

基因组DNA G-四链体的形成及功能

潘海晨, 薛红月, 徐雅楠*

(内蒙古民族大学生命科学与食品学院, 通辽 028000)

摘要: G-四链体(G-quadruplex, G4)是广泛存在于基因组中的非典型核酸二级结构, 是富含鸟嘌呤的核酸通过碱基间Hoogsteen配对形成的四联体堆积。G4结构在DNA端粒、基因启动子区、核糖体DNA和重组热点区等成簇存在, 在DNA复制调控、转录调控、维持基因组的完整性和重组过程调控等方面起作用。近年来, 体内G4结构相关的功能及调控机理研究层出不穷。本文将围绕G4在基因组内的分布、转录调节及DNA复制相关的功能研究进展进行阐述, 总结目前对G4结构和功能, 解释G4调控的方式, 从而为肿瘤、病毒感染和神经退行性病变等相关疾病的机理探究、治疗和药物研发提供思路与方向。

关键词: G-四链体; 转录调节; DNA复制; 基因组稳定

Biological functions of G-quadruplex in genome

PAN Haichen, XUE Hongyue, XU Ya'nan*

(College of Life Sciences and Food Engineering, Inner Mongolia Minzu University, Tongliao 028000, China)

Abstract: G-quadruplex (G4), one of non-canonical DNA secondary structures, is widespread in genomes. It is formed by guanines and Hoogsteen base pairs creating ring-like formations known as G-quartets. Guanine-rich DNAs can fold into four-stranded structures that contain stacks of G-quartets. G4 structure-forming sequences (G4 motifs) are significantly enriched at key functional units like telomeres, promoters, ribosomal DNA and recombination hot spots, pointing to a variety of critical cellular functions including replication regulation (including cell cycle regulation), transcriptional regulation, telomere maintenance, maintenance of genomic integrity and recombination process regulation. More recent work has demonstrated that the functions and regulatory mechanisms are related to the G4 structure *in vivo*. This review focused on the research progress related to the distribution of G4 within the genome, transcriptional regulation and DNA replication; summarized the current understanding of the G4 structure and function; and tried to explain the mode of regulation by G4. This paper tried to provide ideas and guidance for the mechanism, treatment and drug development for related diseases including tumors, viral infections and neural degenerative diseases.

Key Words: G-quadruplex; transcriptional regulation; DNA replication; genomic stability

G-四链体结构(G-quadruplex, G4)是一种常见的核酸二级结构, 可以在富含G的核酸(DNA/

RNA)中由多个G-四联体(G-quartet, 由四个鸟苷酸通过Hoogsteen配对形成的环状平面)堆积形成(图

收稿日期: 2021-11-20

基金项目: 国家自然科学基金项目(12104246); 内蒙古自然科学基金项目(2018LH03019); 内蒙古自治区高等学校科学研究项目(NJZY18157); 内蒙古民族大学博士基金项目(BS409)

第一作者: E-mail: haichenpan@163.com

*通信作者: E-mail: xuyan@imu.edu.cn

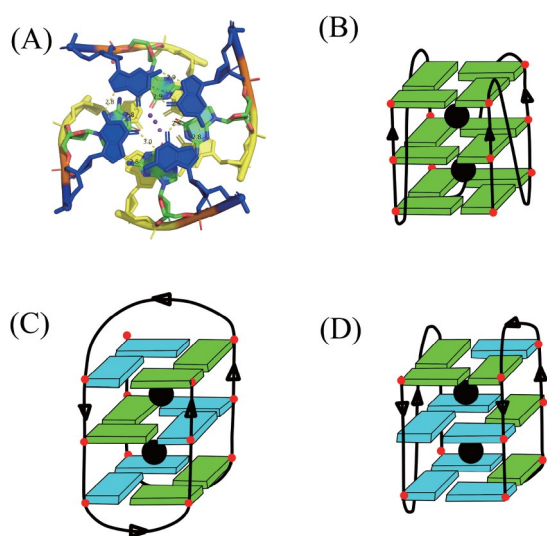
1A)。早在1910年, 在对鸟苷酸的研究中就发现鸟苷酸浓溶液可以自发形成胶体。半个世纪后, 通过对鸟苷酸纤维的特性研究发现它们具有自发形成四链体螺旋结构的倾向^[1]。随后, 在体外研究中也证实了保守的端粒DNA序列可以形成G-四链体结构^[2]。当时, 科学家Aaron Klug就预言既然G4可以在体外如此稳定, 那么自然界中就有可能找到利用它的方式^[3]。后来发现, 在活细胞内的DNA分子中也确实存在G4结构。G4序列存在于基因组DNA端粒、基因启动子区、核糖体DNA和重组热点区等^[4], 并不断被证实与肿瘤的发生、基因调控、病毒感染等过程相关。

