

DOI: 10.13376/j.cblls/2021034

文章编号: 1004-0374(2021)03-0303-09



程红, 中国科学院分子细胞科学卓越创新中心研究组长、研究员、博士生导师。国家杰出青年科学基金、中国科学院“百人计划”和上海市“浦江人才计划”获得者。1998年毕业于中国医科大学获医学学士学位, 2003年毕业于日本神户大学获博士学位。2003年至2008年在美国哈佛大学医学院细胞生物系从事博士后研究。2008年回国, 担任中国科学院分子细胞科学卓越创新中心(原中国科学院生物化学与细胞生物学研究所)研究员。主要从事RNA转运与降解的机制和功能研究, 旨在揭示新生RNA出核转运或核内降解命运决定的机制、调控和生物学意义。重点研究方向包括: RNA的质量控制和核内降解及其调控; 细胞核亚结构调控RNA出核或降解的生理病理机制。自2008年组建独立实验室以来获得了一系列前沿性进展, 先后在*Mol Cell*、*Genes Dev*、*EMBO J*、*PNAS*、*JCB*、*NAR*、*Cell Death Dis*等国际著名学术期刊上发表多篇具有重要影响力的论文。近期发表的论文被F1000Prime推荐, 并被国际知名期刊*Nat Rev Genet*等正面引用和评述。

RNA结合蛋白研究技术

陈冠霖[#], 王怡旻[#], 范 静[#], 陈素丽[#], 王建姝[#], 王常收[#],
张 力, 徐 林, 朱 怡, 张 强, 童 登, 程 红*
(中国科学院分子细胞科学卓越创新中心, 上海 200031)

摘 要: RNA结合蛋白在细胞中与RNA互作, 广泛参与RNA的转录、加工、降解和出核转运等众多过程, 从而调控RNA的生物学功能。随着RNA功能研究的逐渐深入, 理解RNA与特定RNA结合蛋白的具体作用机制变得尤为重要, 由此推动了RNA结合蛋白与RNA互作研究技术的迅猛发展。该文对经典常用的RNA结合蛋白与RNA互作研究技术进行综述, 并根据其纯化中心(蛋白质或RNA)将其分成两大类, 主要介绍相应的技术原理、应用领域及研究进展, 并分析相关技术存在的问题, 为RNA相关领域的研究提供参考。

关键词: RNA; RNA结合蛋白; RNA-蛋白质互作组

中图分类号: Q591.4; Q946.5 **文献标志码:** A

Technologies of studying RNA binding proteins

CHEN Guan-Lin[#], WANG Yi-Min[#], FAN Jing[#], CHEN Su-Li[#], WANG Jian-Shu[#], WANG Chang-Shou[#],
ZHANG Li, XU Lin, ZHU Yi, ZHANG Qiang, TONG Deng, CHENG Hong*
(Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China)

收稿日期: 2021-02-01

基金项目: 国家重点研发计划(2017YFA0504400, 2017YFA0102700); 国家自然科学基金项目(31925008, 31770880, 31800686, 32071287, 32000913); 中国科学院战略性先导科技专项(XDB19000000); 中国博士后科学基金项目(2019TQ0337, 2020M671251); 中国科学院特别研究助理项目(29202000003)

[#]共同第一作者

*通信作者: E-mail: hcheng@sibcb.ac.cn

Abstract: RNA-binding proteins are widely involved in many biological processes by interacting with cellular RNA, such as RNA transcription, processing, degradation, and nuclear export, and thus play key roles in regulating the biological functions of targeted RNA. With the deepening understanding of RNA functions, it is particularly important to explain the mechanisms of the interaction between RNAs and RNA binding proteins, which promotes the rapid development of technologies about studying RNA binding protein-RNA interaction. This review summarizes the classical and commonly used technologies in RNA binding protein-RNA interaction research, which are divided into two major categories based on different purification centers, protein or RNA. We also introduce their technical principles, application fields and research progress, and analyze the shortages of related technologies. This review will provide important references for RNA-related researches.

