

· 综 述 ·

同义密码子使用偏嗜性对 mRNA 半衰期及翻译调控的影响

李易聪^{1,2}, 蒲飞洋^{1,2}, 王慧慧^{1,2}, 程燕^{1,2}, 李倬^{1,2}, 马忠仁¹, 周建华^{1,3}

1 西北民族大学 生物医学研究中心, 甘肃 兰州 730030

2 西北民族大学 生命科学与工程学院, 甘肃 兰州 730010

3 中国农业科学院兰州兽医研究所, 甘肃 兰州 730046

李易聪, 蒲飞洋, 王慧慧, 程燕, 李倬, 马忠仁, 周建华. 同义密码子使用偏嗜性对 mRNA 半衰期及翻译调控的影响. 生物工程学报, 2022, 38(3): 882-892.

LI YC, PU FY, WANG HH, CHENG Y, LI Z, MA ZR, ZHOU JH. Effects of synonymous codon usage bias on mRNA half-life and translational regulation. Chin J Biotech, 2022, 38(3): 882-892.

摘 要: 随着基因组学和转录组学在不同生物体遗传和细胞生物学领域的广泛应用, 同义密码子使用的偏嗜性逐渐被接受, 并且在研究生物进化与生物表型之间的深层联系时, 同义密码子使用模式受到相关领域研究人员的重视。信使 RNA (messenger RNA, mRNA) 最终表达出具有正常生物活性的蛋白产物是生命活动的重要环节。被称为“第二遗传密码”的同义密码子使用模式, 可以通过精微调控翻译选择压力等分子机制, 从转录调控、翻译调控及代谢活动等水平表达其承载的遗传信息。研究表明, mRNA 半衰期的长短对 mRNA 活性以及转录和翻译过程有显著的影响。因此, 系统地归纳同义密码子使用模式在基因转录、翻译调控及翻译后修饰等生命活动中所扮演的角色, 将有助于全方位审视生物体如何巧妙利用密码子使用模式所产生的遗传效应来精准合成不同种类蛋白质, 并以此保障生长或分化的特定基因表达程序顺利执行、维持正常的生命周期。

关键词: 同义密码子使用模式; mRNA; 精微调控翻译选择压力; 转录; 翻译

Received: May 18, 2021; **Accepted:** October 8, 2021

Supported by: Natural Science Foundation of Gansu Province, China (20JR5RA505); the Fundamental Research Funds for the Central Universities, China (31920200073)

Corresponding author: ZHOU Jianhua. E-mail: zhoujianhua@xbmu.cn

基金项目: 甘肃省自然科学基金 (20JR5RA505); 中央高校基本科研业务费 (31920200073)

Effects of synonymous codon usage bias on mRNA half-life and translational regulation

LI Yicong^{1,2}, PU Feiyang^{1,2}, WANG Huihui^{1,2}, CHENG Yan^{1,2}, LI Zhuo^{1,2}, MA Zhongren¹, ZHOU Jianhua^{1,3}

1 Biomedical Research Center, Northwest Minzu University, Lanzhou 730030, Gansu, China

2 College of Life Science and Engineering, Northwest Minzu University, Lanzhou 730010, Gansu, China

3 Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou 730046, Gansu, China

Abstract: With the widespread application of genomics and transcriptomics in the genetics and cell biology of different species, synonymous codon usage bias has been gradually accepted and used to study the deep connection between biological evolution and biological phenotypes. It is an important part of the life activities that mRNA is expressed into proteins with normal biological activities. The synonymous codon usage patterns, which were named as ‘the second genetic codon’, can express genetic information carried by themselves at the levels of transcriptional regulations, translational regulations and metabolic activities through molecular mechanisms such as fine-tune translation selection. Some studies have shown that the length of mRNA half-life has significant impacts on mRNA activity and the process of transcription and translation. This review summarized the roles of synonymous codon usage patterns in transcription, translational regulation and post-translational modification, with the aim to better understand how organisms skillfully utilize the genetic effects caused by codon usage patterns to accurately synthesize different types of proteins, so as to ensure the growth or differentiation of the specific gene expression procedures to carry out smoothly and maintain the normal life cycle.

Keywords: synonymous codon usage patterns; mRNA; fine-tune translation selection; transcription; translation

密码子 (codon) 是由核苷酸三联体组成的遗传密码, 在生物体将遗传信息通过特定的核苷酸序列储存最终以活性蛋白表达的过程中发挥着“桥梁”作用。目前, 自然界中存在 64 种常规密码子, 其中 61 种能够编码 20 种氨基酸, 另外 3 个即 TAA、TGA 和 TAG 是终止密码子。这种由一种以上密码子编码同一个氨基酸的遗传现象称为密码子的简并性, 对应于同一氨基酸的密码子称为同义密码子, 同义密码子使用模式 (synonymous codon usage pattern) 这一遗传学概念也由此产生。虽然同义密码子编码同

一种氨基酸, 但在自然界中同义密码子使用具有偏嗜性^[1]。在不同生物体内, 密码子使用频率的高低可能受细胞内核糖体识别密码子的速率以及转运 RNA (transfer RNA, tRNA) 丰度所影响。研究人员按照同义密码子使用模式中密码子使用频率的强弱将其划分为优势、普通以及劣势密码子; 在研究不同微生物密码子使用模式过程中, 笔者团队又根据同义密码子使用偏嗜性将其划分为优势密码子 (overrepresented codon)、普通使用密码子以及稀有密码子 (underrepresented codon)^[2-4]。同义密码子使用偏

嗜性的存在意味着同义密码子使用模式不仅能够丰富生物遗传信息的内涵,还可将基因结构功能从单一核苷酸序列提升到核苷酸三联体层次,这使得生物基因组、mRNA、核糖体和 tRNA 联系得更紧密。

