轻微修改, 比较细胞的条形码来重建每个细胞的“家谱”, 使用类似的方法可以跟踪新冠病毒的演变。另一种方法是基因组中整合转基因的多个副本, Junker团队^[41]开发出通过核酸酶激活对泛在序列的编辑进行谱系追踪系统 (lineage tracing by nuclease-activated editing of ubiquitous sequences, LINNAEUS), 能够让人们确定细胞类型和细胞谱系。然而, 这些追踪系统记录信息的能力有限, 因为随着不断地编辑, 可编辑的目标序列会丢失, 为了能够在人

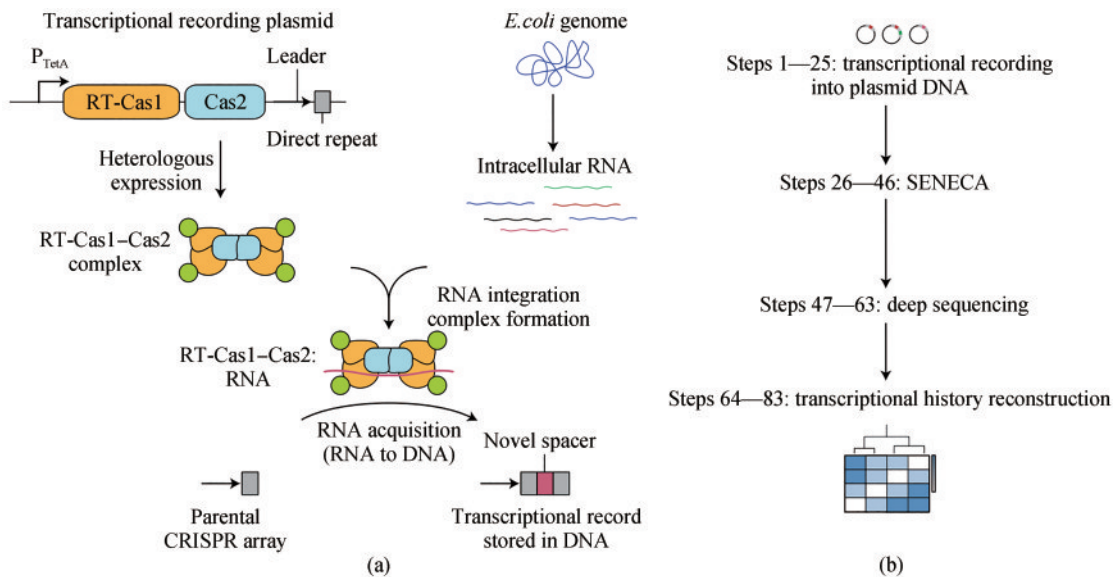


图4 从RNA中获取CRISPR间隔区的转录记录^[59]

(a) Record-Seq使用从食糖梭菌获得RNA的RT-Cas1-Cas2复合体将转录信息编码到质粒携带的CRISPR序列中。转录记录是通过直接从细胞内RNA获得CRISPR间隔区, 然后通过FsRT-Cas1-Cas2的RT结构域逆转录RNA原间隔区来产生的。(b) 提取质粒DNA, 然后选择性扩增扩展CRISPR阵列 (SENECA) 和深度测序, 进行转录历史的重建

Fig. 4 Transcriptional record of RNA extracted from CRISPR spacer acquisition^[59]

(a) Record-seq uses the RNA-acquiring RT-Cas1-Cas2 complex from *Fusicatenibacter saccharivorans* to encode transcriptional information into plasmid-borne CRISPR arrays. The transcriptional record is generated by CRISPR spacer acquisition directly from intracellular RNAs followed by reverse transcription of RNA protospacers through the RT domain of FsRT-Cas1-Cas2. (b) Extraction of plasmid DNA followed by the selective amplification of expanded CRISPR arrays (SENECA) and deep sequencing enable the reconstruction of transcriptional histories

类细胞中对种群水平的记忆进行持续编码，需要建立一个可以重复写入的模块化存储单元。Perli 等^[42]开发了一种模拟记忆系统，能够以 DNA 突变的形式记录人类细胞群体中的细胞事件：整合生物事件的哺乳动物合成细胞记录仪（mammalian synthetic cellular recorder integrating biological events, mSCRIBE），使用 CRISPR-Cas 的活性修改转录 sgRNA 的 DNA 序列，在编码 SDS 的区域下游即包括 5'-NGG-3' PAM，从而得到的 PAM 修饰的 stgRNA，能将 Cas9 内切酶活性定向到 stgRNA 自己的 DNA 位点。在 SDS 编码区引入 dsDNA 断裂并通过 NHEJ 修复途径修复后，产生的从头突变的 stgRNA 基因座应该继续转录为原始 stgRNA 的突变版本，并参与另一轮自我靶向突变，从而产生一个连续的、自我进化的 Cas9 stgRNA 系统。通过将该系统的开启与感兴趣的调控事件的发生，在生物学上联系起来，mSCRIBE 除了可以作为一种记忆设备，以 DNA 突变的形式记录信息，还可用于谱系跟踪和生成条形码以同时唯一地标记多个细胞。研究人员在哺乳动物细胞中用这一装置将 Cas9 靶向基因的活性与细胞内炎症途径的活性相结合，在暴露于更多炎症因子 TNF- α 的细胞中，在靶序列中记录了更多轮的 Cas9 突变。将修饰的细胞注射到动物体内，并将促炎分子注射到一些细胞中，与未治疗的小鼠相比，注射了修饰细胞的小鼠中 CRISPR 靶向序列发生了显著变化。该方法可应用于体内以模拟方式记录生理相关的生物信号；记录肿瘤微环境中癌细胞的刺激情况，并在疾病发展过程中追踪细胞内特定途径的活性；在大脑研究中记录基本记忆所涉及通路的活动，将临时过程转变为永久记录，并将其呈现给大脑中的所有神经元细胞。

He 团队^[43]开发出 iTracer 技术，包含静态序列标记原件，即构建一个高复杂度的条码序列文库、转染并整合到干细胞中，从而标记起始时间点的不同干细胞，也包括基于 CRISPR 编辑系统的动态序列标记，结合带有可诱导 Cas9 蛋白基因的干细胞，可在特定时间点产生额外的随机突变，从而得到第二层细胞谱系信息。这一技术既能获得多个时间点的谱系层次信息，也能降低标记冲突问题造成的影响。将 iTracer 技术用于大脑类器官培养系统中，对人类脑部早期发育过程进行模拟和

