



R-loop功能及检测技术研究进展

何倩楠^{1,2}, 姚红杰^{1,3*}

1. 广州国家实验室基础研究部, 广州 510005

2. 中国科学院广州生物医药与健康研究院, 广州 510530

3. 广州医科大学附属第一医院, 呼吸疾病全国重点实验室, 广州 510120

* 联系人, E-mail: yao_hongjie@gzlab.ac.cn

收稿日期: 2025-05-08; 接受日期: 2025-05-20; 网络版发表日期: 2025-06-12

广州国家实验室重点专项(批准号: GZNL2023A02010)、国家重点研发计划(批准号: 2021YFA1100300)和国家自然科学基金(批准号: 32430016)资助

摘要 R-loop是一种由DNA:RNA杂合链与互补单链DNA共同形成的三链核酸结构. 作为基因组中普遍存在的调控元件, R-loop在原核和真核生物中参与调控DNA损伤修复及转录调控等关键生物学过程, 同时对组蛋白修饰及染色质高级结构等表观遗传景观具有调控作用. 然而, 当R-loop动态平衡被打破导致其异常积累时, 将造成DNA损伤、基因组不稳定及染色质结构异常, 进而与多种肿瘤性疾病和遗传性疾病的发病机制密切相关. 目前, 研究人员已开发出多种R-loop检测技术, 主要包括基于S9.6抗体或RNase H1杂合链结合结构域的亲和富集技术, 以及基于单分子标记的spKAS-seq或核酸酶切割的RIAN-seq等非亲和富集方法. 本文系统介绍了R-loop的形成分子机制、生理病理功能及其检测技术的研究进展, 旨在为深入理解R-loop的生物学意义及其在疾病中的作用提供理论参考.

关键词 R-loop, DNA:RNA杂合链, 表观遗传, R-loop检测技术

R-loop是由一条DNA:RNA杂合链以及一条被置换的单链DNA共同组成的三链核酸结构(图1). 这种结构最早于1976年由Thomas等人^[1]通过电子显微镜在体外实验中观察到. R-loop的形成受多种因素调节: (i) 负超螺旋结构, 转录过程产生的DNA负超螺旋通过增加新生RNA和模板链结合的概率, 从而促进R-loop的形成与稳定^[2,3]; (ii) GC含量, 高GC含量区域由于能形成更稳定的RNA-DNA杂合链(GC碱基对含3个氢键, 而AT仅2个氢键)更易形成R-loop^[4]; (iii) 转录水平, 转录水平的升高会显著增加RNA-DNA杂交的概率^[5,6]等.

现有研究表明, R-loop在多种生物学过程中发挥着关键的调控作用, 包括基因转录^[7-10]、DNA损伤修复^[11-13]、端粒维持^[14]、抗体类别转换重组^[15]以及造血干细胞分化^[16]等. 在表观遗传调控层面, R-loop影响DNA甲基化模式^[17]、组蛋白修饰^[18,19]以及染色质高级构象^[20-22]等多种表观遗传信息, 具有关键的表观调控作用. 此外, R-loop生物学功能的揭示离不开R-loop检测技术的开发, 这些技术为精准解析R-loop的动态分布及其生物学功能提供了重要工具. 本综述将阐述

引用格式: 何倩楠, 姚红杰. R-loop功能及检测技术研究进展. 中国科学: 生命科学, 2025, 55: 1121–1133

He Q N, Yao H J. Research advances in R-loop functions and detection technologies (in Chinese). Sci Sin Vitae, 2025, 55: 1121–1133, doi: 10.1360/SSV-2025-0151

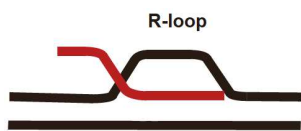


图1 R-loop的结构

Figure 1 The architecture of R-loop

R-loop在生理和病理状态下的调控功能, 深入剖析R-loop对多层次表观遗传信息的调控作用, 并系统介绍当前主要的R-loop检测技术及其应用进展。

1 R-loop调控多种生理功能

R-loop在特定基因组位点的时空特异性形成与及时消解, 使其成为维持细胞稳态的关键分子开关, 在多种生理过程中发挥核心调控作用。R-loop主要参与调控的生理活动包括以下方面。

(1) 转录调控. 在转录终止区域形成的R-loop可特异性募集解旋酶SETX(senataxin), 通过消解R-loop结构促进外切酶XRN2对下游转录本的降解, 从而实现高效的转录终止^[7,8]。此外, R-loop可通过诱导反义链转录产生双链RNA(dsRNA), 进而招募RNA干扰相关因子(如DICER和AGO2)形成RNA诱导的转录沉默复合体, 最终导致转录抑制和终止^[10]。另外, 位于启动子近端的R-loop可被SOSS-INTAC(sensor of single-strand-integrator complex)稳定复合体特异性识别, 该复合体通过其内源性核酸内切酶活性诱导新生转录本的剪切, 实现转录的提前终止^[9]。这些发现表明, R-loop参与从启动子到终止子的全转录过程调控网络。

(2) DNA损伤修复. 在DNA损伤修复过程中, R-loop具有两面性, 其瞬时形成可促进修复, 而持续性存在则会导致修复异常。研究发现, 在同源重组(HR)修复途径中, RAD52介导的R-loop可促进DNA断端桥接, 并利用RNA作为同源模板指导精确修复^[11]。在脆性X综合征模型中, *FMR1*基因CGG三核苷酸重复序列的异常扩张会形成病理性R-loop, 而通过抑制MEK/BRAF激酶信号通路可促进R-loop依赖性修复机制, 选择性切除扩增的CGG重复序列, 从而恢复*FMR1*基因的正常表达^[12,13]。这些发现揭示了R-loop在DNA损伤修复中精细的平衡机制。

(3) 端粒维持. RNA聚合酶II(RNAPII)在端粒区域转录产生的长链非编码RNA(telomeric repeat-contain-

ing RNA, TERRA)能够与互补的端粒DNA序列形成R-loop结构。这种特殊的核酸结构能够作为分子支架选择性募集BRCA1和RAD51等DNA修复因子至端粒区域, 并促进端粒末端损伤的识别与修复^[14]。这一机制表明, R-loop的存在对端粒完整性具有保护作用。

(4) 调控抗体类别转换重组(class switch recombination, CSR). CSR作为一种独特的DNA重组机制, 通过介导抗体恒定区(C区)基因片段的重排, 实现抗体效应功能的多样性。研究发现, 在转换区(S region)形成稳定的R-loop结构能够特异性募集复制相关解旋酶(minichromosome maintenance, MCM)复合体, 通过改变染色质空间构象缩短重组区域的空间距离, 从而促进CSR过程^[15]。该发现表明R-loop在适应性免疫应答中具有重要作用。

(5) 调控造血干细胞分化. 在胚胎期, 造血干细胞(HSPCs)的快速增殖需要精确的基因组稳定性调控。色氨酸-天冬氨酸重复蛋白5(Wdr5)介导的组蛋白H3K4甲基化通过抑制HSPCs中异常R-loop的形成, 从而维持基因组稳定性。当Wdr5缺失时, H3K4甲基化水平降低, 导致R-loop异常积累, 进而引发DNA损伤加剧和p53依赖性凋亡, 最终造成HSPCs数量显著减少^[16]。该发现表明R-loop的动态在平衡造血干细胞分化中发挥关键作用。

