

G-四链体与染色质结构相互作用关系的研究进展

杨 婧¹, 侯 越^{1,2}, 舒慧灵¹, 孙 啸^{1*}

(1. 东南大学 生物科学与医学工程学院 生物电子学国家重点实验室, 中国江苏 南京 210096; 2. 西安交通大学 生命科学与技术学院 生物医学信息工程教育部重点实验室, 中国陕西 西安 710049)

摘 要: G-四链体(G-quadruplex, G4)是最早在体外条件下观察到的一类特殊的核酸二级结构, 基因组 DNA 上的 G4 以多种方式调控染色质结构与状态, 其在 DNA 复制、基因转录、染色质构象变化等生物学过程中发挥着重要作用。本文概述了 G-四链体 DNA (G4-DNA) 的基本结构特征和功能, 重点综述了 G4 与染色质结构之间的相互作用关系, 包括 G4 在染色质开放区域的富集现象, G4 与表观修饰之间的相互作用关系, G4 对染色质高级结构的影响。对 G4 与染色质结构相互关系的研究不仅能够帮助我们更深入地探究 G4 的功能, 还能进一步地了解 G4 与疾病的关系, 并以 G4 为靶标开展新药研究。

关键词: G-四链体(G4); 染色质结构状态; 染色质空间结构

中图分类号: Q75

文献标识码: A

文章编号: 1007-7847(2022)02-0175-07

Research Progress in Interaction Between G-quadruplex and Chromatin Structure

YANG Jing¹, HOU Yue^{1,2}, SHU Hui-ling¹, SUN Xiao^{1*}

(1. State Key Laboratory of Bioelectronics, School of Biological Sciences and Medical Engineering, Southeast University, Nanjing 210096, Jiangsu, China; 2. Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China)

Abstract: G-quadruplex (G4) is a special secondary structure of nucleic acid first observed *in vitro*. G4 of genomic DNA regulates chromatin structure and state in a variety of ways. It plays an important role in DNA replication, gene transcription, chromatin conformation change and other biological processes. Herein, the basic structural characteristics and functions of G4 DNA (G4-DNA) were summarized. The interaction between G4 and chromatin structure was mainly reviewed, including the enrichment of G4 in the open region of chromatin, the interaction between G4 and epigenetic modification, and the effect of G4 on the higher structure of chromatin. The research in the relationship between G4 and chromatin structure can help further explore the functions of G4, understand the relationship between G4 and diseases, and carry out new drug research with G4 as a target.

Key words: G-quadruplex (G4); chromatin structural state; chromatin spatial structure

(*Life Science Research*, 2022, 26(2): 175~181)

G-四链体(G-quadruplex, G4)是一种特殊的核酸二级结构, 多形成于富含鸟嘌呤的 DNA 或 RNA 序列中。G4 结构首次被发现于体外环境, 广泛存在于基因组, 通常标志着具有调控功能的染

色质区域, 在生物体内具有重要的调控功能, 与染色质结构、表观调控、基因组稳定性、人类疾病等密切相关。

1910 年, Bang^[1]首次发现鸟嘌呤核苷酸在高

收稿日期: 2021-04-10; 修回日期: 2021-07-22; 网络首发日期: 2021-08-12

基金项目: 国家自然科学基金资助项目(61972084)

作者简介: 杨婧(1996—), 女, 山西阳泉人, 硕士研究生; * 通信作者: 孙啸(1962—), 男, 江苏南京人, 博士, 东南大学教授, 主要从事生物信息学研究, Tel: 025-83795174, E-mail: xsun@seu.edu.cn。

浓度下能够形成凝胶, 这表明 DNA 中富含鸟嘌呤的序列可能形成更高阶的结构。1962 年, Ge-llert 等^[2]使用 X 射线衍射实验证明, 鸟嘌呤可以组装成四聚体结构。随着高通量测序技术的发展, G4 检测技术经历了从体外到体内检测的发展过程。2015 年, 剑桥大学 Balasubramanian 团队提出 G4-seq^[3]方法, 并利用该方法在人类基因组水平上得到了超过 70 万个能够在体外形成 G4 的序列。为探索 G4 结构在内源性染色质环境中是否真实形成, 该团队在 2018 年提出了染色质免疫沉淀测序方法(G4 ChIP-seq)^[4], 利用 G4 特异性抗体 BG4, 在 K562 细胞系中鉴定出了约 8 000 个能够在体内形成 G4 的序列。最近, 中国科学院的研究团队基于 G4 结合蛋白 DHX36 的结合结构域, 设计出一种人工 G4 探针蛋白^[5], 结合 ChIP-seq 技术, 在 A549 细胞系中测得了超过 10 万个 G4 结构。基因组中众多 G4 预示着 G4 可能具有复杂的生物学功能, 如表观调控、与染色质的相互作用等。