能是ORC识别复制起始位点的关键^[6]。而且现有研究已证实, 内源性DNA G4s富集在开放染色质区域和高表达癌基因的启动子中, 这些G4s最近被发现与乳腺癌的潜在转录因子程序相关^[7]。人的线粒体DNA中也有G4结构, 并在线粒体结构不稳定性中发挥作用^[8]。除此之外, 细菌和病毒中也可以形成G4结构^[9,10], 如在席卷全球的COVID-19中也存在大量形成G4结构的潜在位点^[11,12]。核酸二级结构对生物过程产生影响的证明已经彻底改变了现代生物学, 尤其是对G4结构与功能的关注。

1 G4的结构特征及基因组内的分布特征

G4结构具有多态性, 可以由多聚鸟苷酸链内或链间产生。形成链内G4需要核酸单链具有四个或更多的多聚G区, 最典型的序列是含有四个保守的鸟苷酸区, 中间由相对短的环境间隔($G_3 \cdot N_x G_{3-5} N_x G_{3-5} N_x G_{3-5}$, N可以是任意核苷酸, x为1~7)^[13,14]。间隔序列长度、核苷酸种类及位置会使G4稳定性发生变化。当G4序列中间的间隔序列低于3个核苷酸时, G4结构的稳定性下降。同时, 间隔序列长度的增加也会使G4结构的稳定性降低。而链间G4可以由两条或四条单链形成平行或反平行的G4结构^[15]。一价阳离子可以稳定G4结构, K^+ 比 Na^+ 、 NH_4^+ 对G4结构的稳定性效果更好^[16]。 K^+ 在进入G-triplets堆积形成的中心孔道时引发的能量损耗更低^[17]。能够形成G4的单链DNA的最小长度是多少? 在体外, 这个问题的答案相当简单: 一个热力学稳定的G4可以最少只需要在G4序列中间隔一个核苷酸($G_3N_1)_3G_3$, 只有15 bp。而在体内还需要综合考虑其他因素, 包括邻近DNA的状态、局部作用于DNA的扭转和纵向力, 以及DNA结合蛋白和核小体的存在等。

结构研究显示, G4的构象具有多样性, 主要的两种构象为平行构象和反平行构象(图1B、图1C)。在平行构象中, 形成G4的核酸链的方向一致, 连接环结构如螺旋状分布在G4结构周围, 连接着一个链的末端和另一个链的起始。而在反平行构象中, 形成G4的链的方向相反, 连接环结构使相邻的链连接在一起。在一些情况下, G4也可能出现混合构象(图1D)。与DNA G4结构的多样性



A: G-四链体中的G-四联体, K^+ , d(A(GGGTTA)3GGG, PDB ID: 1KF1^[19]; B: G-四链体的平行构象; C: G-四链体的反平行构象; D: G-四链体的混合构象(蓝色为顺式, 绿色为反式)

图1 G-四链体的结构形式

G4结构在基因复制、转录、翻译和表观遗传调控中均起作用。虽然细胞内的大多数基因组DNA都是典型的双链B型构象, 但DNA的转录和复制等过程需要解开双链。形成的单链DNA就有机会折叠成各种非B构象, 其中一些比双链DNA结构稳定性更高, G-四链体结构是其中最为重要的一种。G4结构大量存在于细胞分裂的S期^[5]。除了广为人知的端粒中的G4结构阻止DNA聚合酶的延伸从而影响复制外, 更有趣的是, 在90%的人类DNA复制起始位点附近都含有G4序列。而复制起始复合体(origin recognition complex, ORC)特异性识别复制起始位点附近形成的G4结构, 提示G4可

相比, RNA G4结构不依赖溶液环境而具有更高的热稳定性。稳定性的差异主要源于核糖骨架上的2'-OH。G4结构的多样性本质上是由G4的一级结构(即碱基排列顺序)决定的, 但与G-四联体的数量、环的长度、链的走向和手性等都相关^[15,18]。现在, 我们已经对G4结构的多样性有深入的了解, 但是这些瞬时产生的二级结构的热动力学特性与其生物学功能的关系仍有待进一步探索。