机理研究，研究大脑类器官中不同类型神经元之间的谱系关联。此外，在 iTracer 的基础上额外引入一个针对靶点基因的 gRNA，还能在记录细胞谱系信息的同时，对靶点基因进行基于 CRISPR 技术的基因敲除，并据此对基因敲除如何影响细胞谱系的发展分化进行研究。携带这些遗传记录器的生物受到不同的神经刺激，产生的突变特征可以用来推断整个动物大脑的平均活动时间。或者，编码在可移动遗传元件上的 DNA 编写器可以通过突触，用 DNA 条形码来独特地标记神经连接，并以高分辨率和高通量的方式绘制连接体的图谱。尽管存在许多技术挑战，但应用分子记录技术来破译大脑的功能架构和研究神经系统疾病，将是这些技术进步的强大推动力，特别是在可扩展性、记录能力以及时间和空间分辨率方面。

2.4.3 碱基编辑

Tang 等^[39]使用两种不同的 CRISPR 介导的 DNA 修饰机制来记录活细胞中感兴趣的事件：Cas9 核酸酶催化的双链 DNA（double-stranded DNA, dsDNA）切割和碱基编辑介导的点突变。CAMERA 系统（CRISPR-mediated analog multi-event recording apparatus）永久记录在活细胞 DNA 中已知时间尺度上刺激的幅度，或具有已知幅度的刺激的持续时间（图 5）。这两种记录系统的模拟特性允许对感兴趣的信号进行连续监视，并且与标准数字存储设备相比提供了更多信息。在第一种被称为“CAMERA 1”的细胞记录系统中，研究人员将两种彼此之间略有不同的质粒注射到细菌细胞中，随后在接触到所需的刺激物期间，利用 CRISPR-Cas9 对这两种质粒中的一种进行切割，改变两种质粒在细胞中存在的比例并被记录下来，且该系统可以重复记录和擦除。然而，CRISPR 核酸酶造成 DNA 双链断裂的修复过程是通过非同源末端连接（non-homologous end joining, NHEJ）途径进行的，可能导致随机的有害突变，如插入、缺失和易位。在改进的“CAMERA 2”细胞记录系统中，当所需的信号在细胞中发生时，利用碱基编辑器改变遗传密码中的单个碱基，融合在 dCas9 上的胞苷脱氨酶（cytidine deaminase）能够在不引起 DNA 断裂的情况下将靶向序列的 A-T 碱基对修改为 C-G 碱基对，这种特性降低了随机错误

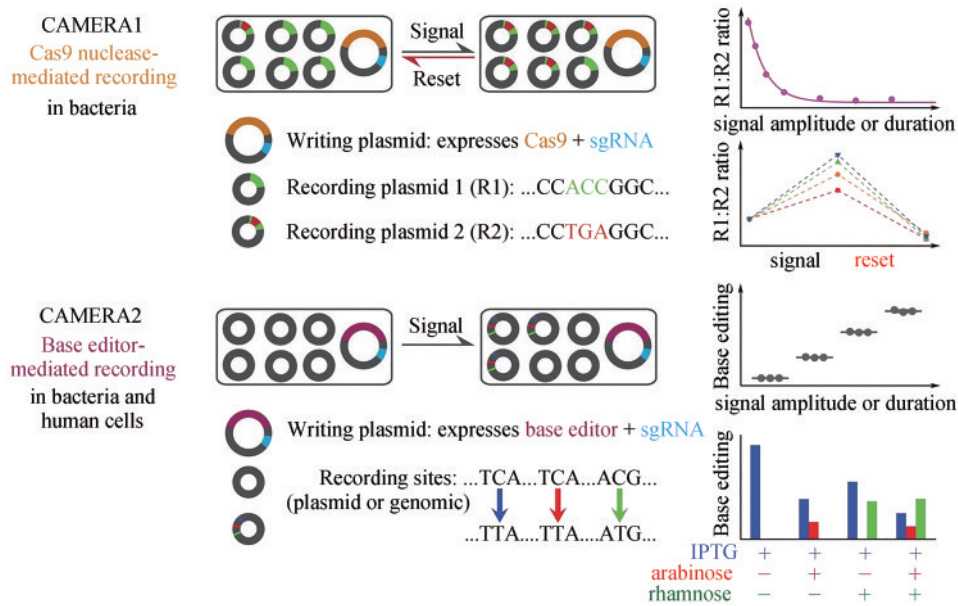


图5 CAMERA在细菌和哺乳动物细胞中的多路模拟细胞记录^[39]

CAMERA 1将刺激记录为相互排斥的DNA序列比例的变化。CAMERA 2使用碱基编辑器记录单核苷酸变化时信号的持续时间或振幅。这两个系统都可以复用，独立记录多个事件，包括暴露于抗生素、营养物质、病毒、光以及Wnt信号

Fig. 5 Multiple analog of cellular recording by CAMERA systems in bacteria and mammalian cells^[39]

CAMERA 1 records stimuli as changes in the ratio of mutually exclusive DNA sequences. CAMERA 2 uses base editors to record the duration or amplitude of signals as single-nucleotide changes. Both systems can be repeatedly used to independently record multiple events, including exposure to antibiotics, nutrients, viruses, and light, as well as Wnt signaling

写入事件的频率，提高了记录的准确性。使用碱基编辑以与单核苷酸编辑效率线性相关的形式，记录在已知时间尺度上的外源信号的强度，或已知强度的信号的持续时间以及出现的顺序。研究人员利用细菌细胞记录了暴露于病毒、光线、抗生素和营养物质等刺激物时产生的信号，在人类HEK293T细胞中也实现了多路复合记录，并且仅观察到不到0.07%的脱靶率，其精确度更高，在包含仅10个细胞的样本也能可靠记录信号。针对该系统需要进一步证实这种CAMERA技术在活的动物体内的细胞中也能够发挥作用，将来可应用于记录低丰度细胞外和细胞内信号的存在，构建复杂的细胞状态图，进一步实时监测细胞状态，获取包括细胞分裂、谱系分化、代谢衰老、癌症发生和疾病发展等过程的重要信息。