G4 在多个层次与染色质结构存在相互作用, 包括染色质可及性、染色质局部修饰状态和染色质全局空间构象等。染色质是由核小体压缩折叠形成的高级结构, 当 DNA 复制、转录等生物学过程发生时, 致密的核小体结构将会被破坏, 染色质由紧密变为开放的状态, 这段区域即被称为开放染色质(open chromatin), 这种允许调控蛋白与之结合的特性称为染色质的可及性(chromatin accessibility)。G4 与核小体缺失及染色质开放结构存在着密切的联系^[6], 同时, G4 也可能招募蛋白质因子进行染色质修饰, 包括 DNA 甲基化和组蛋白修饰^[7]。我们前期的研究结果也表明, G4 对染色质空间构象有影响^[8], 染色质空间构象包括染色质环结构、染色质拓扑关联域(topologically associating domain, TAD)及染色质区室结构。

本文将从 3 个方面详细讨论 G4 与染色质结构的相互作用关系, 包括: G4 在开放染色质区域的富集现象, G4 通过招募功能蛋白质调控染色质局部区域的表观修饰, 以及 G4 对染色质空间结构的影响。

1 G4 结构及其生物学功能

G4 结构由多个 G-四分体(G-tetrad)平面构成, G-四分体是 G4 序列中的鸟嘌呤(G)互相作用形成的稳定的平面正方形阵列, 通过 Hoogsteen 氢键碱基相互配对^[9], 三层及三层以上的 G-四分体堆叠即形成稳定的 G4 结构(图 1)。每个鸟嘌呤的电负性羧基指向平面中心, 可以容纳一价金属阳离子(M⁺)以稳定 G4 结构^[9](图 1 左)。其中, K⁺的作用最为明显, 人体富含 K⁺、Na⁺离子的生理环境有利于 G4 结构形成^[10]。

G4 结构具有多种拓扑结构^[11](图 2), 根据 DNA 链的方向可分为平行结构、反向平行结构和杂合结构。此外, G4 结构可以形成于单链 DNA 内部(intramolecular)或几条链之间(intermolecular), 同时也能够形成各种环状结构(loop structures)。结构参数和外在因素如化学修饰^[12]、分子聚集^[13]等, 会直接影响 G4 分子结构的几何形态和构象, 使其表现出广泛性和多样性。

$G_xN_{1-7}G_xN_{1-7}G_xN_{1-7}G_x$ 是可能形成 G4 结构的模式序列(G4 motif)。“N”表示碱基“ATCGU”中的一个, 长度范围为 1 至 7, 称为环(loop); 序列中 G 的数目“x”大于等于 3, 称为茎(G-tract)。但随着研究的深入, 研究人员发现 G4 序列存在很多特例, 结构也可以更加灵活, 如在茎的长度大于 3 的体外环境下, 环长可以达到 30 nt^[14]; G 茎上的碱基可以发生凸起或错配^[15], 从而增加 G4 序列和结构的复杂性。序列分析结果表明, 人类基因组包含的 G4

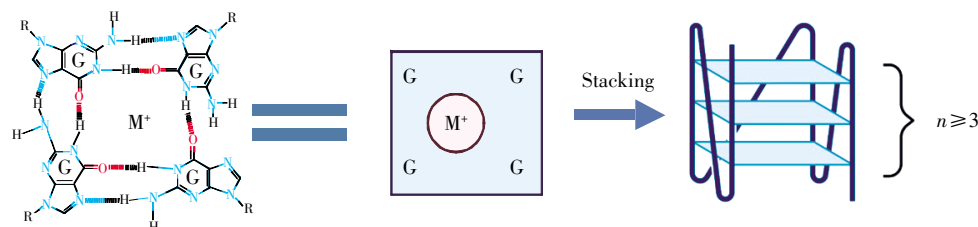


图 1 G-四链体结构示意图(使用 BioRender 绘图)

从左至右分别为 G-四分体的化学结构式示意图、G-四分体的简化示意图和 G4 结构示意图。“n”为 G-四分体的个数。

Fig.1 G4 structure diagram (created with BioRender)

From left to right are diagram of the chemical structure of G-tetrads, simplified diagram of G-tetrads, and diagram of G4 structure. “n” is the number of G-tetrads.