2 R-loop调控多层次表观遗传景观

R-loop不仅参与瞬时性的基因表达调控, 还能够通过建立可遗传的表观遗传记忆实现更持久的生物学功能调控。R-loop对DNA甲基化、RNA甲基化、组蛋白修饰以及染色质高级结构等表观遗传景观均具有调控作用。这种多层次、精细化的特征体现了R-loop在表观遗传调控中的枢纽作用。

2.1 R-loop调控DNA甲基化与RNA甲基化修饰

R-loop常富集在低甲基化的CpG岛(CGI)上, 提示R-loop可能与DNA甲基化修饰负相关^[17,23,24]。R-loop中的RNA/DNA杂合链阻碍从头DNA甲基转移酶DNMT3B1在启动子区域的甲基化作用(图2), 从而维持低甲基化水平的CGIs^[17]。此外, 生长阻滞与DNA损伤蛋白45A(GADD45A)可特异性识别并结合启动子区的R-loop, 招募DNA去甲基化酶TET家族, 促进DNA发

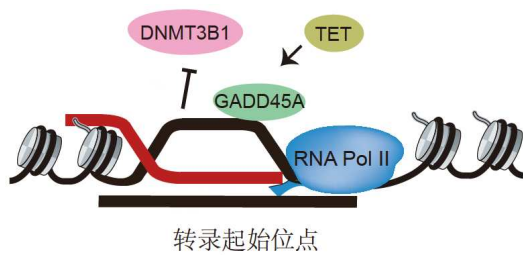


图 2 R-loop维持启动子区域DNA低甲基化的调控机制
Figure 2 Mechanism of R-loop in maintaining DNA hypomethylation at promoters

生去甲基化(图2)^[12,25]. R-loop通过以上机制, 抑制DNA甲基转移酶结合并招募去甲基化酶, 维持人类基因组中数千个基因启动子区域的低甲基化^[26].

m⁶A修饰作为真核生物RNA中最普遍的转录后修饰, 与R-loop存在交互调控作用. 一方面R-loop可促进m⁶A修饰, R-loop通过招募解旋酶DDX21至基因末端, 促进甲基转移酶METTL3在终止位点附近RNA上添加m⁶A修饰, 进而引导XRN2介导的转录终止^[27]. 另一方面, R-loop受到m⁶A及其相关因子调控, 如转录终止阶段m⁶A促进R-loop形成, 该步骤可以防止RNAPII通读, 缺失m⁶A甲基转移酶METTL3会显著降低基因转录终止区域的R-loop水平(图3)^[28]. 相反, 缺失m⁶A识别蛋白YTHDF2会导致核内R-loop水平急剧上升^[29]. YTHDF2和METTL3的缺失导致R-loop水平变化截然相反, 但均会造成DNA双链断裂, 从而破坏基因组的完整性^[28,29]. 此外, ALKBH2和ALKBH3等去甲基化酶可以去除R-loop上的1-meA和3-meC, 从而使RNA不稳定并抑制R-loop的形成^[30]. 这些发现表明, RNA甲基化修饰与R-loop之间存在密切联系.

2.2 R-loop与组蛋白修饰之间的交互调控

由于R-loop大多数是由新生RNA与双链DNA原位杂交形成, 其基因组分布与激活型表观遗传修饰呈现出显著的相关性. 例如, 启动子区域的R-loop与组蛋白修饰H3K4me1, H3K27ac和H3K36me3等共定位^[31]. 启动子有R-loop分布的基因比不含R-loop的基因富集更高水平的H3K36me3^[31]. 这种关联性在不同生物学过程中均得到验证, 如在B细胞类别转换过程中以及发生B细胞特异的激活诱导型胞苷脱氨酶的脱靶效应区域中, R-loop富集区域呈现出高转录及激活型组蛋白修饰的富集^[32]. 转录终止区的R-loop被发现富集

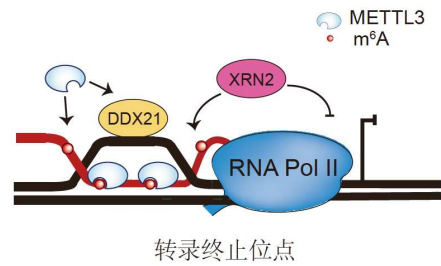


图 3 R-loop与m⁶A共同介导转录终止
Figure 3 R-loop and m⁶A cooperatively regulate transcription termination

H3K4me1标记, 并招募p300乙酰转移酶、CTCF和黏连蛋白^[24]. 另外, 在人胚胎干细胞向多种不同类型细胞定向分化的过程中, 大部分R-loop伴随谱系特异基因表达增加而产生, 与组蛋白修饰H3K27ac和H3K4me3等具有更高的共定位水平^[18].

R-loop与抑制型组蛋白修饰之间存在更加复杂的调控模式. 过表达RNase H1降低R-loop水平会使基因启动子区域招募更多的PRC2, 增加组蛋白H3K27me3修饰水平, 进而导致基因沉默^[19]. 而R-loop水平升高则会抑制PRC2复合物的结合, 造成R-loop区域的组蛋白H3K27me3修饰水平明显低于无R-loop区域^[33]. 组蛋白去甲基化酶LSD1可以去除活跃基因启动子区域的H3K4甲基化修饰, 促进PRC2复合物对组蛋白H3K27位点进行甲基化修饰, 该位点进一步招募RNA解旋酶DDX19A去除启动子上的R-loop^[34]. 催化组蛋白H2AK119ub修饰的PRC1^[35]、组蛋白乙酰基转移酶KAT8^[36]以及负责组蛋白去乙酰化的SIN3A复合物^[37]均抑制R-loop的形成. 然而, 在环境压力刺激下, R-loop能够招募PRC1, PRC2和G9a等因子, 促进抑制型组蛋白修饰H2AK119ub, H3K27me3和H3K9me2/3的沉积, 维持局部转录抑制(图4)^[38]. 转录终止区域的R-loop还通过招募组蛋白赖氨酸甲基转移G9a催化H3K9me2修饰, 以促进异染色质形成, 进而执行转录终止功能^[10].

