

评论

DOI: 10.12211/2096-8280.2021-034

Z-基因组的生物合成奥秘被揭示

金交羽, 周佳海

(中国科学院深圳理工大学, 深圳合成生物学创新研究院, 中国科学院定量工程生物学重点实验室, 广东 深圳 518055)

摘要: *Science* 期刊于2021年4月30日刊登了3篇关于Z-基因组的研究论文。本文将重点评论其中赵素文、张雁和赵惠民三个实验室的合作论文: 介绍多酶系统介导的Z-基因组生物合成、降低宿主菌中dATP浓度的dATPase和DUF550发现, 并阐明Z-基因组的测序鉴定和功能意义。44年前, 苏联科学家首次发现二氨基嘌呤(Z)存在于蓝藻噬菌体(cyanophage) S-2L的基因组中。Z是一种特殊的碱基, 它完全取代了腺嘌呤并与胸腺嘧啶形成三个氢键, Z-基因组的生物合成通路一直是未解之谜。上海科技大学赵素文实验室、天津大学张雁实验室和伊利诺伊大学厄巴纳-香槟分校/新加坡科技研究局的赵惠民实验室组成的合作团队, 通过生物信息学、计算生物学和生物化学手段揭示了负责Z-基因组生物合成的多酶系统。研究发现此通路可能也存在于数十个分布在全球各地的噬菌体中, 包括在上海被发现和分离的 *Acinetobacter* phage SH-Ab 15497。合作团队使用HPLC-UV、质谱技术和纳米孔测序验证了Z碱基存在于 *Acinetobacter* phage SH-Ab 15497中, 且完全取代了A碱基, 识别位点中含有A碱基的限制性核酸内切酶通常无法切割Z-DNA, 因此Z-DNA赋予了噬菌体逃避宿主限制性核酸内切酶攻击的进化优势。Z基因组生物合成通路的解析, 可促进新型核酸产品和相关新DNA技术的开发。

关键词: 噬菌体; Z-基因组; Z碱基; 新型核酸; 生物合成

中图分类号: Q931 **文献标志码:** A

The mystery of Z-genome biosynthesis has been elucidated

JIN Jiaoyu, ZHOU Jiahai

(Shenzhen Institute of Advanced Technology, Chinese Academy of Science, Shenzhen Institute of Synthetic Biology, Chinese Academy of Science Key; Laboratory of Quantitative Engineering Biology, Shenzhen; 518055, Guangdong, China)

Abstract: On April 30, 2021, three Z-genome research papers were published in *Science*. This article will comment on the collaborative work performed at laboratories led by Zhao Suwen at Shanghai Technology University, Zhang Yan at Tianjin University, and Zhao Huimin at University of Illinois at Urbana-Champaign and Agency for Science, Technology and Research (A*STAR) Singapore. First, we introduce the multi-enzyme system mediated Z-genome biosynthesis. Then we address the discovery of dATPase and DUF550 that reduce the concentration of dATP in the host

收稿日期: 2021-03-18 修回日期: 2021-04-30

基金项目: 国家重点研发计划“合成生物学”重点专项(2018YFA0901900)

引用本文: 金交羽, 周佳海. Z-基因组的生物合成奥秘被揭示[J]. 合成生物学, 2022, 3(1): 1-5

Citation: JIN Jiaoyu, ZHOU Jiahai. The mystery of Z-genome biosynthesis has been elucidated[J]. Synthetic Biology Journal, 2022, 3(1): 1-5