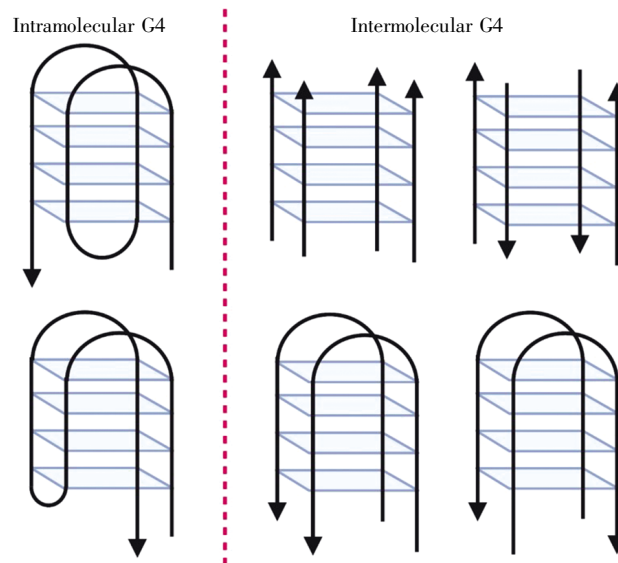


图 2 分子内 G-四链体(左)和分子间 G-四链体(右)的结构示意图(参照文献[11], 使用 BioRender 绘图)

Fig.2 Intramolecular G4 (left) and intermolecular G4 (right) structure diagrams (created with BioRender, according to Reference [11])

序列超过 30 万个^[6]。G4 序列具有形成 G4 的潜力, 因此也将其称为 PG4 (potential G-quadruplex)^[11]。

G4 在端粒维持、DNA 复制和转录调控等生物学过程中发挥着重要作用。早期针对 G4 功能的研究主要集中在端粒区域, 端粒由 3' 端尾部富含鸟嘌呤的单链 DNA 区和双链区组成, 这两个区域都极易形成 G4 结构。G4 会影响端粒酶活性, 端粒区域形成的反向平行 G4 可以阻碍端粒延伸^[17], 进而阻止癌细胞持续增殖分裂, 因此 G4 可作为靶点进行癌症治疗, 并由此引申出了 G4 的小分子配体的研究, 用以结合和稳定体内 G4^[18]。G4 在 DNA 复制中有双重作用: 一方面是可以作为复制的障碍导致突变基因组的不稳定性增加, 另一方面是后动物复制起源的组成部分^[10]。G4 结构在启动子、增强子等调控区域的富集, 可能对基因表达产生正向或负向的影响, 从而引发转录组的改变。例如: Siddiqui-Jain 等^[19]最早对致癌基因 *c-MYC* 的研究表明, G4 序列的突变或 G4 稳定配体的添加会影响体内转录事件; 华盛顿大学的团队针对人类细胞的基因组分析表明, 转录解旋酶 XPB 和 XPD (xeroderma pigmentosum group B and D genes)与 G4 序列明显重叠, XPB 和 XPD 能够结合 G4, 且 XPD 还能够解旋 G4, 因此它们可能被招募到 G4 结构上帮助其完成转录^[20]。此外, G4 结合蛋白也会影响转录, 如核仁蛋白与 *c-MYC* NHE III1 元件的结合能够诱导 G4 折叠, 并降低癌基因 *c-MYC* 的转录水平^[21]。

深入了解 G4 与染色质结构信息之间的关联, 认识 G4 与染色质结构的相互作用, 有助于理解 G4 的生物学功能。图 3 概括了 G4 与染色质结构的相互作用关系, 从染色质的线性结构、表观调控和空间结构方面展示了 G4 结构对染色质结构的影响。在染色质线性结构方面, G4 主要表现为在染色质开放区域的富集(图 3A); 在表观调控方面, G4 通常通过招募蛋白质因子调控或影响表观调控(图 3B~C); 在染色质空间结构方面, G4 能够影响染色质空间分割能力和远程相互作用(图 3D)。

2 G4-DNA 与染色质开放结构之间的关系

开放染色质区域具有转录活性, 与基因表达调控过程密切相关。已有研究证明, G4 能够影响基因转录调控功能^[16], 因此 G4 必定与染色质开放结构存在关联。

由于 G4-DNA 检测技术复杂, 成本高, 早期 G4-DNA 相关研究多采用 PG4 作为研究对象。2006 年, Segal 等^[22]观察了核小体占位率(nucleosome occupancy probability, NOP)与 PG4 的分布, 发现与低 NOP 区域相比, 高 NOP 区域的 PG4 基序密度较低。2009 年, Halder 等^[9]绘制了全基因组尺度下 PG4 以及核小体在 1 号染色体中的分布, 结果显示 PG4 密度高的区域常表现出对核小体的排斥, 且核小体和 PG4 很少有重叠, 其他染色体中也存在相似的情况。由此可知, PG4 基序与核小体存在竞争关系。染色质的调控区域以核小体

