

doi: 10.3969/j.issn.0490-6756.2019.04.026

长非编码 RNA 与 SAFB1 相互作用中 RNA 二级结构的作用

李晓杰, 缪辉, 王迪, 詹浩淼, 王琳琳, 吴传芳

(四川大学生命科学院 功能基因组实验室, 成都 610064)

摘要: 为探究 RNA 与 SAFB1 蛋白相互作用过程中 RNA 二级结构的作用, 采用足迹法确定 SAFB1 与 LINC-PINT-1 特定的结合区域, 在此基础上根据 RNA 结构分析研究 SAFB1 结合区域所具有特定的二级结构, 最后通过凝胶迁移率实验论证该二级结构在 RNA 与 SAFB1 蛋白结合中的作用。结果显示: 在针对 LINC-PINT-1 的亚克隆发现只有包含 30~55 nt 区域的亚克隆才能维持 RNA 结合区域的二级结构, 并且证实了 LINC-PINT-1 与 SAFB1 蛋白相互作用依赖于特定的 RNA 二级结构。结果表明: 长非编码 RNA 与 SAFB1 蛋白相互作用或依赖于特定的 RNA 二级结构, 本研究为 RNA 与蛋白质相互作用的分子机制研究提供了有力的理论实验支撑。

关键词: RNA; SAFB1 蛋白; RNA 二级结构; 亚克隆; 凝胶迁移率

中图分类号: Q71 **文献标识码:** A **文章编号:** 0490-6756(2019)04-0743-06

The Role of RNA secondary structure in the interaction between lncRNA and SAFB1

LI Xiao-Jie, MIAO Hui, WANG Di, ZHAN Hao-Miao, WANG Lin-Lin, WU Chuan-Fang

(Department of Functional Genome, College of Life Science, Sichuan University, Chengdu 610064, China)

Abstract: To explore the role of RNA secondary structure in the interaction between RNA and SAFB1 protein, structure analysis and RNA footprinting analysis were adopted. It revealed that SAFB1 binds to a specific region of LINC-PINT-1 and the SAFB1 binding region has a specific secondary structure. In addition, subcloning of LINC-PINT-1 confirmed that subclones contained the 30~55 nt region could maintain this secondary structure. Electrophoretic mobility shift assay(EMSA) found that this secondary structure was involved in the interaction between RNA and SAFB1. In this paper, software prediction and necessary experimental verification have indicated that the interaction between LINC-PINT-1 and SAFB1 depends on the specific RNA secondary structure, which provides a powerful theoretical experimental support for the study of the molecular mechanism of RNA-protein interaction.

Keywords: lncRNA; LINC-PINT; SAFB1; RNA secondary structure; Electrophoretic mobility shift assay

收稿日期: 2018-05-18

基金项目: 国家自然科学基金(81490752)

作者简介: 李晓杰(1988-), 女, 四川乐山人, 硕士研究生, 研究方向医学分子生物学。E-mail: lixiaojiedzy@163.com

通讯作者: 吴传芳。E-mail: wuchuanfang@gmail.com

1 引言

RNA 能够通过分子内的碱基配对形成二级结构及生物大分子,而分子结构决定分子的性质和功能, RNA 的各种功能与其结构紧密相连,为进一步挖掘其功能,须了解 RNA 的结构.随着分子生物学的不断发展,越来越多非编码基因及其功能被识别和揭示,人们逐渐认识到长非编码 RNA(long noncoding RNA, lncRNA)和蛋白质分子一样重要,甚至是主要的功能性分子. RNA 二级结构对于人们熟知的 RNA 类群都是必须的,如 rRNA^[1], snRNAs^[2], snoRNAs^[3], miRNAs^[4] 及 mRNA^[5]. 目前,对 RNA 的结构和功能的研究已成为当今生物信息学研究的一个重要方向.