G4结构最初被认为的特殊结构, 但现在已经被普遍认为可为体内存在的调控生物学过程的基本机制。G4在细胞内大量、广泛的存在也预示着它们参与细胞内多个过程, 是细胞内的重要调节机制, 如端粒维持、基因表达调控^[20]等。G4结构在细胞内的瞬时精确调控和在不同细胞状态时可能发生的不同变化, 使G4与癌症、神经退行性疾病、肿瘤、病毒感染等人类疾病的发病相关。因此, 本文进一步探讨基因组DNA中的G4的分布特征以及相关的调控功能, 将为相关疾病的治疗提供更多的思路。

G4广泛存在于各种生物中。大肠杆菌中, G4潜在形成序列在转录基因、次级代谢生物合成和信号传导的调控区含量较多。人类基因组中, G4潜在形成序列在基因组的很多重要区域都有分布, 包括复制起始位点、癌基因启动子区、端粒及免疫球蛋白的转换区等, 分别在复制调控、转录调控、维持基因组的完整性和重组过程调控等方面起作用。通过不同的算法, 已经确定除了端粒序列含有的重复序列可以形成20 000多G4结构以外, 人类基因组中有可能形成G4结构的区域超过350 000, 甚至更多^[21,22]。随着G4结构广泛存在被证明, 越来越多的研究人员开始在现有技术上重新考量, 评价G4结构形成可能产生的影响, 从而进一步确证了基因组内G4结构的广泛形成^[23]。在最近的研究中, 来自中国科学院的研究团队利用G4结合蛋白DHX36结合结构域设计出一种人工G4探针蛋白, 结合ChIP-seq技术, 在A549细胞系中测得了超过10万个G4结构^[24]。病毒基因组中, SARS-CoV-2基因组中有四个G4潜在形成序列, 其中有一个存在于SARS-CoV-2的核衣壳蛋白编码区中(RG-1), 相关研究为抗病毒药物的发展提供了新的视角, 包括现在影响全球的COVID-19^[25]。

多物种基因组分析表明, G4结构在多物种中普遍存在^[26]。在进化过程中, 从低等的单细胞生物到高等哺乳动物和人的基因组, G4潜在形成序列的数量、长度、密度和多样性呈现增加趋势。随着物种复杂性的增加, 整个基因组中G4潜在形成序列的密度和长度比例也随之增加。小鼠和人的基因组中G4潜在形成序列最多, 但是G4潜在形成序列的数量却与GC含量无关^[27]。富含GC的区域不利于核小体形成, 而核小体缺失则可能有利于G4的形成, 形成G4结构的DNA区域以排除核小体的开放构象存在, 从而提高转录^[28]。G4潜在形成序列在染色体上成簇存在, 并富集于基因区域。在高等生物中, G4结构更加密集地出现在基因的调控区中。在大多数物种中, 短环G4序列比较保守, 但是在哺乳动物中环的长度变化较多。在酵母菌基因组比对研究中, G4潜在序列的保守性较高, 且在G4基序中突变导致G4 DNA结构破坏的关键位点保守性更高。进化过程中, 含有G4的基因的比例逐渐增加, 尤其是在转录因子中的增加更加明显。通过对人类基因组分析发现, 近20 000个基因的启动子区存在G4潜在形成序列, 并集中于转录起始位点(transcription start site, TSS)上游1 kb和下游500 nt区域内^[28], 并与已确定的增强子、转录因子结合位点和核酸酶超敏位点重叠, G4潜在序列在原癌基因中形成G4的能力较高, 如c-MYC、BCL2、c-KIT、VEGF、KRAS等。不同物种中, G4潜在形成序列在基因启动子区域富集程度有较大差异, 如酵母基因组中启动子区1 kb范围内G4潜在形成序列是哺乳动物的六倍。所以, 从以上基因组内的分布情况可以看出, G4结构相关的转录调控是从酵母菌到人类普遍存在的调控方式, 且保守性极高。

2 基因组内形成G4结构是转录调控的关键因素

G4结构对转录起始过程的影响是双向的, 既有激活, 也有抑制。G4结构对转录调控的作用首先体现在可以被转录因子特异性识别上。随着G4-染色质免疫沉淀测序(G4-chromatin immunoprecipitation sequencing, G4-ChIP-seq)技术这一分辨G4结构的高通量测序技术的发展, 人类染色质中G4