bacteria. Finally, we highlight the sequencing and identification of the Z-genome and its functional significance. Forty-four years ago, Soviet scientists discovered for the first time that diaminopurine (Z) was presented in the DNA of Cyanophage S-2L. Z is a modified unique base that replaces adenine (A), which form three hydrogen bonds with thymine (T), but the biosynthetic pathway of the Z-genome had been an unsolved mystery for the past decades. Very recently, the multi-enzyme system to biosynthesize Z-genome has been characterized by various methods including bioinformatics, computational biology and biochemistry methods through collaborations among Professors Zhao Suwen, Zhang Yan and Zhao Huimin. They concluded that this pathway exists in dozens of phages distributed around the world, including *Acinetobacter* phage SH-Ab 15497, which was discovered and isolated in Shanghai. The team used HPLC-UV, MS and nanopore sequencing to verify that Z exists in *Acinetobacter* phage SH-Ab 15497 and completely replaced adenine. Restriction endonucleases usually cannot digest Z-DNA at recognition sites containing A. Therefore, Z-DNA gives phage an evolutionary advantage to evade attack by host restriction endonucleases. The findings of Z-genome biosynthesis will shed light on the development of new nucleic acid agents and novel DNA technologies.

Keywords: phage; Z-genome; Z-base; new nucleic acid agents; biosynthesis

DNA由4种不同的脱氧单磷酸核苷(dNMPs)构成, dNMPs的区别在于分子结构中碱基的不同。这4种碱基分别为:胞嘧啶(Cytosine, C)、胸腺嘧啶(Thymine, T)、腺嘌呤(Adenine, A)和鸟嘌呤(Guanine, G)。DNA的碱基上可以被添加各种化学基团,其中包括甲基、羟基、氨基酸、多胺、单糖和双糖等^[1]。目前报道的碱基修饰已经超过了50种^[2]。在真核生物中, DNA的碱基修饰可以在不改变基因型的情况下影响基因的表达表型,部分碱基被修饰过的DNA可以通过与蛋白质间的形成相互作用使DNA免于被降解,并参与基因的调控^[3]。比如研究得比较充分的胞嘧啶甲基化,甲基转移酶基因能够与限制性内切酶基因形成配对,在破坏外源DNA的同时可保护自身的DNA免受限制性内切酶的切割^[4]。尽管DNA的碱基修饰种类繁多,但大多数遵循经典的Watson-Crick碱基互补配对原则^[4-5]。

1977年, Ivan Khudyakov等^[6]从噬菌体(cyanophage)S-2L的基因组分离和鉴定出一种特殊的碱基——2,6-二氨基嘌呤(Z)。与典型的腺嘌呤碱基(A)相比, Z碱基在嘌呤分子的2号碳原子上多出一个氨基基团。在经典的Watson-Crick碱基互补配对原则中, A和T形成2根氢键[图1(a)], G和C形成3根氢键^[7]。研究发现, S-2L的基因组中的Z完全取代了A,并与T互补配对形成3根氢键, Z

碱基额外的氨基可与T形成第三根氢键[图1(b)],从而增加DNA双螺旋结构的热稳定性^[8]。此后,多家研究单位,包括法国巴斯德研究所,都向Ivan Khudyakov索要噬菌体S-2L及其宿主。2002年,法国巴斯德研究所的研究者将噬菌体S-2L进行测序,并把它的序列及潜在的应用申请了专利(WO2003093461)。他们在专利中推测,噬菌体基因组中腺苷琥珀酸合成酶(PurA)的同源蛋白可能参与Z碱基的生物合成。但多年来,此蛋白的功能一直没有被阐释清楚,且dZTP是如何被选择性地整合到噬菌体S-2L的DNA上等生物学问题也没有得到更深入研究。

2021年4月30日,上海科技大学赵素文副教授、天津大学张雁教授和伊利诺伊大学厄巴纳-香槟分校/新加坡科技研究局赵惠民教授合作在*Science*上发表了题为“A widespread pathway for substitution of adenine by diaminopurine in phage genomes”的论文,报道了参与Z碱基生物合成的多酶系统,阐明了以dGTP为起点,一步步转化为可被DNA聚合酶利用的dZTP的完整路径,并提出噬菌体的Z-基因组能够帮助逃避宿主限制性内切酶攻击的基础^[9]。

