

DOI: 10.13376/j.cblls/2023042

文章编号: 1004-0374(2023)03-0347-05

睿 ● 观 ● 家

打开“生命之书”——DNA双螺旋发现70周年记

吴家睿^{1,2}

(1 中国科学院分子细胞科学卓越创新中心, 上海 200031; 2 上海交通大学安泰经济与管理学院, 上海 200030)

生物体与非生物体之根本差异在于, 生物体能够保存和利用信息, 并把自身的信息一代又一代地传递下去。从古至今, 人们一直试图阐明生物体这种承载和传递信息的能力。这个梦想在 1953 年终于得到了实现——美国科学家沃森 (Watson J) 和英国科学家克里克 (Crick F) 提出了 DNA 双螺旋模型, 借此在分子水平上清晰地解释了生物体是如何保存和传递信息; 而分子生物学也在此基础上诞生, 现代生命科学的“帷幕”也就此正式拉开。70 年过去了, 人们从 DNA 双螺旋的研究中取得了哪些重要的突破?

1 “生命之书”的阅读

研究者从 DNA 双螺旋模型中认识到, 生命的底层逻辑是“文字”, 即生命用“A”“T”“G”“C”4 种碱基作为基本单元连接成长长的多核苷酸链, 这样的两条多核苷酸链相互缠绕而形成 DNA 双螺旋。这 4 种碱基正是记录生命遗传信息的“字母”, 其中每 3 个碱基组成一个类似“单词”的遗传密码子, 对应一个特定的氨基酸。从 4 种碱基中选择 3 种来构成 1 个密码子的总数为 64, 所以生物体拥有 64 个密码子“单词”。由于生物体用来合成蛋白质的天然氨基酸仅有 20 种, 因此除了甲硫氨酸和色氨酸分别由一个密码子决定以外, 其他 18 种氨基酸分别对应 2~6 个密码子。生物体内的“基因”正是一段由多种密码子连接起来的碱基“句子”, 用来指导一种蛋白质的氨基酸组成和排序。

由此可见, 大自然在创造生命时采用了写书的方式, 自然界的万千生命种类就如同万千本书——只需要 4 个碱基作为基本的字母, 就可以创作出无数的作品。因此, 研究者的主要工作通常就是阅读“生命之书”中的一段段碱基序列组成的基因“句子”, 并揭示这些句子的含义和可能的生物学功能。

1.1 识文断字——编码基因的认识

研究者最初认为, 一个编码基因用于指导一种蛋白质的合成。但随着“阅读”工作的深入, 研究者发现, 虽然这种“一基因一蛋白”的观点在大肠杆菌等简单生物体上基本成立, 但在动植物等复杂生物体上情况就明显不一样了——这类生物体拥有大量的“断裂基因”, 即在一个编码蛋白质的碱基序列之间插有若干段不编码蛋白的碱基序列, 其基因内编码蛋白质的碱基序列称为“外显子”(Exon), 而不编码的碱基序列则称为“内含子”(Intron)。换句话说, “断裂基因”就好比一个完整的“语句”被拆开, 在中间加入了一些无意义的文字。研究者还注意到, 生命形式越高等, 断裂基因就越多。如在单细胞真核生物“芽殖酵母”里, 只有 4% 的基因拥有内含子; 而在小鼠或人的基因组内, 绝大部分基因都拥有内含子。

对拥有众多断裂基因的复杂生物体而言, 一个断裂基因内多个外显子通常会受到不同的剪切和拼接, 称为可变剪接; 如果把不同的外显子用可变剪接方式进行连接, 一个基因就能制造出多种蛋白质。据统计, 人的编码蛋白质的转录链平均含有 8.7 个外显子, 小鼠的转录链则平均含有 8.4 个外显子。举一个极端的例子, 在小鼠的基因组内有一个称为 *DSCAM* 的基因, 它有 6.1 万个碱基, 通过剪接可以形成含有 24 个外显子, 长为 7800 个碱基的 mRNA。但这个基因的外显子有许多不同的剪接方式, 如果把所有可能的剪接方式都考虑进去, 这个基因能够产生的 mRNA 和蛋白质的种类, 可以超过 3.8 万种 (小鼠基因组的全部基因数也不过是 2.2 万左右)^[1]