2.3 R-loop调控染色质高级结构

近年的研究发现, R-loop与调控染色质高级结构的因子密切相关. 在小鼠胚胎干细胞中, G四联体(G4)相关的R-loop结构促进CTCF结合到染色质, 进而促进染色质环形成^[39]. 黏连蛋白复合物的核心亚基STAG能够通过R-loop互作, 介导黏连蛋白以非经典途径

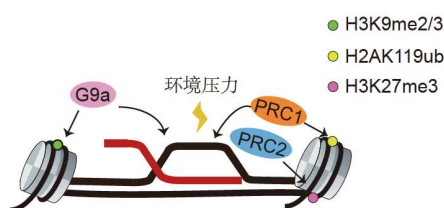


图4 R-loop响应环境压力塑造转录抑制的状态

Figure 4 R-loop orchestrates stress-responsive transcriptional repression

(不依赖NIPBL)装载至染色质^[40]。R-loop还可以增强CTCF染色质边界活性, 阻断黏连蛋白介导的DNA压缩, 从而在染色质拓扑相关结构域(TAD)的边界处发挥界定作用(图5)^[41,42]。这些研究结果表明, R-loop有助于染色质结构相关蛋白的招募或稳定, 参与调控染色质高级结构。

3 R-loop异常积累导致疾病发生

异常的R-loop积累会增加基因组不稳定性, 进而引发多种疾病^[43-45]。基因组不稳定性主要源于转录和复制过程异常, 以及DNA损伤累积与修复机制受损^[19,45,46]。因此, R-loop相关疾病可分为两种主要类型:

(1) 转录或复制异常引发的疾病。骨髓增生异常综合征(myelodysplastic syndromes, MDS)患者携带突变的编码剪接因子SRSF2和U2小核RNA辅助因子1基因, 导致R-loop的积累, 进而损害造血祖细胞增殖, 促进恶性转化^[47,48]。幽门螺杆菌感染通过激活NF- κ B信号通路, 增加转录水平, 引起R-loop依赖性复制应激, 最终导致胃癌发生^[49]。在神经系统疾病中, SETX解旋酶基因突变导致R-loop水平降低, 激活TGF- β 通路, 最终诱发IV型肌萎缩侧索硬化症(ALS4)特征性的神经元功能障碍和死亡^[26]。

(2) DNA损伤或损伤修复机制受损引起的疾病。在遗传性乳腺癌中, BRCA1/2基因功能缺陷会引起R-loop积累, 进而造成DNA损伤^[50]。在部分嗜酸性粒细胞白血病患者中检测到聚腺苷酸化因子FIP1与PDGFR α 发生融合, 使FIP1蛋白功能受损, 导致R-loop累积和DNA损伤^[51]。II型精神性视觉失明共济失调(AOA2)是一种罕见的常染色体隐性遗传小脑共济失调疾病, AOA2患者的iPS细胞模型表现出R-loop积累增加、氧化损伤水平增加以及DNA损伤诱导细胞死亡等表型^[52]。尤

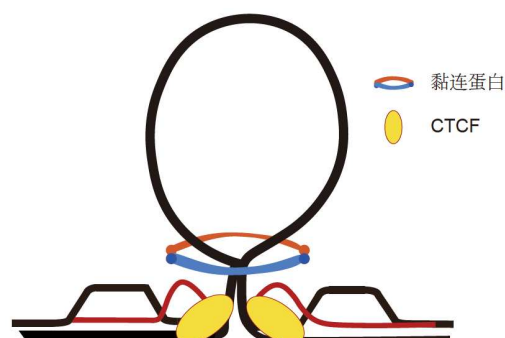


图5 R-loop在细胞中阻碍黏连蛋白介导的环挤压

Figure 5 R-loop acts as barriers to cohesin-mediated loop extrusion in cells

因肉瘤中, 染色质易位导致EWSR1与FLI1蛋白发生融合, 抑制EWSR1的正常功能从而导致R-loop积累, 同时招募BRCA1至转录延伸复合物, 破坏同源重组修复途径, 最终导致基因组不稳定^[53]。

综上所述, R-loop在多项生命过程中发挥正常生理作用的同时, 也会导致基因组不稳定而驱动多种疾病的发生。因此, 细胞内通过多种机制维持R-loop的平衡, 以保证其积极效应并减少不利影响, 来维持生命活动的正常进行。

4 R-loop检测方法的开发

随着二代测序技术的发展, 研究人员开发出了多种研究R-loop的高通量检测方法(表1)。现有R-loop的高通量检测技术可分为亲和富集和非亲和富集两类。亲和富集主要依赖于S9.6抗体和RNase H的DNA:RNA杂合链结合结构域(hybrid binding domain, HBD)(图6和7)。S9.6是一种单克隆抗体, 它是由 Φ X174噬菌体DNA体外转录产生的DNA:RNA杂合链作为抗原免疫小鼠后, 经过多轮杂交瘤细胞筛选而生成, 对DNA:RNA杂合链展现出了很高的亲和力^[54,55]。RNase H1是一类保守的核糖核酸酶, 包含了由一个三链反平行 β 折叠和两个短螺旋组成的HBD结构域。HBD能够特异性识别DNA:RNA杂合链。除这两类技术外, 研究人员还开发出了非亲和富集的R-loop检测技术, 主要依赖于化学标记和核酸酶切割。

4.1 基于S9.6抗体开发的技术

S9.6抗体是一种能够特异性识别DNA:RNA杂合

表 1 R-loop检测技术
Table 1 R-loop detecting technologies

策略	检测方法	基本原理	优点	不足	提出时间
基于S9.6抗体的检测手段	DRIP-seq ^[17]	利用S9.6抗体特异性识别DNA:RNA杂合链, 并通过免疫沉淀和高通量测序定位R-loop的分布	普适性、R-loop信息可靠	分辨率有限	2012
	DRIPc-seq ^[31,56]	在DRIP-seq基础上引入逆转录步骤进行链特异性测序	具有链特异性信息	分辨率有限	2016
	S1-DRIP-seq ^[58]	通过S1核酸酶进行酶切, 在超声前去除R-loop的非模板单链DNA(ssDNA), 从而得到R-loop的分布信息	信噪比高	对AT富集的区域有偏好性	2016
	ssDRIP-seq ^[57]	通过限制性内切酶组合打断基因组DNA后, 利用S9.6抗体富集DNA:RNA杂合链, 并对其中的单链DNA进行检测	保留链特异性信息	分辨率较低、依赖抗体富集效率	2017
	bisDRIP-seq ^[59]	将DRIP和亚硫酸盐印迹技术相结合, 精确定位R-loop	增加了单链中胞嘧啶被转换的信息	亚硫酸盐处理易导致DNA降解	2017
	S9.6-based CUT&TAG ^[60]	S9.6抗体引导Tn5转座酶原位切割R-loop位点, 获得R-loop信息	高灵敏度、所需细胞少	对抗体的要求高	2018
基于HBD的检测手段	DRIVE-seq ^[17,61]	利用HBD体外捕获DNA:RNA杂合链并进行建库测序	特异性高	实验起始投入量高	2012
	R-ChIP ^[23,62]	利用丧失切割活性的突变型RNase H1 (D210N)进行染色质免疫共沉淀, 实现R-loop的富集	实现原位R-loop位点的精确鉴定	需构建稳定表达突变型RNase H1的细胞系, 限制其应用	2017
	MapR ^[63,64]	通过HBD与微球菌核酸酶(MNase)融合蛋白靶向结合并切割R-loop, 建库测序	实现原位R-loop位点的精确鉴定, 获得瞬时R-loop信息	获得的R-loop信号不具有链特异性信息	2019
	BisMapR ^[66]	在MapR技术基础上整合亚硫酸盐处理, 通过单链DNA修饰差异保留链方向信息	保留R-loop的链方向信息	需亚硫酸盐转化, 操作复杂	2021
	HBD-based CUT&TAG ^[67]	将HBD与Tn5转座酶融合, 在固定细胞中原位标记R-loop位点	高灵敏度、所需细胞少、特异性高	依赖于HBD与GST-His ₆ 的结合效率及稳定性	2021
非亲和富集方法	spKAS-seq ^[68]	基于N3-酮醛(N3-kethoxal)能够快速且特异标记ssDNA, 建库测序	无需抗体、操作简便并能快速且特异标记ssDNA	无法区分非典型DNA结构中的单链	2022
	RIAN-seq ^[69]	利用核酸酶选择性消化单链RNA(ssRNA)、单链DNA(ssDNA)和双链DNA(dsDNA), 保留DNA:RNA杂合链, 并通过Tn5转座酶精准建库, 实现R-loop的检测	单碱基分辨率、高灵敏度	丢失瞬时R-loop信息、依赖于核酸酶消化效率	2025