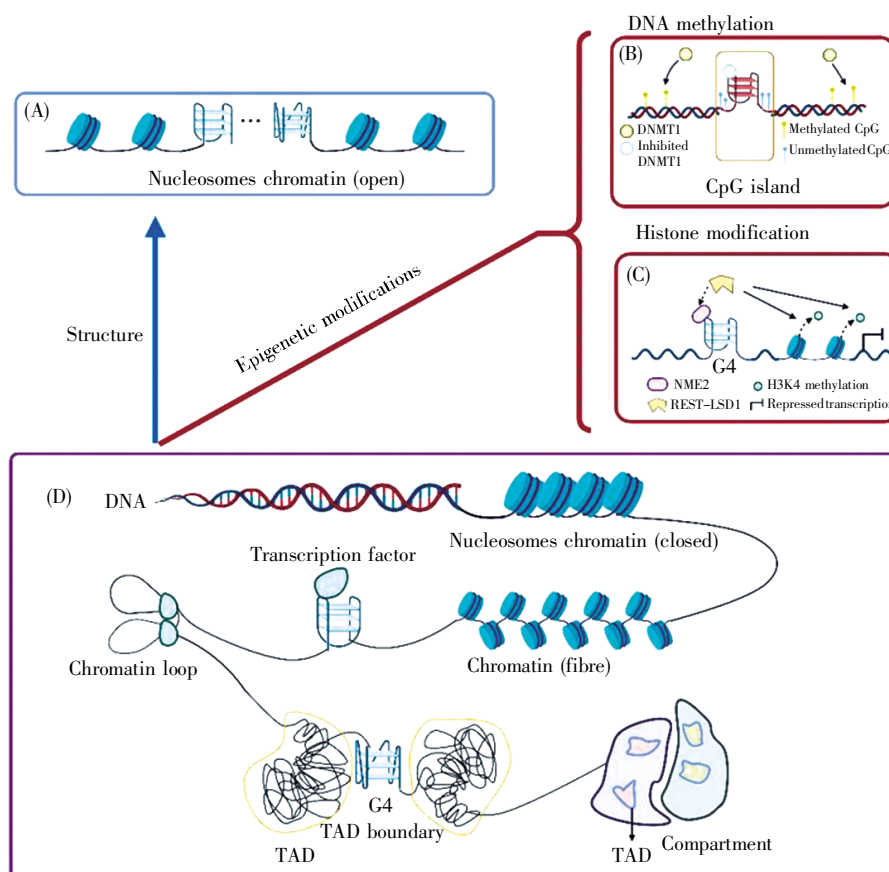


图3 G-四链体与染色质结构的相互作用关系(使用 BioRender 绘图)

(A) G4 在染色质开放区域内有富集现象; (B) CpG 岛上的 G4 能够招募 DNMT1 并抑制其活性, 调控 DNA 甲基化水平; (C) *hTERT* 启动子上的 NME2 依赖 G4 结构招募阻遏复合物 REST-LSD, 去除 H3K4 甲基化, 抑制基因表达; (D) 染色质空间结构示意图。G4 与转录因子相互作用有助于染色质环的形成; G4 在 TAD 边界显著富集, 影响 TAD 边界的分割能力。

Fig.3 Interaction between G4 and chromatin structure (created with BioRender)

(A) G4 is enriched on the open region of chromatin; (B) G4 on CpG islands can recruit DNMT1, inhibit its activity and regulate the level of DNA methylation; (C) NME2 on the *hTERT* promoter relies on the G4 structure to recruit the repressor complex REST-LSD, remove H3K4 methylation and inhibit gene expression; (D) Schematic diagram of the spatial structure of chromatin. G4 interacts with transcription factors to contribute to the formation of chromatin loops. The significant enrichment of G4 at the TAD boundary indicates that G4 can affect the segmentation ability of the TAD boundary.

缺失为特征, 这使得调控基因转录、复制和表观遗传的蛋白质得以进入开放染色质区域。根据 PG4 基序与核小体存在竞争关系的结论, 以及 PG4 具有形成 G4 的可能性, 可以推测出 G4 与染色质开放区域必然存在某种特定的关联。

2016年, Balasubramanian 团队着手阐明人类细胞中 G4 结构和染色质开放性之间的关系^[23]。研究人员通过 G4 ChIP-seq 技术绘制了 HaCaT 细胞染色质中 G4 结构的全基因组位置图, 并利用测序技术 ATAC-seq/FAIRE-seq 绘制了染色质开放区域。结果显示, G4 ChIP-seq 峰的 98% 与 ATAC-seq 和 FAIRE-seq 确定的区域重叠, 这说明在染色质开放区域, 即核小体缺失区域, 存在 G4 结构的富集现象(图 3A)。这预示着, G4 结构的