由于NCBI数据库中并无噬菌体S-2L的基因组注释,合作团队首先对噬菌体S-2L的基因组进行了注释,顺利发现了此前专利中所提到的PurA同

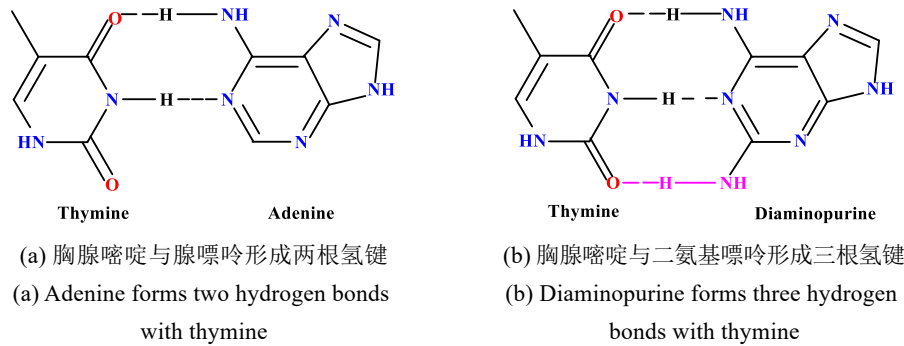


图1 胸腺嘧啶与腺嘌呤、二氨基嘌呤形成的氢键作用

Fig. 1 Thymine forms hydrogen bonds with adenine and diaminopurine

源蛋白，并将该酶命名为PurZ。接下来以*EcPurA*为模板，构建*CpPurZ*的结构模型，并与*EcPurA*的活性口袋进行比对，发现*EcPurA*中的催化残基Asp13在*CpPurZ*中对应的是Ser15，这个残基正好对着嘌呤的2号碳，因此他们推测Asp到Ser的突变正好可以容纳比腺嘌呤更大的片段。这一推测被接下来的分子对接所验证，分子对接的结果表明，*CpPurZ*及其在其他噬菌体中的同源蛋白(*ApPurZ*、*SbPurZ*、*VpPurZ*和*SpPurZ*)的底物可能是dGMP/GMP和ATP。他们利用质谱实验对几种PurZs酶活性反应的中间体和产物进行定性和定量的检测，最终发现4种PurZs在镁离子存在时，以dGMP、ATP和L-天冬氨酸为底物生成2-氨基脱氧单磷酸腺苷琥珀酸(ADAS)。由于这些噬菌体中并不存在嘌呤从头合成通路中的腺苷琥珀酸裂解酶(PurB)，因此他们推测宿主菌中的PurB应该具有一定的泛杂性，可以将ADAS裂解为dZMP和延胡索酸。随后他们使用大肠杆菌的*EcPurB*进行了假设验证，发现它的确可以将ADAS裂解为dZMP和延胡索酸(图2)。

合作团队通过构建PurZ编码基因的基因组邻

近区相似性网络，发现PurZ编码基因周围存在2个功能相关基因。其中一个基因编码的蛋白属于金属依赖的磷酸水解酶HD家族。该蛋白的功能随后被证实为dATP磷酸水解酶(因此被命名为dATPase)，它可在钴离子的存在下，将dATP/dADP/dAMP水解为磷酸和2'-脱氧腺苷(dA)。dATPase的生物学功能是不可逆地降低宿主菌中的dATP的浓度。而另一个基因编码了属于DUF550家族的蛋白。经实验证明，DUF550为脱氧嘌呤核苷三磷酸的焦磷酸水解酶，可在钴离子的存在下，将dATP/dGTP水解为dAMP/dGMP和焦磷酸。因此，DUF550一方面可以降低宿主细胞中的dATP浓度，另一方面可以为PurZ提供反应底物dGMP。综合来看，dATPase和DUF550都发挥了降低宿主菌中dATP浓度的作用，这使得噬菌体的DNA聚合酶更容易优先选择热力学上更稳定的dZTP，来进行噬菌体DNA的合成。