Farzadfard等^[44]构建的活细胞DNA分子记录仪——基于DNA的有序记忆和迭代网络运算符(DNA-based ordered memory and iteration network operator, DOMINO)，利用基因编辑技术构建了一个单核苷酸分辨率的读写头，这个读写头由与胞苷脱氨酶融合的Cas9切口酶(Cas9 nickase, nCas9)

与尿嘧啶DNA糖基化酶抑制剂组成，用于在细菌和真核细胞中编码逻辑和记忆。nCas9用来读取，由gRNA指向特定的DNA靶标并将其切割，胞苷脱氨酶用来写入，而尿嘧啶DNA糖基化酶抑制剂可以通过阻断细胞修复机制提高写入效率。一旦12nt的gRNA序列定位到靶标，编辑器模块可以让靶标序列5'末端附近引入C-T突变，从而在DNA中产生永久记录；只有细胞中存在特定DNA序列时突变才会发生，通过测量这种突变即可得到细胞遇到的信号。输入信号量与输出结果呈线性关系，当输出结果为激活GFP表达时，通过定量GFP的表达水平，即可得知细胞内发生了多少突变，这就避免了测序破坏细胞。DNA中可以记录各种信号(具有生物或生理相关性)，包括糖[阿拉伯糖(Ara)]、重金属(Cu^{2+})和光，以及几种胃肠道炎症的生物标志物[血液(血红素)、过氧化氢和一氧化氮]，将为众多生物技术和生物医学应用建立计算和存储基因线路奠定基础。在研究中DOMINO可以设计出检测疾病(如癌症)相关基因的信号，输出激活治疗的基因(抗癌分子)，使系统能够检测和治疗疾病。

2.4.4 基因片段编辑

现阶段开发的DNA分子记录器都无法还原细胞内分子事件的顺序，或者局限于预先编程好的逻辑线路，无法实现多路复用记录。最近的报道中，Choi团队^[40]开发出一种新型DNA存储系统(DNA typewriter)(图6)，具有以下优点：①高度复用，可与至少数千个不同信号或事件类型的记录元件兼容；②将事件记录到DNA中的顺序和方向，因此能够明确地捕获所记录事件的准确顺序；

③在哺乳动物细胞中活跃。根据CRISPR-Cas系统切割^[60]和pgeRNA^[61]进行基因编辑的原理，作者设计了一个空白的“DNA磁带”(TAPE-1)，由5'端的key(GGA)和三个串联重复的单体组成。其中只有第一个位点的spacer是完整的20 bp(称为Type guide)，之后的两个spacer都在5'端截短至14 bp并与Type guide的4~17 bp重复，14 bp中4~6位的3个碱基为PAM。Type guide被pegRNA识别，通过引导编辑插入一段5 bp的序列，前2 bp

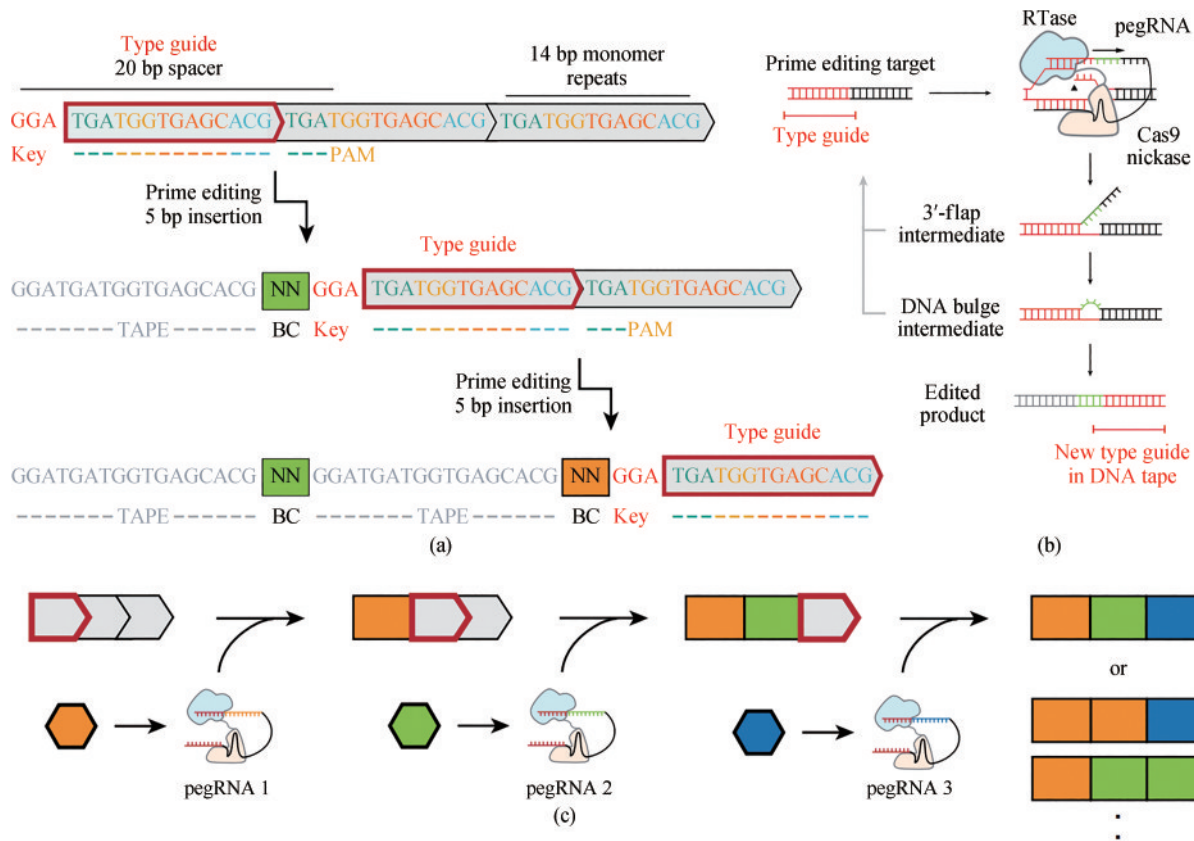


图6 用DNA typewriter进行序列基因组编辑^[40]