链的单克隆抗体, 自其发现以来已成为R-loop研究的重要工具^[54,55]。基于S9.6抗体的DRIP-seq(DNA:RNA immunoprecipitation sequencing)通过免疫沉淀结合高通量测序, 实现了全基因组水平R-loop的定位分析^[17]。近年来, 研究人员在DRIP-seq基础上不断优化技术体系, 开发了多种改进方法, 如结合逆转录的DRIPc-seq(DNA:RNA immunoprecipitation followed by cDNA conversion and sequencing)^[31,56]、保留链特异性信息的ssDRIP-seq(single-strand DNA:RNA immunoprecipitation sequencing)^[57]等, 为深入解析R-loop的生物学功能提供了强有力的技术支撑。

4.1.1 DRIP-seq

DRIP-seq的核心原理是S9.6抗体能够精准地识别并结合DNA:RNA杂合链, 从而将含有R-loop结构的核酸片段特异性富集出来。再通过高通量测序技术对富集的核酸片段进行测序分析, 进而准确鉴定R-loop在基因组中的具体位置^[17]。DRIP-seq在技术上具有普适性, 能够提供可靠的R-loop信息。然而, 受限于限制性内切酶切割打断的分辨率, 该技术难以精确定位R-loop边界, 同时S9.6抗体存在与dsRNA潜在的交叉反应以及缺乏链特异性信息等问题限制了其在R-loop精细研究中的应用^[54]。因此, 研究人员对此技术进行了一

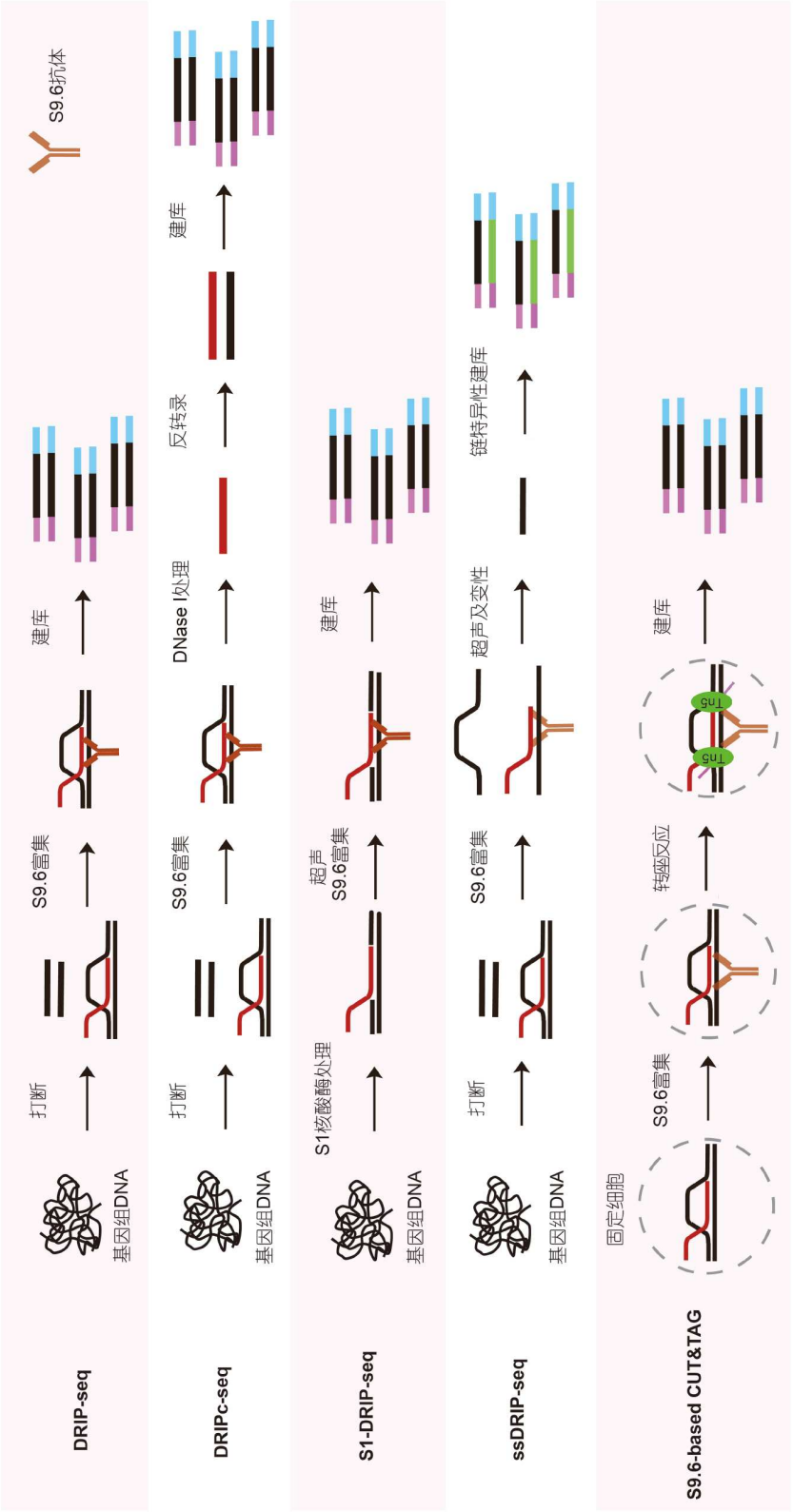


图 6 基于S9.6抗体的R-loop高通量测序技术
Figure 6 S9.6 antibody-based R-loop high-throughput sequencing technology