研究者最初认为非编码的内含子序列没有什么功能。但许多实验表明, 内含子在生命活动中也发挥着重要的作用。例如, 有研究发现, 内含子能够促进拥有它的断裂基因的活性; 此外, 部分内含子序列可以用来产生微小 RNA (microRNA, miRNA)

等非编码 RNA 调控元件。不久前, 加拿大科学家通过对酵母细胞基因组内所有内含子的逐一敲除研究, 发现多数内含子都可以用来调控细胞对营养匮乏的响应^[2]。美国科学家也同时揭示出, 酵母的 34 个内含子在剪切后继续稳定地存在于细胞内, 并在外部压力条件下用来帮助细胞生存^[3]。

转录产生的 mRNA 剪切还导致了一类环状 RNA (circular RNA) 的形成, 即前体 mRNA (pre-mRNA) 除了采用常规剪切产生成熟的 mRNA 之外, 还可以在特异性反向剪切作用下, 把下游 3' 剪接位点在一个或多个外显子上以相反的顺序与上游 5' 剪接位点连接, 形成环状 RNA^[4]。这类环状 RNA 广泛存在于各种生物体内; 有实验表明: 一个 HeLa 细胞内通常就含有总数为 1 万个拷贝数的环状 RNA^[5]。重要的是, 环状 RNA 还广泛地参与到各种生命活动, 包括调控基因转录和蛋白质合成, 影响细胞内的多种生物学过程, 并参与到机体的免疫调控等多种生理或病理活动中^[4]。

1.2 察言观色——表观遗传的发现

“生命之书”拥有众多编码各种蛋白质的基因, 生物体必须精准地控制这些基因, 以便在需要时制造出特定的蛋白质。为此, 生命在基因到蛋白质合成的过程中增加了一个中间步骤, 称为“转录”——把基因的碱基序列之信息转录到一条称为“信使 RNA” (mRNA) 的碱基序列之上, 然后用 mRNA 作为模板指导蛋白质合成。也就是说, 生命通过转录方式选择性地“阅读”基因。近年来的研究发现, DNA 上的许多碱基序列通常被化学基团进行修饰, 其中最常见的是甲基化修饰, 即 DNA 甲基转移酶在多核苷酸链的胞嘧啶“C”上用共价键结合一个甲基基团 (CH₃)。而 DNA 甲基化修饰的主要功能正是调控基因的转录活动, 或促进某个基因的转录, 或抑制其转录。以色列科学家最近发布了目前最大规模的正常人 39 种细胞类型的 DNA 甲基化图谱; 他们发现, 同种细胞类型的 DNA 甲基化模式在不同个体之间高度一致; 这意味着 DNA 甲基化模式影响着特定的细胞分化和细胞类型的维持^[6]。

染色质上的组蛋白修饰同样在基因转录调控中发挥着重要作用。组蛋白 H3 和 H4 的乙酰化修饰可使得染色质结构更为开放, 从而有利于基因的转录; 而它们的去乙酰化则会抑制转录的进行。由于 DNA 和组蛋白的化学修饰通常受到机体内外环境的调节, 从而成为了生命响应环境变化来控制基因转录的重要手段。例如, 研究者发现, 一种组蛋白

去甲基化酶 KDM6B 调控一种海龟的温度依赖型性别决定——在温度高时 (如 32 °C), 这个酶不活跃, 组蛋白 H3 保持着甲基化状态, 这种海龟发育为雌龟; 在温度低时 (如 26 °C), 该酶被激活, 然后将雄性性腺分化的关键基因 *Dmrt1* 启动子区域组蛋白 H3 上的甲基化修饰基团去除, 使得 *Dmrt1* 转录, 进而导致了雄龟的形成^[7]。

最重要的是, 这些控制基因转录活动的化学修饰往往可以通过细胞分裂的过程传递给子代细胞。显然, 在这种基因的化学修饰的传递过程中, DNA 碱基序列本身并没有发生改变。人们把这种 DNA 碱基序列的化学修饰信息传递之现象称为表观遗传; 由此产生了一门新学科: 表观遗传学 (Epigenetics)。如果把 DNA 碱基序列的各种化学修饰视为不同的颜色, 那么生命之“书”就不再是最初人们认识到单色印刷本, 而是一本五颜六色的彩色图书——不仅用碱基序列写出的“文字”可以被复印和传递, 而且用化学基团涂抹在这本“书”里的各种“颜色”也可以被复印和传递。