稳定形成能够阻碍核小体组装, 因此有利于染色质局部排斥核小体以保持开放构象, 并增强转录起始和转录速率。

2020年, 中国科学院的研究团队使用 G4 结合蛋白探针及 G4-ChIP 测序技术, 探究了转录起始位点(transcription start site, TSS)功能区域的 G4 形成, 发现 G4 集中在 TSS 的两侧, 与其他功能区域相比, TSS 功能区域能够形成更多稳定的 G4^[5], 这意味着在转录驱动的启动子中, G4 的形成更为活跃和稳定。上述结果提示, 转录过程与 G4 的形成可能是相互影响的。

G4 在染色质开放区域的富集阻碍了核小体的组装形成, 使局部染色质结构能够保持开放状态, 从而使与基因表达调控相关的蛋白质因子得

以进入染色质开放区域发挥作用,增强转录起始和转录速率,G4在TSS功能区域两侧的稳定形成也证实了G4与转录过程的关联关系。

3 G4-DNA与染色质状态之间的关系

G4能够作用于DNA甲基化及组蛋白修饰,引起染色质结构、DNA构象、DNA稳定性及DNA与蛋白质相互作用方式的改变,从而调控表观遗传修饰。

2016年,奥塔哥大学的Cree等^[24]首次在体外环境下测量了G4对DNA甲基转移酶1(DNA methyltransferase 1, DNMT1)、DNMT3A、DNMT3B的亲合力。表面等离子共振分析结果表明,G4与DNMT3A和DNMT3B具有较高的亲合力,其亲合力与其他已知的G4结合蛋白相当,表明G4在DNA甲基化过程中可以通过调控甲基转移酶来发挥作用。DNMT3A和DNMT3B的N末端调节区均包含1个ADD域和1个PWWP域。已有研究证明,含有ADD域的ATRX蛋白能够与G4相互作用^[25]。上述分析提示, DNMT3A/B可能通过ADD域与G4结构结合从而发挥作用,但DNMT是通过何种方式调控DNA甲基化的还需要进一步研究。

2018年, Balasubramanian团队提出了G4对DNA甲基化的重塑作用^[7]。实验通过全基因组重亚硫酸盐测序(whole genome bisulfite sequencing, WGBS)方法对K562细胞系的甲基化水平进行了全面表征,并利用G4 ChIP-seq方法生成G4的全基因组数据库,比较了BG4峰(即G4位点)与CpG岛的重叠部分,发现大部分BG4峰与CpG岛重叠,即大部分G4形成在低甲基化水平的CpG岛区域,同时实验人员注意到BG4峰的区域表现出显著的甲基化缺失,这有力地证明了CpG岛的低甲基化现象与G4结构密切相关。为了探究DNMT与G4之间是否有某种形式的相互作用, Balasubramanian团队以DNMT1为研究对象,探究DNMT1与CpG岛的低甲基化是否存在关联,发现DNMT1在低甲基化水平的CpG岛上显著富集,形成双峰的分布模式,随后研究人员使用酶联免疫吸附分析实验(ELISA)证实,G4具有结合DNMT1的能力。因此,G4可能通过识别并结合DNMT1,在CpG岛局部隔绝DNMT1,导致特定基因组位点的甲基化缺失(图3B)。

2020年, Jara-Espejo等^[26]分析了PG4与CpG岛低甲基化的关联,并确定了可能调节这种现象

的PG4的内在和外特性。实验人员使用全基因组甲基化数据探索了PG4在CpG岛上的作用,发现PG4对甲基化的作用是由稳定性、位置和染色质可及性驱动的:开放染色质中的高稳定性PG4具有高转录活性和形成潜能,可更有效地促进低甲基化。

G4还会影响组蛋白修饰。人类端粒结合蛋白能够结合体内端粒重复RNA上的G4和体外DNA上的G4,从而招募对端粒和亚端粒异染色质维持起重要作用的组蛋白甲基转移酶,促进端粒的组蛋白三甲基化^[27]。端粒的表观遗传修饰与端粒长度调节和端粒相关疾病的发生发展过程及衰老密切相关,对于基因组的完整性至关重要,因此,G4调控端粒表观遗传修饰的研究有助于端粒生物功能及作用机制的研究。

相关研究报告,人类端粒酶逆转录酶(human telomerase reverse transcriptase, hTERT)依赖于G4结构的启动子的转录抑制,该途径通过肿瘤转移抑制因子NME2 (metastasis suppressor non-metastatic 2)招募RE1沉默转录因子-赖氨酸特异性组蛋白脱甲基酶1 (RE1-silencing transcription factor-lysine-specific histone demethylase 1, REST-LSD1)阻遏复合物,去除组蛋白H3K4激活基因的单甲基化和二甲基化(即H3K4me1和H3K4me2),并抑制hTERT基因表达^[28](图3C)。hTERT在90%的肿瘤细胞中表达,但在绝大多数正常细胞中不表达^[29],因此,依赖于G4与NME2的抑制hTERT基因表达的机制,能够为癌症提供新的治疗思路,有助于开辟控制癌细胞中hTERT激活的潜在方法。

