

## 特约评述

DOI: 10.12211/2096-8280.2020-007

## 基因密码子拓展技术的方法原理和前沿应用研究进展

付宪<sup>1,2</sup>, 林涛<sup>1</sup>, 张帆<sup>3</sup>, 张惠铭<sup>1</sup>, 章文蔚<sup>1</sup>, 杨焕明<sup>1,4</sup>, 朱师达<sup>1,5</sup>, 徐讯<sup>1,2</sup>, 沈玥<sup>1,2,5,6</sup>

(<sup>1</sup> 深圳华大生命科学研究院, 广东 深圳 518083; <sup>2</sup> 广东省高通量基因组测序与合成编辑应用重点实验室, 深圳华大生命科学研究院, 广东 深圳 518120; <sup>3</sup> 中国科学院大学华大教育中心, 广东 深圳 518083; <sup>4</sup> (广东省) 华大基因合成基因组院士工作站, 深圳华大基因科技有限公司, 广东 深圳 518120; <sup>5</sup> 深圳创新分子诊断技术工程实验室, 深圳华大生命科学研究院, 广东 深圳 518120; <sup>6</sup> 深圳国家基因库, 广东 深圳 518120)

**摘要:** 自然界生物根据其高度保守的密码子表来对20种天然氨基酸进行基因编码, 这些种类有限的氨基酸构成了天然蛋白质合成的基本构筑单元。生物在漫长进化中通过改变蛋白质中氨基酸的排列顺序来丰富其结构与功能, 但上述过程是随机的, 缺乏可控性。拓展用于蛋白质合成的氨基酸种类亦可实现对蛋白质结构和功能的改变与操纵。通过对中心法则中翻译系统的设计与改造, 基因密码子拓展技术可将非天然氨基酸特异性地引入到细胞内目标蛋白的指定位点, 利用非天然氨基酸中特殊的官能团赋予目标蛋白新的物理化学性质, 最终达到蛋白质功能创新的目的。本文主要介绍基因密码子拓展技术中翻译工具开发和适配底盘改造研究相关的原理、技术和前沿进展, 讨论其在蛋白质功能调控、生物医药、生物防控等新兴领域应用中的成果进展与未来展望。

**关键词:** 非天然氨基酸; 密码子拓展系统; 翻译工具; 适配底盘细胞; 正交性

**中图分类号:** Q755 **文献标志码:** A

## Progress in the study of genetic code expansion related methods, principles and applications

FU Xian<sup>1,2</sup>, LIN Tao<sup>1</sup>, ZHANG Fan<sup>3</sup>, ZHANG Huiming<sup>1</sup>, ZHANG Wenwei<sup>1</sup>, YANG Huanming<sup>1,4</sup>, ZHU Shida<sup>1,5</sup>, XU Xun<sup>1,2</sup>, SHEN Yue<sup>1,2,5,6</sup>

(<sup>1</sup> BGI-Shenzhen, Shenzhen 518083, Guangdong, China; <sup>2</sup> Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen 518120, Guangdong, China; <sup>3</sup> BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, Guangdong, China; <sup>4</sup> Guangdong Provincial Academician Workstation of BGI Synthetic Genomics, BGI-Shenzhen, Shenzhen 518120, Guangdong, China; <sup>5</sup> Shenzhen Engineering Laboratory for Innovative Molecular Diagnostics, BGI-Shenzhen, Shenzhen 518120, Guangdong, China; <sup>6</sup> China National GeneBank, BGI-Shenzhen, Shenzhen 518120, Guangdong, China)

**Abstract:** All life on the earth uses a set of 20 amino acids to synthesize proteins according to the highly conservative

收稿日期: 2020-02-28 修回日期: 2020-03-25

基金项目: 国家重点研发计划 (2018YFA0900100) 国家自然科学基金 (31901029、31800078)

引用本文: 付宪, 林涛, 张帆, 张惠铭, 章文蔚, 杨焕明, 朱师达, 徐讯, 沈玥. 基因密码子拓展技术的方法原理和前沿应用研究进展[J]. 合成生物学, 2020, 1(1): 103-119

Citation: FU Xian, LIN Tao, ZHANG Fan, ZHANG Huiming, ZHANG Wenwei, YANG Huanming, ZHU Shida, XU Xun, SHEN Yue. Progress in the study of genetic code expansion related methods, principles and applications[J]. Synthetic Biology Journal, 2020, 1(1): 103-119

codon table, and these limited kinds of amino acids serve as the building blocks for the natural protein synthesis. During the long-term evolution, nature is able to expand the structure and function of cellular proteins *via* changing the sequence order of amino acids. However, the evolution process is random and lack of controllability. Manipulating the structure and function of target proteins can also be realized by incorporating an expanded set of building blocks with new chemical and physical properties. Genetic code expansion for synthesis of proteins containing unnatural amino acids at any designed position can be achieved *via* the manipulation of the cellular components responsible for the translation step of the central dogma, which could endow target proteins with new and expanded properties. This review will be focused on the introduction of principles, strategies, techniques to engineer and rewire translational machinery and chassis underpinning genetic code expansion technology. Furthermore, emerging applications in the field of protein function regulation, innovative biomedicine and biocontainment relying on this technology will also be discussed.

Most life on the earth use a set of 20 naturally occurring amino acids as the building blocks for protein synthesis, according to the highly conserved codon table. Natural evolution modulates the structure and function of cellular proteins *via* random mutations among the limited, canonical collection of basic units. Instead, an expanded set of unnatural amino acids with new chemical and physical properties can be incorporated into proteins by synthetic biologists with molecular precision. Genetic code expansion for proteins can be achieved by engineering the translation step of the central dogma. This review will first introduce the principles, strategies, and techniques underpinning genetic code expansion technology, which targets the translational machinery in model chassis. Furthermore, related, emerging applications including protein function regulation, innovative biomedicine, and enhanced biocontainment will be discussed. We conclude with future perspectives.

**Key words:** unnatural amino acid; genetic codon expansion system; translational tool; adaptive cell chassis; orthogonality

遗传密码（决定DNA中三个核苷酸为一组的密码子转译为蛋白质中氨基酸序列的规则）曾被认为是不可改变的，且可被基因编码的氨基酸种类局限于20种。随后的研究打破了这一结论，表明生物的遗传密码在不同物种间也存在一定的差异性。例如，在酿酒酵母的线粒体中，终止密码子UGA也能编码色氨酸<sup>[1]</sup>。在包括人类在内的许多物种中，UGA亦可被用于编码常规氨基酸之外的第21种氨基酸，即硒代半胱氨酸<sup>[2]</sup>。这些发现均暗示着遗传密码在生物体内具有可拓展和被重编的潜力。

受到自然界中密码子重编案例的启示，在过去的20多年中，研究人员基于对中心法则的理解运用，成功开发出在生物体内进行密码子拓展的方法，实现了常规20种氨基酸以外的非天然氨基酸的编码。除 $\beta$ -氨基酸和D- $\alpha$ -氨基酸外，非天然氨基酸还包括针对特殊氨基酸侧链基团（R）设计改造的氨基酸。这样的设计改造能极大地拓展天然编码系统中有限的蛋白质构筑基元种类，对精

细改造或调控蛋白质的结构功能将起到重要的作用。

基因密码子拓展技术的开发优化主要围绕翻译工具和适配底盘细胞两部分。开发高效且正交的翻译工具，使其既能特异性识别非天然氨基酸又能与生物内源的翻译系统兼容，从而保证目标蛋白质合成过程的正交性。构建适配的底盘细胞，需要改造细胞基因组使其具有被翻译工具识别的空白密码子，并对翻译系统等诸多细胞途径进行后续改造，从而保证遗传信息传递和解码过程的正交性。借助正交的翻译工具和适配底盘细胞，基因密码子拓展技术可实现将特殊用途的非天然氨基酸定点引入到目标蛋白的指定位点，为人类在医药、健康、能源、材料等多个重要领域的发展注入新的活力。本文作者将从翻译工具和适配底盘的开发两方面展开，系统讨论基因密码子拓展技术的发展，并介绍其在科研与产业中的相关应用。

## 1 密码子拓展系统的翻译工具

自然界生物按照其对应的密码子表来基因编码蛋白质合成所需的20种天然氨基酸（不包括硒代半胱氨酸和吡咯赖氨酸）。在生物体内，如磷酸化、乙酰化、泛素化等翻译后修饰是丰富天然氨基酸结构进而拓展蛋白质功能的主要途径，在许多生命过程的调控中发挥着关键的作用。蛋白质氨基酸种类的拓展亦可以通过向目标蛋白中引入非天然氨基酸的方式来实现。技术手段包括往细胞中注射化学合成的氨酰化 tRNA<sup>[3]</sup>，或改造营养缺陷型菌株使其可利用氨基酸类似物<sup>[4]</sup>。此外，以 Peter G. Schultz 等为代表的研究人员开发了一种基因密码子拓展方法，通过往细胞内引入改造后的翻译工具来实现基因编码非天然氨基酸，为拓展蛋白质的结构和功能开辟了新的路径<sup>[5]</sup>。密码子拓展技术的核心是需要引入一套正交的外源翻译工具，即能特异性识别非天然氨基酸的氨酰-tRNA 合成酶以及能识别空白密码子的配对 tRNA [图 1 (a)]。同时，该工具配对不能与宿主细胞中的内源性氨酰-tRNA 合成酶或 tRNA 发生交叉反应 [图 1 (b)]。此外，通过对其他核心翻译元件（如延伸因子和核糖体）进行设计和改造，可进一步实现对密码子拓展系统的优化，从而提高非天然氨基酸被特异性地引入到目标蛋白中的效率<sup>[6-8]</sup>。本节将重点围绕以氨酰-tRNA 合成酶/tRNA 配对为代表的翻译元件正交性以及多种其他翻译元件的改造方法进行介绍讨论。