(a) Type guide上连续两个编辑序列件的示意图，它随着每个编辑序列的位置移动。DNA带由CRISPR-Cas9靶点(灰色方框)的串联序列组成，除第一个外，所有的靶点在其5'端被截断，因此不起作用。5 bp的插入包括一个2 bp的pegRNA特异条形码以及一个3 bp的key来激活下一个单体。因为基因组编辑在这个系统中是按顺序的，所以记录事件的时间顺序可以简单地通过它们沿着序列的物理顺序来读出。(b)用DNA typewriter引导编辑的示意图，引导编辑识别CRISPR-Cas9目标，并使用pegRNA指定的编辑对其进行修改。对于DNA typewriter，插入编辑序列会在随后的单体上生成新的编辑目标。(c)DNA typewriter顺序记录原理图，单个pegRNA与PE2酶一起，可以是事件驱动或者结构性表达的

Fig. 6 Sequential genome editing with DNA typewriter^[40]

(a) Schematic of two successive editing events at the type guide, which shifts in position with each editing event. The DNA tape consists of a tandem array of CRISPR - Cas9 target sites (grey boxes), all but the first of which are truncated at their 5' ends and therefore inactive. The 5-bp insertion includes a 2-bp pegRNA-specific barcode as well as a 3-bp key that activates the next monomer. Because genome editing is sequential in this scheme, the temporal order of recorded events can simply be read out by their physical order along the array. (b) Schematic of prime editing with DNA typewriter. Prime editing recognizes a CRISPR - Cas9 target and modifies it with the edit specified by the pegRNA. With DNA typewriter, an insertional editing event generates a new prime editing target at the subsequent monomer. (c) Schematic of ordered recording *via* DNA typewriter. Individual pegRNAs are potentially event-driven or constitutively expressed, together with the PE2 enzyme

的NN是pegRNA特有的barcode,后3bp是Type guide的key(GGA),这将破坏前一位点并激活下一位点,因为插入的key与随后一个单体,以及下一个单体的前6bp组成一个完整的20bp spacer和PAM,成为新的靶向位点。因此,每个连续的编辑记录了介导编辑的pegRNA的身份,同时还沿序列移动了一个单位的活性靶点的位置。在任何时刻,完整的spacer和邻近基序(PAM)只存在于序列中的一个位置,类似于打字机的“打字向导”。这项技术未来可以监测癌症患者在治疗期间接受连续的化疗剂量的反应,了解癌细胞如何对化疗造成的伤害作出反应,或者跟踪早期胚胎的发育进程。

3 CRISPR-Cas 系统在分子记录中的优势

CRISPR-Cas系统仅含两个功能元件:发挥核酸酶作用的Cas蛋白和引导RNA序列。该系统因靶向范围广、设计简单而迅速成为哺乳动物细胞基因编辑的首选方法^[62]。捕获外源基因不断扩大的新型和工程化的Cas蛋白不仅使哺乳动物基因组编辑功能日益强大,还为合成生物学家提供了创建基因线路的工具,对细胞表型和功能具有前所未有的控制能力^[63]。有条件的转录或转录后激活DNA编写组件,可以将给定DNA编写器的活动与感兴趣的信号联系起来^[64-65],例如,通过使用信号响应启动子,生物因子或环境因素的存在、持续时间、强度、顺序和时间的信息可以记录在DNA中^[39, 42, 66]。

Liu团队^[67]在基因特异性调控方面进行了深入研究,在利用合成生物学的思维构建基因线路,为分子记录仪在疾病监测和基础研究中的应用奠定了基础。哺乳动物miRNA不仅可以靶向细胞质中编码蛋白质的mRNA,还可以调节细胞核中的非编码RNA。利用天然miRNA支架开发人造miRNAs(amiRNA),以靶向不常见的RNA,不会扰乱细胞的正常活动状态,并且与短发夹状RNA(shRNAs)相比,提供更安全的RNAi表达载体^[67]。Huang等^[68]合成开发了一种amiRNA,通过与sgRNA结合,并联合RNAi增强剂依诺沙星,

对CRISPR系统进行定量抑制。该技术可以对CRISPR系统进行时空特异性调控,同时提高sgRNA靶向性并增强CRISPR系统的信息记录功能。

传统的癌症基因治疗策略特异性和效率有限,Liu等^[69]通过两个特异性启动子构建“与”门的逻辑线路,只有当两个启动子感应到的不同转录信号输入在被测试的细胞系中都活跃时,才会激活下游基因的输出,有效地特异性抑制膀胱癌细胞的生长,诱导细胞凋亡,大幅提高了识别记录肿瘤遗传信息的特异性和稳定性,为膀胱癌提供了一个检测和治疗的平台。进一步将sgRNA与识别特定信号的核糖开关融合,创建了基于CRISPR-Cas9的“信号导体”,调节内源基因的转录,以响应感兴趣的外部或内部信号(图7)。这些设备可以在哺乳动物细胞中执行逻辑信号操作,不需要多个遗传线路的分层。还可以用来通过构建连接不同信号通路的合成线路重新连接细胞信号事件。此外,这种方法可以通过控制同时的双向(开-关)基因转录来重定向致癌信号转导,从而使改变癌细胞的命运成为可能^[70]。

构建可控、灵敏的基因线路是实现细胞生理活动模拟调控的重要手段^[71-72],核酶开关是近年来发展起来的一种人工基因调控装置,由三个模块组成:传感器(检测特定的输入)、连接子(传递信息)和执行器(激活靶基因表达)^[73]。核酶与配体结合后改变其构象,诱导下游报告基因的表达,如荧光信号^[74-75],这种开关通过实现更好的基因表达、结构组装和基因特异性的可控性,改进了以前的蛋白质响应RNA设备^[76],通过串联多个引导RNA增加靶向目标和效果^[77]。基于CRISPR-Cas系统在基因线路中多变的调节方式,与感兴趣的特定事件联系起来,以创建生物记忆的合成基因线路是研究调控细胞信号的宝贵工具,尤其是在编码生物记忆以及进行谱系追踪和正向遗传筛选方面^[78]。