SAFB1(scaffold attachment factor B1)^[6,7] 主要包含一个位于 N 端的 SAF-box,其中间区域包含一个 RRM(RNA recognition region) 以及一个 NSL(nuclear localization signal),C 端主要为一个可转移的转录抑制结合域^[8,9]. 作为一种多功能蛋白,SAFB1 可以结合 DNA 和 RNA,参与 RNA 加工和应激反应.在参与 RNA 代谢过程方面,由于自身具有 RRM 和蛋白相互作用区域(Glu/Arg rich domain and Gly rich domain),SAFB1 不仅可以与 RNA 代谢相关的蛋白相互作用,如:RNA polymerase II C 末端^[10]、作为 RNA 剪切因子的 SRrp86^[11] 和 SLM-1^[12] 等,此外,Caroline Rivers^[13]、Elaine Hong^[14] 等人的研究表明 SAFB1 还能直接与 RNA 相互作用.与此同时,近年来越来越多的研究证实 lncRNA 在肿瘤的发生、发展、侵袭和转移过程中扮演着重要角色^[15],在 SAFB1 与 lncRNA 互作而发挥生物学作用的相关研究中,对 SAFB1 与 lncRNA 结合方式的探究及分析具有重要意义.

我们实验室前期的研究采用 SELEX(Systematic Evolution of Ligands by Exponential Enrichment)技术,通过多轮富集 SAFB1 特异性结合的 RNA,发现了一系列与 SAFB1 相结合的 lncRNA,其中包含一条名为 *LINC-PINT*(long intergenic non-protein coding RNA, p53 induced transcript)的长非编码 RNA,为了研究 RNA 与 SAFB1 作用的 RNA 二级结构基础,我们将 *LINC-PINT* 与 SAFB1 高亲和性区域命名为 *LINC-PINT-1* 并克隆该区域用于本研究.

2 材料和方法

2.1 材料

2.1.1 主要质粒、菌株、细胞系 pcDNA3.1(+)质粒来源于 invitrogen 公司,由本实验室保存使用;pMD19-T 载体来源于 TaKaRa 公司;*E. coli* DH5 α 菌株由本实验室保存;HEK-293T 来源于 ATCC,由本实验室保存使用.

2.1.2 主要试剂 酵母抽提物、胰蛋白胍均购自 OXOID 公司;DMEM 高糖培养基(Hyclone SH30243.01B);胎牛血清(GIBCO 10099-141);青霉素链霉素混合溶液(Hyclone SV30010);0.25% 胰酶(HyClone SH30042.01B);质粒大提试剂盒:PureLink[®] HiPure Plasmid Filter Purification Kits(Invitrogen K2100-17);转染试剂:Lipofectamine[®] 2000 Reagent(Invitrogen 11668-019);ANTI-FLAG M2 Magnetic Beads(SIGMA-ALDRICH M8823-5 ML);FLAG[®] Peptide(SIGMA F3290-25 MG);蛋白酶抑制剂:Protease inhibitor cocktail tablets(Roche 11 697 498 001);rNTP(NEB # N0450L);T7 RNA Polymerase(Thermo Fisher Scientific Inc EP0112);Thermo scientific Ribolock RNase inhibitor(Thermo Fisher Scientific Inc E00381);RNA Probe Purification Kit(Omega R6249-01);RNase A(Thermo Fisher Scientific Inc AM2274);RNase T1(Thermo Fisher Scientific Inc AM2283);Pierce[™] RNA 3' End Biotinylation Kit(Thermo Fisher Scientific Inc 20160);Chemiluminescent Nucleic Acid Detection Module Kit(Thermo Fisher Scientific Inc 89880);Biotin-16-UTP(Roche 11388908910);Urea(SIGMA-ALDRICH U5378-500 G);T Vector pMD 19(TaKaRa 3271);PrimeSTAR[®] Max DNA Polymerase(TaKaRa R045A);胶回收试剂盒(FOREGENE DE-02011);PCR 产物纯化试剂盒(FOREGENE DE-03011);其他化学试剂除文中注明外,均来自国药公司生产且为分析纯.

2.2 方法

2.2.1 电泳迁移率实验 将标记后的探针于 80 °C 变性 5 min,随后缓慢降温至室温复性,加入含有 SAFB1 蛋白的结合体系于 25 °C 孵育 30 min,随后紫外交联 30 min,加入 RNase A 200 U 和 RNase T1 200 U 于 25 °C 反应 10 min,加入 5 \times SDS loading buffer 至 1 \times ,充分混匀后于 95 °C 加

热 5 min; 200 V 恒压电泳 4 h. 最后进行放射性显影. 使用 biotin-16-UTP 标记后的探针用于电泳迁移率实验, 则使用 Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific Inc 89880) 进行化学发光显影.