G4不仅能够直接影响表观遗传修饰,还能够间接影响组蛋白修饰的遗传。G4会阻遏复制发生,因此在复制时需要解旋酶解开G4结构,然而当G4结构的解旋暂时延迟时,响应的基因组区域被绕过,导致新合成的组蛋白无法遗传亲本修饰^[30],从而破坏表观修饰的局部遗传。

由此可见,G4结构对染色质状态的影响是多层次、多方面的,二者相互关联、相互影响。因此,G4对生物学过程和表观遗传修饰的影响是复杂多样的。

4 G4-DNA与染色质空间结构的相互作用

大量证据表明,G4的形成会排斥核小体的定位,G4与染色质的基本结构和状态都有关联,但G4与高阶基因组组织之间的关系仍然未知。有研

究报道, G4 可能参与远距离交互, 如在裂殖酵母中, Rif1 蛋白能够结合 G4, 使染色质在细胞核纤层形成环及局部染色体室, 并抑制长距离复制^[31]; 位于乳腺和宫颈恶性肿瘤中的长序列 G4 (long G4, LG4)断裂会导致基因组融合^[32]。

本课题组近期针对 K562 细胞系的 Hi-C^[33]及 G4 ChIP-seq 数据集进行了计算分析, 结果表明, G4 有助于定义被称为拓扑结构域(TAD)的更高层次的染色质域^[8]。我们观察到 G4 在 TAD 边界高度富集的现象(图 3D), 并发现染色质结构蛋白(该蛋白质对 TAD 形成起重要作用)在 TAD 边界周围高度富集。TAD 边界富集大量的高表达基因, 频繁的转录事件能够形成大量的单链 DNA, 有利于 G4 的形成, 这可能是 G4 在 TAD 边界富集的原因。同时, TAD 边界上染色质结构蛋白的含量与 G4 的含量呈正相关。因此, G4 很可能通过影响 TAD 边界区域染色质结构蛋白的结合, 影响边界-边界作用和染色质结构。TAD 边界最重要的作用之一就是阻断两个 TAD 之间的相互作用, 这一功能的实现主要取决于 TAD 边界上大量的绝缘子 CTCF。通过计算我们发现, 存在 G4 的区域的 CTCF 结合位点有更强的染色质分割能力, 即 G4 会影响 TAD 边界的分割能力。此外, G4 还能够参与染色质环的形成, 正链上的 G4 序列显著富集在正向 CTCF 序列上, 负链上的 G4 序列则显著富集在负向 CTCF 序列上。染色质环的形成主要依赖于黏连蛋白识别 CTCF 序列方向性, 但其机制仍不清楚, 因此我们提出模型: 正、负链上的 G4 可以分别富集在正、负向的 CTCF 序列周围, 它们可以阻碍黏连蛋白活动, 从而参与染色质成环。

我们的研究还选用实验数据中的增强子和启动子, 分析了 G4 对染色质的远距离作用是否存在影响, 通过计算我们发现, 包含 G4 的增强子-启动子之间的相互作用频率明显高于不含 G4 的增强子-启动子相互作用对, 因此我们认为, G4 可以通过特殊的结构招募转录因子到增强子或启动子区, 调节长程相互作用。

转录因子 YY1 (Yin Yang 1)是调控染色质远程相互作用的重要蛋白质^[34]。最近, 美国加州大学通过实验证实, G4 能够通过 YY1 进行相互作用调节染色质远距离互作^[35]。实验人员通过无偏定量蛋白质组学方法确定, 转录因子 YY1 能够在体内与 G4 相互作用; 通过对照实验和 Hi-C 方法

证实, YY1-G4 相互作用有助于 DNA 环化以及远程位点相互作用。

越来越多的证据表明, 染色质之间的相互作用能够调节基因表达^[36-37], 即基因的表达调控存在高阶结构特性, 基因表达可能被远程调控。G4 不仅与染色质局部结构有关, 同时也与更高维度的染色质结构相关, 参与调节远程位点相互作用。

5 总结与展望

G4 结构广泛分布于各种生物的基因组和转录组, 它们总体上影响 DNA 二级结构、核小体定位、组蛋白修饰以及染色质的空间结构等, 继而影响 DNA 复制、基因转录、染色质构象变化等生物学过程。