表观遗传现象的发现不仅打破了碱基序列是生命遗传活动唯一载体的教条, 而且被一系列研究证明表观遗传信息在某些情况下可以从亲代遗传给子代。例如, 在斑马鱼受精卵的发育过程中, 来自父本的 DNA 甲基化修饰模式一直保持, 直至胚胎发育的囊胚期才被消除重建; 而来自母本染色体的甲基化修饰模式则在胚胎发育的初期就很快被消除, 然后在这些母本染色体上按照父本 DNA 甲基化修饰模式进行了重建^[8]。也就是说, 这些源自精子的 DNA 甲基化修饰模式可以被遗传到子代, 并用来指导子代的胚胎早期发育。最近的一项研究指出, 在小鼠体内“*Ankrd26*”和“*Ldlr*”两个基因的启动子上人为产生的甲基化状态能够遗传给子代, 且至少可以稳定地遗传到 F6 代; 重要的是, 这两个基因的人工甲基化所诱发的表型也可以在子代中稳定继承^[9]。

由于生物体所在的环境是控制表观遗传修饰的主要因素, 因此表观遗传活动把生命的开放性提升到了一个新的高度, 使得外部环境的信息通过表观遗传修饰与机体内的 DNA 碱基序列上信息进行整合, 不仅能够影响个体的生理和病理活动, 还能够传递给下一代。

2 “生命之书”的编写

人类认识自然的一个主要目的是要改造自然,

为人类服务。同样，人类也不会停留在对生命的“阅读”阶段，而是努力地发展各种新技术去改造生命。为了改写“生命之书”，研究者发展出了二种基本工具：识别“字母”——测定DNA上4种碱基的排列顺序；编辑“文字”——剪切和连接DNA上的碱基序列。

DNA测序技术是用来识别“生命之书”里“碱基文字”最主要的技术。早在1977年，生命科学界在测量DNA碱基序列上就取得了重大突破，其中英国科学家桑格(Sanger F)发明了酶法，美国科学家吉尔伯特(Gilbert W)发明了化学法；同年桑格还利用其技术测定了第一个生物体——噬菌体X174——的全基因组序列，共5375个碱基。桑格的测序技术被称为第一代DNA测序技术。不过，当今生命科学研究的的主流是第二代测序技术，其主要特点是检测通量高；它不仅大大降低了测序成本，而且还明显提高了测序速度，用一代测序技术完成一个人类基因组30亿个碱基的测序需要3年左右的时间，而使用二代测序技术则可能在1周内即可完成。不久前研究者又开发出了能够检测单个多核苷酸分子的第三代测序技术。

在基因编辑领域，研究者开发出来的“基因剪刀”是各式各样的核酸酶，其中最常用的是“限制性内切酶”；这类酶能够识别DNA上特定的碱基序列，从而找到准确的剪切位点并实现DNA链内的定点切割。2012年，美国科学家通过改造细菌的核酸酶系统发展出一种全新的“CRISPR-Cas9”技术，现已成为进行基因编辑最强有力的工具。此外，研究者也有针对性地发展出了若干能够把DNA链间缺口连接起来的DNA连接酶。

2.1 移花接木——基因工程的诞生

20世纪70年代，随着DNA限制性内切酶的发现，研究者开始了“生命之书”的编写工作，并把这类在分子水平上对DNA碱基序列进行操作的技术称为“基因工程”(genetic engineering)——通常是将外源基因转入到受体细胞，从而使其特性发生改变或产生新的性状。在这个过程中，首先是利用限制性内切酶把一种生物体(供体)DNA上的特定基因切下来，将其与质粒或病毒等载体上的DNA在体外人工连接而形成新的重组DNA，然后转送到另一种生物体(受体)中进行扩增和表达。