随着转录组学、蛋白组学以及体外表达技术的发展,研究人员发现同义密码子使用的偏嗜性可能影响着基因的转录、转录后修饰、翻译以及翻译后修饰等关键生命活动环节,这表明被誉为“第二遗传密码”的同义密码子使用模式^[1]所产生的精微调控翻译选择压力 (fine-tune translation selection) 在上述生命过程中也发挥重要作用。与传统对“第二遗传密码”的认知并不矛盾,由同义密码子使用模式所形成的特定的遗传密码以核苷酸三联体的形式指导着氨基酸序列的合成,进而影响着蛋白质的结构。无论是在微生物还是在动物细胞中,同义密码子使用模式形成的最终目标都是保障微生物或者动物细胞体蛋白按照遗传信息的要求实现“时间和空间”合理有序地表达。笔者团队曾在研究不同病毒基因组同义密码子使用模式的过程中发现,宿主细胞的翻译选择压力以及突变压力等遗传因素能够影响病毒同义密码子使用模式的形成,而病毒自身同义密码子使用模式又与其宿主嗜性有一定的相关性,这有利于病毒自身适应宿主细胞的遗传信息^[5-7]。

同义密码子使用偏嗜性与细胞内 tRNA 丰度具有很强的相关性。在翻译过程中,优势密码子往往识别高丰度 tRNA,这有助于核糖体快速扫描翻译以及新生多肽链迅速延伸。由此,如何通过优化密码子来提高目标基因在特定细胞中的表达效率一直都是研究的热点。虽然密码子优化能提升基因表达效率,但大量证据表明,同义密码子使用模式的改变也会影响翻译准确性和新生多肽链的构象。作为“第二遗传

密码”的同义密码子使用模式,其特定模式在特定生物体内所涉及的细胞、组织和不同发育阶段的基因表达的动态微调也逐渐引起研究人员的重视。因此,本文主要将近年来研究人员在同义密码子使用模式所形成的精微调控翻译选择压力对 mRNA 稳定性、tRNA 丰度波动、蛋白质翻译效率和空间折叠等方面影响的成果进行综述。此外,本文还对稀有密码子在众多优势密码子的环境下精微调控转录和翻译过程进行分析。显然,作为核酸与氨基酸之间“桥梁”的密码子,很可能通过同义密码子使用模式赋予基因额外的生物信息,从而影响核糖体扫描翻译 mRNA、多肽链延伸以及蛋白质合成等过程。进一步分析这些问题有助于研究人员解释同义密码子使用模式的成因和了解其调控效应,以便更好地理解密码子生物学与细胞内稳态和疾病之间的联系。

1 对 mRNA 半衰期的影响

由于 mRNA 的结构特征直接影响其本身的稳定性,因此 mRNA 半衰期 (half-life) 的长短对转录和翻译活动有显著的影响,这一生物学现象最早发现于 RNA 结构较为丰富的 5'和 3'非编码 RNA 元件中,其编码形成的 RNA 结构能够直接影响成熟 mRNA 半衰期的长短^[8-9]。自提出同义密码子概念以来,研究人员就一直探索同义密码子使用模式与基因之间的深层次联系。生物学功能相关的 mRNA 通常具有非常相似的同义密码子使用模式,并且它们的衰减速率也相近^[10-11]。在酵母菌中,mRNA 的半衰期短至 1 min 以内、长达 1 h 以上,差别很大。虽然研究人员已经发现了很多影响 mRNA 稳定性或衰减的因素,例如微小 RNA (microRNA, miRNA) 和 mRNA 的 3'或 5'端非编码区域的二级结构,但这些影响 mRNA 半衰期的因素仍停留在

RNA 结构和与翻译调控相关因子的结合能力等方面, 还无法更细致地解释细胞中整体 mRNA 半衰期的波动特征。因此, 在 mRNA 中可能广泛存在一种“内在动力”影响其稳定性。既然 mRNA 碱基成分在其衰减过程中发挥关键作用, 那么 mRNA 对应的同义密码子使用模式会在其中发挥调节作用吗? 胞嘧啶甲基化对 mRNA 半衰期的调节有一定作用, 提高 mRNA 编码序列中优势密码子的使用频率可以改变胞嘧啶甲基化位点的数量, 从而间接影响 mRNA 的稳定性^[12]。有研究发现, 斑马鱼胚胎和非洲爪蟾 (*Xenopus laevis*) 胚胎中的 mRNA 半衰期与稀有密码子的相对富集呈显著正相关^[13]。将稀有密码子插入至一小段基因中会导致 mRNA 稳定性显著下降^[14]。不同于影响 mRNA 半衰期的经典途径 (例如转录中途终止导致的 mRNA 降解、无义翻译介导的 mRNA 降解以及翻译无法终止导致的 mRNA 降解等监控途径), 当 mRNA 富集过多稀有密码子的时候, 会提高其自身脱腺苷酸化 (deadenylation) 的活性以及解除 5'端帽状结构 (decapping) 的能力^[14-16], 这些改变都会促使 mRNA 的半衰期变短。为进一步证实同义密码子使用模式对 mRNA 稳定性的影响, 研究人员利用 RNA 聚合酶的一个热敏等位基因终止转录, 然后在转录停止后的特定时间点用高通量测序来测定 mRNA 的丰度, 分析表明, 优势密码子富集在稳定 mRNA 中, 而稀有密码子则与半衰期短的 mRNA 的形成有关^[10]。密码子稳定系数 (codon stability coefficient, CSC) 用来表示密码子使用频率与 mRNA 稳定性之间的关系。其中酵母菌中的每种密码子的 CSC 值与 tRNA 适应指数 (tRNA adaptation index, tAI, 衡量同义密码子对于全基因组 tRNA 适应程度对翻译速率的影响) 关系密切。基于数学模拟分析的结果, 研究人员利用核苷酸修饰技术来