2.2.2 RNA 足迹分析 将标记后的探针于 80 °C 变性 5 分钟, 随后缓慢降温至室温复性, 加入含有 SAFB1 蛋白的结合体系于 25 °C 孵育 30 min, 加入 RNase A/T1 部分酶切 RNA 室温酶切 15 min. 随后加入 Precipitation/Inactivation Buffer 20 μ L, 涡旋混匀后 -20 °C 沉淀 20 min. 4 °C 17 000 g 离心 15 min, 弃上清液后加入 70% 乙醇溶液洗涤一次, 干燥沉淀, 加入 7 μ L, Gel Loading Buffer II, 充分混匀后于 95 °C 5 min, 取 3 μ L 于 8 mol/L urea 8% 聚丙烯酰胺凝胶分离 RNA 条带. 将分离后的 RNA 片段通过半干转膜仪转印到尼龙膜上, 紫外交联. 最后使用 Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific Inc 89880) 进行后续化学发光显影.

2.2.3 RNA 结构分析 将生物素末端标记后的探针于 80 °C 变性 5 min, 随后缓慢降温至室温复性, 加入 10 \times RSB(RAN Structure Buffer) 至 1 \times , 室温放置 20 min, 使其回复二级结构. 添加梯度量的 RNase A 或 RNase T1 进行切割, 室温酶切 15 min. 加入 Precipitation/Inactivation Buffer 20 μ L, 涡旋混匀后 -20 °C 沉淀 20 min. 于 4 °C 17 000 g 离心 15 min, 弃上清液后加入 70% 乙醇溶液洗涤一次. 空气干燥沉淀, 加入 7 μ L, Gel Loading Buffer II, 充分混匀后于 95 °C 加热 5 min, 随后取 3 μ L 于 8 mol/L urea 8% 聚丙烯酰胺凝胶 200 V 恒压电泳 5 h. 将尿素胶分离后的 RNA 片段通过半干转膜仪转印到尼龙膜上, 并通过紫外交联将 RNA 固定, 最后使用 Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific Inc 89880) 进行后续化学发光显影.

3 结果

3.1 SAFB1 结合于 LINC-PINT-1 的特定区域

由于本文涉及的 RNA 为 LINC-PINT 与 SAFB1 蛋白高亲和区段的亚克隆, 我们暂命名其为 LINC-PINT-1. 为了验证 SAFB1 与 LINC-PINT-1 的相互作用以及 SAFB1 在 LINC-PINT-1 上的定位, 我们在 SAFB1 存在情况下进行 RNA 足迹分析 (图 1), LINC-PINT-1 首先与

SAFB1 孵育结合, 并加入 RNase A 或 RNase T1 进行部分切割, 由于 SAFB1 结合在 RNA 上, 从而被结合区域随后会避免 RNase 切割, 产生“足迹”. 图中虚线框标注了由于 SAFB1 的结合产生的条带消失或减弱 (图 1). 结果可以发现 SAFB1 结合位置位于 30~55 nt 区域. 由于用于足迹分析的 RNA 均为 3' 末端标记的生物素, 因此其中 RNA marker 指示的大小为距离 RNA 3' 末端的距离, 红色虚线标注的区域对应 LINC-PINT-1 的 20~55 nt 区域.

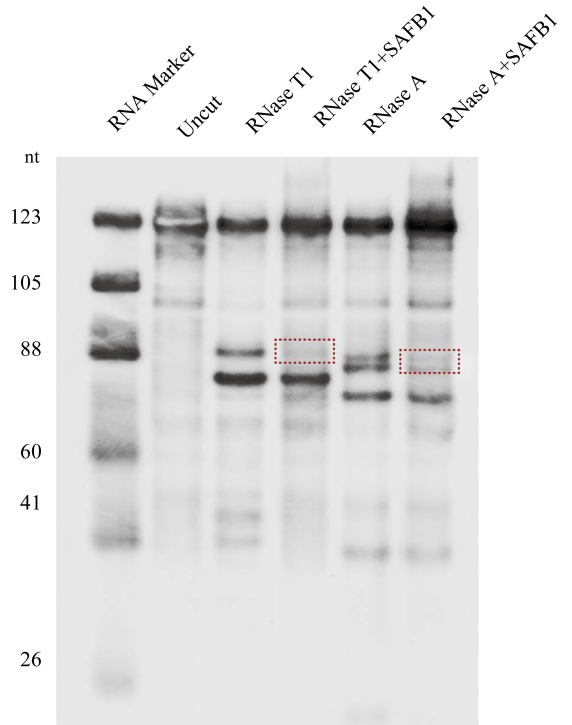


图 1 SAFB1 结合于 LINC-PINT-1 的特定区域

足迹分析 SAFB1 在 LINC-PINT-1 上的结合位置, 其中虚线框标注了出了由于 SAFB1 结合在 RNA 上形成的保护作用, 从而逃避了 RNase 的切割形成的条带减弱区域