基因工程的诞生，为人类的福祉做出了巨大的贡献。生产治疗糖尿病的胰岛素就是一个典型范例。2022年是胰岛素用于临床治疗的第100周年。第一

代医用胰岛素主要是源于牛或猪的胰腺提取物。这种生产方式不仅比较昂贵，而且产量很低，远不能满足临床需求。1981年，美国研究者利用基因工程技术，将人的胰岛素基因导入大肠杆菌，通过大肠杆菌大量生产重组人胰岛素。从此，第一代胰岛素产品完全被这个第二代产品所取代。为了进一步提高胰岛素的疗效和安全性，研究者又对胰岛素基因进行精细的改造，获得了第三代产品——重组胰岛素类似物。

基因工程在农作物领域同样发挥了巨大的作用。当前，“转基因作物”已经深入到人们生活的许多方面。1983年，科学家培育出了第一个转基因作物——转基因烟草；到2002年，世界上大约有550万~600万烟农种植转基因烟草。1996年，美国的农场主开始种植一种转基因大豆——在这种大豆里转入了植物“矮牵牛”的一种抗性基因，从而可以抵抗杀草剂。2021年，美国转基因大豆种植面积占美国大豆种植面积的95%，总产量为1.2亿吨。

2.2 推陈出新——基因编辑的迭代

虽然基于DNA限制性内切酶的基因工程技术取得了许多突出的成果，但是该技术在基因编辑的应用中也表现出来一些明显的不足，如实验流程比较复杂、结果获取需时较长、编写能力不够精准等。为此，研究者一直在努力开发更好的基因编辑技术，直至2012年，美国科学家杜德娜(Doudna JA)和卡彭特(Charpentier E)在细菌的基因编辑系统基础上发展出了一种理想的基因编辑技术——英文的缩写名称为“CRISPR-Cas9”。这种技术的基本原理是：利用一段设计好的RNA序列“sgRNA”引导DNA内切酶“Cas9”至特定的DNA序列上进行剪切。该技术经过不断的“迭代”，目前不仅可以在细胞的基因组特定位置内插入长达36000个碱基的外源DNA片段^[10]，而且可以在细胞内的DNA乃至RNA的序列上进行单个碱基的修改^[11-12]。

这类新型基因编辑技术的出现显著提升了人类抗击疾病的能力，尤其为治疗基因碱基序列异常的遗传性疾病提供了有力的武器。据统计，目前已知的单基因遗传病超过9000种，对人类的健康造成了巨大的危害。例如，血红蛋白基因异常能造成“地中海贫血”，目前全球有近3.5亿“地中海贫血”基因的携带者；中国长江以南各省是该病的高发区，在部分高发地区，这种基因的携带者在人群中超过10%。过去没有好的治疗药物或方法，重度的地中

海贫血患者只能定期进行输血,而现在则有望利用 CRISPR 技术来治疗地中海贫血;目前全球进入临床研究阶段的地中海贫血基因编辑治疗产品已知的有 6 个,其中 5 个是采用 CRISPR 技术。美国科学家杜德娜在关于 CRISPR 技术的最新综述中进一步指出,“CRISPR 在治疗某些特定疾病方面的潜力已经得到了很多的关注,但我更相信,在未来几年我们将会看到 CRISPR 在疾病预防方面有着全新的应用”^[13]。

值得注意的是,CRISPR 技术在改良农作物品种方面也同样有着巨大的优势。它不仅可以利用外源基因来改造作物,而且能够把作物自身的基因按照设计好的目标直接进行“改写”。中国科学家高彩霞 2013 年在世界上首次报道了利用 CRISPR 技术编辑水稻和小麦 DNA 序列的研究工作^[14],并在 2014 年报道了使用 CRISPR 技术修改了一种六倍体小麦的一个基因序列,从而使编辑后的小麦能够抵抗小麦白粉病的侵袭^[15]。可以说,在 CRISPR 等一系列基因编辑技术的推动下,传统的那种不可控的作物自然育种在未来将转变为高度可控的作物设计育种。正是在这样的形势推动下,中国国家自然科学基金委员会和中国科学院不久前联合发起了“未来作物设计项目”^[16]。