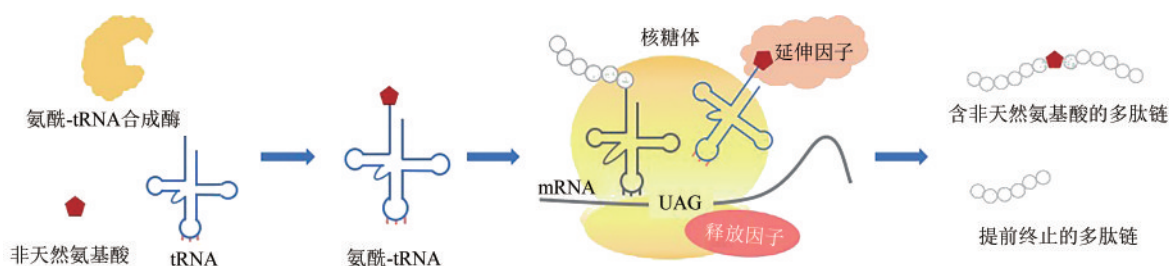
### 1.1 氨酰-tRNA 合成酶/tRNA 配对的正交性

氨酰-tRNA 合成酶/tRNA 配对是翻译过程中的核心工具，保证其正交性是进行技术开发和相应下游应用挖掘的关键。氨酰-tRNA 合成酶/tRNA 配对的正交性包含两个层面：第一，tRNA 的正交性，即外源引入的工具 tRNA（用于匹配空白密码子的 tRNA）只能被对应的工具酶特异性地识别，而不能与内源的其他氨酰-tRNA 合成酶发生交叉反应；第二，氨酰-tRNA 合成酶的正交性，即外源的氨酰-tRNA 合成酶工具能特异性地识别外源添加的非天然氨基酸 [图 1 (b)]<sup>[6-8]</sup>。

#### 1.1.1 tRNA 的正交性

tRNA 是翻译过程中氨基酸的运输工具，每种 tRNA 上带有的被对应氨酰-tRNA 合成酶所识别的特征元素决定了 tRNA 水平的正交性<sup>[9]</sup>。理想的正交氨酰-tRNA 合成酶/tRNA 配对不会与宿主细胞中的天然氨基酸、内源性氨酰-tRNA 合成酶或 tRNA 发生交叉反应。早期研究通过对内源的氨酰-tRNA 合成酶和配对 tRNA 进行设计改造，成功开发出在 tRNA 水平正交的氨酰-tRNA 合成酶/tRNA 配对，使其能独立于内源的工具配对行使特定的功能<sup>[10]</sup>。另一种思路则是从进化关系较远的生物体中选择氨酰-tRNA 合成酶/tRNA 配对进行改造，以降低交叉反应的发生概率。例如，古菌 *Methanococcus janaschii* 的酪氨酰-tRNA 合成酶/酪氨酸 tRNA 配对 (*MjTyrRS/tRNA<sup>Tyr</sup>*) 于 2001 年被引入到大肠杆菌中，用于编码合成含有 *O*-甲基-L-酪氨酸的蛋白质<sup>[11]</sup>。在大肠杆菌中选择上述工具配对的原因包括：① *MjTyrRS* 的氨基酸接纳茎上第一个碱基对为 C-G（大肠杆菌是 G-C），可有效地防止被内源 TyrRS 酶识别<sup>[12]</sup>；② TyrRS 不具有对于底物氨基酸的编辑校对机制<sup>[13]</sup>，有益于非天然氨基酸的引入。这些特征极大地降低了正交工具改造的难度，且改造后的 *MjTyrRS/tRNA<sub>CUA</sub>* 配对因在大肠杆菌中具有良好的正交性和较高的活性，得到了广泛的应用<sup>[14]</sup>。然而，该工具配对在酵母和哺乳细胞等真核系统中不正交，该配对在真核系统中的应用则受到了限制。基于相似原理，可用于真核系统的多个氨酰-tRNA 合成酶/tRNA 配对被相继开发出来。例如，源自大肠杆菌的酪氨酰-tRNA 合成酶/酪氨酸 tRNA 配对 (*EcTyrRS/tRNA<sub>CUA</sub>*) 和亮氨酰-tRNA 合成酶/亮氨酸 tRNA 配对 (*EcLeuRS/tRNA<sub>CUA</sub>*) 被改造为真核生物中的密码子拓展工具<sup>[15-19]</sup>，但由于这些配对工具在大肠杆菌中和一些其他细菌中不正交，故又不适用于原核系统。

吡咯赖氨酰-tRNA 合成酶/吡咯赖氨酸 tRNA 配对 (*PyIRS/tRNA<sup>Pyl</sup>*) 存在于一些产甲烷的古菌和细菌中，是编码吡咯赖氨酸（第 22 种氨基酸）的工具。因其在多种生物系统中（包括细菌和真核生物）都具有高正交性，适合作为在原核系统和真核系统中的通用性密码子拓展工具进行改造。针对 tRNA<sup>Pyl</sup> 的正交性维持机制已积累了较为全面的

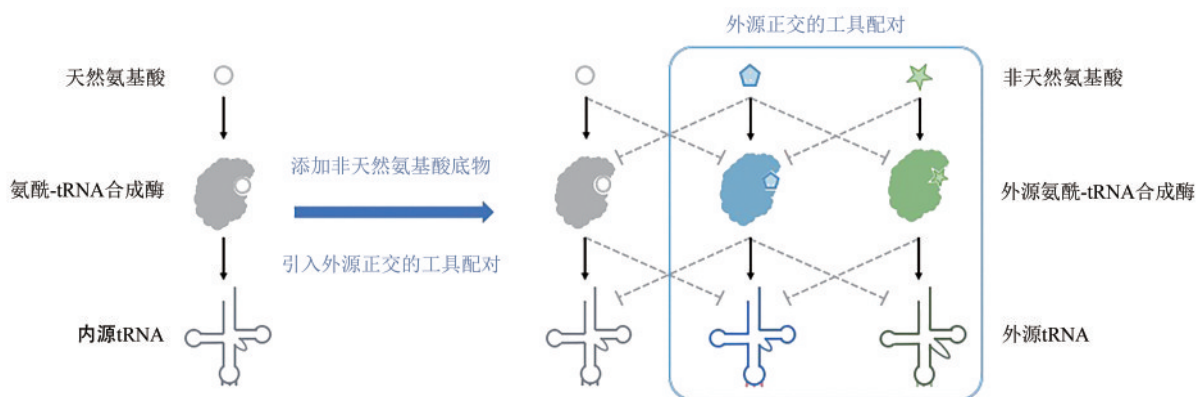


(a) 基因密码子拓展技术示意图

(非天然氨基酸通过氨酰-tRNA合成酶连接至tRNA上, 激活后的氨酰-tRNA被延伸因子带到核糖体, 其反密码子与核糖体中mRNA的指定密码子配对, 最终翻译获得含有非天然氨基酸的多肽链。因释放因子与该氨酰-tRNA的竞争会导致翻译提前终止, 产物中也包括提前终止的多肽链)

(a) The schematic diagram for genetic code expansion technology

[The unnatural amino acid (unAA) is ligated to the engineered tRNA by a dedicated aminoacyl-tRNA synthetase (aaRS). The product (AA-tRNA) is then delivered by an elongation factor (for example, EF-Tu) to the ribosome, where the anticodon of the tRNA matches the UAG codon of the mRNA. The UAG codon will also be recognized by the release factor which results in the termination of translation]



(b) 氨酰-tRNA合成酶/tRNA工具配对的正交性示意图

(外源引入的氨酰-tRNA合成酶只能识别与其对应的非天然氨基酸和外源tRNA, 且不能与内源天然氨基酸和工具配对交叉作用。当引入多对外源工具配对时, 这些外源配对之间也需要保证正交性)

(b) The schematic diagram for the orthogonality of an aaRS/tRNA pair

(The introduced aaRS should specifically recognize its cognate tRNA and unAA, and should not cross-react with the endogenous aaRS/tRNA pair. In addition, if multiple aaRS/tRNA pairs are introduced to incorporate different unAAs into the single polypeptide, these pairs themselves should remain mutually orthogonal)

图1 密码子拓展系统的翻译工具示意图

Fig. 1 Orthogonal translational tools for genetic code expansion

研究基础: 通过体内活性和氨酰动力学测试,  $tRNA^{Pyl}$ 的氨基酸接纳茎中G73位点和第一个碱基对以及T环碱基对G51:C63是重要的识别特征元素<sup>[20]</sup>。PylRS酶氮末端结构域紧密地贴合在由 $tRNA^{Pyl}$ 的T环和极小可变环组成的凹面上, 这一结构特征使得PylRS/ $tRNA^{Pyl}$ 配对自身能够特异性结合, 且能避免其他内源tRNA中较长可变环与PylRS酶的非特异性结合<sup>[21]</sup>。此外, PylRS酶具有氨基酸底物识别可塑性高和不识别tRNA反密码子

环等优良特性, 进一步促使PylRS/ $tRNA_{CUA}$ 配对成为目前应用范围最广泛的编码工具<sup>[20]</sup>。

开发在tRNA水平相互正交的编码工具是同时基因编码多种非标准氨基酸的关键性环节, 也是该领域备受关注的研究方向。早期研究的关注点在于将一种非天然氨基酸引入到蛋白质合成中, 随着近年来新型工具配对的开发和优化方法的改进, 在目标蛋白质中同时基因编码2~3种非天然氨基酸已成为可能<sup>[22-23]</sup>。在细胞中编码多种非天

然氨基酸合成的蛋白质，需要借助相互正交的氨酰-tRNA合成酶/tRNA配对实现。即在不与内源的翻译系统发生交叉反应的前提下，引入的多对外源氨酰-tRNA合成酶/tRNA配对之间亦不能发生相互作用 [图1 (b)]<sup>[8]</sup>。一些已开发的用于编码非天然氨基酸的氨酰-tRNA合成酶/tRNA工具配对自身已经具备相互正交性。例如，古菌来源的 *MjTyrRS/tRNA<sub>CUA</sub>* 和 *PyIRS/tRNA<sup>Pyl</sup>* 配对 (*tRNA<sup>Pyl</sup>* 被改造用于抑制终止密码子 UAA) 可分别将2种非天然氨基酸同时引入大肠杆菌氯霉素乙酰转移酶中<sup>[24]</sup>。进一步配合大肠杆菌衍生工具配对 *EcTrpRS/tRNA*，则可实现同时基因编码3种不同的非天然氨基酸<sup>[22]</sup>。近年来，海量的基因组与宏基因组数据成为开发新型氨酰-tRNA合成酶/tRNA的重要资源，通过生物信息学的深度分析挖掘，一些新型的 *PyIRS/tRNA<sup>Pyl</sup>* 配对被挖掘和发现，可用于开发同类型相互正交的非天然氨基酸编码工具<sup>[25-27]</sup>。

### 1.1.2 氨酰-tRNA合成酶的正交性

作为细胞中最古老的酶之一，氨酰-tRNA合成酶是保证翻译过程严格按照遗传密码子表执行的核心元件。通过长期的进化，氨酰-tRNA合成酶中氨基酸结合口袋这一特定结构能特异性地识别对应的氨基酸，从而保证氨酰-tRNA合成酶对底物的正交性。因此，开发能特异性识别非天然氨基酸的工具酶需要对氨酰-tRNA合成酶中氨基酸结合口袋及其他关键结合区域进行设计改造和定向进化。基于上述思路和策略，不同类型的氨酰-tRNA合成酶被用于改造，多种用于基因密码子拓展的工具酶被成功地开发出来，实现了把超过200种的非天然氨基酸特异性地引入到生物体蛋白中<sup>[28-29]</sup>。例如，包含 $\beta$ -芳香族侧链或 $\gamma$ -芳香族侧链的非天然氨基酸可通过改造 *MjTyrRS* 酶引入<sup>[30]</sup>；特异性改造 *EcLeuRS* 酶可实现 *O*-甲基-L-酪氨酸、 $\alpha$ -氨基辛酸以及光敏感的邻硝基苄基半胱氨酸等的引入<sup>[19]</sup>。