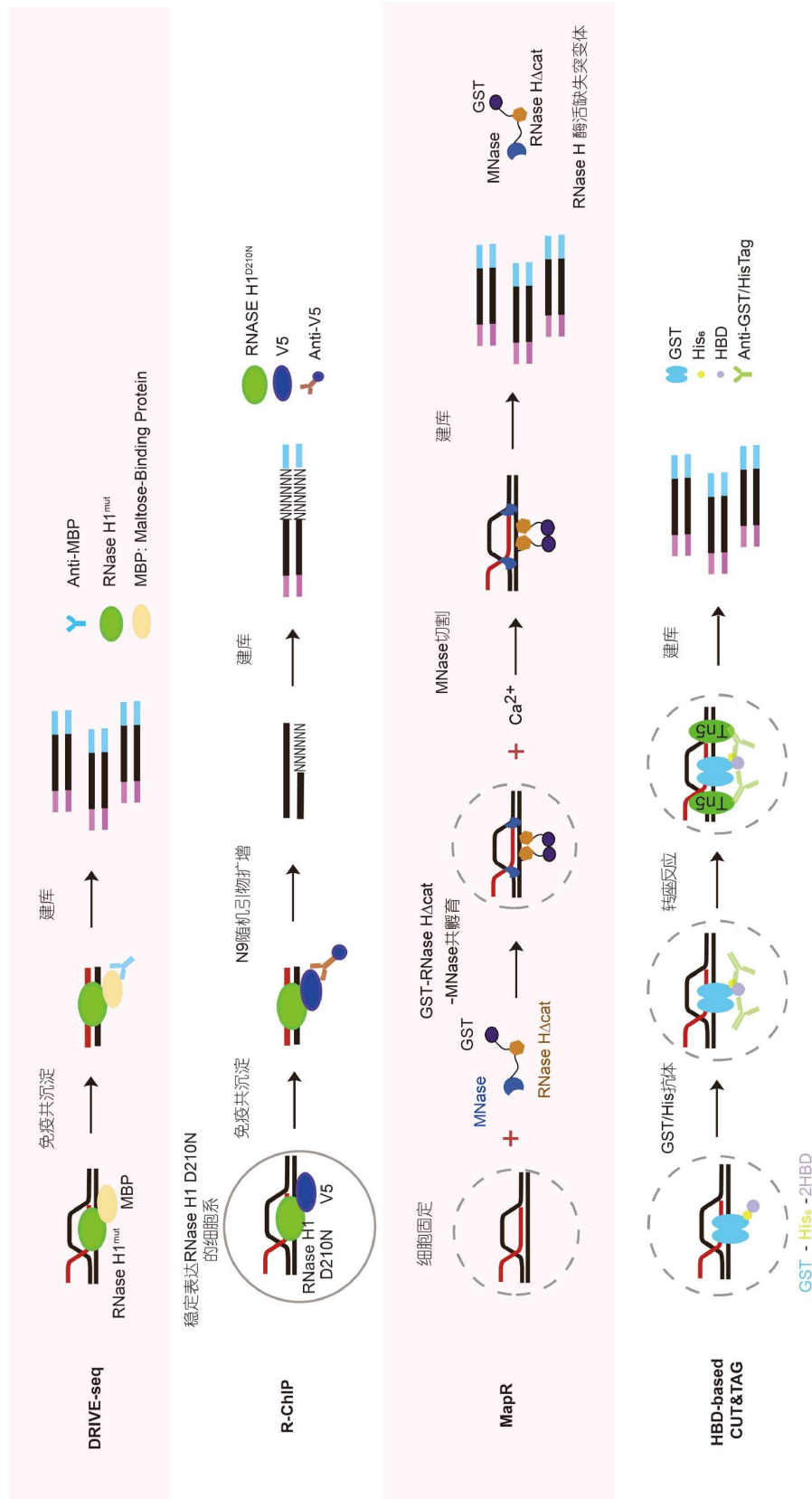


图 7 基于HBD的R-loop高通量测序技术
Figure 7 HBD-based R-loop high-throughput sequencing technology

系列优化.

4.1.2 DRIPc-seq

DRIPc-seq技术是一种基于免疫沉淀和逆转录测序的R-loop检测方法,其核心流程为: S9.6抗体与核酸片段孵育,富集R-loop结构中的DNA:RNA杂合链.随后,加入DNase I对核酸进行处理,特异性消化单链DNA.接着,通过逆转录将DNA:RNA杂合链中的RNA转化为cDNA,进而构建测序文库并进行高通量测序.这一技术通过链特异性分析,能够明确区分R-loop的模板链和非模板链,从而实现对R-loop分布的精确鉴定^[31,56].然而,该技术仍然依赖S9.6抗体,因此可能存在与dsRNA交叉反应的问题,而且其建库流程比标准的DRIP-seq更复杂,可能影响实验的可重复性.

4.1.3 S1-DRIP-seq

S1-DRIP-seq(S1 nuclease DNA:RNA immunoprecipitation sequencing)是通过S1核酸酶对样本进行酶切,在超声前去除R-loop的非模板ssDNA,从而得到较高信噪比的R-loop分布信息,但该技术对AT富集的区域有偏好性^[58].因此,该技术也存在一定的局限性: S1核酸酶对GC富集区域的R-loop敏感性较低,可能导致部分R-loop信号丢失.因此, S1-DRIP-seq更适合用于特定R-loop富集区域的研究,而非全基因组范围的分析.

4.1.4 ssDRIP-seq

ssDRIP-seq技术是一种通过限制性内切酶组合打断基因组DNA后,利用S9.6抗体富集DNA:RNA杂合链,并对其中的DNA单链进行链特异性建库的全基因组R-loop检测技术^[57].该技术通过单链DNA建库策略,不仅有效规避了S9.6抗体对dsRNA的交叉反应,还保留了DNA:RNA杂合链的链特异性信息,同时避免了对染色质状态的依赖性.

4.1.5 bisDRIP-seq

bisDRIP-seq(bisulfite DNA:RNA immunoprecipitation sequencing)技术是将DRIP和亚硫酸盐印迹技术相结合,其核心流程包括: 利用S9.6抗体特异性富集R-loop,随后进行亚硫酸盐处理,将单链DNA中的胞嘧啶(C)转化为尿嘧啶(U).最后通过测序分析来精确定位R-loop,并提供单链DNA链特异性信息^[59].然而,亚

硫酸盐处理可能导致DNA降解,影响文库质量,因此需要优化实验条件以提高数据的可靠性.

4.1.6 S9.6-based CUT&TAG

S9.6-based CUT&TAG(S9.6 antibody-directed cleavage under targets & tagmentation)是一种原位R-loop检测技术,其核心原理是利用融合蛋白(Protein A/G-Tn5转座酶)在S9.6抗体结合的R-loop位点进行靶向切割和测序接头插入,最终通过PCR扩增和高通量测序实现R-loop的高精度定位^[60].该技术的优势在于其低样本需求以及原位检测能力.然而, Tn5转座酶对开放区域的亲和性可能会造成大量的假阳性^[60].

4.2 基于HBD开发的检测技术

尽管S9.6抗体能够特异性识别RNA:DNA杂合链,但其对dsRNA也存在一定的结合活性,这可能影响R-loop检测的准确性和特异性.为了提高检测的准确性和特异性,一系列基于RNase H1的HBD结构域的检测技术被陆续开发.与S9.6抗体相比, HBD能够专门识别DNA:RNA杂合链,它具有更高的特异性,能更准确地识别R-loop结构.

4.2.1 DRIVE-seq

DRIVE-seq(DNA:RNA *in vitro* enrichment and sequencing)是利用体外纯化的酶催化失活型的RNase H1突变体(如人源第210位的天冬氨酸D突变为天冬酰胺N,即D210N)特异性识别并结合细胞内的DNA:RNA杂合链,但不消解DNA:RNA杂合链^[17,61],通过与片段化基因组DNA孵育后富集R-loop,进而进行高通量测序分析.此方法能有效鉴定R-loop,但是其不足之处在于依赖体外纯化的蛋白活性,且不能表征细胞中全基因组R-loop的原位分布.