调控目标基因密码子使用偏嗜性以及 RNA 的拓扑结构后发现, mRNA 半衰期的变化与密码子优化以及编码区核酸序列二级结构的稳定性有着紧密的联系^[17]。此外, 改变 mRNA 的同义密码子使用偏嗜性可使 mRNA 半衰期提升或降低 10 倍左右。这是因为一味追求密码子优化将会破坏目标基因通过同义密码子使用模式储存的遗传信息, 并导致核糖体在扫描翻译 mRNA 的过程中出现多肽链延伸速率异常、翻译提前终止或者蛋白活性位点功能改变等情况^[10]。延续这样的研究思路, 研究人员已经在大肠杆菌、裂殖酵母、斑马鱼以及小鼠源、大鼠源和人源细胞系中发现了同义密码子使用模式与 mRNA 半衰期之间的关联^[18]。mRNA 编码序列中的优势密码子有利于运用自身对应高丰度 tRNA 的特点来吸引更多核糖体与其结合, 核糖体在 mRNA 上的“占位”能够显著阻止核酸内切酶对 mRNA 的降解作用, 并且延长 mRNA 的半衰期^[10,19-20]。此外, 一些特异性探测优势密码子的蛋白感受器 (例如 DEAD-box 蛋白家族中的 Dhh1p 蛋白) 能够通过探测 mRNA 序列中优势密码子的使用程度来介导 mRNA 的降解速率, 从而实现 mRNA 稳定性的调控^[21-22]。在人体细胞中, 作为 RNA 结合蛋白的 ILF2 和 ILF3 能够通过特异性识别末端富集 C/G 或者 A/T 碱基的 mRNA 序列来直接调控 mRNA 的丰度^[23]。根据上述内容可发现, 同义密码子使用模式所施加的生物及遗传学效应会影响 mRNA 半衰期的长短, 并且在转录过程中同义密码子使用模式可以通过调节 mRNA 的稳定性使 mRNA 携带的遗传信息充分体现出来。

2 对新生多肽链合成延伸过程的影响

随着高通量蛋白质组学技术的发展, 人们发现 mRNA 和蛋白质表达水平之间的关系并不

完全匹配, 这表明在翻译过程中可能有其他因素调控蛋白质的合成过程。研究发现, 同义密码子使用模式可以通过影响核糖体扫描翻译活动进而影响多肽链的合成与延伸^[1]。利用放射性标记氨基酸测定翻译速率变化的相关试验表明, 核糖体扫描优势密码子的速率明显较高^[24]。进一步分析发现, 在细胞中这些优势密码子所对应的 tRNA 丰度相对偏高; 相反, 核糖体扫描识别那些对应低丰度 tRNA 的同义密码子(通常是稀有密码子)时会降低自身扫描速率, 这将便于低丰度 tRNA 插入核糖体 A 位点并推进多肽链的延伸。核糖体图谱 (ribosome profiling) 技术为研究密码子使用模式与翻译之间的相关性提供了强有力的技术支撑。早期核糖体图谱研究印证了一个假说——如果核糖体遇到优势密码子时移动加快, 那么当核糖体 A 位点位于优势密码子上时, 核糖体相对密度就会降低^[25]。然而, 当研究人员用翻译延伸抑制剂预处理酿酒酵母 (*Saccharomyces cerevisiae*) 细胞后, 发现 tRNA 丰度和核糖体相对密度之间并不存在明显的相关性^[26]。此后, 研究人员利用该技术在酵母菌上建立的核糖体相对密度数据库进行核糖体相对密度与同义密码子使用模式相关性分析, 发现使用放线菌酮 (cycloheximide, 一种翻译延伸抑制剂) 预处理并不会立即终止 mRNA 上的核糖体扫描活动, 实际上受限制的核糖体仍能继续移动, 致使核糖体占用率计算产生偏差^[27-28]。在不使用放线菌酮预处理的酵母菌核糖体图谱试验中, tRNA 丰度和核糖体占用率呈明显负相关, 这证实核糖体扫描识别优势密码子花费的时间较少^[27,29-30]。不同种类 tRNA 的丰度复杂多变, 核糖体与正确 tRNA 结合之前, 会在胞质丰富的 tRNA 中缓慢地甄选、校对和剔除待进入 A 位点的 tRNA, 因此不同的 tRNA 丰度也会影响核糖体的扫描翻译速率。

由此可见, 细胞内变化的 tRNA 丰度是影响核糖体扫描活动及其速率的众多因素之一。tRNA 反密码子的第一位 (5'端) 碱基和 mRNA 密码子的第 3 位 (3'端) 碱基之间的摆动配对 (与 Watson-Crick 碱基互补配对原则不同) 也影响核糖体扫描翻译活动^[31]。当密码子缺少同源 tRNA 并且必须由同工受体 tRNA 解码, 或者当 tRNA 修饰发生在反密码子的第一位碱基时, 就会出现碱基摆动配对。与 Watson-Crick 碱基配对相比, 摆动配对相对不利, 它在编码序列中不常出现, 却在核糖体图谱试验中占用了较多的核糖体 A 位点, 表明核糖体读取这些摆动配对密码子的速率较慢^[1]。显著影响新生多肽链延伸速率的密码子利用碱基摆动性可至少与一种 tRNA 配对, 导致该类密码子被其他同义密码子替换后会影响到此处 tRNA 的配对, 进一步影响新生多肽链的稳定性 (图 1)。