4 哺乳动物细胞存储在医学研究应用方面的潜力

合成存储设备除了基础研究,更令人期待的

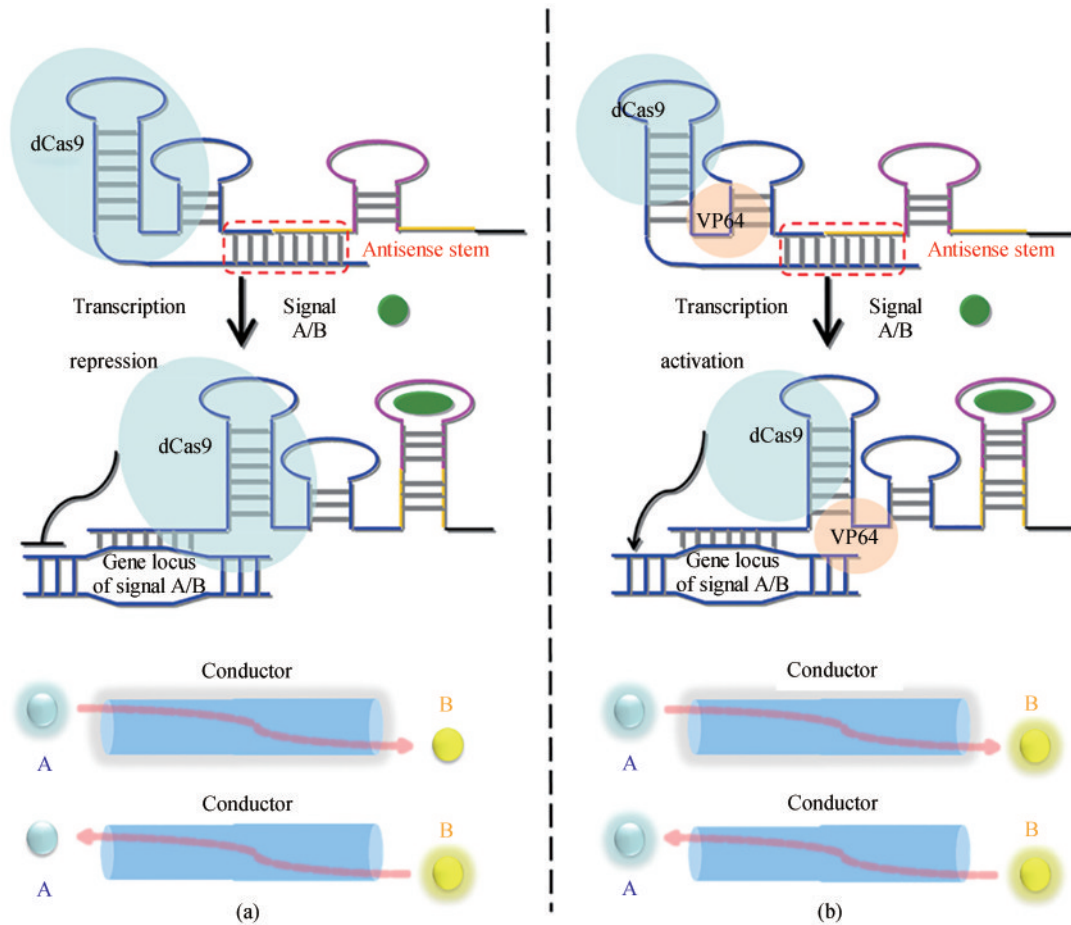


图7 基于CRISPR-Cas9的信号导体的设计^[70]

重新设计的sgRNAs用来灭活(a)或激活(b)基因表达的链置换机制的说明。sgRNA的反义序列显示为蓝色, 适配子茎显示为黄色。在没有信号A/B的情况下, sgRNA的引导区在反义茎内配对, sgRNA处于关闭状态。在信号A/B的存在下, 重新设计的sgRNA的构象被切换到“ON”状态。在这种状态下, sgRNA的导向区与其目标DNA结合, 从而分别通过dCas9-VP64融合蛋白和dCas9蛋白启动和关闭信号B/A的产生

Fig. 7 Design of the CRISPR-Cas9-based signal conductor to link one signal with another^[70]

General illustration of the strand-displacement mechanism by which the redesigned sgRNA acts to deactivate (a) or activate (b) gene expression. The antisense sequence of the sgRNA is shown in blue, and the aptamer stem is shown in yellow. In the absence of signal A/B, the guide region of sgRNA is paired within the antisense stem and the sgRNA is in the ‘off’ state. In the presence of signal A/B, the conformation of the redesigned sgRNA is switched to the ‘on’ state. In this state, the guide region of the sgRNA binds to its target DNA, and thus turns the production of signal B/A on and off through the dCas9-VP64 fusion protein and the dCas9 protein, respectively.

是在医学或工业生物技术等领域的前景。其他类型的合成线路和工程细胞已经被探索为治疗糖尿病、代谢综合征和癌症等疾病的潜在下一代疗法, 诱导对瞬时刺激的持续反应的能力将使新的治疗和诊断形式成为可能, 并对尚未解决的工业监测提供新方法。目前的医学治疗依赖于疾病的诊断, 然后才能实施适当的治疗, 通常, 患者将不得不在治疗过程中反复服用药物^[79]。

转录正反馈环可以耦合到不同的生物信号通路来检测细胞生物学中的瞬时变化, 也可以被设

计成对细胞外刺激做出反应, 例如辐射或缺氧, 这些设备将允许医生知道特定的刺激或条件是否已经发生, 或者曾经发生过。目前存储设备的输出多是容易检测到的报告, 未来可以针对疾病输出特异性的治疗措施, 这将使疾病的检测和治疗实现自动化, 缩短从发现病情开始到第一剂治疗之间的时间^[80]。基因编辑和基因线路工程的快速发展极大地提高了我们编程和研究哺乳动物细胞行为的能力, 这些进展已经推动了新一代合成生物学研究工具和潜在的治疗应用。可编程的DNA

结合结构域和RNA调控基因表达,更小的非破坏性的细胞记录元件的开发^[81],CRISPR-Cas系统在转录调控、碱基编辑和长片段编辑技术的进步,是更新细胞存储器的关键^[66, 82]。在脱离自然环境的情况下重建复杂的生物线路,可以加深我们对细胞信号通路的理解,基于合成生物学的信息实时记录系统还将引领基于细胞疗法、蛋白质药物、疫苗和基因疗法的创新治疗干预^[83]。