4.2.2 R-ChIP

付向东教授实验室^[23,62]运用酶催化失活型的RNase H1开发了一种体内检测R-loop的技术,称为R-ChIP(R-loop chromatin immunoprecipitation).该方法首先构建稳定表达V5表位标记的酶催化失活型RNase H1突变体(D210N)的细胞系.然后,运用V5抗体对RNase H1(D210N)-RNA:DNA杂合链进行免疫共沉淀,从而富集基因组中的R-loop.最后,通过对沉淀的DNA

片段进行建库和高通量测序,实现对哺乳动物细胞全基因组范围内R-loop位点的精确鉴定.但是此方法需要构建稳定表达酶催化失活型RNase H1的细胞系,这使其应用受到了一定的限制.

4.2.3 MapR

MapR是一种依赖于RNase H,并结合CUT&RUN技术来检测基因组中R-loop的方法^[63,64].在MapR中,酶催化失活型的RNase H突变体(RHΔ)与微球菌核酸酶(MNase)形成融合蛋白(RHΔ-MNase),能够特异性结合R-loop,并利用MNase切割邻近DNA,释放R-loop相关片段用于测序^[65].与传统的R-loop检测方法相比,由于MNase具有高效切割特性,因此能够捕获到那些结合较弱或存在时间较短的瞬时R-loop结构,提高了检测灵敏度.然而,其不足之处是依赖于体外纯化的融合蛋白的活性,且获得的R-loop信号不具有链特异性信息.为了获得链特异性信息,研究人员进一步开发了BisMapR^[66]测序方法.该方法是将释放的R-loop在非变性条件下用亚硫酸盐处理,将R-loop中未配对的单链DNA中的胞嘧啶(C)转换为胸腺嘧啶(T),从而保留R-loop的链信息.

4.2.4 HBD-based CUT&TAG

研究人员通过将HBD结构域特异性识别并结合DNA:RNA杂合链,同时结合基于Tn5的CUT&TAG技术,开发了原位R-loop检测技术R-loop CUT&TAG技术^[67].该技术是将细胞与伴刀豆蛋白A(ConA)磁珠固定,并将HBD与谷胱甘肽S转移酶-6×组氨酸标签(GST-His₆)融合表达的蛋白与细胞孵育,利用HBD特异性识别并结合R-loop,然后通过GST或His₆标签抗体将Tn5转座酶复合物精准招募至R-loop,在Mg²⁺激活条件下,转座酶在R-loop区域直接完成DNA切割和测序

接头插入.最后,通过PCR扩增完成文库构建.与其他R-loop定位方法相比,R-loop CUT&TAG不仅能捕获到特异的R-loop信号,而且能捕获到基因间区域和增强子上的瞬时R-loop信号.

4.3 非亲和富集R-loop检测方法

近年来发展的非亲和富集的检测方法突破了传统的依赖S9.6抗体或HBD技术的局限性.spKAS-seq(strand-specific kethoxal-assisted single-stranded DNA sequencing)是利用N3-酮醛标记ssDNA,来检测微量样本中的R-loop,但该方法可能会受其他ssDNA结构干扰.本团队开发的RIAN-seq(R-loop identification assisted by nucleases and sequencing)通过特异性核酸酶消化ssRNA, ssDNA和dsDNA,保留DNA:RNA杂合链,结合Tn5建库,实现了R-loop的单碱基分辨率检测.

4.3.1 spKAS-seq

spKAS-seq是基于N3-酮醛能够快速且特异标记ssDNA来实现对R-loop结构的检测.该技术所需的细胞起始量低,可用于检测少量细胞和临床样本中的R-loop(图8)^[68],但该技术也可能捕获基因组中的其他非典型单链DNA结构,限制了spKAS-seq技术的应用.

4.3.2 RIAN-seq

最近,本团队^[69]开发了高分辨率、高灵敏度、无偏好性的R-loop检测新技术RIAN-seq,可在单碱基分辨率下定位R-loop(图8).该技术通过利用核酸酶P1、T5外切酶和λ外切酶选择性消化ssRNA, ssDNA和dsDNA,同时保留DNA:RNA杂合链,结合Tn5转座酶介导的建库策略,从而可以更加精准鉴定R-loop的位置和长度,检测到的R-loop数量比现有方法高出一个数量级.该技术发现基因组中存在两类不同的R-loops,

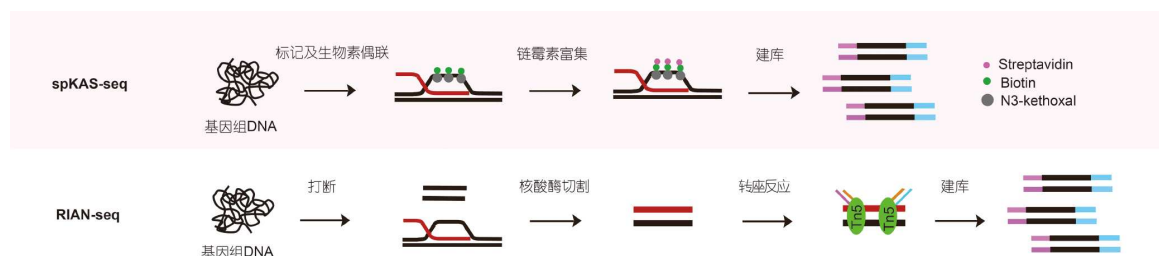


图8 非亲和富集的R-loop高通量测序技术

Figure 8 Non-affinity enriched R-loop high-throughput sequencing technology

即簇状R-loops(clustered R-loops)和DNA损伤易感型R-loops(DNA damage susceptible R-loop, DdsR-loops); 揭示启动子区clustered R-loops以数量依赖的方式调控基因表达和DdsR-loops的序列特征及其物种保守性。

综上所述, 传统基于S9.6抗体的检测方法被广泛应用, 但由于该抗体与dsRNA存在交叉反应, 可能导致假阳性信号。为克服S9.6抗体的局限性, 基于RNase H1的HBD的检测方法应运而生。HBD能特异性识别DNA:RNA杂合链, 可显著提高R-loop检测的特异性, 但仍受限于HBD结合效率低和R-loop捕获数量有限等问题。因此, 现有亲和富集方法普遍难以实现单碱基分辨率检测, 且需要较高的样本起始量。而我们团队最新开发的RIAN-seq技术能够有效突破以上瓶颈, 实现全基因组范围内R-loop的单碱基分辨率检测, 并且对低丰度的R-loop具有更高的检测灵敏度, 检测成本较低、操作简便。同时, 我们运用RIAN-seq技术揭示了R-loop在基因表达调控、基因组稳定性维持等关键生物学过程的分子机制。这不仅深化了人们对R-loop生物学功能的理解, 也为相关疾病的机制研究和临床转化奠定了重要基础。

5 总结与展望

R-loop在维持基因组稳定性、调控基因表达及塑

造表观遗传景观中发挥着关键作用。生理状态下, R-loop通过精确调控DNA损伤修复、转录延伸与终止, 进而参与细胞命运的精细调控。然而, 其动态平衡的破坏会导致DNA损伤累积、基因组不稳定性增加, 进而与多种肿瘤及遗传性疾病的发生密切相关。近年来, R-loop检测技术的快速发展, 为解析其分子特征和功能机制提供了重要工具。基于S9.6抗体与基于RNase H1等亲和富集型方法, 以及spKAS-seq和最新开发的RIAN-seq等非亲和富集型方法, 都极大地推动了R-loop研究的深入发展, 为理解其分子机制提供了有力工具。