当讨论氨酰-tRNA合成酶工具用于密码子拓展的有效性和可行性时，其底物水平的正交性是相对而非绝对的。即相比上文提到的“一对一”识别特异性的案例，还存在“一对多”识别的情况。一些氨酰-tRNA合成酶存在底物多特异性 (polyspecific)，能够同时识别多种天然或者非天然

氨基酸<sup>[31]</sup>。例如，野生型的 *PyIRS* 可识别超过20种赖氨酸衍生物<sup>[32]</sup>，改造后可实现特异性地识别 *N*-乙酰化赖氨酸及其衍生物或者带有脂肪族侧链的非天然氨基酸<sup>[29, 31]</sup>。多特异性现象的存在可能是氨酰-tRNA合成酶在进化过程中缺乏同类型氨基酸底物的竞争所致，该现象的发现对应用某些廉价氨基酸衍生底物进行酶活性测试和定向进化相关研究有着重要的意义。

## 1.2 氨酰-tRNA合成酶/tRNA配对的改造

开发高效且正交的氨酰-tRNA合成酶/tRNA配对是密码子拓展技术的重点研究内容，旨在解决工具翻译效率低、正交性和兼容性差等核心瓶颈。一方面，已被揭示的氨酰化反应原理和已知的晶体结构为工具配对的理性改造奠定了理论基础；另一方面，定向进化技术的快速发展以及海量基因组数据的深度挖掘加速了新型氨酰-tRNA合成酶/tRNA配对的开发。下文将重点讨论氨酰-tRNA合成酶和tRNA改造过程中所涉及的方法和相关的研究进展。

在获得了tRNA水平正交的氨酰-tRNA合成酶/tRNA配对基础之上，针对氨酰-tRNA合成酶的后续改造主要集中在氨基酸底物识别口袋或编辑结构域的活性位点。传统的定向进化方法通常利用连续正反向迭代筛选原理，在建立氨基酸识别口袋突变体文库之后，通过连续正选择（抗生素抗性）和反选择（毒性蛋白产生）来分离可特异性引入非天然氨基酸的氨酰-tRNA合成酶变体<sup>[33]</sup>。氨酰-tRNA合成酶的编辑活性位点决定自身的矫正活性，用于水解错配的氨基酸，对其进行改造亦可有效提高非天然氨基酸的引入效率。例如，对 *EcLeuRS* 酶编辑活性位点的失活突变可降低其识别天然底物亮氨酸的能力，从而提高翻译过程中非天然氨基酸的引入效率<sup>[34]</sup>。除了传统方法，新一代蛋白质定向进化方法加速了氨酰-tRNA合成酶的开发。建立包含更多活性位点的突变体文库或者通过易错PCR获得随机突变等方法，为获得更优质的正交配对提供了可能性<sup>[35]</sup>。此外，多元自动化基因组工程 (multiplex automated genome engineering, MAGE) 和噬菌体辅助持续进化

(phage-assisted continuous evolution, PACE) 技术也在密码子扩展工具开发过程中起到了重要作用<sup>[36-37]</sup>。MAGE 技术可在细胞群内创建组合遗传多样性的大型突变体文库<sup>[38]</sup>。PACE 技术的优势则在于不依赖已知的蛋白质结构, 通过体内随机诱变, 实施工具效率和细胞生长相耦联的筛选方法, 不仅可以在短时间内获得目标突变体, 还有助于揭示工具酶关键残基的未知功能<sup>[28, 39]</sup>。

理性设计和定向进化策略亦可有效地用于 tRNA 的改造和优化。例如, tRNA 的受体茎是被延伸因子 (elongation factor, EF-Tu) 识别的特征元素, 将大肠杆菌 tRNA<sup>Sec</sup> 的受体茎移植到 tRNA<sup>Sec</sup> 上, 建立杂交 tRNA<sup>U<sup>Tu</sup></sup> 成功解决了 EF-Tu 无法识别 tRNA<sup>Sec</sup> 的问题<sup>[40]</sup>。通过后续的定点诱变, 获得了比野生型活性更高的突变体 tRNA<sup>U<sup>TuX</sup></sup>, 从而推动了硒蛋白的高效合成<sup>[41]</sup>。此外, 通过对构建 *MjtRNA<sup>Tyr</sup>* 的突变体库和正反向筛选, 获得了更容易被延伸因子 EF-Tu 识别的突变体 tRNA, 有助于提高目标蛋白的产量<sup>[42]</sup>。基于 PyIRS 酶不识别 tRNA<sup>Pyl</sup> 的反义密码子环特性, tRNA<sup>Pyl</sup> 反义密码子环可被任意改造, 用于读取不同的密码子<sup>[43]</sup>。通过对 tRNA<sup>Pyl</sup> 反义密码茎的定向进化, 可进一步筛选出翻译效率更高的 tRNA<sup>Pyl</sup> 突变体 (tRNA<sup>Pyl-Opt</sup>)<sup>[44]</sup>。未来, 随着生化原理解析的深入和相关技术的快速发展, tRNA 这一核心翻译元件的开发优化有望进一步突破, 实现从源头上推动密码子拓展技术的升级和应用。

### 1.3 其他关键翻译元件的改造

除了氨酰-tRNA 合成酶/tRNA 配对之外, 核糖体、延伸因子、释放因子等翻译元件也在翻译过程中扮演着非常重要的功能, 影响翻译效率及编码氨基酸的特异性。其中, 核糖体是蛋白质合成的分子机器; 延伸因子识别氨酰-tRNA, 并携带其进入核糖体的 A 位点; 多肽链合成完成后通过释放因子 (release factor, RF) 识别终止密码子实现完整肽链和核糖体从 mRNA 上的释放<sup>[45]</sup>。有研究通过对大肠杆菌核糖体中 16S rRNA 的改造来构建正交的核糖体系统<sup>[46]</sup>。也有研究通过对 16S rRNA 和 23S rRNA 之间加以物理束缚以避免工程亚基与天

然亚基发生交叉装配, 保障了工程核糖体和天然核糖体之间的平行功能<sup>[47-48]</sup>。此外, 研究发现工程化的核糖体也可以起到阻碍释放因子的作用, 促进非天然氨基酸的引入<sup>[45]</sup>。部分非天然氨基酸由于其庞大侧链结构或携带某种电荷, 其氨酰-tRNA 无法被延伸因子识别或者受到核糖体阻碍导致该非天然氨基酸的编码效率低下。有相关研究则通过改造延伸因子实现了在大肠杆菌中基因编码带负电荷的磷酸化丝氨酸<sup>[49]</sup>。另外, 当终止密码子被选择用于编码非天然氨基酸, 释放因子的识别也会引起翻译终止的竞争, 从而降低目标蛋白的产量 [图 1 (a)]。研究人员则通过敲除 UAG 重编大肠杆菌中的 RF1 来提高非天然氨基酸的引入效率<sup>[50]</sup>。相比原核生物, 真核释放因子 eRF1 可识别三种终止密码子, 研究人员则采取了 eRF1 改造而非删除的策略, 在哺乳动物细胞中开发出针对 UAG 识别特异性减弱的 eRF1 突变体, 并证实了该策略的可行性<sup>[51]</sup>。

综上所述, 构建正交的翻译体系, 需要综合考虑氨基酸的活化与转运、多肽链合成的起始、延长、终止和释放等多个过程, 从而能精确地操纵和调控参与上述过程的各种翻译元件。未来的研究需要考虑不同翻译元件之间的相互影响, 系统地对多个元件进行迭代升级, 并使其协同合作, 才有望进一步发展更高效的新一代遗传密码子拓展技术所需的翻译系统。

## 2 密码子拓展系统的适配底盘

用于密码子拓展应用的翻译工具需要配套以相应的底盘细胞来承载其功能的实现。原理上, 在未改造的底盘细胞中过量表达密码子拓展翻译工具会导致翻译错误率的提高, 从而造成细胞毒性, 并产生细胞资源的浪费, 不利于后续应用<sup>[52]</sup>。以琥珀终止密码子介导的密码子拓展系统为例, 该过程一方面受到翻译终止的竞争, 导致目标蛋白得率较低; 另一方面造成其他蛋白翻译终止延伸, 引起细胞毒性。开发高适配性底盘细胞, 避免密码子拓展翻译工具引起的细胞功能紊乱, 是建立基于细胞体系的高效密码子拓展系统中必不可少的一环。此外, 也可以通过利用无细胞

体系的方式避免胁迫问题。基于无细胞体系的密码子拓展系统由于摆脱了对适配底盘的依赖，可避免细胞生理胁迫或细胞膜屏障等限制，在如合成毒性蛋白和提高非天然氨基酸利用率方面有其独特优势<sup>[53-54]</sup>。本节将主要围绕基于细胞体系的适配性改造方式展开介绍。

目前，针对体内密码子拓展系统的适配底盘构建主要通过两种思路实现：一是通过对底盘细胞的基因组进行目标密码子的全基因组精简，以释放出空白密码子来编码新的氨基酸，从而实现密码子拓展；二是通过加入非天然碱基对，使得基因组对应的密码子组合数增加，利用新密码子来编码非天然氨基酸，从而实现密码子内涵的拓展。此外，亦有研究人员通过改造 tRNA 和核糖体来实现四联密码子介导基因编码非天然氨基酸过程，达到密码子拓展的目的<sup>[55]</sup>。然而，由于引入的四联密码子会存在于其他 mRNA 中，从而导致目标氨基酸被非特异性引入其他蛋白质中，并引起该密码子后序列的移码错译。为了解决该问题，四联密码子引入的同时也需要将基因组上其他可形成该密码子组合的序列进行整体替换，本质上属于上文提出的第一种思路。另一种解决方案可通过设计和构建正交的核糖体，保证该核糖体只特异性地识别含有四联密码子的 mRNA<sup>[55-57]</sup>。此外，最近研究人员也成功利用相分离的策略在空间上对非天然氨基酸引入系统与细胞内源翻译系统进行分离，实现在目标 mRNA 中的指定位点引入非天然氨基酸，能有效地降低非天然氨基酸引入所带来的细胞功能紊乱程度<sup>[58]</sup>。本节将重点讨论基于全基因组精简和引入非天然核酸这两种策略的适配底盘构建的研究进展。