5 展望

在这些生物存储装置的设计、实施和分析过程中,仍有一些关键的技术挑战和知识差距需要解决。目前记录装置大多存在写入过程敏感性高而导致虚假信号的记录、诱导系统的泄露表达、需要大量细胞群体获取数据、记录信息量少、系统简单、细胞复制过程中DNA的突变等不足。在合成生物学家建立更加严谨调控、正交化的诱导系统后,我们期待灵敏度和复杂度更高的细胞记录装置的开发。这些技术将进一步提高我们以动态、纵向和多路复用的方式操纵生命的自然记忆存储介质的能力。在原核生物和真核细胞中都成功地记录了从小分子到免疫信号再到光的各种生物信号,然而,这些记录主要应用于体外环境并依赖于群体读出,未来需要改进单细胞记录技术,或展示分子记录器在活体动物中的变革性使用。虽然已经取得了实质性的进展,但仍有很大的空间来改进现有的存储体系结构或开发具有所需特征的新的存储体系结构,特别是在记录容量、可变性、稳定性、适应效果、写入周期、时间分辨率和记录动力学方面,以及实现分子记录器在活体条件下的使用。利用DNA重组和CRISPR基因组编辑的方法学正在迅速朝着这一目标改进,这些进展可以作为基础研究、生物技术和医学中对原位生物学研究的新一代强有力的方法。基于CRISPR的基因线路可以作为一种便携式和可改造的存储器,能针对携带在任何可移动遗传元件上的DNA(例如工程化噬菌体^[84-85])进行信息记录,并使用非自然碱基来扩大记录的信息密度和容量^[86]。科学家寄希望于未来可利用其进行长时间的分子事件记录和监控,从而进一步成为人类

研究癌症、衰老和发育发展等的强大工具。未来,人工设计的细胞记录装置可以发展成益生菌,稳定地存在于肠道中^[87],为人体群体进行健康监测和个性化治疗提供服务。

参 考 文 献

- [1] CAMERON D E, BASHOR C J, COLLINS J J. A brief history of synthetic biology[J]. *Nature Reviews Microbiology*, 2014, 12(5): 381-390.
- [2] KHALIL A S, COLLINS J J. Synthetic biology: applications come of age[J]. *Nature Reviews Genetics*, 2010, 11(5): 367-379.
- [3] MOON T S, LOU C B, TAMSIR A, et al. Genetic programs constructed from layered logic gates in single cells[J]. *Nature*, 2012, 491(7423): 249-253.
- [4] STRICKER J, COOKSON S, BENNETT M R, et al. A fast, robust and tunable synthetic gene oscillator[J]. *Nature*, 2008, 456(7221): 516-519.
- [5] FRIEDLAND A E, LU T K, WANG X, et al. Synthetic gene networks that count[J]. *Science*, 2009, 324(5931): 1199-1202.
- [6] WEI P, WONG W W, PARK J S, et al. Bacterial virulence proteins as tools to rewire kinase pathways in yeast and immune cells[J]. *Nature*, 2012, 488(7411): 384-388.
- [7] ROYBAL K T, RUPP L J, MORSUT L, et al. Precision tumor recognition by T cells with combinatorial antigen-sensing circuits[J]. *Cell*, 2016, 164(4): 770-779.
- [8] MAYS Z J, NAIR N U. Synthetic biology in probiotic lactic acid bacteria: at the frontier of living therapeutics[J]. *Current Opinion in Biotechnology*, 2018, 53: 224-231.
- [9] SLOMOVIC S, COLLINS J J. DNA sense-and-respond protein modules for mammalian cells[J]. *Nature Methods*, 2015, 12(11): 1085-1090.
- [10] COURBET A, ENDY D, RENARD E, et al. Detection of pathological biomarkers in human clinical samples *via* amplifying genetic switches and logic gates[J]. *Science Translational Medicine*, 2015, 7(289): 289ra83.
- [11] FENNO L E, MATTIS J, RAMAKRISHNAN C, et al. Targeting cells with single vectors using multiple-feature Boolean logic[J]. *Nature Methods*, 2014, 11(7): 763-772.
- [12] BOGORAD I W, LIN T S, LIAO J C. Synthetic non-oxidative glycolysis enables complete carbon conservation[J]. *Nature*, 2013, 502(7473): 693-697.
- [13] BERNABÉ-ORTS J M, QUIJANO-RUBIO A, VAZQUEZ-VILLAR M, et al. A memory switch for plant synthetic biology based on the phage ϕ C31 integration system[J]. *Nucleic Acids Research*, 2020, 48(6): 3379-3394.
- [14] JIN S, BAE J Y, SONG Y, et al. Synthetic biology on acetogenic bacteria for highly efficient conversion of C_1 gases to bio-