Fig. 1 RNA footprinting assay of LINC-PINT-1 shows the region protected by SAFB1 against RNase digestion

3.2 RNA 结构预测和 RNA 结构分析证实 LINC-PINT-1 具有特定二级结构

越来越多的证据表明 RNA 与蛋白相互作用与 RNA 的二级结构相关, 为了探究 LINC-PINT-1 是否存在特定的二级结构, 并且通过特定的二级结构与 SAFB1 相互作用, 我们使用 RNA 二级结构预测软件 RNA structure 对 LINC-PINT-1 进行二级结构预测, 结果发现该 RNA 的 1~80 nt 和 80~123 nt 形成两个比较独立的茎环结构, 如图 2a 所示. 由于大多数 RNA 二级结构预测基于整条

RNA 链最小自由能原则,而实际的 RNA 折叠几乎与转录同时进行,因此可能会存在预测结构与实际结构的偏差,我们将 *LINC-PINT-1* 的 3' 末端标记生物素后用于 RNA 结构分析,结果表明

SAFB1 结合区域具有与 RNA 结构预测一致的二级结构. 如图 2b、2c 所示,其中图 2b 中标注的 1、2 和 A、B 分别对应 RNase T1 和 RNase A 的切割位置,与图 2c 中的位置对应.

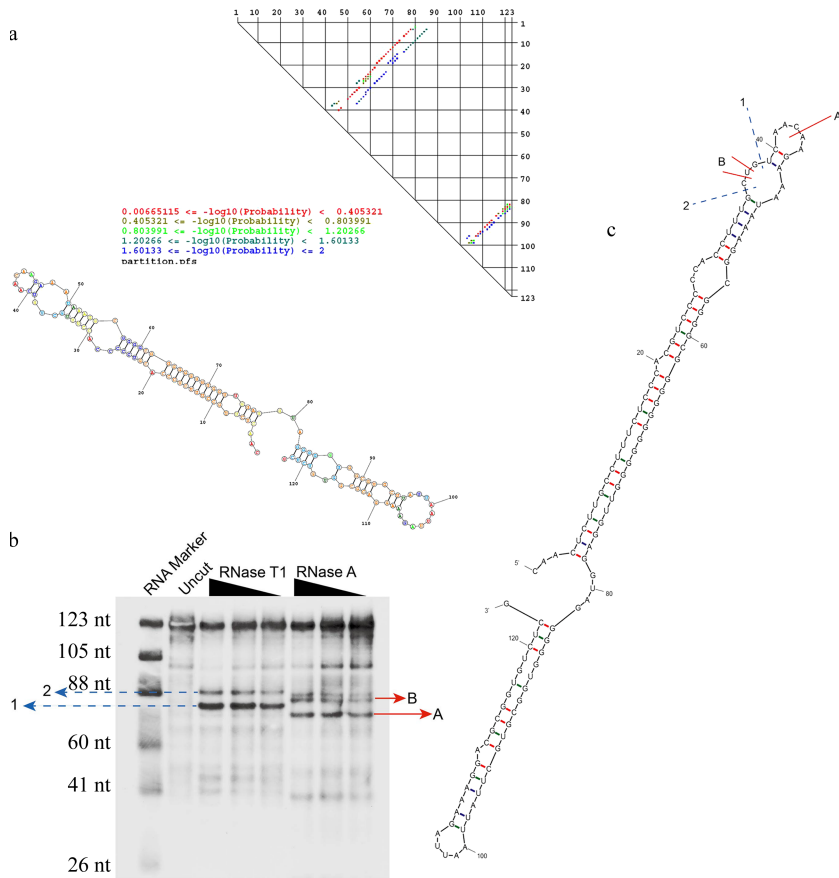


图 2 *LINC-PINT-1* 具有特定的二级结构

a) RNA structure 预测的 RNA 二级结构,图中不同颜色表示各个碱基在整个二级结构中的稳定性. 二级结构中单位碱基形成上述二级结构的可能性用不同颜色标注,横纵轴表示单位碱基位置. b\c) RNA 结构分析 *LINC-PINT-1* 实际的二级结构,其中 B 中的 1\2 和 A\B 对应 C 中的相应切割位置.