## 2.1 全基因组密码子精简的适配底盘构建

使用全基因组密码子精简思路构建用于密码子拓展的适配性底盘需要实现两点：①从有义密码子或终止密码子中选取特定目标密码子，在适配底盘的基因组中进行全局替换，将其替换为其他同义密码子，并引入可特异性识别该密码子的外源翻译工具，用于重新编码的非天然氨基酸；②删除目标密码子原来对应的内源 tRNA（若是终

止密码子，则是删除或改造对应的释放因子），防止原有 tRNA 或释放因子与工具 tRNA 竞争解码目标基因上的目标密码子。

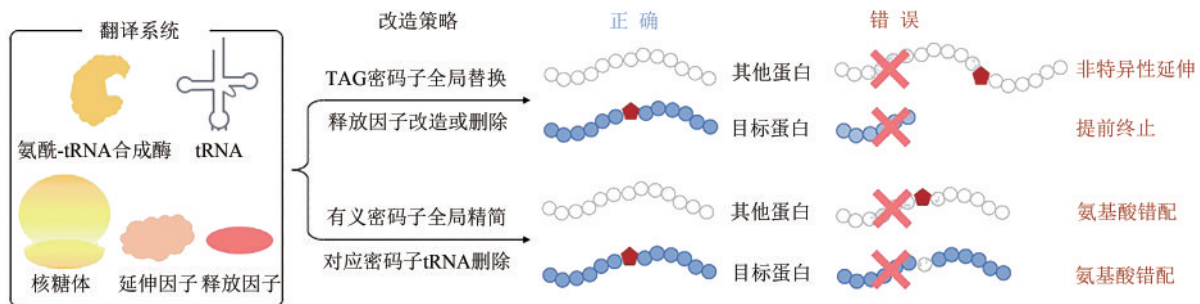
实现全基因组精简的手段主要是通过基因组编辑和基因组合成。基因组编辑的策略适用于基因组较小的底盘细胞，但针对基因组较大的底盘细胞，工作量及技术难度则使得该策略不再适用。而随着 DNA 合成技术与基因组构建技术的快速发展，采用从头合成构建基因组的策略则更为合适。基因组合成技术通过在计算机设计时引入密码子精简设计，通过合成组装设计的基因组实现目标密码子的系统性删除。

目前大部分密码子拓展翻译工具使用 UAG 琥珀终止密码子作为目标，因此，适配底盘的研究主要围绕 UAG 密码子及其对应的释放因子。这种适配底盘的构建，主要解决两方面的问题：一是释放因子 RF1 与翻译工具竞争识别 UAG，使得目标蛋白翻译提前终止，降低蛋白得率；二是翻译工具识别其他基因的琥珀密码子，使得其他蛋白质翻译错误延长。针对问题一，可以通过删除释放因子 1 的基因 *prfA* 来解决。*prfA* 基因之前被认为是必需的基因，无法直接被删除。后续有研究表明，删除 *prfA* 基因只需将大肠杆菌中 7 个必需基因中的 UAG 进行重编<sup>[50]</sup>，或需修复大肠杆菌的释放因子 RF2 基因 *prfB* 中的突变<sup>[59]</sup>。针对问题二，目前大部分非天然氨基酸翻译工具的效率相对较低，过表达外源的翻译工具未见引起严重的细胞胁迫表型。此外，也有研究表明大肠杆菌细胞能在一定程度上耐受翻译错误导致的影响<sup>[60]</sup>。在目前已有的研究基础上来看，针对基因组进行全局系统性精简似乎并非必需，或者只需精简必需基因中的 UAG<sup>[50]</sup>。然而，翻译工具的效率随着研究持续开展逐步在提高，配合底盘细胞面向实际应用时，必然要解决细胞内资源利用最优化的问题。因此，UAG 的系统性精简对构建高效正交的适配底盘在实际应用场景中则尤其必要 [图 2 (a)]。原核与真核系统在系统性精简目标密码子方面均有一定的进展，如 Lajoie 等利用基因组编辑技术 MAGE 和接合组装基因组改造技术 (conjugative assembly genome engineering, CAGE) 将大肠杆菌基因组中所有终止密码子 UAG 替换为 UAA，并删除了

RF1<sup>[61]</sup>, 使UAG密码子能够特异性地仅用于编码非天然氨基酸<sup>[61]</sup>。在真核系统中, 合成基因组里程碑项目人工酵母基因组合成Sc2.0中则设计将全基因组中所有UAG密码子精简为UAA<sup>[62]</sup>, 最终构建的合成型酵母则可通过UAG实现密码子功能拓展。

虽然, UAG精简的底盘能适配于目前大部分的翻译工具, 但若需要同时编码多个非标准氨基酸, 则需要在一个底盘细胞中具有多个空白密码子, 实现的方式可通过对编码同一个或多个氨基酸的同义密码子进行系统性的精简[图2(a)]。同义密码子的选择有几点因素需要考虑: 首先, 原有氨酰-tRNA合成酶不能以目标

反密码子为识别因子, 以保证其翻译工具tRNA的正交性; 其次, 在选择目标密码子时也需要考虑有义密码子的摆动性对翻译工具tRNA识别准确性的影响; 最后, 同义密码子的精简过程中, 还需要综合考虑同义密码子在基因组的特定位置可能存在的特定功能<sup>[63-64]</sup>, 如对基因内/基因间相互作用<sup>[65]</sup>、作为转录和翻译调控元件<sup>[66-67]</sup>、影响核糖体结合能力<sup>[68]</sup>、调节mRNA水平与结构稳定性<sup>[69-75]</sup>、控制翻译速度<sup>[76-77]</sup>, 影响蛋白质折叠与分泌<sup>[68, 78-82]</sup>等与基因转录和翻译相关的功能。其中, 与第三点相关的研究目前还存在较大的空白, 因此实现上仍缺乏足够的理论基础支持。有研究证明, 即便是UAG密

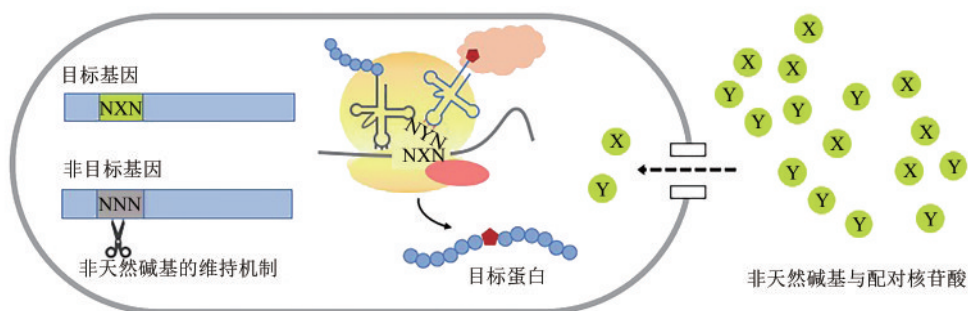


(a) 基于全基因组密码子精简策略构建适配底盘示意图

(全基因组密码子精简适配底盘通过对目标密码子的全局精简及对相应密码子内源tRNA或释放因子的删除或改造, 能消除目标蛋白翻译提前终止、其他蛋白质的非特异性延伸、氨基酸错配等翻译错误)

(a) The schematic diagram for the construction of the adaptive chassis *via* genome recoding

(After comprehensive genome recoding, the 'open' codon will be released by deleting the corresponding tRNAs or release factor. By doing this, many translational errors including premature termination, nonspecific suppression of stop codon and amino acid misincorporation will be eliminated)



(b) 基于非天然碱基策略构建适配底盘示意图

(通过向细胞中引入非天然碱基转运和维持机制的相关元件, 可构建识别和利用非天然碱基对的适配底盘, 利用能特异性识别非天然碱基对的翻译工具, 可实现由非天然核酸介导的基因编码非天然氨基酸过程)

(b) The schematic diagram for the construction of the adaptive chassis *via* unnatural nucleotides mediated replication, transcription, and translation (The chassis that stores and retrieves increased information is constructed by introducing the cellular components for maintaining and transcribing an unnatural base pair (UBP) in its mRNA and DNA. Together with the engineered tRNA with unnatural nucleotides in its anticodon and other translational tools, the semi-synthetic chassis can produce proteins containing unAAs)

图2 密码子拓展系统的适配底盘示意图

Fig. 2 The chassis for genetic code expansion

密码子全基因组去除菌株,在某些条件下仍有生长缺陷,而通过适应性进化的方式研究发现,生长缺陷可能与翻译因素相关,间接说明了密码子可能同时带有其他调控功能以影响翻译过程<sup>[83]</sup>。目前,针对原核系统大肠杆菌中对同义密码子精简的影响因素已有了一定的探索<sup>[72, 84-85]</sup>,研究表明某些位点的同义密码子的精简会对细胞的活性产生严重的影响。基于这些发现,两组研究人员进一步分别通过基因组合成技术构建只含有57个密码子(进行中)<sup>[86]</sup>和61个密码子<sup>[87]</sup>的大肠杆菌。其中后者获得的菌株syn61在倍增时间上比野生菌株慢1.6倍,可以作为一种潜在的多密码子拓展适配底盘。此外,在近期完成的新月柄杆菌基因组合成中<sup>[88]</sup>,也将UUA和UUG密码子在基因组中进行系统性精简作为设计原则之一,结果并未发现对菌株活性有显著影响,亦有潜力作为一种新的适配底盘菌。

随着基因编辑技术和合成基因组学技术的不断发展,通过全基因组密码子精简实现密码子拓展适配底盘的构建取得了不错的进展,并随之建立了具备一定通用性的系列技术流程。但是,值得指出的是,由于目前基因组编辑和基因组合成技术仍存在一定的局限性,且针对同义密码子的研究基础尚不够深入完善,当目标底盘基因组较大较复杂时,相应的理性设计可能会存在诸多不可预测的缺陷,故使用全基因组密码子精简的投入之大使得该策略仍难以被广泛接受。因此,针对这一策略的上游基础性研究还需要进一步投入与积累,以期能为下游的系统性设计改造与应用提供更多技术与规律的支撑。

## 2.2 可识别非天然碱基对的适配底盘改造构建

自然生物中存在的三联密码子实际上是4种碱基(胸腺嘧啶与尿嘧啶并不同时存在于同一类核酸分子中)排列组合的呈现(即 $4^3 = 64$ 个密码子)。由此可知,理论上,若使碱基对的数量增加1对,则会使得潜在的密码子数量理论上可增加至152个(含有非天然核酸的密码子可能具有位置效应,则可用的“空白密码子”将少于理论数量)。目前,研究团队已开发出多对能被生物体利用的