分析核糖体 A 位点同义密码子与 tRNA 相互识别配对所产生的翻译“瞬时”速率变化发现, 处于 A 位点的密码子可调控识别 tRNA 和新生多肽链延伸的速率^[32]。很多研究表明, 同义密码子使用模式也能调控 mRNA 上核糖体的移动速率。应用流式细胞术对特定细胞分选后进行蔗糖密度梯度离心, 可将翻译抑制后的翻译系统划分为核糖核蛋白 (ribonucleoprotein, RNP)、80S 核糖体和相关多聚体复合物 3 种组分, 其中 RNP 复合体组分的转录与翻译机制无关, 而 80S 核糖体和多聚体组分的转录与一个或多个核糖体相关, 说明这两种组分在翻译机制中发挥着重要作用; 另外, 核糖体启动一段时间后, 具有较多优势密码子的 mRNA 离开多聚体和 80S 核糖体, 出现在 RNP 组分中的速率比对照组的非优势密码子快得多, 这进一步表明, 核糖体在这些 mRNA 上移动得更快^[10]。利用新兴的实时监测翻译技术能找到很多优势密

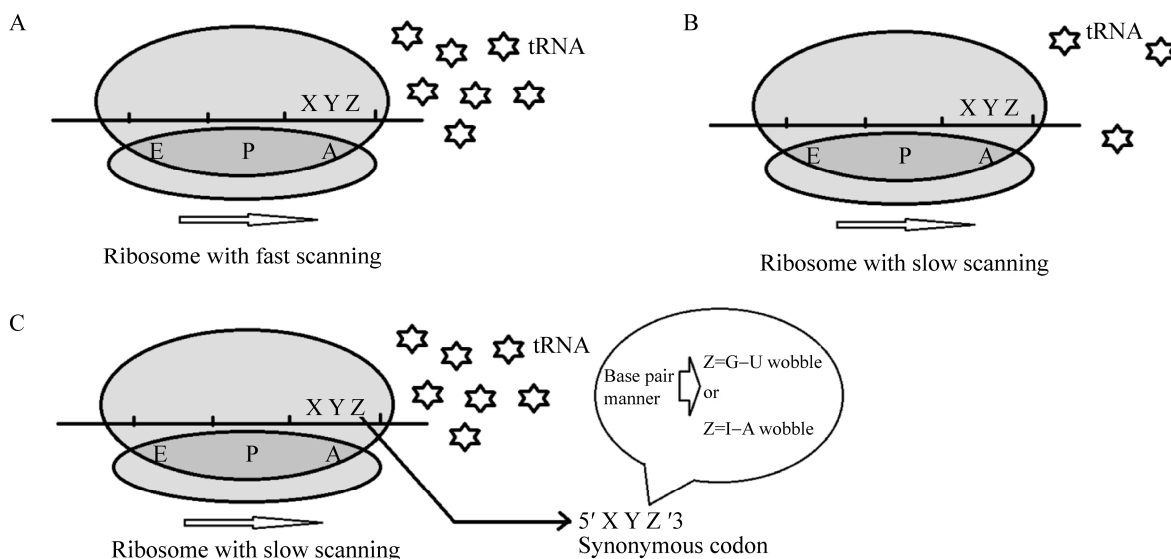


图 1 同义密码子使用模式可以通过 tRNA 识别机制调控核糖体的扫描速率

Figure 1 The synonymous codon usage patterns can regulate the scanning rate of ribosomes via tRNA recognition mechanism. (A) When the specific tRNA abundance is high, the position A in ribosome rapidly accepts the corresponding tRNA via Watson-Crick base-pairing, leading to ribosome with high scanning speed. (B) When the specific tRNA abundance is low, the position A in ribosome probably recognizes the correct tRNA slowly via Watson-Crick base-pairing, leading to ribosome with low scanning speed. (C) When the position A needs to recognize the corresponding tRNA via non-canonical Watson-Crick base-pairing (e. g. G-U wobble or I-A wobble), the ribosome turns down its scanning speed.

密码子含量对新生多肽链延伸速率影响的证据。在体外翻译系统中,将长度为 1 600 个核苷酸的 mRNA 进行密码子优化后比野生型(对照组) mRNA 的翻译用时缩短 1.5 min^[33]。近年来,研究人员利用荧光标记抗体技术来实时监测哺乳动物细胞中单个 mRNA 的翻译,结果表明,密码子优化的 mRNA 翻译速率为 4.9 个密码子/s,而野生型 mRNA 的翻译速率为 3.1 个密码子/s,翻译速率相对提高了 58%^[34]。上述翻译调控机制阐明同义密码子使用模式所表现出来的精微调控翻译选择压力对新生多肽链延伸过程的影响,并且这种精微调控翻译选择压力的作用范围覆盖了基因转录与翻译过程。

3 对翻译效率的影响

在分析密码子使用偏嗜性对翻译效率的影

响时,诸多因素成为人们探寻密码子使用模式的生物学和遗传学效应的“噪声”。mRNA 自身形成的发卡结构显著影响蛋白翻译效率,尤其是在起始密码子附近,因为在翻译启动前发卡结构越是稳定的 mRNA 越需要更多的能量推动核糖体对其进行扫描并保证多肽链顺利地合成^[35]。在大肠杆菌中,稀有密码子与核酸序列折叠能量降低有关,与优势密码子相比,它们往往富含 AT 碱基,因此当稀有密码子位于编码序列 5'端附近时,能提升下游基因的翻译效率^[36]。