结构图谱的描述日渐清晰。有研究进一步证实, 主要的G4-ChIP-seq位点与转录调控相关, 尤其是基因启动子区内的位点^[29]。针对转录起始位点功能区域的G4形成的研究发现, 与其他功能区域相比, TSS功能区域能够形成更多稳定的G4, 意味着在转录驱动的启动子中, G4的形成更为活跃和稳定^[30]。也就是说, 并不是所有的G4潜在序列都会形成G4 DNA, 不同细胞的不同状态下, 转录差异与G4 DNA的形成相互影响, 而导致G4 DNA形成产生差异。形成的G4结构会成为转录相关蛋白质的特异性识别位点而起作用, 如染色质重塑的关键蛋白SMARCA4是直接结合折叠成二级结构的G4 DNA而起到调控作用的。SMARCA4结合G4的位点42%分布在基因的启动子区^[31]。但是, 由于检测技术的限制, 到现在为止的很多相关研究都是通过抗体或特异性结合G4的化学分子稳定G4 DNA得出的结论, 有一些瞬时发生的生物学过程可能会受到干扰, 出现一些假阳性的结果。所以, 如果有更能反映G4 DNA结构, 而又不会影响其瞬时变化的技术出现, 那么相关结果可能有很大的差别。

G4结构对转录过程中RNA聚合酶的作用也会产生影响。由于G4结构稳定性强, 形成后需要特定的解旋酶才能解开, RNA聚合酶在沿着模板链迁移时G4结构会对RNA聚合酶产生阻碍作用, 也称为转录阻滞(transcription blockage)。但是, 实际的影响更复杂。通过体外的单分子检测: G4潜在序列在非模板链上, 分子内形成的G4 DNA与RNA聚合酶间不会产生直接的相互作用, G4的形成将会使模板链更加稳定, 显著提升了RNA聚合酶的起始和延伸效率, 增加转录效率^[32]。这种效应是链依赖的, G4潜在序列在模板链上时, 转录效率并不会增加。但是在模板链上的G4潜在形成序列会与形成的RNA链形成分子间DNA:RNA杂合G4, 这一结构可能会阻滞后续RNA聚合酶的作用, 而抑制连续的转录^[20]。

G4结构对基因的转录调控还出现在细胞的压力环境下。细胞在内源或外源因素影响下时产生由反应活性氧(ROS)导致的鸟嘌呤氧化, 生成8-oxo-7,8-二氢鸟嘌呤(OG)。当OG的生成发生在没有G4潜在序列的启动子中时, 细胞的修复机制会

在OG位点发生糖苷键的断裂形成AP位点, 然后AP位点被APE1识别并被切割修复, 整个修复过程不增加基因转录水平。但是, 当OG的生成发生在具有G4潜在序列的启动子中, 并且这个启动子的G4序列下游还有第五个G重复序列时, OG形成AP后, 原有的G4结构就被掺入第五个替代G序列的G4结构取代, 使AP位点被排挤到一个环结构中, 而APE1结构虽然可以识别环中的AP位点, 但是剪切功能降低, 并刺激了转录因子的招募, 最终增强了转录^[20,33]。

在物种进化过程中, DNA二级结构已成为一种重要的表观遗传调控机制, 对转录调控起作用。在甲基化程度较高的物种如猪和蚕的基因组研究中, 检测胞嘧啶的甲基化水平发现, 含有G4基因的胞嘧啶甲基化频率明显低于所有基因的平均水平。人类基因组中近29 000 CpG岛(CpG islands, CGIs)存在于基因的启动子区, CpG岛的甲基化阻止启动子的激活, 从而实现调控基因表达。DNA甲基化水平的控制对基因调控至关重要, 控制CGIs低甲基化的因素仍未被完全解读。通过G4-ChIP-seq检测出的G4s结构(79%)是在人类基因组中未甲基化的CGIs中形成的。CGIs中形成的G4s结构还与DNMT1显著共定位, 并证实可抑制DNMT1的甲基化活性。所以, G4s结构形成是抑制DNMT1并保护CGIs不被甲基化的机制之一^[34]。至少在哺乳动物和昆虫中, DNA G4和甲基化之间可能存在一种拮抗关系, 而且这种拮抗关系可能是保守的^[34]。

G4结构对基因表达调控的影响方式还有很多是未知的。G4结构在受到调控的基因中更为丰富, 说明G4在级联信号转导的基因表达的精细调控中起作用。例如, 在人类、黑猩猩和小鼠等哺乳动物中, 转录起始位点 ± 5 K的G4密度远高于果蝇和线虫^[35]。这暗示G4结构的可逆变化是高等生物中的基因调控机制之一。同时, 在高等生物中, G4的密集分布也可能是其调控功能发挥作用的关键。G4聚集在人类基因组中很明显, 但在血吸虫和线虫基因组中不存在, 在单细胞酵母的染色体中极少。所以, G4结构的调控功能可能在低等生物中尚未建立, 直到它们在高等动物中发育, 通过形成大量的G4结构发展成为新的表观遗