非天然碱基对<sup>[89-90]</sup>,并证明非天然碱基对可用于构建密码子拓展系统来对非天然氨基酸进行基因编码<sup>[91-93]</sup>,第一次以非天然的形式重现了中心法则,为密码子拓展研究和应用提供了一种非常有潜力的新选择[图2(b)]。

基于现有非天然碱基对的密码子拓展适配底盘菌实际上是一种半合成生物(semisynthetic organism, SSO)<sup>[94]</sup>,构建这种生物除了考虑非天然碱基对本身的性质外,还需要对底盘菌进行针对涉及非天然碱基/核苷酸的转运合成、DNA复制酶、RNA聚合酶、核糖体和与DNA修复(特别是碱基错配修复)相关功能蛋白的改造<sup>[95]</sup>。Romesberg课题组在以大肠杆菌为基础构建半合成生物时,利用*Phaeodactylum tricorutum*三角褐指藻的三磷酸核苷转运蛋白将含有相应非天然碱基的三磷酸核苷转运至细胞内<sup>[94]</sup>。研究组通过对引入的非天然碱基对进行优化,能够达到在一定程度上不被碱基错配修复机制识别,同时能使用胞内的DNA复制酶、T7 RNA聚合酶以及核糖体,以质粒的形式完成DNA复制、转录和翻译<sup>[91, 94]</sup>,最终实现非天然氨基酸的编码。

采用非天然碱基对的优势在于可避免对基因组进行大规模的改造,并能实现多密码子拓展,灵活性更强。然而,该策略目前只在原核生物中以质粒DNA的形式实现<sup>[94]</sup>,且其在体内DNA中长时间稳定存在,仍需要依赖一套维持机制(使用CRISPR/Cas系统去除突变的非天然碱基对)<sup>[94]</sup>,整合至基因组后是否能在体内稳定维持非天然碱基对并稳定行使功能仍未见报道。而该策略若应用于真核系统,目前亦只能通过瞬时转化的方法实现非天然氨基酸的编码<sup>[93]</sup>。该策略最终仍依赖于设计开发一套正交的适应于非天然碱基对的DNA复制酶、RNA聚合酶、核糖体及非天然碱基对/核苷酸的合成或转运机器,以保证底盘的复制、转录和翻译活动高效进行。因此,基于非天然核酸的密码子拓展系统仍需要后续系统的优化和完善。

## 3 基因密码子拓展技术的应用

随着高效正交的翻译工具及适配性底盘被不断地开发出来,基因密码子拓展技术得到了逐

步发展和完善, 可将种类繁多的非天然氨基酸引入到目标蛋白中的指定位点。这些非天然氨基酸可拥有各种类型的特殊基团(例如人工设计的新型氨基酸侧链基团), 从而赋予目标蛋白新的物理化学性质, 实现对蛋白质结构和功能的拓展, 展现出巨大的应用潜能。本节将重点介绍和讨论基因密码子拓展技术在一些重点的应用领域取得的现阶段进展, 如蛋白质功能控制、翻译后修饰等蛋白质调控的上游使能技术, 以及如荧光显影探针、新型治疗和生物防控等可直接用于科研和生产的新型生物技术。

### 3.1 蛋白质功能控制

精准操纵蛋白质功能为细胞生命活动的调控开辟了新的路径, 人工控制蛋白质功能的方法包括表达控制与活性控制。传统的方法通常是利用操纵子等启动元件来实现在转录水平上对蛋白质的表达调控<sup>[96]</sup>, 而密码子拓展技术可用于开发出在翻译水平上精确控制蛋白质表达的方法。通过在目标蛋白的基因内部引入终止密码子, 借助基因密码子拓展工具, 即可通过非天然氨基酸的添加或删减来控制全长蛋白的表达与否。例如, 通过引入正交的非天然氨基酸编码工具, 可实现对基因编辑系统 CRISPR-Cas9 中关键酶 Cas9 的表达调控, 使得有功能的全长 Cas9 的表达受外源添加的非天然氨基酸底物控制, 从而实现细胞内对于基因编辑过程的开关调控<sup>[97]</sup>。

此外, 利用基因密码子拓展技术亦可以开发出具有时间和空间分辨率的蛋白质功能调控方法。通过把酶活性位点的关键氨基酸替换成携带光保护基团的衍生氨基酸, 可实现酶的“光开关”调控。具体而言, 氨基酸侧链上的保护基团起到封闭作用, 经过固定波长的光线照射后, 非天然氨基酸侧链的保护基团被移除, 功能基团得到释放, 从而恢复蛋白质活性。利用半胱氨酸和赖氨酸作为目标位点, 研究证明“光开关”调控策略可以用于调控细胞中烟草蚀刻病毒蛋白酶和 Cas9 的活性<sup>[98]</sup>。该调控策略同样适用于 DNA 重组酶活性的精确调控, 并被后续应用于斑马鱼胚胎发育过程中细胞谱系的追踪研究<sup>[99-100]</sup>。上述方法需要选择

目标蛋白中的特定活性残基, 对各种类型的蛋白质均需要进行定制化的策略设计, 具有一定的局限性。最近, 研究人员发展了更具普适性的蛋白质功能调控技术。具体而言, 该新型技术基于可遗传编码非天然氨基酸的“邻近脱笼”策略, 结合计算机辅助设计与筛选, 可在活体细胞或动物内瞬时激活各类型蛋白质, 为许多重要细胞生理过程的研究提供了新的化学生物学工具<sup>[101]</sup>。

### 3.2 翻译后修饰

翻译后修饰参与调控细胞内众多关键的生命活动调控过程, 这些修饰使蛋白质的结构和功能更为丰富, 调节更精细, 作用更专一。因此, 基于翻译后修饰的蛋白质结构和功能研究是该领域的一个重点方向。得益于质谱技术的发展, 许多翻译后修饰的种类和位点被不断发掘, 但是, 在体内和体外精确地合成带有修饰的功能蛋白仍然挑战巨大, 严重制约了对其进行分子和生化原理的深度研究。以体内合成磷酸化蛋白为例, 传统的方法是将丝氨酸/苏氨酸突变为谷氨酸来模拟其磷酸化的作用, 但具有作用效果不真实的劣势。此外, 磷酸化是一个体内高度动态变化的瞬时过程, 单个蛋白中存在众多潜在的修饰位点, 且作用相关的激酶和磷酸酶常常未知, 也加大了体内和体外合成磷酸化蛋白的难度<sup>[102]</sup>。利用基因密码子拓展技术可实现对目标蛋白中指定位点进行真实的翻译后修饰, 已经成功实现丝氨酸、色氨酸及酪氨酸的磷酸化修饰。通过利用古菌中磷酸化丝氨酸合成酶 SepRS 和改造的 tRNA<sup>Sep</sup>, 并对延伸因子 EF-Tu 进行定向进化, 科学家在大肠杆菌中实现了由终止密码子 UAG 介导的磷酸化丝氨酸合成<sup>[103]</sup>。后续实验表明, 通过对上述翻译元件的进一步改造和定向进化, 以及利用 UAG 密码子被系统重编的底盘细胞, 密码子拓展技术可以更为高效地基因编码携带磷酸化丝氨酸的蛋白质<sup>[104-105]</sup>。此外, 磷酸化的丝氨酸可以转化为脱氢丙氨酸, 因其具有不饱和键, 可与带有各种翻译后修饰的侧链基团连接, 用于体外合成具有真实修饰的蛋白质, 应用前景广阔<sup>[106]</sup>。通过构建磷酸化苏氨酸生物合成通路, 并对 SepRS/tRNA<sup>Sep</sup> 工具配对开展

连续的定向进化, 辅助以深度测序分析介导的并行正向筛选策略, 高效基因编码磷酸化苏氨酸的方法亦被开发出来<sup>[107]</sup>。此外, 利用基于非天然氨基酸的脱保护和前肽策略, 不同的团队成功地开发了提高大肠杆菌体内磷酸化酪氨酸和类似物的浓度的方法, 并实现了合成指定位点上携带磷酸化酪氨酸及其类似物的蛋白质<sup>[108]</sup>。除了用于合成磷酸化蛋白, 基因密码子拓展技术亦可合成乙酰化、泛素化和甲基化修饰的赖氨酸<sup>[7]</sup>。上述翻译后修饰的氨基酸既可以被直接引入到蛋白质中的指定位点, 也可以通过后续的选择性化学反应来实现<sup>[109]</sup>。综上, 基因密码子拓展技术已经成功地应用于各类型的翻译后修饰研究, 为进一步理解其作用机制和生理功能的科研或者临床研究奠定了基础。

### 3.3 荧光显影探针

荧光显影在蛋白质追踪和定位研究中有着非常广泛的应用, 通过荧光显影可以在细胞甚至细胞器层面上对目标蛋白进行准确的定位。常规的荧光显影方法主要使用荧光蛋白与小分子化合物探针。虽然借助各种类型的荧光蛋白可有效地用于荧光成像和追踪<sup>[110-111]</sup>, 但与荧光蛋白的融合可能会改变目标蛋白的构象或结构, 影响后续的结果分析。虽然小分子化合物探针灵敏且稳定性高, 但是通常无法特异性地结合在目标蛋白上。通过基因密码子拓展技术可以实现对目标蛋白的显影示踪。具体而言, 在不影响目标蛋白功能的情况下, 可将携带高化学反应性基团的非天然氨基酸引入到目标蛋白的指定位点, 通过特定的化学反应(如点击化学)可实现蛋白质与荧光探针的生物正交结合, 用于后续荧光成像。例如, 含降冰片烯的赖氨酸衍生物可以与四嗪类荧光探针快速结合<sup>[112]</sup>, 反式环辛烯赖氨酸则可以与类罗丹明的荧光探针产生正交反应<sup>[113]</sup>, 用于蛋白质显影追踪。密码子拓展技术在蛋白质显影追踪方面有两点显著的优势: 经过设计与筛选后的非天然氨基酸侧链基团与荧光探针具有高选择性, 保证两者可以进行高效准确的生物正交反应; 密码子拓展技术可以将非天然氨基酸引入到目标蛋白的非关