当前R-loop研究领域仍然充满机遇与挑战。在技术层面, 亟需突破现有检测方法的局限性, 开发高灵敏度、低样本需求的单细胞R-loop检测技术, 这将为揭示细胞异质性背景下R-loop的动态调控规律提供关键工具, 并有助于构建跨细胞类型、组织和物种的R-loop调控网络。在机制研究方面, 需要建立生理性与病理性R-loop的分子特征谱系, 系统阐明二者在形成、维持和清除过程中的调控差异及分子基础。值得期待的是, 随着空间多组学、单分子成像等技术的协同应用, R-loop研究有望在基因表达调控、表观遗传学和疾病机制等方面取得突破性进展, 这些发现将深化人们对R-loop的认识, 并为疾病的早期诊断和精准治疗提供潜在分子靶点和干预策略。

参考文献

- 1 Thomas M, White R L, Davis R W. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc Natl Acad Sci USA*, 1976, 73: 2294–2298
- 2 Massé E, Drolet M. *Escherichia coli* DNA topoisomerase I inhibits R-loop formation by relaxing transcription-induced negative supercoiling. *J Biol Chem*, 1999, 274: 16659–16664
- 3 Kuzminov A. When DNA topology turns deadly—RNA polymerases dig in their R-loops to stand their ground: new positive and negative (Super)Twists in the replication-transcription conflict. *Trends Genet*, 2018, 34: 111–120
- 4 Lang K S, Hall A N, Merrikh C N, et al. Replication-transcription conflicts generate R-loops that orchestrate bacterial stress survival and pathogenesis. *Cell*, 2017, 170: 787–799.e18
- 5 Abraham K J, Khosraviyani N, Chan J N Y, et al. Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature*, 2020, 585: 298–302
- 6 Chédin F. Nascent connections: R-loops and chromatin patterning. *Trends Genet*, 2016, 32: 828–838
- 7 Skourti-Stathaki K, Proudfoot N J, Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell*, 2011, 42: 794–805
- 8 Morales J C, Richard P, Patidar P L, et al. XRN2 links transcription termination to DNA damage and replication stress. *PLoS Genet*, 2016, 12: e1006107
- 9 Xu C, Li C, Chen J, et al. R-loop-dependent promoter-proximal termination ensures genome stability. *Nature*, 2023, 621: 610–619

- 10 Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot N J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature*, 2014, 516: 436–439
- 11 Iannelli F, Galbiati A, Capozzo I, et al. A damaged genome's transcriptional landscape through multilayered expression profiling around *in situ*-mapped DNA double-strand breaks. *Nat Commun*, 2017, 8: 15656
- 12 Lee H G, Imaichi S, Kraeutler E, et al. Site-specific R-loops induce CGG repeat contraction and fragile X gene reactivation. *Cell*, 2023, 186: 2593–2609.e18
- 13 Malik I, Kelley C P, Wang E T, et al. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat Rev Mol Cell Biol*, 2021, 22: 589–607
- 14 Vohhodina J, Goehring L J, Liu B, et al. BRCA1 binds TERRA RNA and suppresses R-Loop-based telomeric DNA damage. *Nat Commun*, 2021, 12: 3542
- 15 Wiedemann E M, Peycheva M, Pavri R. DNA replication origins in immunoglobulin switch regions regulate class switch recombination in an R-loop-dependent manner. *Cell Rep*, 2016, 17: 2927–2942
- 16 Wang X, Liu M, Zhang Y, et al. Wdr5-mediated H3K4 methylation facilitates HSPC development via maintenance of genomic stability in zebrafish. *Proc Natl Acad Sci USA*, 2025, 122: e2420534122
- 17 Ginno P A, Lott P L, Christensen H C, et al. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell*, 2012, 45: 814–825
- 18 Yan P, Liu Z, Song M, et al. Genome-wide R-loop landscapes during cell differentiation and reprogramming. *Cell Rep*, 2020, 32: 107870
- 19 Skourti-Stathaki K, Torlai Triglia E, Warburton M, et al. R-loops enhance polycomb repression at a subset of developmental regulator genes. *Mol Cell*, 2019, 73: 930–945.e4
- 20 Huang H, Zhu Q, Jussila A, et al. CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains. *Nat Genet*, 2021, 53: 1064–1074
- 21 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. *Cell*, 2014, 159: 1665–1680
- 22 Dixon J R, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 2012, 485: 376–380
- 23 Chen L, Chen J Y, Zhang X, et al. R-ChIP using inactive RNase H reveals dynamic coupling of R-loops with transcriptional pausing at gene promoters. *Mol Cell*, 2017, 68: 745–757.e5
- 24 Ginno P A, Lim Y W, Lott P L, et al. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res*, 2013, 23: 1590–1600
- 25 Arab K, Karaulanov E, Musheev M, et al. GADD45A binds R-loops and recruits TET1 to CpG island promoters. *Nat Genet*, 2019, 51: 217–223
- 26 Grunseich C, Wang I X, Watts J A, et al. Senataxin mutation reveals how R-loops promote transcription by blocking DNA methylation at gene promoters. *Mol Cell*, 2018, 69: 426–437.e7
- 27 Hao J D, Liu Q L, Liu M X, et al. DDX21 mediates co-transcriptional RNA m⁶A modification to promote transcription termination and genome stability. *Mol Cell*, 2024, 84: 1711–1726.e11
- 28 Yang X, Liu Q L, Xu W, et al. m⁶A promotes R-loop formation to facilitate transcription termination. *Cell Res*, 2019, 29: 1035–1038
- 29 Abakir A, Giles T C, Cristini A, et al. N⁶-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nat Genet*, 2020, 52: 48–55
- 30 Falnes P Ø, Bjørås M, Aas P A, et al. Substrate specificities of bacterial and human AlkB proteins. *Nucleic Acids Res*, 2004, 32: 3456–3461
- 31 Sanz L A, Hartono S R, Lim Y W, et al. Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals. *Mol Cell*, 2016, 63: 167–178
- 32 Wang I X, Core L J, Kwak H, et al. RNA-DNA differences are generated in human cells within seconds after RNA exits polymerase II. *Cell Rep*, 2014, 6: 906–915
- 33 Chen P B, Chen H V, Acharya D, et al. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat Struct Mol Biol*, 2015, 22: 999–1007
- 34 Pinter S, Knodel F, Choudalakis M, et al. A functional LSD1 coregulator screen reveals a novel transcriptional regulatory cascade connecting R-loop homeostasis with epigenetic regulation. *Nucleic Acids Res*, 2021, 49: 4350–4370