同时，他们还发现细菌中的GMP激酶能够将dZMP磷酸化生成dZDP及dZTP，dZTP从而可能被噬菌体的DNA聚合酶利用，与dTTP、dGTP及dCTP一起参与Z-基因组的合成(图3)。

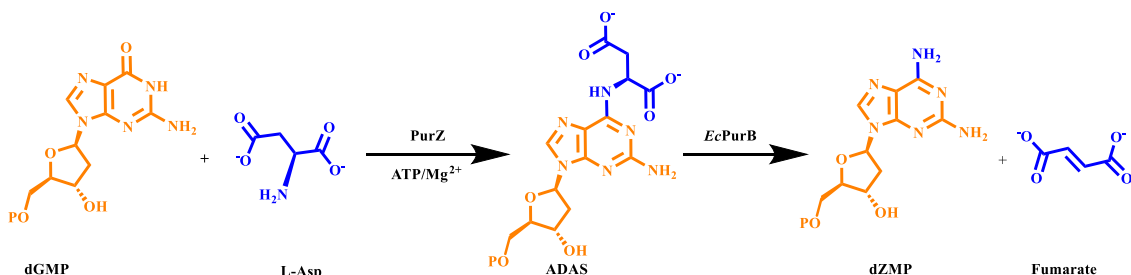


图2 dZMP的生物合成通路

Fig. 2 Biosynthetic pathway for dZMP

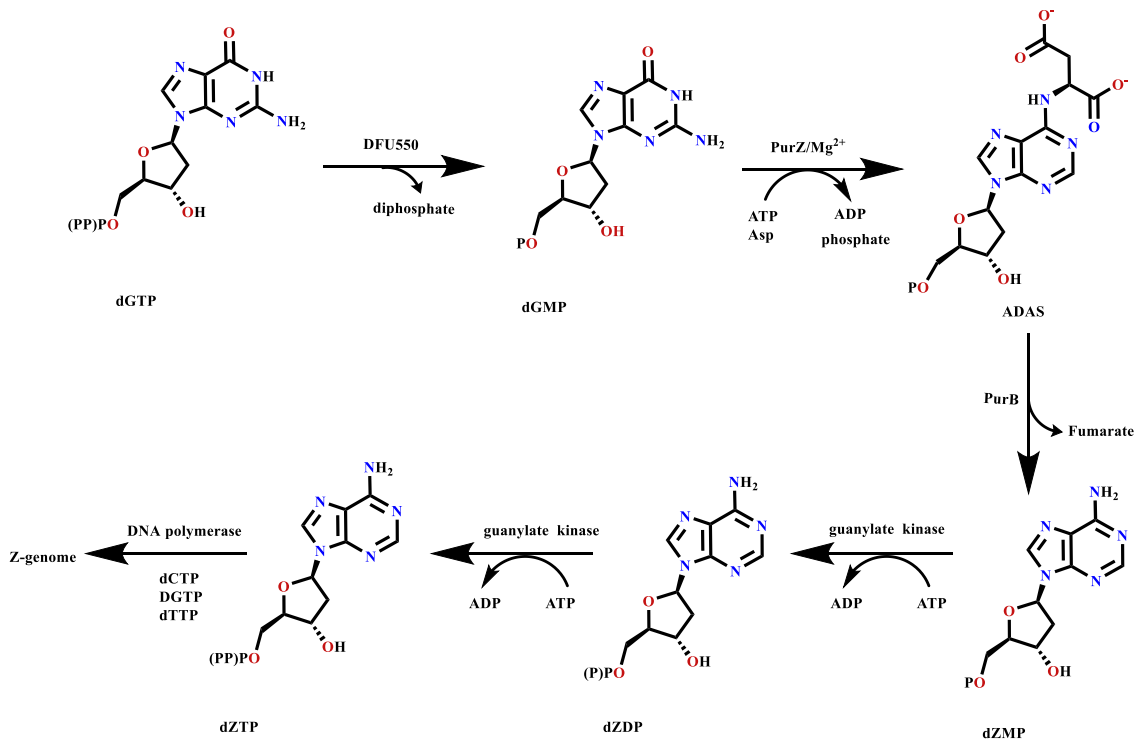


图3 推测的Z-基因组生物合成通路

Fig. 3 Proposed biosynthetic pathway for Z-genome in phage

合作团队通过酶解法水解含有Z-基因组的 *Acinetobacter* phage SH-Ab 15497 (由上海交通大学医学院的何平教授提供) DNA, 并使用HPLC-MS检测水解产物发现, 噬菌体DNA中的确存在2'-脱氧2-氨基腺苷(dZ), MS/MS结果也可以观察到Z碱基的存在。同样的水解产物用HPLC-UV检测发现, dZ/dT和dC/dG的比值分别为0.99和1.05。他们使用纳米孔测序法对含有Z-基因组的 *Acinetobacter* phage SH-Ab 15497进行测序, 结果显示dZTP被整合到噬菌体的DNA上。对 *Acinetobacter* phage SH-Ab 15497的DNA进行的限制性内切酶实验表明, 识别序列中含有A的限制性内切酶 *EcoRI*、*PstI*和 *Sau3AI*并不能切开噬菌体的DNA, 其中 *Sau3AI*在该噬菌体的DNA上含有超过200个识别位点^[10]。这些结果表明, Z完全取代了A, 并且Z-基因组赋予噬菌体能够躲避宿主限制性内切酶攻击的优势。

Z-基因组生物合成通路的解析, 为科学家研究DNA的修饰、结构、组装与调控等打开了一扇新的大门。多个含Z-基因组的活性生物体发