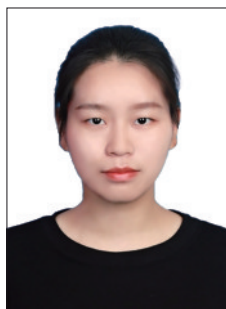
- chemicals[J]. *International Journal of Molecular Sciences*, 2020, 21(20): 7639.
- [15] BIRD L J, KUNDU B B, TSCHIRHART T, et al. Engineering wired life: synthetic biology for electroactive bacteria[J]. *ACS Synthetic Biology*, 2021, 10(11): 2808-2823.
- [16] CUBILLOS-RUIZ A, GUO T X, SOKOLOVSKA A, et al. Engineering living therapeutics with synthetic biology[J]. *Nature Reviews Drug Discovery*, 2021, 20(12): 941-960.
- [17] LIU Z H, WANG J Y, NIELSEN J. Yeast synthetic biology advances biofuel production[J]. *Current Opinion in Microbiology*, 2022, 65: 33-39.
- [18] SHENDURE J, JI H. Next-generation DNA sequencing[J]. *Nature Biotechnology*, 2008, 26(10): 1135-1145.
- [19] HEATHER J M, CHAIN B. The sequence of sequencers: the history of sequencing DNA[J]. *Genomics*, 2016, 107(1): 1-8.
- [20] FARZADFARD F, LU T K. Emerging applications for DNA writers and molecular recorders[J]. *Science*, 2018, 361(6405): 870-875.
- [21] SONG X, REIF J. Nucleic acid databases and molecular-scale computing[J]. *ACS Nano*, 2019, 13(6): 6256-6268.
- [22] CEZE L, NIVALA J, STRAUSS K. Molecular digital data storage using DNA[J]. *Nature Reviews Genetics*, 2019, 20(8): 456-466.
- [23] HAM T S, LEE S K, KEASLING J D, et al. Design and construction of a double inversion recombination switch for heritable sequential genetic memory[J]. *PLoS One*, 2008, 3(7): e2815.
- [24] CHURCH G M, GAO Y, KOSURI S. Next-generation digital information storage in DNA[J]. *Science*, 2012, 337(6102): 1628.
- [25] FARZADFARD F, LU T K. Synthetic biology. Genomically encoded analog memory with precise *in vivo* DNA writing in living cell populations[J]. *Science*, 2014, 346(6211): 1256272.
- [26] OISHI K, KLAVINS E. Framework for engineering finite state machines in gene regulatory networks[J]. *ACS Synthetic Biology*, 2014, 3(9): 652-665.
- [27] AALIPOUR A, CHUANG H Y, MURTY S, et al. Engineered immune cells as highly sensitive cancer diagnostics[J]. *Nature Biotechnology*, 2019, 37(5): 531-539.
- [28] JIANG K Y, KOOB J, DAWN CHEN X, et al. Programmable eukaryotic protein synthesis with RNA sensors by harnessing ADAR[J]. *Nature Biotechnology*, 2022: 1-10.
- [29] 杨璐, 吴楠, 白茸茸, 等. 基因回路型全细胞微生物传感器的设计、优化与应用[J]. *合成生物学*, 2022, 3(6): 1061-1080.
- YANG L, WU N, BAI R R, et al. Design, optimization and application of whole-cell microbial biosensors with engineered genetic circuits[J]. *Synthetic Biology Journal*, 2022, 3(6): 1061-1080.
- [30] KRETZSCHMAR K, WATT F M. Lineage tracing[J]. *Cell*, 2012, 148(1/2): 33-45.
- [31] WAGNER D E, KLEIN A M. Lineage tracing meets single-cell omics: opportunities and challenges[J]. *Nature Reviews Genetics*, 2020, 21(7): 410-427.
- [32] MEACHAM C E, MORRISON S J. Tumour heterogeneity and cancer cell plasticity[J]. *Nature*, 2013, 501(7467): 328-337.
- [33] WEINREB C, RODRIGUEZ-FRATICELLI A, CAMARGO F D, et al. Lineage tracing on transcriptional landscapes links state to fate during differentiation[J]. *Science*, 2020, 367(6479): eaaw3381.
- [34] SNIPPERT H J, VAN DER FLIER L G, SATO T, et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells[J]. *Cell*, 2010, 143(1): 134-144.
- [35] MCKENNA A, FINDLAY G M, GAGNON J A, et al. Whole-organism lineage tracing by combinatorial and cumulative genome editing[J]. *Science*, 2016, 353(6298): aaf7907.
- [36] MO C Y, GUO J X, QIN J C, et al. Single-cell transcriptomics of LepR-positive skeletal cells reveals heterogeneous stress-dependent stem and progenitor pools[J]. *The EMBO Journal*, 2022, 41(4): e108415.
- [37] HE L J, LI Y, LI Y, et al. Enhancing the precision of genetic lineage tracing using dual recombinases[J]. *Nature Medicine*, 2017, 23(12): 1488-1498.
- [38] TANNA T, SCHMIDT F, CHEREPKOVA M Y, et al. Recording transcriptional histories using Record-seq[J]. *Nature Protocols*, 2020, 15(2): 513-539.
- [39] TANG W X, LIU D R. Rewritable multi-event analog recording in bacterial and mammalian cells[J]. *Science*, 2018, 360(6385): eaap8992.
- [40] CHOI J, CHEN W, MINKINA A, et al. A time-resolved, multi-symbol molecular recorder *via* sequential genome editing[J]. *Nature*, 2022, 608(7921): 98-107.
- [41] SPANJAARD B, HU B, MITIC N, et al. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars[J]. *Nature Biotechnology*, 2018, 36(5): 469-473.
- [42] PERLI S D, CUI C H, LU T K. Continuous genetic recording with self-targeting CRISPR-Cas in human cells[J]. *Science*, 2016, 353(6304): aag0511.
- [43] HE Z S, MAYNARD A, JAIN A, et al. Lineage recording in human cerebral organoids[J]. *Nature Methods*, 2022, 19(1): 90-99.
- [44] FARZADFARD F, GHARAEI N, HIGASHIKUNI Y, et al. Single-nucleotide-resolution computing and memory in living cells[J]. *Molecular Cell*, 2019, 75(4): 769-780.e4.
- [45] SHETH R U, WANG H H. DNA-based memory devices for recording cellular events[J]. *Nature Reviews Genetics*, 2018, 19(11): 718-732.
- [46] AUSLÄNDER S, FUSSENEGGER M. Synthetic biology. Dynamic genome engineering in living cells[J]. *Science*, 2014, 346(6211): 813-814.