键位点, 从而尽量避免对蛋白质的结构和功能产生潜在影响。

### 3.4 新型治疗策略

通过基因密码子拓展技术向细胞因子、抗体、受体蛋白等免疫相关蛋白中引入各类型的非天然氨基酸为新型治疗策略开辟了新的领域。抗体作为一种具有高特异性识别能力的蛋白质, 可以将药物分子靶向性地带到靶向作用细胞, 以减少对正常细胞的伤害。早期抗体耦联药物通过化学修饰后与抗体的赖氨酸或半胱氨酸进行反应耦联, 实现药物分子的靶向传输<sup>[114-115]</sup>。但非特异性耦联方式会导致抗体上携带的药物分子数量不均匀, 进而影响药效和导致细胞毒性。在新一代抗体耦联药物研发过程中, 非天然氨基酸(例如对乙酰苯丙氨酸和叠氮苯丙氨酸)可作为正交的耦联反应位点, 与药物发生特异性的结合, 从而精确控制药物抗体比例。使用脒连接与无铜点击化学可以将药物准确地耦联在抗体中非天然氨基酸的指定位点<sup>[116-117]</sup>。此外, 一些条件更为温和的耦联反应被相继开发出来。例如, 增加皮克特-斯宾格勒(Pictet-Spengler, P-S)反应使脒连接反应可以在中性pH条件下完成, 或是使用咪唑基的非天然氨基酸进行光交联反应实现小分子化合物的耦联<sup>[118-119]</sup>。嵌合抗原受体T细胞(chimeric antigen receptor-T, CAR-T)免疫疗法是一种治疗肿瘤的新型精准靶向疗法, 避免细胞因子释放综合征是目前开发高安全性CAR-T免疫疗法的研究重点<sup>[120]</sup>。基因密码子拓展技术可用于控制工程化T细胞的特异性和活性, 通过利用开关分子来连接T细胞和肿瘤细胞, 为开发新一代CAR-T免疫疗法奠定基础<sup>[121]</sup>。此外, 基因密码子拓展技术还可以用于新一代疫苗的研发。传统的疫苗制作需要对活病毒进行减活或者灭活处理, 病毒蛋白结构在灭活过程中一旦被破坏, 则可能难以达到理想的接种效果<sup>[122]</sup>。科学家把病毒基因中不易返祖的位点突变成UAG终止密码子, 利用密码子拓展技术实现病毒在特殊细胞系的增殖, 从而制备与正常病毒相似免疫原性的非天然病毒<sup>[123]</sup>。这些病毒不能在正常细胞中扩增, 可显著提升疫苗的安全性,

具有巨大的应用潜力。

### 3.5 生物防控

近些年来,合成生物学突取得了突飞猛进的发展,基因改造生物在生物技术领域发挥着越来越举足轻重的作用,科学家与大众对生物安全的关注也被提升至新的高度。基于非天然氨基酸的蛋白质表达调控和蛋白质从头设计策略为生物防逃逸技术打开了一个崭新的研发视角。生物防控手段可依赖于遗传隔离、营养缺陷型调控、基因回路设计等思路<sup>[124-126]</sup>,但上述方法的有效性易被自然突变、环境补充和水平基因转移等因素打破<sup>[127]</sup>。基因密码子拓展技术可用于控制细胞中必需蛋白的表达或者功能,进而构建出依赖于非天然氨基酸的生命体。例如,通过把甘露糖-6-磷酸异构酶这一必需蛋白中关键的组氨酸替换成为组氨酸类似物,科学家成功地构建出生长依赖于外源氨基酸的大肠杆菌<sup>[128]</sup>。上述策略虽然有效,但是只适用于特定类型的蛋白质(如前文提到的金属结合蛋白)。通过对目标蛋白的关键区域进行重新设计或者定向进化,研究人员利用密码子拓展技术成功获得了功能依赖于某种非天然氨基酸的新型蛋白,该策略具有更高的普适性,并可有效地避免基因改造生物逃逸到自然环境中<sup>[129-130]</sup>。值得指出的是,上述策略将必需基因中特定氨基酸对应的密码子突变成为终止密码子,若对应的mRNA发生回复突变等情况会使得防逃逸策略的有效性丧失。利用基因组重编的大肠杆菌作为基因密码子拓展技术的底盘,研究团队系统地测试了大肠杆菌众多必需基因中不同位点引入非天然氨基酸后对于生物防控策略的有效性,并发现多个位点和基因的组合可有效地降低非天然氨基酸依赖型生长的逃逸概率<sup>[131]</sup>。此外,通过往细胞中引入毒素-抗毒素表达系统,并利用非天然氨基酸的加减作为抗毒素蛋白表达的开关,亦可构建出有效的生物防控方法<sup>[132]</sup>。综上所述,利用基因密码子拓展技术,通过巧妙地选择目标蛋白的种类并对其进行设计或改造,基于非天然氨基酸的新型生物防控策略为本领

域的发展带来了新的动能。

## 4 总结与展望

蛋白质是生命的物质基础,也是生命活动的主要承担者。蛋白质功能化策略创新被认为是推动生命科学研究与产业应用转化的重要引擎之一。通过基因密码子拓展技术进行蛋白质功能创新在过去20多年中发展迅速,在翻译工具改造及底盘细胞开发方面均取得了可喜的进展。然而,目前已开发的基因密码子拓展系统(尤其是基于真核生物的系统)普遍存在翻译效率低、正交性和兼容性差等核心瓶颈。针对翻译系统中多种翻译元件的系统性优化改造乃至从头设计,构建具有多个空白密码子的底盘细胞,针对翻译工具和底盘细胞的相互适配原则的探索与优化改造,以及结合这些研究内容实现同时基因编码多种非天然氨基酸,将是该领域未来研究的重点和难点。

随着对于蛋白质合成机制理解的不断深入、合成生物学技术及定向进化技术的应用、基因组与宏基因组数据的持续积累及解读工具的迭代,更多高效、正交的翻译工具及适配底盘有望被开发,促进基因密码子拓展技术的快速发展。利用基因密码子拓展技术实现基因编码非天然氨基酸,将成为继无细胞体系、化学合成等体外合成方法之后实现蛋白质和多肽功能创新的一种新策略,为相关的下游应用的发展提供了新契机,推动人类在医药健康、能源、材料等关键领域发展。

### 参 考 文 献

- [1] MACINO G, CORUZZI G, NOBREGA F G, et al. Use of the UGA terminator as a tryptophan codon in yeast mitochondria[J]. PNAS, 1979, 76(8): 3784-3785.
- [2] BOCK A and STADTMAN T C. Selenocysteine, a highly specific component of certain enzymes, is incorporated by a UGA-directed co-translational mechanism[J]. Biofactors, 1988, 1(3): 245-250.
- [3] NOREN C, ANTHONY-CAHILL S, GRIFFITH M, et al. A general method for site-specific incorporation of unnatural amino acids into proteins[J]. Science, 1989, 244(4901): 182-188.

- [4] BUDISA N, MINKS C, ALEFELDER S, et al. Toward the experimental codon reassignment *in vivo*: protein building with an expanded amino acid repertoire [J]. *FASEB J.*, 1999, 13(1): 41-51.
- [5] CORNISH V W, MENDEL D, SCHULTZ P G. Probing protein structure and function with an expanded genetic code [J]. *Cheminform.* 1995, 34(6): 621-633.
- [6] CHIN J W. Expanding and reprogramming the genetic code of cells and animals[J]. *Annu. Rev. Biochem.*, 2014, 83: 379-408.
- [7] CHIN J W. Expanding and reprogramming the genetic code[J]. *Nature*, 2017, 550(7674): 53-60.
- [8] ARRANZ-GIBERT P, PATEL J R, ISAACS F J. The role of orthogonality in genetic code expansion[J]. *Life*, 2019, 9(3): 58.
- [9] GIEGÉ R, SISSLER M, FLORENTZ C. Universal rules and idiosyncratic features in tRNA identity[J]. *Nucleic Acids Res.*, 1998, 26(22): 5017-5035.
- [10] LIU D R, MAGLIERI T J, PASTRNAK M, et al. Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins *in vivo*[J]. *PNAS*, 1997, 94(19): 10092-10097.
- [11] WANG L, BROCK A, HERBERICH B, et al. Expanding the genetic code of *Escherichia coli* [J]. *Science*, 2001, 292(5516): 498-500.
- [12] STEER B A, SCHIMMEL P. Major anticodon-binding region missing from an archaeobacterial tRNA synthetase[J]. *Journal of Biological Chemistry*, 1999, 274(50): 35601-35606.
- [13] JAKUBOWSKI H, GOLDMAN E. Editing of errors in selection of amino acids for protein synthesis[J]. *Microbiology and Molecular Biology Reviews*, 1992, 56(3): 412-429.
- [14] XIE J, SCHULTZ P G. An expanding genetic code [J]. *Methods*, 2005, 36(3): 227-238.
- [15] CHIN J W, CROPP T A, ANDERSON J C, et al. An expanded eukaryotic genetic code [J]. *Science*, 2003, 301(5635): 964-967.
- [16] CHIN J W, CROPP T A, CHU S, et al. Progress toward an expanded eukaryotic genetic code [J]. *Chemistry & Biology*, 2003, 10(6): 511-519.
- [17] SAKAMOTO K, HAYASHI A, SAKAMOTO A, et al. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells [J]. *Nucleic Acids Research*, 2002, 30(21): 4692-4699.
- [18] LIU W R, SCHULTZ P G. Genetic incorporation of unnatural amino acids into proteins in mammalian cells [J]. *Nature Methods*, 2007, 4(3): 239-244.
- [19] WU N, DEITERS A, CROPP T A, et al. A genetically encoded photocaged amino acid[J]. *Journal of the American Chemical Society*, 2004, 126(44): 14306-14307.
- [20] AMBROGELLY A, GUNDLLAPALLI S, HERRING S, et al. Pyrrolysine is not hardwired for cotranslational insertion at UAG codons[J]. *PNAS*, 2007, 104(9): 3141-3146.
- [21] SUZUKI T, MILLER C, GUO L T, et al. Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase [J]. *Nature Chemical Biology*, 2017, 13(12): 1261-1266.
- [22] ITALIA J S, ADDY P S, ERICKSON S B, et al. Mutually orthogonal nonsense-suppression systems and conjugation chemistries for precise protein labeling at up to three distinct sites[J]. *Journal of the American Chemical Society*, 2019, 141(15): 6204-6212.
- [23] NEUMANN H, PEAK-CHEW S Y, CHIN J W. Genetically encoding *N*<sup>ε</sup>-acetyllysine in recombinant proteins [J]. *Nature Chemical Biology*, 2008, 4(4): 232-234.
- [24] WAN W, HUANG Y, WANG Z, et al. A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli* [J]. *Angewandte Chemie International Edition*, 2010, 49(18): 3211-3214.
- [25] CHATTERJEE A, XIAO H, SCHULTZ P G. Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli* [J]. *PNAS*, 2012, 109(37): 14841-14846.
- [26] WILLIS J C W, CHIN J W. Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs [J]. *Nat. Chem.*, 2018, 10(8): 831-837.
- [27] MEINEKE B, HEIMGÄRTNER J, LAFRANCHI L, et al. Methanomethylophilus alvus Mx1201 provides basis for mutual orthogonal pyrrolysyl tRNA/aminoacyl-tRNA synthetase pairs in mammalian cells [J]. *ACS Chemical Biology*, 2018, 13(11): 3087-3096.
- [28] VARGAS-RODRIGUEZ O, SEVOSTYANOVA A, SÖLL D, et al. Upgrading aminoacyl-tRNA synthetases for genetic code expansion [J]. *Current Opinion in Chemical Biology*, 2018, 46: 115-122.
- [29] GUO L, WANG Y, NAKAMURA A, et al. Polyspecific pyrrolysyl-tRNA synthetases from directed evolution [J]. *PNAS*, 2014, 111(47): 16724-16729.
- [30] WANG L, SCHULTZ P G. Expanding the genetic code [J]. *Angewandte Chemie International Edition*, 2005, 44(1): 34-66.
- [31] WAN W, THARP J M, LIU W R. Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool [J]. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 2014, 1844(6): 1059-1070.
- [32] DUMAS A, LERCHER L, SPICER C D, et al. Designing logical codon reassignment-expanding the chemistry in biology [J]. *Chemical Science*, 2015, 6(1): 50-69.
- [33] LIU C C, SCHULTZ P G. Adding new chemistries to the genetic code [J]. *Annual Review of Biochemistry*, 2010, 79: 413-444.
- [34] DULIC M, CVETESIC N, ZIVKOVIC I, et al. Kinetic origin of