现, 使通过生物方法制备含Z碱基的新型核酸产品得以实现, 也将促进含Z碱基的DNA纳米技术开发。在用核酸构建人工设计的结构中, 一旦A碱基被Z碱基所替代, 将会在对应的碱基配对处增加一对氢键的键能, 这对于核酸链的稳定性和结构产生明显的影响, 许多新的DNA折纸结构将被设计出来。在基于DNA的数据存档研究中^[11], 当Z碱基被用于替代A碱基后, 能够躲避宿主限制性内切酶攻击, 提高了储存数据在活体细胞中的稳定性, 而数据的读取可以通过纳米孔测序技术获得且质量不受影响。另外, 这项工作也有望应用到合成生物学领域, 如通过将Z-基因组生物合成酶结合到工程噬菌体中, 特别是已成功应用于多耐药细菌的临床治疗并挽救了患者生命的那些噬菌体中^[12-13], 或许可以扩大这些噬菌体的宿主范围。

*Science*上同期报道了两项来自法国的Z-基因组相关研究工作, 其中一篇文章的主要发现与赵素文等的结果类似, 但是用了较多的篇幅在描述PurZ蛋白及其系列复合物的晶体结构信息^[14]; 另

一篇文章则侧重于 DNA 聚合酶 DpoZ 的发现和鉴定, 研究结果表明在复制阶段该酶催化 dZTP 加到链上的活性比 dATP 强^[15]。Z-基因组生物中所发现的 DNA 聚合酶能够容忍非经典碱基对的能力是非常值得关注的, 在 Watson-Crick 经典碱基配对原则得到扩充的同时保持 DNA 复制的高保真度, 这对维持生命有机体遗传信息的完整性具有非常重要的意义。值得一提的是, 2021 年 4 月 23 日 *Nature Communications* 发表的一篇文章报道了 dATP 磷酸水解酶 DatZ 的系列晶体结构和一个疑似聚合酶 PrimPol 的晶体结构^[16], 不过 PrimPol 对催化 dZTP 与 dATP 加到链上的活性没有区分度, 这一点与前面提到的 DpoZ 不同。