Fig. 2 The special secondary structure of *LINC-PINT-1*

As is shown in the figure, 4 colors represent the 4 sequence motifs respectively.

3.3 30~55 nt 是 *LINC-PINT-1* 维持特定二级结构的核心区域

为了进一步验证 SAFB1 在 *LINC-PINT-1* 结合区域形成的二级结构对于 RNA 与 SAFB1 结合的关系,我们将 *LINC-PINT-1* 进行分段亚克隆后再进行结构预测,我们发现只要包含 30~55 nt 的区域,其中一个特有的茎环结构就会被保留,反之会导致该结构的解体(图 3a、b),值得注意的是 30~55 nt 正好是 SAFB1 结合区域,如图 1 所示. 其中 1~80 nt 包含前半部分茎环结构,81~123 nt 包含后半部分茎环结构,1~50 nt 由于缺少原有茎环结构的茎部分序列,因而二级结构发生了

较大的变化,如图 3a、3b 所示,1~60 nt 相对于 1~50 nt 保留了形成茎的序列,因此茎环结构与全长 RNA 的二级结构一致,最后 30~55 nt 仅仅保留了最小的一个茎环结构. 所有亚克隆结构预测如图 3a、b 所示.

3.4 *LINC-PINT-1* 30~55 nt 形成的二级结构是 SAFB1 识别的最小结构

根据上述的预测结果,我们发现所有针对 *LINC-PINT-1* 的亚克隆中,只要包含 30~55 nt 区域,*LINC-PINT-1* 与 SAFB1 结合区域的二级结构均会与全长的二级结构保持一致,反之则会破坏该二级结构,为了验证该二级结构对于 RNA 与

SAFB1 结合的重要性,我们将上述亚克隆分别体外转录并标记后通过凝胶迁移率实验验证与 SAFB1 的结合能力,结果发现只要包含 30~55 nt 区域就会有结合信号,并且仅有 30~55 nt 的转录本信号最强(图 4)。

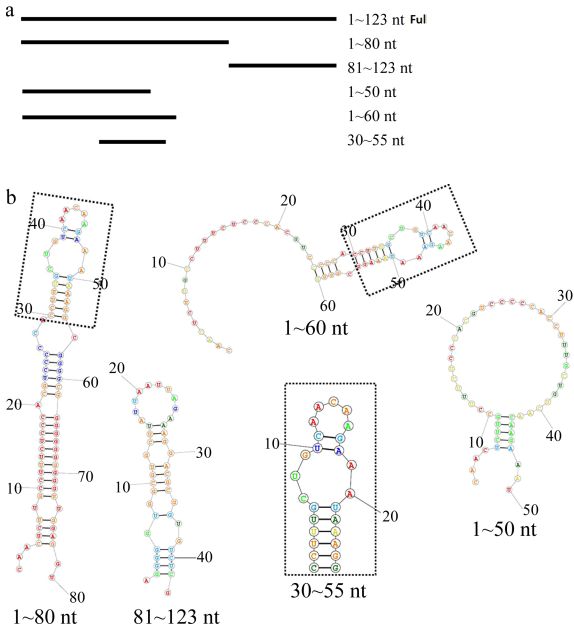


图 3 30~55 nt 区域是 *LINC-PINT-1* 维持特定二级结构的核心区域

a) *LINC-PINT-1* 亚克隆示意图; b) *LINC-PINT-1* 亚克隆二级结构预测结果其中 1~80 nt、1~60 nt、30~55 nt 中虚线标注的区域各个碱基相对位置一致保持不变,均能维持原有 LLS 结构。