- [47] ATKINSON M R, SAVAGEAU M A, MYERS J T, et al. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*[J]. *Cell*, 2003, 113(5): 597-607.
- [48] GARDNER T S, CANTOR C R, COLLINS J J. Construction of a genetic toggle switch in *Escherichia coli*[J]. *Nature*, 2000, 403(6767): 339-342.
- [49] BIANCALANI T, ASSAF M. Genetic toggle switch in the absence of cooperative binding: exact results[J]. *Physical Review Letters*, 2015, 115(20): 208101.
- [50] BHOMKAR P, MATER I W, WISHART D S. The bacterial nanorecorder: engineering *E. coli* to function as a chemical recording device[J]. *PLoS One*, 2011, 6(11): e27559.
- [51] KOTULA J W, KERNS S J, SHAKET L A, et al. Programmable bacteria detect and record an environmental signal in the mammalian gut[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, 111(13): 4838-4843.
- [52] MIMEE M, TUCKER A C, VOIGT C A, et al. Programming a human commensal bacterium, *Bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota[J]. *Cell Systems*, 2015, 1(1): 62-71.
- [53] BONNET J, YIN P, ORTIZ M E, et al. Amplifying genetic logic gates[J]. *Science*, 2013, 340(6132): 599-603.
- [54] SIUTI P, YAZBEK J, LU T K. Synthetic circuits integrating logic and memory in living cells[J]. *Nature Biotechnology*, 2013, 31(5): 448-452.
- [55] ROQUET N, SOLEIMANY A P, FERRIS A C, et al. Synthetic recombinase-based state machines in living cells[J]. *Science*, 2016, 353(6297): aad8559.
- [56] LAMPSON B C, INOUE M, INOUE S. Retrons, msDNA, and the bacterial genome[J]. *Cytogenetic and Genome Research*, 2005, 110(1/2/3/4): 491-499.
- [57] LIM D B, MAAS W K. Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B[J]. *Cell*, 1989, 56(5): 891-904.
- [58] SCHMIDT F, CHEREPKOVA M Y, PLATT R J. Transcriptional recording by CRISPR spacer acquisition from RNA[J]. *Nature*, 2018, 562(7727): 380-385.
- [59] SCHMIDT F, ZIMMERMANN J, TANNA T, et al. Noninvasive assessment of gut function using transcriptional recording sentinel cells[J]. *Science*, 2022, 376(6594): eabm6038.
- [60] AMITAI G, SOREK R. CRISPR-Cas adaptation: insights into the mechanism of action[J]. *Nature Reviews Microbiology*, 2016, 14(2): 67-76.
- [61] ANZALONE A V, RANDOLPH P B, DAVIS J R, et al. Search-and-replace genome editing without double-strand breaks or donor DNA[J]. *Nature*, 2019, 576(7785): 149-157.
- [62] DOUDNA J A, CHARPENTIER E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9[J]. *Science*, 2014, 346(6213): 1258096.
- [63] MANGHWAR H, LINDSEY K, ZHANG X L, et al. CRISPR/cas system: recent advances and future prospects for genome editing[J]. *Trends in Plant Science*, 2019, 24(12): 1102-1125.
- [64] ZHAN H J, XIAO L L, LI A L, et al. Engineering cellular signal sensors based on CRISPR-sgRNA reconstruction approaches[J]. *International Journal of Biological Sciences*, 2020, 16(8): 1441-1449.
- [65] NISSIM L, PERLI S D, FRIDKIN A, et al. Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells[J]. *Molecular Cell*, 2014, 54(4): 698-710.
- [66] ZHAN H J, LI A L, CAI Z M, et al. Improving transgene expression and CRISPR-Cas9 efficiency with molecular engineering-based molecules[J]. *Clinical and Translational Medicine*, 2020, 10(6): e194.
- [67] NIU Q W, LIN S S, REYES J L, et al. Expression of artificial microRNAs in transgenic *Arabidopsis thaliana* confers virus resistance[J]. *Nature Biotechnology*, 2006, 24(11): 1420-1428.
- [68] HUANG X B, CHEN Z C, LIU Y C. RNAi-mediated control of CRISPR functions[J]. *Theranostics*, 2020, 10(15): 6661-6673.
- [69] LIU Y C, ZENG Y Y, LIU L, et al. Synthesizing AND gate genetic circuits based on CRISPR-Cas9 for identification of bladder cancer cells[J]. *Nature Communications*, 2014, 5: 5393.
- [70] LIU Y C, ZHAN Y H, CHEN Z C, et al. Directing cellular information flow via CRISPR signal conductors[J]. *Nature Methods*, 2016, 13(11): 938-944.
- [71] SLOMOVIC S, PARDEE K, COLLINS J J. Synthetic biology devices for *in vitro* and *in vivo* diagnostics[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2015, 112(47): 14429-14435.
- [72] HAELLMAN V, FUSSENEGGER M. Synthetic biology—toward therapeutic solutions[J]. *Journal of Molecular Biology*, 2016, 428(5 Pt B): 945-962.
- [73] WALTER J G, STAHL F. Aptazymes: expanding the specificity of natural catalytic nucleic acids by application of *in vitro* selected oligonucleotides[J]. *Advances in Biochemical Engineering/Biotechnology*, 2020, 170: 107-119.
- [74] YOKOBAYASHI Y. Aptamer-based and aptazyme-based riboswitches in mammalian cells[J]. *Current Opinion in Chemical Biology*, 2019, 52: 72-78.
- [75] WANG X Y, DONG K L, KONG D Q, et al. A far-red light-inducible CRISPR-Cas12a platform for remote-controlled genome editing and gene activation[J]. *Science Advances*, 2021, 7(50): eabh2358.
- [76] KAVITA K, BREAKER R R. Discovering riboswitches: the past and the future[J]. *Trends in Biochemical Sciences*, 2023, 48(2): 119-141.

- [77] LIU Y C, HAN J H, CHEN Z C, et al. Engineering cell signaling using tunable CRISPR - Cpf1-based transcription factors[J]. Nature Communications, 2017, 8: 2095.
- [78] KIM T, LU T K. CRISPR/Cas-based devices for mammalian synthetic biology[J]. Current Opinion in Chemical Biology, 2019, 52: 23-30.
- [79] WIELAND M, FUSSENEGGER M. Engineering molecular circuits using synthetic biology in mammalian cells[J]. Annual Review of Chemical and Biomolecular Engineering, 2012, 3: 209-234.
- [80] INNIS M C, SILVER P A. Building synthetic memory[J]. Current Biology, 2013, 23(17): R812-R816.
- [81] ZHAN H J, ZHOU Q, GAO Q J, et al. Multiplexed promoterless gene expression with CRISPRReader[J]. Genome Biology, 2019, 20(1): 113.
- [82] ANZALONE A V, KOBLAN L W, LIU D R. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors[J]. Nature Biotechnology, 2020, 38(7): 824-844.
- [83] LIENERT F, LOHMUELLER J J, GARG A, et al. Synthetic biology in mammalian cells: next generation research tools and therapeutics[J]. Nature Reviews Molecular Cell Biology, 2014, 15(2): 95-107.
- [84] LU T K, COLLINS J J. Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy[J]. Proceedings of the National Academy of Sciences of the United States of America, 2009, 106(12): 4629-4634.
- [85] ORTIZ M E, ENDY D. Engineered cell-cell communication via DNA messaging[J]. Journal of Biological Engineering, 2012, 6(1): 16.
- [86] ZHANG Y, PTACIN J L, FISCHER E C, et al. A semi-synthetic organism that stores and retrieves increased genetic information[J]. Nature, 2017, 551(7682): 644-647.
- [87] HO J M L, BENNETT M R. Improved memory devices for synthetic cells[J]. Science, 2018, 360(6385): 150-151.



通讯作者: 刘宇辰(1988—),男,副研究员,研究生导师、博士后合作导师。主要研究方向:(1)肿瘤生物治疗和医学合成生物学;(2)创新肿瘤治疗新方法;(3)构建人工基因线路,肿瘤精准治疗。
E-mail: liuyuchenmdcg@163.com



第一作者: 马孟丹(1997—),女,硕士研究生。主要研究方向为医学合成生物学。
E-mail: mamengdan7@163.com