在体内直接测量翻译速率^[34,37-38]并通过计算机模拟核糖体的起始与延伸能够说明翻译起始通常是蛋白质合成的限速步骤。如果翻译起始速率受到限制,那么其延伸率的任何变化都只会导致核糖体径流更快,而不会使每分子 mRNA 产生蛋白质的速率发生任何改变。缓慢启动才

能确保 mRNA 上不会出现核糖体拥堵现象,否则会引起核糖体之间的冲突并降低翻译效率^[39]。如果翻译起始阶段核糖体扫描受到限制,那么密码子使用模式和翻译效率之间明显的联系则意味着翻译起始和延伸存在内在联系。稀有密码子上任何核糖体的减速信息都可能传递至同一 mRNA 上其他核糖体,然后通过降低核糖体移离翻译起始区的速率来限制翻译启动速率^[40]。用翻译动力学模型解释基因表达:稀有密码子分散分布在 mRNA 序列中,可能会显著降低基因整体翻译效率;而稀有密码子聚集在翻译起始区,则可能有助于基因在核糖体扫描翻译起始阶段正向调控基因的表达。

4 对蛋白质折叠的影响

多肽合成速率决定蛋白折叠构象的观点由来已久,密码子使用模式与蛋白质二级结构形

成之间的关系表明,密码子使用偏嗜性很可能通过调控多肽合成速率从而实现蛋白构象的准确折叠(图 2)。近年来,用核磁共振法测定编辑密码子使用模式对蛋白结构影响的研究揭示:不同密码子使用模式的 mRNA 所表达的蛋白质空间构象明显不同、半胱氨酸氧化和体外蛋白质稳定性也存在显著差异^[41]。同时,tRNA 丰度的变化对蛋白质折叠有重要影响,人为提升细胞 tRNA 丰度后会导致多肽空间错误折叠率增加。研究人员利用粗糙脉孢菌(*Neurospora crassa*)翻译系统研究上述问题时发现,尽管优化密码子的荧光素酶 mRNA 比野生型 mRNA 翻译延伸更快并产生更多蛋白质,但大部分蛋白不具有功能活性,因此使用密码子优化的 mRNA 表达功能蛋白质的实际效率要低^[33]。优化荧光素酶报告基因关键功能结构域中的 200 个密码子后会导致荧光素酶活性降低,说明在这一关

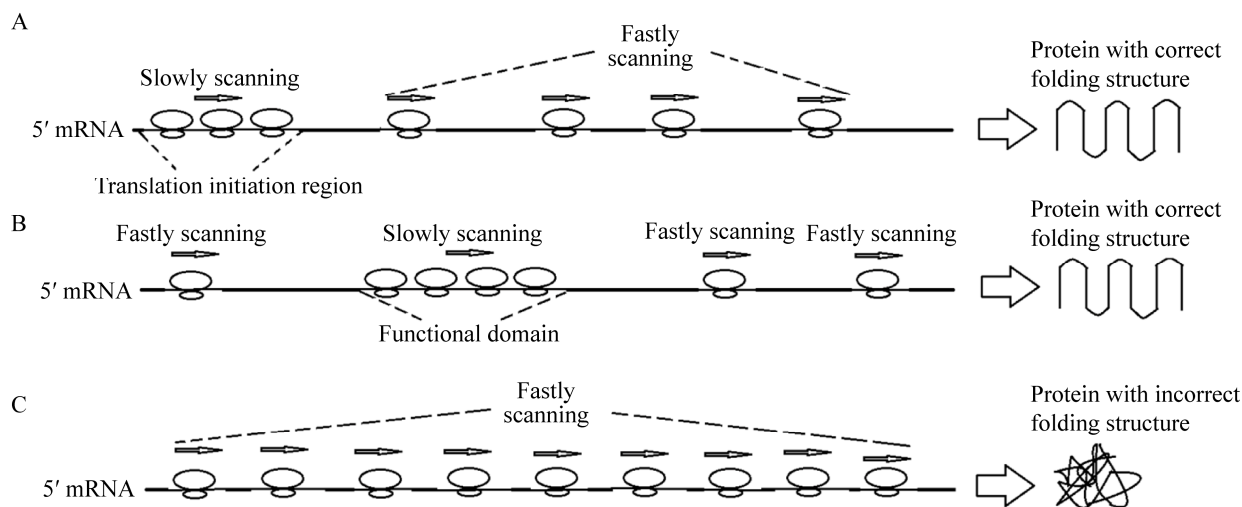


图 2 同义密码子使用模式调节核糖体移动速率的相关机制以及所导致蛋白质折叠的效应

Figure 2 Relevant mechanisms involving the unequal usage of synonymous codons and their effect on folding protein structure. (A) 'Non-preferred codons', which are selected in high frequencies at the translation initiation region of the mRNA, are able to get down scanning speeds of ribosomes and lead to decreasing ribosome density at the downstream. (B) 'Non-preferred codons', which cluster at the protein functional domain of mRNA, are able to mediate correct folding structure of the functional domain. (C) Lack of 'non-preferred codons' in mRNA can result in high density of ribosomes with high scanning speed along the whole mRNA and finally contribute to protein with incorrect folding structure.