Fig. 3 30~55 nt iscore region of the *LINC-PINT-1* secondary structure

4 讨论

生物体内 RNA 二级结构的形成均遵守 Watson-Crick, Hoogsteen, sugar-edge 三种模氢键配对模式,并且受到诸多因素的影响,如 RBPs(RNABinding Proteins)、分子伴侣(chaperone)、RNA polymerase 的延伸速度、离子浓度、温度条件等。在体外转录的 RNA 回复其自然状态下的二级结构时,虽然不能完全重现其细胞内的环境,但我们尽量使用与体内环境近似的缓冲环境:RSB(RNA Structure Buffer),从而避免不同的环境造成偏差较远的二级结构的出现。同时 RNA 二级结构的软件预测基本上均基于 MFE(minimizing free energy)原则。尽管算法上的优化以及广泛的运用,但是很多与 RNA 二级结构形成密切相关的因素被排除在外,例如蛋白质与 RNA 的相互作用、进化上序列的保守性以及 RNA 自身的动力学等,尽管在环境较为简单的体外环境中(*in vitro*),多种

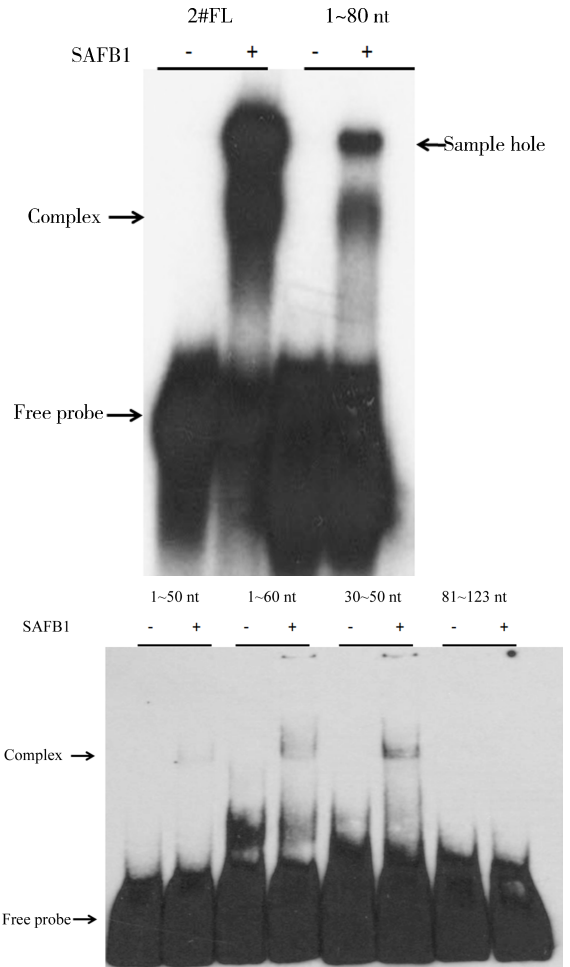


图 4 *LINC-PINT-1* 亚克隆 RNA 与 SAFB1 在体外相互作用

RNA 与 SAFB1 形成的复合体如图标注。

Fig. 4 *LINC-PINT-1* subclones interact with SAFB1 *in vitro*

As is shown in the figure, black arrows represent the free probe and complex of RNA and SAFB1.

盐离子强度的变化以及温度的波动也对导致软件预测结果的不可靠。对于造成这种不可靠的原因很可能由于软件预测过程中无法将与 RNA 相互作用的 SAFB1 作为一个二级结构折叠的辅助因素考虑进去,据已有的经验,RNA 结合蛋白对于 RNA 二级结构具有辅助折叠、约束或稳定二级结构等作用。

研究表明,RNA 序列和结构存在进化上的保守性,但现有的 RNA 二级结构预测软件仅有 Rfam 等少数在考虑 RNA 结构在进化上的保守性方面做得较好,但同其他软件预测类似,随着 RNA 序列的增长,软件预测的可靠性会显著降低。因此针对软件预测的结果需要实验证实。

通过依赖核酸酶的 RNA 结构分析(RNA structure analysis)可以推测 RNA 的实际结构,从

而弥补或者验证软件预测的结果. RNase A、T1、U2 等均能在单链 RNA 的特定位点进行切割, 虽然有众多切割单链 RNA 的 RNase, 但是切割双链 RNA 的 RNase 到目前为止只发现了一种: 来源于 *Naja oxiana* 毒液中的 RNase V1, 由于现在全球唯一供货商将其停产, 虽然没有办法通过 RNase 来确定双链的位置, 但本研究仍能采用解析 RNA 二级结构中的单链部分来推断 RNA 的整体二级结构.