参 考 文 献

- [1] GOMMERS-AMPT J H, BORST P. Hypermodified bases in DNA [J]. *The FASEB Journal*, 1995, 9(11): 1034-1042.
- [2] SOOD A J, VINER C, HOFFMAN M M. DNAmdb: the DNA modification database[J]. *Journal of Cheminformatics*, 2019, 11(1): 30.
- [3] WU H, ZHANG Y. Reversing DNA methylation: Mechanisms, genomics, and biological functions[J]. *Cell*, 2014, 156(1/2): 45-68.
- [4] WEIGELE P, RALEIGH E A. Biosynthesis and function of modified bases in bacteria and their viruses[J]. *Chemical Reviews*, 2016, 116(20): 12655-12687.
- [5] WARREN R A. Modified bases in bacteriophage DNAs[J]. *Annual Review of Microbiology*, 1980, 34: 137-158.
- [6] KIRNOS M D, KHUDYAKOV I Y, ALEXANDRUSHKINA N I, et al. 2-Amino adenine is an adenine substituting for a base in S-2L cyanophage DNA[J]. *Nature*, 1977, 270(5635): 369-370.
- [7] KHUDYAKOV I Y, KIRNOS M D, ALEXANDRUSHKINA N I, et al. Cyanophage S-2L contains DNA with 2, 6-diaminopurine substituted for adenine[J]. *Virology*, 1978, 88(1): 8-18.
- [8] CRISTOFALO M, KOVARI D, CORTI R, et al. Nanomechanics of diaminopurine-substituted DNA[J]. *Biophysical Journal*, 2019, 116(5): 760-771.
- [9] ZHOU Y, XU X X, WEI Y F, et al. A widespread pathway for substitution of adenine by diaminopurine in phage genomes[J]. *Science*, 2021, 372(6541): 512-516.
- [10] SZEKERES M, MATVEYEV A V. Cleavage and sequence recognition of 2, 6-diaminopurine-containing DNA by site-specific endonucleases[J]. *FEBS Letters*, 1987, 222(1): 89-94.
- [11] CEZE L, NIVALA J, STRAUSS K. Molecular digital data storage using DNA[J]. *Nature Reviews Genetics*, 2019, 20(8): 456-466.
- [12] SCHOOLEY R T, BISWAS B, GILL J J, et al. Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant acinetobacter baumannii infection[J]. *Antimicrobial Agents and Chemotherapy*, 2017, 61(10): e00954-e00917.
- [13] DEDRICK R M, GUERRERO-BUSTAMANTE C A, GARLENA R A, et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*[J]. *Nature Medicine*, 2019, 25(5): 730-733.
- [14] SLEIMAN D, GARCIA P S, LAGUNE M, et al. A third purine biosynthetic pathway encoded by amino adenine-based viral DNA genomes[J]. *Science*, 2021, 372(6541): 516-520.
- [15] PEZO V, JAZIRI F, BOURGUIGNON P Y, et al. Noncanonical DNA polymerization by amino adenine-based siphoviruses[J]. *Science*, 2021, 372(6541): 520-524.
- [16] CZERNECKI D, LEGRAND P, TEKPINAR M, et al. How cyanophage S-2L rejects adenine and incorporates 2-amino adenine to saturate hydrogen bonding in its DNA[J]. *Nature Communications*, 2021, 12:2420.

通讯作者: 周佳海(1972—), 男, 研究员, 博士生导师。研究方向是工业生物技术、合成生物学相关的微生物酶学以及活性天然产物的化学生物学, 通过会聚酶学、结构生物学及蛋白质理性设计等技术方法, 重点研究酶的结构与催化机理、定向进化酶的科学规律、分子智能设计等。

E-mail: jiahai@siat.ac.cn



第一作者: 金交羽(1995—), 女, 硕士研究生, 研究助理。研究方向为酶结构的机制研究。

E-mail: jy.jin@siat.ac.cn