键区域使用稀有密码子才能调节核糖体扫描速率并保证蛋白质正确折叠。分析大肠杆菌基因发现,其稀有密码子聚集区位于能够编码且控制蛋白结构相关序列的下游^[42],并且稀有密码子与蛋白质二级结构正确形成的关系更密切^[43]。核糖体在编码蛋白结构域连接区的 mRNA 上移动较慢,进一步验证了同义密码子使用偏嗜性、核糖体翻译速率与新生多肽链结构域的正确折叠之间存在相关性。在大肠杆菌中优化稀有密码子虽然可以提高蛋白质总产量,却也导致蛋白质错误折叠率上升。为保障蛋白质空间结构正确,稀有密码子会减缓关键氨基酸基序的翻译速率。生物信息学分析蛋白质二级和三级结构附近的密码子使用也同样证实了以上结果^[30,44-45],表明稀有密码子协同蛋白构象的形成而参与翻译环节。

随后,有研究人员利用这些现象阐述一系列与蛋白构象相关的保守密码子的使用情况。尽管形成蛋白质结构域连接区的氨基酸主要由稀有密码子编码,但与结构域的关键基序对应的 mRNA 保守区倾向于选择优势密码子进行翻译编码,这反映出优势密码子可能对蛋白某些重要结构的准确合成具有重要作用。分子伴侣等因子在 mRNA 翻译过程中起监控作用,生物体将同义密码子使用模式与 tRNA 匹配来确保翻译正确率,为蛋白质正确折叠增加了“双保险”。而且优势密码子与正确的 tRNA 结合的能力比其他因子更强,受到类似 tRNA 与正确 tRNA “竞争上岗”方面的影响小。另外,生物体优选的同义密码子通过提高蛋白质折叠准确率从而减少细胞不必要的代谢耗损。与此同时,研究人员发现稀有密码子也能维持蛋白质的生物学活性,例如阻止不溶性多肽中间体和蛋白质聚集体的形成^[46],或者调节多肽在核糖体轨道上形成螺旋结构的能力^[47-49]。此外,笔者团

队曾报道编码病毒蛋白关键结构(尤其是不同折叠单元交界处)的同义密码子使用模式具有很强的偏嗜性^[50-52]。使用强偏嗜密码子有利于病毒基因在宿主细胞中开展自身基因的转录与翻译活动,从而成功实现子代病毒在宿主细胞内的增殖。将病毒作为最直接且高效的外源基因表达载体,其同义密码子使用模式的形成将利于病毒适应宿主细胞内环境,从而高效利用细胞的资源实现自身基因的高效复制、转录及翻译^[53]。

5 在基因转录及翻译中存在极化现象

同义密码子优化(synonymous codon optimality)能够影响多种细胞活动,而这些影响直接体现在基因转录的活跃程度以及蛋白质表达产量等生命现象中^[54-55]。研究表明,优化基因密码子虽然提高了翻译效率,但是错误折叠率大大增加,生成的无活性产物对细胞还具有毒性^[56]。优势密码子和稀有密码子各自排布在 mRNA 上特定的位置,才能更好地提升翻译的效率和准确度。笔者团队曾对一些病毒基因组翻译起始区密码子使用模式进行分析,发现稀有密码子倾向于在翻译起始区域聚集^[57-58]。这一生物学现象也反映出稀有密码子在维持和调控新生蛋白表达效率及功能活性方面发挥着翻译调控功能。在基因组、转录组以及蛋白组学的基础上进行研究发现,同义密码子使用模式对 mRNA 的 5'端翻译效率的影响程度要强于其他区段^[59]。用果蝇 S2 细胞(*Drosophila* S2 cells)研究热休克蛋白 70(heat shock protein 70, Hsp70)启动子下游密码子使用模式对基因转录翻译的影响,发现临近启动子的密码子使用偏嗜性对转录和翻译活动的影响明显强于其他区域^[60]。同义密码子使用偏嗜性对不同基因区段转录翻译效率的影响不同,揭示其在转录翻

译过程中存在极化现象,并且这种由密码子使用产生的极化效应在细胞基因组精密、有序及高效的转录和翻译水平上发挥着重要的作用。

6 总结与展望

本文主要将近几年来围绕同义密码子使用模式对生物体遗传进化活动影响的研究结果进行归纳总结。旨在强调看似无法直接影响蛋白质一级结构的同义密码子竟然能够充当一只“无形的”指挥棒,影响生命活动的方方面面。随着研究方法的不断改进,研究人员通过基因组学和蛋白组学技术、检测报告基因技术、遗传和生物信息学分析技术,已逐渐将同义密码子使用模式从不会影响多肽链合成和蛋白质生物学功能的观点纠正为密码子使用偏嗜性可以通过精微调控翻译选择压力来指导蛋白的空间构象、稳定性和功能的正确形成。mRNA中不同密码子的位置、密码子的排序规则以及其他尚不明确的因素影响着同义密码子使用模式对各个表达环节的调节作用。因此,还需要研究人员对同义密码子使用偏嗜性与不同模式生物各种表型之间的相关性进行系统地研究和归纳。同义密码子使用偏嗜性直接影响外源基因在宿主细胞表达水平的高低,并且密码子优化能够促进不同外源基因通过宿主细胞表达系统提高其自身翻译表达效率。这样,若想实现外源基因在工程菌/细胞系中的高效表达,适度改造目标基因的同义密码子使用模式是一个不错的选择。另外,研究人员也可以尝试对目的基因翻译起始区的密码子进行编辑,即加入一定数目宿主细胞对应的稀有密码子,使得核糖体启动扫描翻译消耗少量能量并且避免核糖体在mRNA序列上“拥堵”。这为人们今后在基因工程领域中“人为调控”目的基因在工程细菌/细胞系中的翻译表达提供了参考依据。因此,如

何建立敏感、稳定及高通量的实验技术平台是当前密码子使用领域研究人员的主要任务之一。相信随着实验技术的不断改进以及人们对同义密码子使用模式在遗传学和生物学中调控机制的深入研究,同义密码子偏嗜性将可以更好地服务于相关基因突变导致的一些人类疾病的研究工作。