传机制。

3 G4在DNA复制和损伤修复通路中的调控作用

在DNA复制过程中，G4结构的形成影响基因组稳定性。G4结构对DNA复制的影响分成两个方面：首先是G4结构对复制体的阻碍作用。无论是前导链还是滞后链都会受G4结构形成的影响，包括滞后链中端粒形成的G4，都是天然的复制叉屏障。多种人类癌细胞中的比对研究都发现G4结构增加，这可能是G4加工不正常导致的。具体原因可能是调节基因组稳定性的细胞过程的改变或癌组织中G4位点染色质状态的改变^[36]。其次，G4结构的形成也是DNA复制必需的。有研究发现，当细胞内缺失G4结合蛋白Zou1时，G4结构的形成减少，会导致细胞生长缓慢及UV敏感^[37]。Zou1结合G4并招募核苷酸切除修复因子(nucleotide excision repair factor, NER factor)，从而增加基因组的稳定性。胰腺癌细胞内的研究可以看到核仁内的rDNA在G₀/G₁期瞬时G4形成较少，而在G₁/S期DNA复制依赖的G4结构形成明显增加^[38]。这说明在DNA复制相关的过程中，端粒中的G4结构导致复制阻滞的同时，部分区域G4结构的形成对于复制过程反而是必要的。

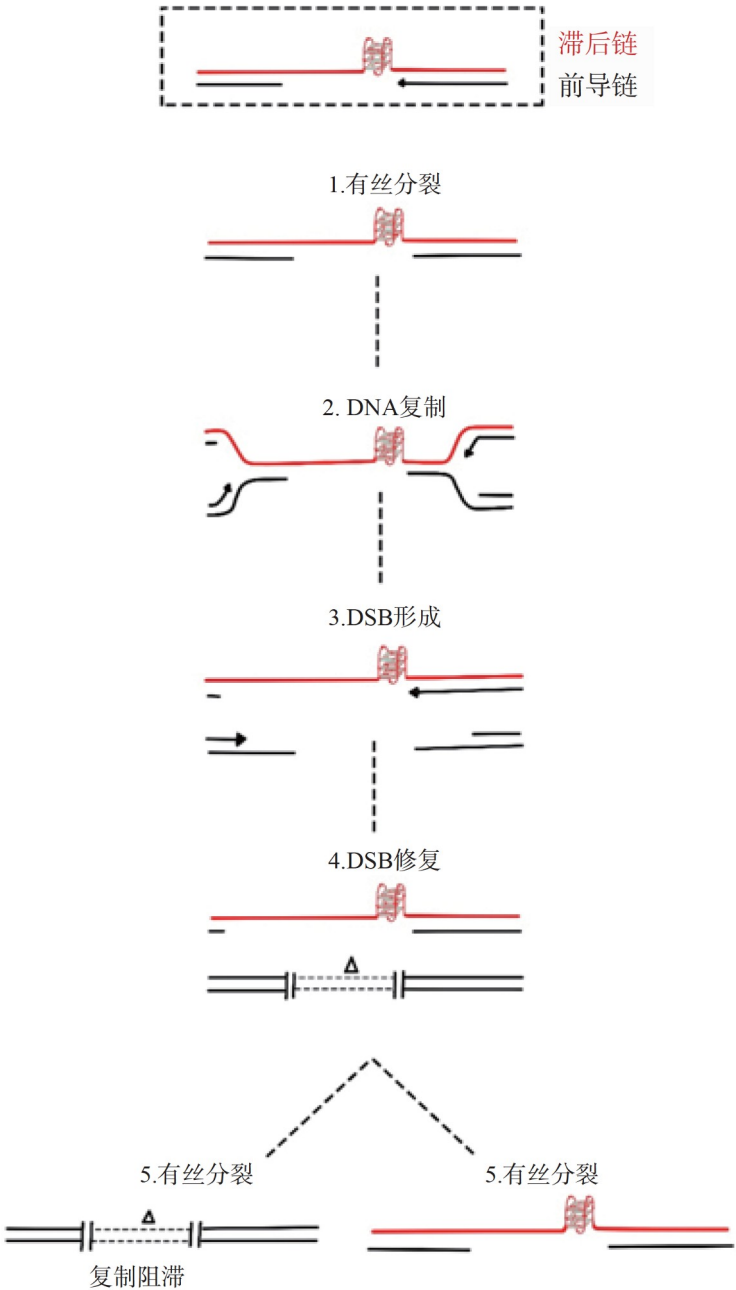
在复制过程中，细胞依赖特定的聚合酶和解旋酶来确保G4结构的精确调控，从而进行有效复制，如FANCI、BLM和Pif1等^[39]。人类细胞中，FANCI解旋酶解开G4结构保证DNA复制。FANCI具有一个G4特异性的识别位点，同时还与REV1和PCNA共同起作用帮助DNA修复和复制^[40]。布鲁姆综合征(Bloom syndrome)是由BLM解旋酶基因突变引发的癌症易患性疾病，其患者的细胞会出现大量的姐妹染色单体交换(sister chromatid exchange events, SCEs)。在没有BLM的情况下，人类和小鼠细胞中的SCEs不会在整个基因组中随机发生，而是在编码区显著富集，特别是在转录基因中的G4位点。所以，BLM通过抑制G4结构位点的重组，特别是基因组转录区域的重组，来保护基因组的稳定性^[41]。在酿酒酵母中，Pif1解旋酶对不同特征的G4结构敏感性存在差异，但是能够处理基因组中存在的大部分G4结构，从而减少了G4结构