G4 具有调节中心法则的每个步骤的潜力, 其与染色质结构和状态的关联研究能够揭示 G4 结构与染色质相互作用的原理, 为我们理解 G4 在生物学过程中的功能提供帮助。G4 在染色质开放区域和核小体缺失区域呈现富集状态, 当它们与基因转录调控区域重叠时, 就会影响基因的转录调控。G4 不仅能够协同转录因子识别并隔绝 DNMT1, 导致关键基因组位点的甲基化缺失, 从而影响转录过程, 还能够通过形成复合物并招募相关酶, 影响组蛋白修饰, 从而调控染色质的表观修饰。但染色质并非简单的线性模型, 高度折叠和压缩的空间结构势必对染色质结构和基因表达产生影响。从空间结构角度出发, G4 影响 TAD 边界的分割能力, 参与染色质环的形成, 除此之外, G4 还能够通过招募蛋白质影响长程相互作用。

G4 既然能够影响基因调控, 自然也与人类疾病相关, 其今后有望成为新药开发的靶标。端粒 G4 结构影响端粒酶表达, 可能会导致癌症发生发展; G4 影响基因调控过程, 包括转录调控、转录后及翻译调控, 从而影响基因表达的结果, 导致疾病发生; 表观遗传修饰也受到 G4 的调控, 包括 DNA 甲基化、组蛋白修饰等, G4 影响表观遗传修饰可能会引起表观遗传修饰异常疾病; G4 过度稳定和积累也会导致疾病。总之, 深入研究 G4 生物学功能, 有助于更好地了解疾病的发生机制, 制定相应的治疗方案。而且, 通过了解 G4 的结构和功能, 可以利用其生化特性开发新的 G4 配体, 控制 G4 的形成及其生物学作用, 从而针对疾病相关的 G4 开展特异性的新药研发, 推进 G4 的靶向治疗, 促进药物治疗新方法和新技术的发展。

参考文献(References):