REFERENCES

- [1] Hanson G, Collier J. Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol*, 2018, 19(1): 20-30.
- [2] Cao XA, Hu W, Shang YJ, et al. Analyses of nucleotide, synonymous codon and amino acid usages at gene levels of *Brucella melitensis* strain QY1. *Infect Genet Evol*, 2018, 65: 257-264.
- [3] Han SY, Hu W, Kan W, et al. Analyses of genetics and pathogenesis of *Salmonella enterica* QH with narrow spectrum of antibiotic resistance isolated from yak. *Infect Genet Evol*, 2020, 82: 104293.
- [4] Wang YN, Ji WH, Li XR, et al. Unique features of nucleotide and codon usage patterns in mycoplasmas revealed by information entropy. *Biosystems*, 2018, 165: 1-7.
- [5] Zhou JH, Gao ZL, Zhang J, et al. The analysis of codon bias of foot-and-mouth disease virus and the adaptation of this virus to the hosts. *Infect Genet Evol*, 2013, 14: 105-110.
- [6] Ge ZY, Li XR, Cao XA, et al. Viral adaption of staphylococcal phage: a genome-based analysis of the selective preference based on codon usage Bias. *Genomics*, 2020, 112(6): 4657-4665.
- [7] Zhou JH, Li XR, Lan X, et al. The genetic divergences of codon usage shed new lights on transmission of hepatitis E virus from swine to human. *Infect Genet Evol*, 2019, 68: 23-29.
- [8] Lee S, Lykke-Andersen J. Emerging roles for ribonucleoprotein modification and remodeling in controlling RNA fate. *Trends Cell Biol*, 2013, 23(10): 504-510.
- [9] Geisberg J, Moqtaderi Z, Fan X, et al. Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. *Cell*, 2014, 156(4): 812-824.

- [10] Presnyak V, Alhusaini N, Chen YH, et al. codon optimality is a major determinant of mRNA stability. *Cell*, 2015, 160(6): 1111-1124.
- [11] Wang YL, Liu C, Storey J, Tibshirani R, et al. Precision and functional specificity in mRNA decay. *PNAS*, 2002, 99(9): 5860-5865.
- [12] Xing YZ, Gong RQ, Xu YC, et al. Codon usage bias affects α -amylase mRNA level by altering RNA stability and cytosine methylation patterns in *Escherichia coli*. *Can J Microbiol*. 2020, 66(9): 521-528.
- [13] Bazzini A, Del Viso F, Moreno-Mateos M, et al. Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *EMBO J*. 2016, 35(19): 2087-2103.
- [14] Sweet T, Kovalak C, Collier J. The DEAD-box protein Dhh1 promotes decapping by slowing ribosome movement. *PLoS Biol*, 2012, 10(6): e1001342.
- [15] Qian W, Zhang J. Codon usage bias and nuclear mRNA concentration: correlation vs. causation. *PNAS*, 2021, 118(20): e2104714118.
- [16] Shoemaker C, Green R. Translation drives mRNA quality control. *Nat Struct Mol Biol*, 2012, 19(6): 594-601.
- [17] Mauger DM, Cabral BJ, Presnyak V, et al. mRNA structure regulates protein expression through changes in functional half-life. *PNAS*, 2019, 116(48): 24075-24083.
- [18] Hia F, Takeuchi O. The effects of codon Bias and optimality on mRNA and protein regulation. *Cell Mol Life Sci*, 2021, 78(5): 1909-1928.
- [19] Hanson G, Alhusaini N, Morris N, et al. Translation elongation and mRNA stability are coupled through the ribosomal A-site. *RNA*, 2018, 24(10): 1377-1389.
- [20] Buschauer R, Matsuo Y, Sugiyama T, et al. The Ccr4-Not complex monitors the translating ribosome for codon optimality. *Science*, 2020, 368(6488): eaay6912.
- [21] Zlotorynski E. RNA decay: Dhh1p condemns mRNAs with non-optimal codons to decay. *Nat Rev Mol Cell Biol*, 2016, 17(11): 675.
- [22] Radhakrishnan A, Chen Y, Martin S, et al. The DEAD-box protein Dhh1p couples mRNA decay and translation by monitoring codon optimality. *Cell*, 2016, 167(1): 122-132.
- [23] Hia F, Yang S, Shichino Y, et al. Codon bias confers stability to human mRNAs. *EMBO Rep*, 2019, 20(11): e48220.
- [24] Sørensen M, 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 Mol Biol*, 1991, 222(2): 265-80.
- [25] Ingolia N, Ghaemmaghami S, Newman JR, et al. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science*, 2009, 324(5924): 218-23.
- [26] Qian WF, Yang JR, Pearson N, et al. Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet*, 2012, 8(3): e1002603.
- [27] Hussmann JA, Patchett S, Johnson A, et al. Understanding biases in ribosome profiling experiments reveals signatures of translation dynamics in yeast. *PLoS Genet*, 2015, 11(12): e1005732.
- [28] Gerashchenko MV, Gladyshev VN. Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res*, 2014, 42(17): e134.
- [29] Gardin J, Yeasmin R, Yurovsky A, et al. Measurement of average decoding rates of the 61 sense codons *in vivo*. *eLife*, 2014, 3: e03735.
- [30] Weinberg D, Shah P, Eichhorn S, et al. Improved ribosome-footprint and mRNA measurements provide insights into dynamics and regulation of yeast translation. *Cell Rep*, 2016, 14(7): 1787-1799.
- [31] 金志忠. 密码子与反密码子的本质与拓展. *生物学通报*, 2020, 55(11): 13-16.
- Jin ZZ. The essence and extension of codon and anticodon. *Bull Biol*, 2020, 55(11): 13-16 (in Chinese).
- [32] Gamble C, Brule C, Dean K, et al. Adjacent codons act in concert to modulate translation efficiency in yeast. *Cell*, 2016, 166(3): 679-690.
- [33] Yu CH, Dang YK, Zhou ZP, et al. Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Mol Cell*, 2015, 59(5): 744-754.
- [34] Yan XW, Hoek TA, Vale RD, et al. Dynamics of translation of single mRNA molecules *in vivo*. *Cell*, 2016, 165(4): 976-989.
- [35] Tsao D, Shabalina S, Gauthier J, et al. Disruptive mRNA folding increases translational efficiency of catechol-*O*-methyltransferase variant. *Nucleic Acids Res*, 2011, 39(14): 6201-6212.
- [36] Goodman D, Church G, Kosuri S. Causes and effects of N-terminal codon bias in bacterial genes. *Science*, 2013, 342(6157): 475-479.
- [37] Morisaki T, Lyon K, DeLuca K, et al. Real-time