对基因组稳定性的影响^[41]。

这些酶功能的缺失或G4配体的加入都会阻碍DNA的复制和修复^[42,43]。DNA G4可以持续通过多次有丝分裂而不改变构象，在有丝分裂增殖期间持续存在，在后续的每个分裂期(S期)进行复制阻断。遇到G4障碍后，DNA复制失败导致产生单链DNA缺口，单链DNA缺口在随后的细胞分裂中导致子代细胞中出现DNA双链断裂(DNA double-strandbreaks, DSBs)，这是由聚合酶 θ 介导的替代末端连接(alternative end joining)处理的(图2)。而另一个子代细胞中DNA的G4复制障碍依然存在，直至被相关解旋酶处理。单个复制阻断的G4结构可以在子代细胞中产生多个独特的缺失，从而在器官发育过程中引起遗传异质性^[44]。所以，在不断复制的癌细胞群体中，G4结构的形成会导致子代细胞间产生遗传改变，而单细胞分辨率上展示G4结构的方法如snG4-CUT&Tag也应运而生^[45]。

细胞复制叉崩溃时由于G4产生的DNA双链断裂是通过同源重组(homologous recombination, HR)和非同源末端连接(nonhomologous end joining, NHEJ)修复的。在缺乏HR的细胞中，富含G4的端粒重复序列的复制效率显著降低^[46]。BRCA1、Rad51和Ku80分别在HR或NHEJ过程中与G4结构相互作用^[37]。G4稳定化合物吡啶他汀(PDS)增加了BRCA2缺陷细胞的端粒脆性，表明G4形成导致端粒不稳定。PDS通过诱导DSBs积累、检查点激活、G₂/M进程失调和复制缺陷来减少HR缺陷细胞的增殖^[46]。

除了以上的修复途径，复制后的修复蛋白如翻译合成(TLS)蛋白Rev1也与G4结构的形成有关。Rev1可以通过置换四分体鸟嘌呤来展开DNA中的G4促进分叉进程，帮助DNA聚合酶延伸。同时，Rev1还可以防止DNA G4结构重新折叠^[47]。Zou1蛋白结合并稳定G4结构，阻断Rev1与DNA结合，并召集核苷酸切除修复(nucleotide excision repair, NER)因子Rad4/23复合体，进而稳定基因组。如果缺乏核酸切除修复功能(Rad4突变)或Zou1突变都会导致细胞跨损伤复制(translesion synthesis)。此外，参与转录调控和核苷酸切除修复(NER)的解旋酶XPD和XPB已被证明在体外和体内调节G4结构，但具体作用方式仍待研究^[48]。G4结构增加了DNA



虚线框显示了一个复制阻断的G-四链体，在包含G-四链体的链上产生一个ssDNA缺口

图2 持久性G-四链体的复制模型^[44]

断裂的风险，会对基因组的稳定性产生影响。与之相反，G4结构诱导的损伤在类开关重组、抗原变异或氧化鸟嘌呤修复过程中也有利于细胞。这些相反的发现表明，G4结构诱导的基因组不稳定性与G4结构促进的修复过程之间一定存在微妙的平衡。

4 总结与展望

G4在生物体的广泛分布预示着其生物功能的

多样化，但是全面阐明G4生物学功能仍需要时间。从DNA代谢层面上，现在可以看出G4结构的调控作用是平衡作用下受到相互作用的蛋白质的精准调控，而不仅是通过结构产生阻滞作用来起到调节功能。G4为探讨二级结构对生物功能的响应提供了一个范例。虽然目前还不可能完全确定G4在体内的精确构象，只能通过控制二级结构的形成使结构和功能之间的相关性显现出来，但随