- substrate specificity in post-transfer editing by leucyl-tRNA synthetase[J]. *Journal of Molecular Biology*, 2018, 430(1): 1-16.
- [35] O'DONOGHUE P, LING J, WANG Y S, et al. Upgrading protein synthesis for synthetic biology [J]. *Nature Chemical Biology*, 2013, 9(10): 594-598.
- [36] AMIRAM M, HAIMOVICH A D, FAN C, et al. Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids [J]. *Nature Biotechnology*, 2015, 33(12): 1272.
- [37] BRYSON D I, FAN C, GUO L T, et al. Continuous directed evolution of aminoacyl-tRNA synthetases [J]. *Nature Chemical Biology*, 2017, 13(12): 1253.
- [38] PACKER M S, LIU D R. Methods for the directed evolution of proteins [J]. *Nature Reviews Genetics*, 2015, 16 (7) : 379-394.
- [39] BADRAN A H, LIU D R. Development of potent *in vivo* mutagenesis plasmids with broad mutational spectra [J]. *Nature Communications*, 2015, 6: 8425.
- [40] ALDAG C, BRÖCKER M J, HOHN M J, et al. Rewiring translation for elongation factor Tu - dependent selenocysteine incorporation [J]. *Angewandte Chemie International Edition*, 2013, 52(5): 1441-1445.
- [41] MILLER C, BRÖCKER M J, PRAT L, et al. A synthetic tRNA for EF-Tu mediated selenocysteine incorporation *in vivo* and *in vitro* [J]. *FEBS letters*, 2015, 589(17): 2194-2199.
- [42] GUO J, MELANÇON III C E, LEE H S, et al. Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids [J]. *Angewandte Chemie International Edition*, 2009, 48(48): 9148-9151.
- [43] THARP J M, EHNBOM A, LIU W R. tRNAPyl: structure, function, and applications [J]. *RNA Biology*, 2018, 15(4/5): 441-452.
- [44] FAN C, XIONG H, REYNOLDS N M, et al. Rationally evolving tRNAPyl for efficient incorporation of noncanonical amino acids [J]. *Nucleic Acids Research*, 2015, 43(22): e156.
- [45] PETRY S, BRODERSEN D E, MURPHY IV F V, et al. Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon [J]. *Cell*, 2005, 123(7): 1255-1266.
- [46] WANG K, NEUMANN H, PEAK-CHEW S Y, et al. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion [J]. *Nature Biotechnology*, 2007, 25 (7): 770-777.
- [47] ORELLE C, CARLSON E D, SZAL T, et al. Protein synthesis by ribosomes with tethered subunits [J]. *Nature*, 2015, 524 (7563): 119-124.
- [48] FRIED S D, SCHMIED W H, UTTAMAPINANT C, et al. Ribosome subunit stapling for orthogonal translation in *E. coli* [J]. *Angewandte Chemie International Edition*, 2015, 54(43): 12791-12794.
- [49] PARK H S, HOHN M J, UMEHARA T, et al. Expanding the genetic code of *Escherichia coli* with phosphoserine [J]. *Science*, 2011, 333(6046): 1151-1154.
- [50] MUKAI T, HAYASHI A, IRAHA F, et al. Codon reassignment in the *Escherichia coli* genetic code [J]. *Nucleic Acids Res.*, 2010, 38(22): 8188-8195.
- [51] SCHMIED W H, ELSÄSSER S J, UTTAMAPINANT C, et al. Efficient multisite unnatural amino acid incorporation in mammalian cells *via* optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1 [J]. *Journal of the American Chemical Society*, 2014, 136(44): 15577-15583.
- [52] ARRANZ-GIBERT P, VANDERSCHUREN K, ISAACS F J. Next-generation genetic code expansion [J]. *Current Opinion in Chemical Biology*, 2018, 46: 203-211.
- [53] SHIMIZU Y, INOUE A, TOMARI Y, et al. Cell-free translation reconstituted with purified components [J]. *Nature Biotechnology*, 2001, 19(8): 751-755.
- [54] 高伟, 卜宁, 卢元. 无细胞体系非天然蛋白质合成研究进展 [J]. *生物工程学报*, 2018, 34(9): 1371-1385.
- GAO W, BU N, LU Y. Recent advances in cell-free unnatural protein synthesis [J]. *Chinese Journal of Biotechnology*, 2018, 34(9): 1371-1385.
- [55] NEUMANN H, WANG K H, DAVIS L, et al. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome [J]. *Nature*, 2010, 464(7287): 441-444.
- [56] RACKHAM O, CHIN J W. A network of orthogonal ribosome-mRNA pairs [J]. *Nature Chemical Biology*, 2005, 1 (3) : 159-166.
- [57] SCHMIED W H, TNIMOV Z, UTTAMAPINANT C, et al. Controlling orthogonal ribosome subunit interactions enables evolution of new function [J]. *Nature*, 2018, 564 (7736) : 444-448.
- [58] REINKEMEIER C D, GIRONA G E, LEMKE E A. Designer membraneless organelles enable codon reassignment of selected mRNAs in eukaryotes [J]. *Science*, 2019, 363 (6434) : eaaw2644.
- [59] JOHNSON D B, XU J, SHEN Z, et al. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites [J]. *Nature Chemical Biology*, 2011, 7(11): 779-786.
- [60] RUAN B, PALIOURA S, SABINA J, et al. Quality control despite mistranslation caused by an ambiguous genetic code [J]. *PNAS*, 2008, 105(43): 16502-16507.
- [61] LAJOIE M J, ROVNER A J, GOODMAN D B, et al. Genomically recoded organisms expand biological functions [J]. *Science*, 2013, 342(6156): 357-360.
- [62] RICHARDSON S M, MITCHELL L A, STRACQUADANIO G, et al. Design of a synthetic yeast genome [J]. *Science*, 2017, 355(6329): 1040-1044.

- [63] KOMAR A A. The Yin and Yang of codon usage [J]. *Human Molecular Genetics*, 2016, 25(R2): R77.
- [64] CHANEY J L, CLARK P L. Roles for synonymous codon usage in protein biogenesis [J]. *Annual Review of Biophysics*, 2015, 44: 143-166.
- [65] QUAX T E, CLAASSENS N J, SOLL D, et al. Codon bias as a means to fine-tune gene expression [J]. *Molecular Cell*, 2015, 59(2): 149-161.
- [66] CHO B K, ZENGLER K, QIU Y, et al. The transcription unit architecture of the *Escherichia coli* genome [J]. *Nat. Biotechnol.*, 2009, 27(11): 1043-1049.
- [67] LI G W, OH E, WEISSMAN J S. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria [J]. *Nature*, 2012, 484(7395): 538-541.
- [68] ZHOU M, GUO J, CHA J, et al. Non-optimal codon usage affects expression, structure and function of clock protein FRQ [J]. *Nature*, 2013, 495(7439): 111-115.
- [69] PRESNYAK V, ALHUSAINI N, CHEN Y H, et al. Codon optimality is a major determinant of mRNA stability [J]. *Cell*, 2015, 160(6): 1111-1124.
- [70] BOËL G, LETSO R, NEELY H, et al. Codon influence on protein expression in *E. coli* correlates with mRNA levels [J]. *Nature*, 2016, 529(7586): 358-363.
- [71] MISHIMA Y, TOMARI Y. Codon usage and 3' UTR length determine maternal mRNA stability in Zebrafish [J]. *Molecular Cell*, 2016, 61(6): 874-885.
- [72] NAPOLITANO M G, LANDON M, GREGG C J, et al. Emergent rules for codon choice elucidated by editing rare arginine codons in *Escherichia coli* [J]. *PNAS*, 2016, 113(38): E5588-E5597.
- [73] KUDLA G, MURRAY A W, TOLLERVEY D, et al. Coding-sequence determinants of gene expression in *Escherichia coli* [J]. *Science*, 2009, 324(5924): 255-258.
- [74] ZHOU Z, DANG Y, ZHOU M, et al. Codon usage is an important determinant of gene expression levels largely through its effects on transcription [J]. *PNAS*, 2016, 113(41): E6117-E6125.
- [75] CHEN S, LI K, CAO W, et al. Codon-resolution analysis reveals a direct and context-dependent impact of individual synonymous mutations on mRNA level [J]. *Molecular Biology and Evolution*, 2017, 34(11): 2944-2958.
- [76] SORENSEN M A, KURLAND C G, PEDERSEN S. Codon usage determines translation rate in *Escherichia coli* [J]. *Journal of Molecular Biology*, 1989, 207(2): 365-377.
- [77] SORENSEN M A, PEDERSEN S. Absolute *in vivo* translation rates of individual codons in *Escherichia coli*. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate [J]. *Journal of Molecular Biology*, 1991, 222(2): 265-280.
- [78] TSAI C J, SAUNA Z E, KIMCHI-SARFATY C, et al. Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima [J]. *Journal of Molecular Biology*, 2008, 383(2): 281-291.
- [79] BUHR F, JHA S, THOMMEN M, et al. Synonymous codons direct cotranslational folding toward different protein conformations [J]. *Molecular Cell*, 2016, 61(3): 341-351.
- [80] ZALUCKI Y M, BEACHAM I R, JENNINGS M P. Biased codon usage in signal peptides: a role in protein export [J]. *Trends Microbiol.*, 2009, 17(4): 146-150.
- [81] ZALUCKI Y M, BEACHAM I R, JENNINGS M P. Coupling between codon usage, translation and protein export in *Escherichia coli* [J]. *Biotechnol. J.*, 2011, 6(6): 660-667.
- [82] PECHMANN S, CHARTRON J W, FRYDMAN J. Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP *in vivo* [J]. *Nature Structure & Molecular Biology*, 2014, 21(12): 1100-1105.
- [83] WANNIER T M, KUNJAPUR A M, RICE D P, et al. Adaptive evolution of genomically recoded *Escherichia coli* [J]. *PNAS*, 2018, 115(12): 3090-3095.
- [84] WANG K, FREDENS J, BRUNNER S F, et al. Defining synonymous codon compression schemes by genome recoding [J]. *Nature*, 2016, 539(7627): 59-64.
- [85] GOODMAN D B, CHURCH G M, KOSURI S. Causes and effects of N-terminal codon bias in bacterial genes [J]. *Science*, 2013, 342(6157): 475-479.
- [86] OSTROV N, LANDON M, GUELL M, et al. Design, synthesis, and testing toward a 57-codon genome [J]. *Science*, 2016, 353(6301): 819-822.
- [87] FREDENS J, WANG K, DE LA TORRE D, et al. Total synthesis of *Escherichia coli* with a recoded genome [J]. *Nature*, 2019, 569(7757): 514-518.
- [88] VENETZ J E, DEL MEDICO L, WÖLFLE A, et al. Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality [J]. *PNAS*, 2019, 116(16): 8070-8079.
- [89] HOSHIKA S, LEAL N A, KIM M J, et al. Hachimoji DNA and RNA: a genetic system with eight building blocks [J]. *Science*, 2019, 363(6429): 884-887.
- [90] MALYSHEV D A, DHAMI K, LAVERGNE T, et al. A semi-synthetic organism with an expanded genetic alphabet [J]. *Nature*, 2014, 509: 385-388.
- [91] 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: 644.
- [92] DIEN V T, HOLCOMB M, FELDMAN A W, et al. Progress toward a semi-synthetic organism with an unrestricted expanded genetic alphabet [J]. *Journal of the American Chemical Society*, 2018, 140(47): 16115-16123.