3 “生命之书”的创作

人工合成生命始终是研究者在探索生命奥秘过程中一个不灭的梦想。早在 1828 年,德国化学家维勒 (Wolher F) 就在实验室中利用氰酸铵合成了尿素,首次证明了人们可以在实验室里利用简单的无机分子合成源自生物体内的有机化合物。20 世纪 60 年代,中国科学家在世界上首次人工合成了第一个具有生物活性的蛋白质——由 51 个氨基酸组成的牛胰岛素;于 1981 年又在世界上首次人工合成了具有生物活性的多核苷酸链——由 76 个核苷酸组成的酵母丙氨酸转移核糖核酸 (酵母 tRNA^{Ala})。随着科学技术的发展,研究者近年来在实验室合成生命的能力也在显著的提高,他们甚至试图创作出自然界不存在的“生命之书”。

3.1 按图索骥——试管里合成的生命

2010 年 5 月,美国生物学家文特尔 (Venter C) 发布了世界上首个“人造生命”——科学家依据一种最简单的细菌基因组序列之信息,首次全人工化学合成了这个完整的、长度超过 100 万个碱基的基因组;且含有这一人造基因组的人工细菌能够展现

出相应的生命活动^[17]。不久前,美国和中国等多国科学家联合发起的“合成酵母基因组计划”(Synthetic Yeast Genome Project, Sc2.0),提出了一个更为宏大的目标——把拥有大约 2 400 万个碱基的芽殖酵母基因组序列用化学合成方式复制出来;芽殖酵母基因组共有 16 条染色体,Sc2.0 在 2014 年完成了一条酵母染色体的化学合成,至 2017 年又完成了 5 条染色体的化学合成,中国科学家完成了其中 4 条。

20 世纪 90 年代启动的“人类基因组计划”(Human Genome Project, HGP) 的目标是,把人类自身这部由 30 亿个碱基构成的“天书”通过测序技术完整地“阅读”一遍。它的实施为生命科学和医学带来了革命性的变化。而在 2016 年 5 月,100 多位专家学者在美国哈佛大学开会,提出了要把人类基因组这部“天书”用化学合成的方式在试管里“写”出来,称为“人类基因组计划——编写版”(HGP-Write)^[18]。可以想见,未来一旦 HGP-Write 正式实施,必将对人类认识和控制自身带来更为深远的影响和意义。

3.2 奇思妙想——人造新生命的序曲

研究者的脚步并没有停留在按照自然界已有的“生命之书”进行复制,而是开始按照人的意愿来创作全新的“生命之书”。例如,美国生物学家文特尔在其化学全合成的首个“人造生命”上进行了设计和改造,把他认为冗余的或非必需的基因“句子”从该基因组上删除,化学合成了“微型细菌基因组”,原基因组序列 1 079 kb 在改造后被缩减至 531 kb,而基因数目也减至 473 个;这种压缩导致含有这个微型基因组的细胞自我繁殖的速度比具有全长基因组序列的细胞快了几乎 5 倍^[19]。

自然状态下的“生命之书”通常拥有 64 个“单词”——遗传密码子,用来指导 20 种天然氨基酸合成蛋白质,其中有一些密码子是冗余的,称为同义密码子。2019 年,英国科学家发布了一个“人造”大肠杆菌,它含有一个人工设计并化学合成的人造大肠杆菌基因组,其中全新设计并合成的人工基因“句子”大约包含 400 万个碱基对;在这些人工设计的基因里,只保留了 61 个遗传密码子,删除了编码丝氨酸的 2 个同义密码子 (TCG、TCA) 和 1 个终止密码子 (TAG);而这些改造过的基因序列仍然可以制造出正常的蛋白质^[20]。研究者进而将这个化学合成的大肠杆菌里对应 TCG 和 TCA 的 tRNA 以及相关的释放因子 (Release factor-1) 移除,使该人造菌株拥有了对噬菌体的完美抵抗力^[21];此外,研