着今后相关研究的进一步深入, G4相关的二级结构对功能影响的全景必将更为清晰地展现出来。

参考文献

- [1] Gellert M, Lipsett MN, Davies DR. Helix formation by guanylic acid. *Proc Natl Acad Sci USA*, 1962, 48(12): 2013-2018
- [2] Lipps HJ, Gruijssem W, Prescott DM. Higher order DNA structure in macronuclear chromatin of the hypotrichous ciliate *Oxytricha nova*. *Proc Natl Acad Sci USA*, 1982, 79(8): 2495-2499
- [3] Sundquist WI, Klug A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature*, 1989, 342(6251): 825-829
- [4] Varshney D, Spiegel J, Zyner K, et al. The regulation and functions of DNA and RNA G-quadruplexes. *Nat Rev Mol Cell Biol*, 2020, 21(8): 459-474
- [5] Antonio M D, Ponjavic A, Radzevicius A, et al. Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat Chem*, 2020, 12(9): 832-837
- [6] Eladl A, Yamaoki Y, Hoshina S, et al. Investigation of the interaction of human origin recognition complex subunit 1 with G-Quadruplex DNAs of human *c-myc* promoter and telomere regions. *Int J Mol Sci*, 2021, 22(7): 3481
- [7] Hänsel-Hertsch R, Simeone A, Shea A, et al. Landscape of G-quadruplex DNA structural regions in breast cancer. *Nat Genet*, 2020, 52(9): 878-883
- [8] Dahal S, Siddiqua H, Katapadi VK, et al. Characterization of G4 DNA formation in mitochondrial DNA and their potential role in mitochondrial genome instability. *FEBS J*, 2022, 289(1): 163-182
- [9] Beaume N, Pathak R, Yadav VK, et al. Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of *D. radiodurans* involves G4 DNA-mediated regulation. *Nucleic Acids Res*, 2013, 41(1): 76-89
- [10] Sundquist WI, Heaphy S. Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA. *Proc Natl Acad Sci USA*, 1993, 90(8): 3393-3397
- [11] Cui H, Zhang L. G-quadruplexes are present in human coronaviruses including SARS-CoV-2. *Front Microbiol*, 2020, 11: 567317
- [12] 黄保月, 梁晓燕, 杨帆, 等. 新型冠状病毒基因组G-四链体结构的初步研究. *病毒学报*, 2021, 37(1): 9
- [13] Guédin A, De Cian A, Gros J, et al. Sequence effects in single-base loops for quadruplexes. *Biochimie*, 2008, 90(5): 686-696
- [14] Guédin A, Gros J, Alberti P, et al. How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res*, 2010, 38(21): 7858-7868
- [15] Burge S, Parkinson GN, Hazel P, et al. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*, 2006, 34(19): 5402-5415
- [16] Yang Y, Yang Y, Wang S, et al. Detecting the formation of human *c-KIT* oncogene promoter G-quadruplex by taylor dispersion analysis. *Talanta*, 2021, 233: 122533
- [17] Hud NV, Smith FW, Anet FAL, et al. The selectivity for K^+ versus Na^+ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by 1H NMR. *Biochemistry*, 1996, 35(48): 15383-15390
- [18] Chung WJ, Heddi B, Schmitt E, et al. Structure of a left-handed DNA G-quadruplex. *Proc Natl Acad Sci USA*, 2015, 112(9): 2729-2733
- [19] Marchand A, Gabelica V. Native electrospray mass spectrometry of DNA G-quadruplexes in potassium solution. *J Am Soc Mass Spectrom*, 2014, 25(7): 1146-1154
- [20] Robinson J, Raguseo F, Nuccio SP, et al. DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Res*, 2021, 49(15): 8419-8431
- [21] Bedrat A, Lacroix L, Mergny JL. Re-evaluation of G-quadruplex propensity with G4 hunter. *Nucleic Acids Res*, 2016, 44(4): 1746-1759
- [22] Chambers VS, Marsico G, Boutell JM, et al. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol*, 2015, 33(8): 877-881
- [23] Tu J, Duan M, Liu W, et al. Direct genome-wide identification of G-quadruplex structures by whole-genome resequencing. *Nat Commun*, 2021, 12(1): 6014
- [24] Zheng KW, Zhang JY, He Y, et al. Detection of genomic G-quadruplexes in living cells using a small artificial protein. *Nucleic Acids Res*, 2020, 48(20): 11706-11720
- [25] Zhao C, Qin G, Niu J, et al. Targeting RNA G-quadruplex in SARS-CoV-2: a promising therapeutic target for COVID-19? *Angew Chem Int Ed*, 2021, 60(1): 432-438
- [26] Wu F, Niu K, Cui Y, et al. Genome-wide analysis of DNA G-quadruplex motifs across 37 species provides insights into G4 evolution. *Commun Biol*, 2021, 4(1): 98
- [27] Marsico G, Chambers VS, Sahakyan AB, et al. Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res*, 2019, 47(8): 3862-3874
- [28] Kim N. The interplay between G-quadruplex and transcription. *Curr Med Chem*, 2019, 26(16): 2898-2917
- [29] Hänsel-Hertsch R, Beraldi D, Lensing SV, et al. G-quadruplex structures mark human regulatory chromatin.