- quantification of single RNA translation dynamics in living cells. *Science*, 2016, 352(6292): 1425-1429.
- [38] Wu B, Eliscovich C, Yoon YJ, et al. Translation dynamics of single mRNAs in live cells and neurons. *Science*, 2016, 352(6292): 1430-5.
- [39] Mitarai N, Sneppen K, Pedersen S. Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. *J Mol Biol*, 2008, 382(1): 236-45.
- [40] Potapov I, Mäkelä J, Yli-Harja O, et al. Effects of codon sequence on the dynamics of genetic networks. *J Theor Biol*, 2012, 315: 17-25.
- [41] Buhr F, Jha S, Thommen M, et al. Synonymous codons direct cotranslational folding toward different protein conformations. *Mol Cell*, 2016, 61(3): 341-351.
- [42] 邢燕子. 密码子偏好在翻译前水平对基因表达的影响研究[D]. 上海: 华东理工大学, 2020.
Xing YZ. Study the role of codon bias in regulating gene expression pre-translationally[D]. Shanghai: East China University of Science and Technology, 2020 (in Chinese).
- [43] Chaney J, Steele A, Carmichael R, et al. Widespread position-specific conservation of synonymous rare codons within coding sequences. *PLoS Comput Biol*, 2017, 13(5): e1005531.
- [44] Zhou M, Wang T, Fu JJ, et al. Nonoptimal codon usage influences protein structure in intrinsically disordered regions. *Mol Microbiol*, 2015, 97(5): 974-987.
- [45] Alexaki A, Kames J, Holcomb D, et al. Codon and codon-pair usage tables (CoCoPUTs): facilitating genetic variation analyses and recombinant gene design. *J Mol Biol*, 2019, 431(13): 2434-2441.
- [46] Jacobson G, Clark P. Quality over quantity: optimizing co-translational protein folding with non-'optimal' synonymous codons. *Curr Opin Struct Biol*, 2016, 38: 102-110.
- [47] Hetz C, Zhang K, Kaufman R. Mechanisms, regulation and functions of the unfolded protein response. *Nat Rev Mol Cell Biol*, 2020, 21(8): 421-438.
- [48] Chaney J, Clark P. Roles for synonymous codon usage in protein biogenesis. *Annu Rev Biophys*, 2015, 44: 143-166.
- [49] Angov E. Codon usage: nature's roadmap to expression and folding of proteins. *Biotechnol J*, 2011, 6(6): 650-659.
- [50] Ma XX, Wang YN, Cao XA, et al. The effects of codon usage on the formation of secondary structures of nucleocapsid protein of peste des petits ruminants virus. *Genes Genomics*, 2018, 40(9): 905-912.
- [51] Zhou JJ, You YN, Chen HT, et al. The effects of the synonymous codon usage and tRNA abundance on protein folding of the 3C protease of foot-and-mouth disease virus. *Infect Genet Evol*, 2013, 16: 270-274.
- [52] Zhou JH, Zhang J, Chen HT, et al. The codon usage model of the context flanking each cleavage site in the polyprotein of foot-and-mouth disease virus. *Infect Genet Evol*, 2011, 11(7): 1815-1819.
- [53] Sirihongthong T, Jitobaom K, Phakaratsakul S, et al. The relationship of codon usage to the replication strategy of parvoviruses. *Arch Virol*, 2019, 164(10): 2479-2491.
- [54] Fu JJ, Dang YK, Counter C, et al. Codon usage regulates human KRAS expression at both transcriptional and translational levels. *J Biol Chem*, 2018, 293(46): 17929-17940.
- [55] Zhou ZP, Dang YK, Zhou M, et al. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *PNAS*, 2016, 113(41): E6117-E6125.
- [56] Jitobaom K, Phakaratsakul S, Sirihongthong T, et al. Codon usage similarity between viral and some host genes suggests a codon-specific translational regulation. *Heliyon*, 2020, 6(5): e03915.
- [57] Zhou JH, Zhang J, Sun DJ, et al. Potential roles of synonymous codon usage and tRNA concentration in hosts on the two initiation regions of foot-and-mouth disease virus RNA. *Virus Res*, 2013, 176(1/2): 298-302.
- [58] Zhou JH, Su JH, Chen HT, et al. Clustering of low usage codons in the translation initiation region of hepatitis C virus. *Infect Genet Evol*, 2013, 18: 8-12.
- [59] Chu D, Kazana E, Bellanger N, et al. Translation elongation can control translation initiation on eukaryotic mRNAs. *EMBO J*, 2014, 33(1): 21-34.
- [60] Yang Q, Lyu XL, Zhao FZ, et al. Effects of codon usage on gene expression are promoter context dependent. *Nucleic Acids Res*, 2021, 49(2): 818-831.

(本文责编 郝丽芳)