由于 RNA 二级结构的形成基础为分子内碱基配对, 因此我们对 LINC-PINT-1 做了一系列的缺失, 目的是缺失维持二级结构必须的互补序列, 从而破坏已有的二级结构. 结果发现 2 # 克隆 RNA 缺失掉 50~60 nt 区域时不能形成原有的二级结构, 而亚克隆 30~55 nt、1~60 nt、1~80 nt 均能维持原有的二级结构, (图 3 a、b). 缺失后的 RNA 片段与 SAFB1 的电泳迁移率实验结果显示二级结构对于 RNA 与 SAFB1 相结合的重要性, 这个实验结果暗示序列与二级结构共同作用于 RNA 与 SAFB1 的识别并结合.

参考文献:

- [1] Yusupova G, Yusupov M. High-resolution structure of the eukaryotic 80S ribosome [J]. *Annu Rev Biomed*, 2014, 83: 467.
- [2] Madhani H D. snRNA catalysts in the Spliceosome's AncientCore [J]. *Cell*, 2013, 155: 1213.
- [3] Kiss T. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions [J]. *Cell*, 2002, 109: 145.
- [4] Carthew R W, Sontheimer E J. Origins and mechanisms of miRNAs and siRNAs [J]. *Cell*, 2009, 136: 642.
- [5] Gosai S J, Foley S W, Wang D, *et al.* Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the Arabidopsis nucleus [J]. *Mol Cell*, 2015, 57: 376.
- [6] Renz A, Fackelmayer F O. Purification and molecular cloning of the scaffold attachment factor B (SAF-B), a novel human nuclear protein that specifically binds to S/MAR-DNA [J]. *Nucleic Acids Res*, 1996, 24: 843.
- [7] Oesterreich S, Lee A V, Sullivan T M, *et al.* Novel nuclear matrix protein HET binds to and influences activity of the HSP27 promoter in human breast cancer cells [J]. *J Cell Biomed*, 2015, 67: 275.
- [8] Garee J, Meyer R, Oesterreich S. Sumoylation of corepressor SAFB1-mechanism of action and biological relevance in breast cancer [J]. *Cancer Res*, 2010, 69: S4142.
- [9] Jiang S, Meyer R, Britton O, *et al.* Analysis of SAFB1 and SAFB2 knockout mice reveals non-redundant functions of the two estrogen receptor Corepressor paralogs [J]. *Cancer Res*, 2010, 69: S2162.
- [10] Riva S, Biamonti G, Cobianchi F, *et al.* A novel hnRNP protein (HAP/SAF-B) enters a subset of hnRNP complexes and relocates in nuclear granules in response to heat shock [J]. *J Cell Sci*, 1999, 112: 1465.
- [11] Li J, Hawkins I C, Harvey C D, *et al.* Regulation of alternative splicing by SRrp86 and its interacting proteins [J]. *Mol Cell Biol*, 2003, 23: 7437.
- [12] Stoss O, Novoyatleva T, Gencheva M, *et al.* p59 (fyn)-mediated phosphorylation regulates the activity of the tissue-specific splicing factor rSLM-1 [J]. *Mol Cell Neurosci*, 2004, 27: 8.
- [13] Caroline R, Jalilah I, Helen S, *et al.* iCLIP identifies novel roles for SAFB1 in regulating RNA processing and neuronal function [J]. *BMC Biol*, 2015, 13: 111.
- [14] Hong E, Best A, Gautreg H, *et al.* Unravelling the RNA-binding properties of SAFB proteins in breast cancer cells [J]. *Biomed Res Int*, 2015 (10): 395816.
- [15] 王承恩, 罗玉, 贺正池, 等. 长非编码 RNA LINC00941 在结直肠癌中的表达及对细胞增殖的影响 [J]. *四川大学学报: 自然科学版*, 2017, 54: 1301.

引用本文格式:

中文: 李晓杰, 缪辉, 王迪, 等. 长非编码 RNA 与 SAFB1 相互作用依赖 RNA 二级结构 [J]. *四川大学学报: 自然科学版*, 2019, 56: 743.

英文: Li X J, Miao H, Wang D, *et al.* The interaction between lncRNA and SAFB1 depends on RNA secondary structure [J]. *J Sichuan Univ: Nat Sci Ed*, 2019, 56: 743.