- Nat Genet*, 2016, 48(10): 1267-1272
- [30] Li L, Williams P, Ren W, et al. YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat Chem Biol*, 2021, 17(2): 161-168
- [31] Zhang X, Spiegel J, Martínez Cuesta S, et al. Chemical profiling of DNA G-quadruplex-interacting proteins in live cells. *Nat Chem*, 2021, 13(7): 626-633
- [32] Lee CY, McNerney C, Ma K, et al. R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nat Commun*, 2020, 11(1): 3392
- [33] Fleming AM, Zhu J, Ding Y, et al. 8-Oxo-7,8-dihydro-guanine in the context of a gene promoter G-quadruplex is an on-off switch for transcription. *ACS Chem Biol*, 2017, 12(9): 2417-2426
- [34] Mao SQ, Ghanbarian AT, Spiegel J, et al. DNA G-quadruplex structures mold the DNA methylome. *Nat Struct Mol Biol*, 2018, 25(10): 951-957
- [35] Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 2009, 462(7271): 315-322
- [36] Lowran K, Campbell L, Popp P, et al. Assembly of a G-quadruplex repair complex by the FANCDJ DNA helicase and the REV1 polymerase. *Genes*, 2020, 11(1): 5
- [37] De Magis A, Götz S, Hajikazemi M, et al. Zuo1 supports G4 structure formation and directs repair toward nucleotide excision repair. *Nat Commun*, 2020, 11(1): 3907
- [38] Lu YJ, Hu DP, Zhang K, et al. New pyridinium-based fluorescent dyes: a comparison of symmetry and side-group effects on G-quadruplex DNA binding selectivity and application in live cell imaging. *Biosens Bioelectron*, 2016, 81: 373-381
- [39] Dahan D, Tsirkas I, Dovrat D, et al. Pif1 is essential for efficient replisome progression through lagging strand G-quadruplex DNA secondary structures. *Nucleic Acids Res*, 2018, 46(22): 11847-11857
- [40] van Wietmarschen N, Merzouk S, Halsema N, et al. BLM helicase suppresses recombination at G-quadruplex motifs in transcribed genes. *Nat Commun*, 2018, 9(1): 271
- [41] Wang L, Wang QM, Wang YR, et al. DNA-unwinding activity of *Saccharomyces cerevisiae* Pif1 is modulated by thermal stability, folding conformation, and loop lengths of G-quadruplex DNA. *J Biol Chem*, 2018, 293(48): 18504-18513
- [42] Lerner LK, Sale JE. Replication of G quadruplex DNA. *Genes*, 2019, 10(2): 95
- [43] Deiana M, Obi I, Andreasson M, et al. A minimalistic coumarin Turn-On probe for selective recognition of parallel G-quadruplex DNA structures. *ACS Chem Biol*, 2021, 16(8): 1365-1376
- [44] Lemmens B, van Schendel R, Tijsterman M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. *Nat Commun*, 2015, 6(1): 8909
- [45] Hui WWI, Simeone A, Zyner KG, et al. Single-cell mapping of DNA G-quadruplex structures in human cancer cells. *Sci Rep*, 2021, 11(1): 23641
- [46] Zimmer J, Tacconi EMC, Folio C, et al. Targeting *BRCA1* and *BRCA2* deficiencies with G-quadruplex-interacting compounds. *Mol Cell*, 2016, 61(3): 449-460
- [47] Eddy S, Ketkar A, Zafar MK, et al. Human Rev1 polymerase disrupts G-quadruplex DNA. *Nucleic Acids Res*, 2014, 42(5): 3272-3285
- [48] Gray LT, Vallur AC, Eddy J, et al. G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nat Chem Biol*, 2014, 10(4): 313-318