究者还把该菌株里有意义密码子重新排布，合成了非天然的氨基酸聚合物和大环化合物^[21]。

研究者不仅试着修改基因“句子”和密码“单词”，而且在碱基“字母”上也做起了文章。2014年，美国研究者宣布他们在大肠杆菌的DNA序列里加入了两种非天然碱基——dNaM和dTPT3（被简称为X和Y碱基），这两个碱基也能够体内通过复制的方式传递信息；研究者认为，这种6个碱基系统理论上可以指导合成的氨基酸数目将从20种变成172种^[22]。而在2017年的论文里，这些研究者进一步证明，这两个插入到天然基因序列里的非天然碱基，能够在蛋白质合成中引入特定的天然氨基酸或非天然氨基酸^[23]。不久前，另外一组美国研究人员进一步开发出来添加了4个人工碱基的“8碱基”DNA双螺旋，并且通过结构分析技术证明：“新的DNA晶体结构完整，保持了正确的碱基配对，同时没有丢失天然DNA的关键性特征”；研究者同时还发展出了具有8个碱基的RNA分子^[24]。

4 小结

DNA双螺旋模型的发现让我们认识到，生命最根本的基石就是4个碱基“字母”及其相应的“书写”规则。这一发现为我们打开了“生命之书”的阅读之门，引导我们进入了比世界上任何一个图书馆藏书都要丰富的“生命之书”殿堂。

[参 考 文 献]

- [1] Black DL. Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell*, 2000, 103: 367-70
- [2] Parenteau J, Maignon L, Berthoumieux M, et al. Introns are mediators of cell response to starvation. *Nature*, 2019, 565: 612-7
- [3] Morgan JT, Fink GR, Bartel DP. Excised linear introns regulate growth in yeast. *Nature*, 2019, 565: 606-11
- [4] Liu CX, Chen LL. Circular RNAs: characterization, cellular roles, and applications. *Cell*, 2022, 185: 2016-34
- [5] Liu CX, Li X, Nan F, et al. Structure and degradation of circular RNAs regulate PKR activation in innate immunity. *Cell*, 2019, 177: 865-80
- [6] Loyfer N, Magenheimer J, Peretz A, et al. A DNA methylation atlas of normal human cell types. *Nature*, 2023, 613: 355-64
- [7] Ge C, Ye J, Weber C, et al. The histone demethylase KDM6B regulates temperature-dependent sex determination in a turtle species. *Science*, 2018, 360: 645-8
- [8] Jiang L, Zhang J, Wang J, et al. Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell*, 2013, 153: 773-84
- [9] Takahashi Y, Valencia MM, Yu Y, et al. Transgenerational inheritance of acquired epigenetic signatures at CpG islands in mice. *Cell*, 2023, 186: 715-31.e19
- [10] Yarnall MTN, Ioannidi EI, Schmitt-Ulms C, et al. Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat Biotechnol*, 2022, DOI: 10.1038/s41587-022-01527-4
- [11] Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, 2017, 551: 464-71
- [12] Cox DBT, Gootenberg JS, Abudayyeh OO, et al. RNA editing with CRISPR-Cas13. *Science*, 2017, 358: 1019-27
- [13] Wang JY, Doudna JA. CRISPR technology: a decade of genome editing is only the beginning. *Science*, 2023, 379: eadd8643
- [14] Shan Q, Wang Y, Li J, et al. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol*, 2013, 31: 686-88
- [15] Wang Y, Cheng X, Shan Q, et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol*, 2014, 32: 947-52
- [16] Tian Z, Wang J, Li J, et al. Designing future crops: challenges and strategies for sustainable agriculture. *Plant J*, 2021, 105: 1165-78
- [17] Gibson DG, Glass JI, Lartigue C, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 2010, 329: 52-6
- [18] Boeke JD, Church G, Hesse A, et al. The genome project-write. *Science*, 2016, 353: 126-7
- [19] Hutchison CA 3rd, Chuang RY, Noskov VN, et al. Design and synthesis of a minimal bacterial genome. *Science*, 2016, 351: aad6253
- [20] Fredens J, Wang K, Torre D, et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature*, 2019, 569: 514-8
- [21] Robertson WE, Funke LF, Torre D, et al. Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science*, 2021, 372: 1057-62
- [22] Malyshev DA, Dhami K, Lavergne T, et al. A semi-synthetic organism with an expanded genetic alphabet. *Nature*, 2014, 509: 385-8
- [23] Zhang Y, Ptacin JL, Fischer EC, et al. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature*, 2017, 551: 644-7
- [24] Hoshika S, Leal NA, Kim MJ, et al. Hachimoji DNA and RNA: a genetic system with eight building blocks. *Science*, 2019, 363: 884-7