- 35 Klusmann I, Wohlberedt K, Magerhans A, et al. Chromatin modifiers Mdm2 and RNF2 prevent RNA:DNA hybrids that impair DNA replication. *Proc Natl Acad Sci USA*, 2018, 115: E11311–E11320
- 36 Singh D K, Pandita R K, Singh M, et al. MOF suppresses replication stress and contributes to resolution of stalled replication forks. *Mol Cell Biol*, 2018, 38: e00484-17
- 37 Salas - Armenteros I, Pérez - Calero C, Bayona - Feliu A, et al. Human THO-Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability. *EMBO J*, 2017, 36: 3532–3547
- 38 Marnef A, Legube G. R-loops as Janus-faced modulators of DNA repair. *Nat Cell Biol*, 2021, 23: 305–313
- 39 Wulfridge P, Yan Q, Rell N, et al. G-quadruplexes associated with R-loops promote CTCF binding. *Mol Cell*, 2023, 83: 3064–3079.e5
- 40 Porter H, Li Y, Neguembor M V, et al. Cohesin-independent STAG proteins interact with RNA and R-loops and promote complex loading. *eLife*, 2023, 12: e79386
- 41 Zhang H, Shi Z, Banigan E J, et al. CTCF and R-loops are boundaries of cohesin-mediated DNA looping. *Mol Cell*, 2023, 83: 2856–2871.e8
- 42 Luo H, Zhu G, Eshelman M A, et al. HOTTIP-dependent R-loop formation regulates CTCF boundary activity and TAD integrity in leukemia. *Mol Cell*, 2022, 82: 833–851.e11
- 43 Bhatia V, Barroso S I, García-Rubio M L, et al. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature*, 2014, 511: 362–365
- 44 Groh M, Albulescu L O, Cristini A, et al. Senataxin: genome guardian at the interface of transcription and neurodegeneration. *J Mol Biol*, 2017, 429: 3181–3195
- 45 Groh M, Lufino M M P, Wade-Martins R, et al. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet*, 2014, 10: e1004318
- 46 Alecki C, Chiwara V, Sanz L A, et al. RNA-DNA strand exchange by the *Drosophila* Polycomb complex PRC2. *Nat Commun*, 2020, 11: 1781
- 47 Chen L, Chen J Y, Huang Y J, et al. The augmented R-loop is a unifying mechanism for myelodysplastic syndromes induced by high-risk splicing factor mutations. *Mol Cell*, 2018, 69: 412–425.e6
- 48 Nguyen H D, Leong W Y, Li W, et al. Spliceosome mutations induce R loop-associated sensitivity to ATR inhibition in myelodysplastic syndromes. *Cancer Res*, 2018, 78: 5363–5374
- 49 Bauer M, Nascakova Z, Mihai A I, et al. The ALPK1/TIFA/NF-κB axis links a bacterial carcinogen to R-loop-induced replication stress. *Nat Commun*, 2020, 11: 5117
- 50 Tan S L W, Chadha S, Liu Y, et al. A class of environmental and endogenous toxins induces BRCA2 haploinsufficiency and genome instability. *Cell*, 2017, 169: 1105–1118.e15
- 51 Stirling P C, Chan Y A, Minaker S W, et al. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev*, 2012, 26: 163–175
- 52 Becherel O J, Sun J, Yeo A J, et al. A new model to study neurodegeneration in ataxia oculomotor apraxia type 2. *Hum Mol Genet*, 2015, 24: 5759–5774
- 53 Gorthi A, Romero J C, Loranc E, et al. EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature*, 2018, 555: 387–391
- 54 Boguslawski S J, Smith D E, Michalak M A, et al. Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. *J Immunol Methods*, 1986, 89: 123–130
- 55 Bou-Nader C, Bothra A, Garboczi D N, et al. Structural basis of R-loop recognition by the S9.6 monoclonal antibody. *Nat Commun*, 2022, 13: 1641
- 56 Sanz L A, Chédin F. High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. *Nat Protoc*, 2019, 14: 1734–1755
- 57 Xu W, Xu H, Li K, et al. The R-loop is a common chromatin feature of the *Arabidopsis* genome. *Nat Plants*, 2017, 3: 704–714
- 58 Wahba L, Costantino L, Tan F J, et al. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev*, 2016, 30: 1327–1338
- 59 Dumelie J G, Jaffrey S R. Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. *eLife*, 2017, 6: e28306
- 60 Hartono S R, Malapert A, Legros P, et al. The affinity of the S9.6 antibody for double-stranded RNAs impacts the accurate mapping of R-loops in

- fission yeast. *J Mol Biol*, 2018, 430: 272–284
- 61 Nowotny M, Gaidamakov S A, Ghirlando R, et al. Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol Cell*, 2007, 28: 264–276
- 62 Chen J Y, Zhang X, Fu X D, et al. R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1. *Nat Protoc*, 2019, 14: 1661–1685
- 63 Skene P J, Henikoff J G, Henikoff S. Targeted *in situ* genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc*, 2018, 13: 1006–1019
- 64 Skene P J, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife*, 2017, 6: e21856
- 65 Yan Q, Shields E J, Bonasio R, et al. Mapping native R-loops genome-wide using a targeted nuclease approach. *Cell Rep*, 2019, 29: 1369–1380. e5
- 66 Wulfridge P, Sarma K. A nuclease- and bisulfite-based strategy captures strand-specific R-loops genome-wide. *eLife*, 2021, 10: e65146
- 67 Wang K, Wang H, Li C, et al. Genomic profiling of native R loops with a DNA-RNA hybrid recognition sensor. *Sci Adv*, 2021, 7: eabe3516
- 68 Wu T, Lyu R, He C. spKAS-seq reveals R-loop dynamics using low-input materials by detecting single-stranded DNA with strand specificity. *Sci Adv*, 2022, 8: eabq2166
- 69 Li Y, Sheng Y, Di C, et al. Base-pair resolution reveals clustered R-loops and DNA damage-susceptible R-loops. *Mol Cell*, 2025, 85: 1686–1702. e5

Research advances in R-loop functions and detection technologies

HE QianNan^{1,2} & YAO HongJie^{1,3*}

¹ Department of Basic Research, Guangzhou National Laboratory, Guangzhou 510005, China

² Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China

³ State Key Laboratory of Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China

* Corresponding author; E-mail: yao_hongjie@gzlab.ac.cn

R-loops are three-stranded nucleic acid structures formed by a DNA:RNA hybrid paired with a displaced single-stranded DNA. As ubiquitous regulatory elements in genomes, R-loops participate in crucial biological processes including DNA damage repair and transcriptional regulation in both prokaryotes and eukaryotes, while also modulating epigenetic landscapes through the regulation on histone modifications and higher-order chromatin structure. However, disruption of R-loop homeostasis leading to aberrant accumulation can induce DNA damage, genomic instability, and chromatin structural abnormalities, thereby contributing to the pathogenesis of various cancers and genetic disorders. Researchers have developed multiple R-loop detection methodologies, primarily including affinity-based enrichment techniques utilizing S9.6 antibodies or RNase H1 hybrid-binding domains, as well as non-affinity approaches such as single-molecule labeling via spKAS-seq or nuclease cleavage-based RIAN-seq. This review systematically summarizes current understanding of R-loop formation mechanisms, physiological/pathological functions, and detection technologies, providing a theoretical framework for elucidating R-loop biology and its disease relevance.

R-loop, DNA:RNA hybrid, Epigenetics, R-loop detection technologies

doi: [10.1360/SSV-2025-0151](https://doi.org/10.1360/SSV-2025-0151)