- [93] ZHOU A X Z, SHENG K, FELDMAN A W, et al. Progress toward eukaryotic semisynthetic organisms: translation of unnatural codons [J]. *Journal of the American Chemical Society*, 2019, 141(51): 20166-20170.
- [94] ZHANG Y, LAMB B M, FELDMAN A W, et al. A semisynthetic organism engineered for the stable expansion of the genetic alphabet[J]. *PNAS*, 2017, 114(6): 1317-1322.
- [95] ROMESBERG F E. Synthetic biology: the chemist's approach [J]. *Israel Journal of Chemistry*, 2019, 59(1/2): 91-94.
- [96] BROWNING D F, GODFREY R E, RICHARDS K L, et al. Exploitation of the *Escherichia coli* lac operon promoter for controlled recombinant protein production [J]. *Biochemical Society Transaction*, 2019, 47(2): 755-763.
- [97] SUZUKI T, ASAMI M, PATEL S G, et al. Switchable genome editing *via* genetic code expansion [J]. *Scientific Report*, 2018, 8(1): 10051.
- [98] HEMPHILL J, BORCHARDT E K, BROWN K, et al. Optical control of CRISPR/Cas9 gene editing [J]. *Journal of the American Chemical Society*, 2015, 137(17): 5642-5645.
- [99] EDWARDS W F, YOUNG D D, DEITERS A. Light-activated Cre recombinase as a tool for the spatial and temporal control of gene function in mammalian cells [J]. *ACS Chemistry Biology*, 2009, 4(6): 441-445.
- [100] BROWN W, LIU J, TSANG M, et al. Cell-lineage tracing in zebrafish embryos with an expanded genetic code [J]. *ChemBioChem*, 2018, 19(12): 1244-1249.
- [101] WANG J, LIU Y, LIU Y, et al. Time-resolved protein activation by proximal decaging in living systems [J]. *Nature*, 2019, 569(7757): 509-513.
- [102] MOK J, KIM P M, LAM H Y, et al. Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs [J]. *Sci. Signal.*, 2010, 3(109): ra12.
- [103] PARK H S, HOHN M J, UMEHARA T, et al. Expanding the genetic code of *Escherichia coli* with phosphoserine [J]. *Science*, 2011, 333(6046): 1151-1154.
- [104] PIRMAN N L, BARBER K W, AERNI H R, et al. A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation [J]. *Nature Communications*, 2015, 6: 8130.
- [105] ROGERSON D T, SACHDEVA A, WANG K, et al. Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog [J]. *Nature Chemical Biology*, 2015, 11(7): 496-503.
- [106] YANG A, HA S, AHN J, et al. A chemical biology route to site-specific authentic protein modifications [J]. *Science*, 2016, 354(6312): 623-626.
- [107] ZHANG M S, BRUNNER S F, HUGUENIN-DEZOT N, et al. Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing [J]. *Nature Methods*, 2017, 14(7): 729-736.
- [108] LUO X, FU G, WANG R E, et al. Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria [J]. *Nature Chemical Biology*, 2017, 13(8): 845-849.
- [109] WANG Z A, KURRA Y, WANG X, et al. A versatile approach for site-specific lysine acylation in proteins [J]. *Angewandte Chemie International Edition*, 2017, 56(6): 1643-1647.
- [110] TSIEN R Y. The green fluorescent protein [J]. *Annu. Rev. Biochem.*, 1998, 67: 509-544.
- [111] CAMPBELL R E, TOUR O, PALMER A E, et al. A monomeric red fluorescent protein [J]. *PNAS*, 2002, 99(12): 7877-7882.
- [112] LANG K, DAVIS L, TORRES-KOLBUS J, et al. Genetically encoded norbornene directs site-specific cellular protein labelling *via* a rapid bioorthogonal reaction [J]. *Nat. Chem.*, 2012, 4(4): 298-304.
- [113] LUKINAVICIUS G, UMEZAWA K, OLIVIER N, et al. A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins [J]. *Nat. Chem.*, 2013, 5(2): 132-139.
- [114] SIEVERS E L, SENTER P D. Antibody-drug conjugates in cancer therapy [J]. *Annu. Rev. Med.*, 2013, 64: 15-29.
- [115] LAMBERT J M. Drug-conjugated antibodies for the treatment of cancer [J]. *Br. J. Clin. Pharmacol.*, 2013, 76(2): 248-262.
- [116] WANG L, ZHANG Z, BROCK A, et al. Addition of the keto functional group to the genetic code of *Escherichia coli* [J]. *PNAS*, 2003, 100(1): 56-61.
- [117] KOLB H C, FINN M G, SHARPLESS K B. Click chemistry: diverse chemical function from a few good reactions [J]. *Angewandte Chemie International Edition*, 2001, 40(11): 2004-2021.
- [118] AGARWAL P, KUDIRKA R, ALBERS A E, et al. Hydrazino-Pictet-Spengler ligation as a biocompatible method for the generation of stable protein conjugates [J]. *Bioconjug. Chem.*, 2013, 24(6): 846-851.
- [119] SCHMIDT M J, WEBER A, POTT M, et al. Structural basis of furan-amino acid recognition by a polyspecific aminoacyl-tRNA-synthetase and its genetic encoding in human cells [J]. *ChemBioChem*, 2014, 15(12): 1755-1760.
- [120] TEY S K. Adoptive T-cell therapy: adverse events and safety switches [J]. *Clin. Transl. Immunology*, 2014, 3(6): e17.
- [121] MA J S, KIM J Y, KAZANE S A, et al. Versatile strategy for controlling the specificity and activity of engineered T cells [J]. *PNAS*, 2016, 113(4): E450-E458.
- [122] JANG Y H, SEONG B L. Principles underlying rational design of live attenuated influenza vaccines [J]. *Clin. Exp. Vaccine Res.*, 2012, 1(1): 35-49.
- [123] SI L, XU H, ZHOU X, et al. Generation of influenza A viruses as live but replication-incompetent virus vaccines [J].

- Science, 2016, 354(6316): 1170-1173.
- [124] LI Q, WU Y J. A fluorescent, genetically engineered microorganism that degrades organophosphates and commits suicide when required [J]. Appl. Microbiol. Biotechnol., 2009, 82(4): 749-756.
- [125] RONCHEL M C, RAMOS J L. Dual system to reinforce biological containment of recombinant bacteria designed for rhizoremediation [J]. Appl. Environ. Microbiol., 2001, 67(6): 2649-2656.
- [126] WRIGHT O, STAN G B, ELLIS T. Building-in biosafety for synthetic biology[J]. Microbiology, 2013, 159(Pt 7): 1221-1235.
- [127] DIWO C, BUDISA N. Alternative biochemistries for alien life: basic concepts and requirements for the design of a robust biocontainment system in genetic isolation [J]. Genes (Basel), 2019, 10(1):17.
- [128] GAN F, LIU R, WANG F, et al. Functional replacement of histidine in proteins to generate noncanonical amino acid dependent organisms [J]. Journal of American Chemical Society, 2018, 140(11): 3829-3832.
- [129] MANDELL D J, LAJOIE M J, MEE M T, et al. Biocontainment of genetically modified organisms by synthetic protein design[J]. Nature, 2015, 518(7537): 55-60.
- [130] KOH M, YAO A, GLEASON P R, et al. A general strategy for engineering noncanonical amino acid dependent bacterial growth[J]. Journal of American Chemical Society, 2019, 141(41): 16213-16216.
- [131] ROVNER A J, HAIMOVICH A D, KATZ S R, et al. Recoded organisms engineered to depend on synthetic amino acids[J]. Nature, 2015, 518(7537): 89-93.
- [132] KATO Y. An engineered bacterium auxotrophic for an unnatural amino acid: a novel biological containment system [J]. PeerJ., 2015, 3: e1247.



**通讯作者:**沈玥(1986—),女,博士,研究员,研究方向为合成生物学、合成基因组学、DNA合成技术与工具开发。

E-mail:shenyue@genomics.cn



**第一作者:**付宪(1989—),男,博士,副研究员,研究方向为合成生物学、合成基因组学、蛋白质定向进化。

E-mail:fuxian1@genomics.cn

---

广告索引:武汉国家生物产业基地(后彩一)/中山大学达安基因股份有限公司(后彩二)/杜邦工业生物科技部(后彩三)/Cell Signaling Technology(后彩四)/诚志生命科技有限公司(封三)