- [1] BANG I. Untersuchungen über die guanylsäure[J]. *Biochemische Zeitschrift*, 1910, 26: 293–311.
- [2] GELLERT M, LIPSETT M N, DAVIES D R. Helix formation by guanylic acid[J]. *Proceedings of the National Academy of Sciences USA*, 1962, 48(12): 2013–2018.
- [3] CHAMBERS V S, MARSICO G, BOUTELL J M, *et al.* High-throughput sequencing of DNA G-quadruplex structures in the human genome[J]. *Nature Biotechnology*, 2015, 33(8): 877–881.
- [4] HÄNSEL-HERTSCH R, SPIEGEL J, MARSICO G, *et al.* Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing[J]. *Nature Protocols*, 2018, 13(3): 551–564.
- [5] ZHENG K W, ZHANG J Y, HE Y D, *et al.* Detection of genomic G-quadruplexes in living cells using a small artificial protein[J]. *Nucleic Acids Research*, 2020, 48(20): 11706–11720.
- [6] HALDER K, HALDER R, CHOWDHURY S. Genome-wide analysis predicts DNA structural motifs as nucleosome exclusion signals[J]. *Molecular BioSystems*, 2009, 5(12): 1703–1712.
- [7] MAO S Q, GHANBARIAN A T, SPIEGEL J, *et al.* DNA G-quadruplex structures mold the DNA methylome[J]. *Nature Structural & Molecular Biology*, 2018, 25(10): 951–957.
- [8] HOU Y, LI F Y, ZHANG R X, *et al.* Integrative characterization of G-quadruplexes in the three-dimensional chromatin structure[J]. *Epigenetics*, 2019, 14(9): 894–911.
- [9] FAY M M, LYONS S M, IVANOV P. RNA G-quadruplexes in biology: principles and molecular mechanisms[J]. *Journal of Molecular Biology*, 2017, 429(14): 2127–2147.
- [10] RHODES D, LIPPS H J. G-quadruplexes and their regulatory roles in biology[J]. *Nucleic Acids Research*, 2015, 43(18): 8627–8637.
- [11] BOCHMAN M L, PAESCHKE K, ZAKIAN V A. DNA secondary structures: stability and function of G-quadruplex structures[J]. *Nature Reviews Genetics*, 2012, 13(11): 770–780.
- [12] SACCÀ B, LACROIX L, MERGNY J L. The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides[J]. *Nucleic Acids Research*, 2005, 33(4): 1182–1192.
- [13] MIYOSHI D, NAKAO A, SUGIMOTO N. Molecular crowding regulates the structural switch of the DNA G-quadruplex[J]. *Biochemistry*, 2002, 41(50): 15017–15024.
- [14] GUÉDIN A, GROS J, ALBERTI P, *et al.* How long is too long? Effects of loop size on G-quadruplex stability[J]. *Nucleic Acids Research*, 2010, 38(21): 7858–7868.
- [15] MUKUNDAN V T, PHAN A T. Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences[J]. *Journal of the American Chemical Society*, 2013, 135(13): 5017–5028.
- [16] HÄNSEL-HERTSCH R, DI ANTONIO M, BALASUBRAMANIAN S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential[J]. *Nature Reviews. Molecular Cell Biology*, 2017, 18(5): 279–284.
- [17] ZÄHLER A M, WILLIAMSON J R, CECH T R, *et al.* Inhibition of telomerase by G-quartet DNA structures[J]. *Nature*, 1991, 350(6320): 718–720.
- [18] NEIDLE S. Human telomeric G-quadruplex: the current status of telomeric G-quadruplexes as therapeutic targets in human cancer[J]. *The FEBS Journal*, 2010, 277(5): 1118–1125.
- [19] SIDDIQUI-JAIN A, GRAND C L, BEARSS D J, *et al.* Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription[J]. *Proceedings of the National Academy of Sciences USA*, 2002, 99(18): 11593–11598.
- [20] GARY L T, VALLUR A C, EDDY J, *et al.* G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD[J]. *Nature Chemical Biology*, 2014, 10(4): 313–318.
- [21] SUTHERLAND C, CUI Y X, MAO H B, *et al.* A mechanosensor mechanism controls the G-quadruplex/i-motif molecular switch in the MYC promoter NHE III [J]. *Journal of the American Chemical Society*, 2016, 138(42): 14138–14151.
- [22] SEGAL E, FONDUFE-MITTENDORF Y, CHEN L Y, *et al.* A genomic code for nucleosome positioning[J]. *Nature*, 2006, 442(7104): 772–778.
- [23] HÄNSEL-HERTSCH R, BERARDI D, LENSING S V, *et al.* G-quadruplex structures mark human regulatory chromatin[J]. *Nature Genetics*, 2016, 48(10): 1267–1272.
- [24] CREE S L, FREDERICKS R, MILLER A, *et al.* DNA G-quadruplexes show strong interaction with DNA methyltransferases *in vitro*[J]. *FEBS Letters*, 2016, 590(17): 2870–2883.
- [25] LAW M J, LOWER K M, VOON H P J, *et al.* ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner[J]. *Cell*, 2010, 143(3): 367–378.
- [26] JARA-ESPEJO M, LINE S R. DNA G-quadruplex stability, position and chromatin accessibility are associated with CpG island methylation[J]. *The FEBS Journal*, 2020, 287(3): 483–495.
- [27] TAKAHAMA K, TAKADA A, TADA S, *et al.* Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS[J]. *Chemistry & Biology*, 2013, 20(3): 341–350.
- [28] SAHA D, SINGH A, HUSSAIN T, *et al.* Epigenetic suppression of human telomerase (*hTERT*) is mediated by the metastasis suppressor NME2 in a G-quadruplex-dependent fashion[J]. *The Journal of Biological Chemistry*, 2017, 292(37): 15205–15215.
- [29] KANAYA T, KYO S, TAKAKURA M, *et al.* hTERT is a critical determinant of telomerase activity in renal-cell carcinoma[J]. *International Journal of Cancer*, 1998, 78(5): 539–543.
- [30] JASENCAKOVA Z, SCHARF A N, ASK K, *et al.* Replication stress interferes with histone recycling and predeposition marking of new histones[J]. *Molecular Cell*, 2010, 37(5): 736–743.
- [31] KANOY Y, MATSUMOTO S, FUKATSU R, *et al.* Rif1 binds to G quadruplexes and suppresses replication over long distances[J]. *Nature Structural & Molecular Biology*, 2015, 22(11): 889–897.
- [32] WILLIAMS J D, HOUSEROVA D, JOHNSON B R, *et al.* Characterization of long G4-rich enhancer-associated genomic regions engaging in a novel loop: loop ‘G4 kissing’ interaction[J]. *Nucleic Acids Research*, 2020, 48(11): 5907–5925.
- [33] RAO S S P, HUNTLEY M H, DURAND N C, *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping[J]. *Cell*, 2014, 159(7): 1665–1680.
- [34] WEINTRAUB A S, LI C H, ZAMUDIO A V, *et al.* YY1 is a structural regulator of enhancer-promoter loops[J]. *Cell*, 2017, 171(7): 1573–1588.e28.
- [35] LI L, WILLIAMS P, REN W D, *et al.* YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression[J]. *Nature Chemical Biology*, 2021, 17(2): 161–168.
- [36] KONG N, JUNG I. Long-range chromatin interactions in pathogenic gene expression control[J]. *Transcription*, 2020, 11(5): 211–216.
- [37] MORRISON A J. Chromatin-remodeling links metabolic signaling to gene expression[J]. *Molecular Metabolism*, 2020, 38: 100973.