Key words: RNA; RNA binding protein; RNA-protein interactome

RNA在生物体遗传信息表达的过程中扮演重要角色,其加工和成熟的过程往往与基因转录后调控相关。RNA结合蛋白(RNA binding proteins, RBPs)种类繁多、功能多样,通过与RNA结合参与各种生物学过程,例如在转录后水平调控RNA加工、定位、出核和蛋白质翻译等过程,并且与人类多种疾病密切相关^[1-2]。因此, RNA-蛋白质相互作用网络的鉴定和研究对理解细胞内复杂的生理调控系统具有重要意义。

RNA结合蛋白研究技术种类繁多,根据纯化中心的不同可以将其分成两大类:(1)通过纯化特定蛋白质来分析其结合的RNA底物的类型和特征;(2)通过纯化特定种类的RNA来分析与之结合的蛋白质的特性和功能。这两大类方法往往具有相同或相似的基本原理,但具体的技术细节又有很大的不同,导致它们在实际应用中各具优势,在具体的RNA-蛋白质互作研究中通常被综合运用,互为补充。本文将综合介绍几种最为经典和常用的RNA-蛋白质互作研究手段,并对一些前沿的分析技术进行总结。

1 以蛋白质为中心鉴定其结合RNA的分析技术

以蛋白质为中心鉴定其特异性结合的RNA底物是研究RNA与蛋白质相互作用最为重要的手段,其往往依赖于目标蛋白与对应抗体或配体的特异性结合,从而将与目标蛋白结合的RNA共沉淀下来进行分析。这类方法不仅可以鉴定不同蛋白特有的RNA结合底物,还可以分析同一个蛋白对不同类型RNA结合能力的强弱变化。

目前已有多种不同的方法可以在体内鉴定与目标蛋白相互作用的RNA,其中应用最为广泛的是RNA结合蛋白免疫沉淀(RNA-binding-protein immunoprecipitation, RIP)和紫外交联免疫沉淀

(ultraviolet crosslinking and immunoprecipitation, CLIP);由于研究需要,更多改进的CLIP技术也逐渐被发明应用,将蛋白质与RNA的相互作用研究推至特定序列,乃至单核苷酸的层面。

为了在体外研究蛋白质与RNA的相互作用,研究人员通常依靠凝胶迁移(electrophoretic mobility shift assay, EMSA)、硝酸纤维素膜过滤、等温滴定量热法等技术对蛋白质-RNA的特异性结合进行定量分析^[3-5]。但由于蛋白质与RNA相互作用方式和底物种类的多样性,这些技术通常存在通量低、操作复杂等缺点。为了克服这些局限性,进一步拓宽蛋白质与RNA互作的研究领域,近年来几种高通量体外分析蛋白质-RNA结合能力的方法被开发出来^[6-7],以更高的准确率及通量被用于研究更为复杂的生物反应过程。

1.1 RIP

RNA结合蛋白免疫沉淀(RIP)是一种基于抗原-抗体杂交原理的技术手段,能够对细胞内天然条件下单个蛋白质和RNA分子间的物理结合情况进行研究,是了解转录后调控网络动态过程的有力工具。RIP技术利用靶向目标蛋白的抗体免疫沉淀相应的RNA-蛋白质复合物,后续经过分离纯化就可以对复合物中的RNA进行分析。如果没有目标蛋白相应的抗体,也可以在目标蛋白上融合标签蛋白,从而改用标签蛋白的抗体即可^[8]。RIP技术下游通常结合定量PCR(RIP-qPCR)、基因芯片技术(RIP-Chip)或高通量测序(RIP-seq)等方法定性定量分析RNA,从而帮助我们从整体水平上了解RNA结合底物的类型和特征^[9]。

纯化相应的RNA-蛋白质复合物是RIP中最为关键的技术步骤,很大程度上决定了实验的成败。为了获得完整的RNA-蛋白质复合物,在RIP操作过程中应尽量避免RNA的降解。免疫沉淀磁珠的选择同

样非常重要, 应取决于抗体的种属及亚型。进行有效的超声处理也是获得高质量免疫沉淀的关键一步, 对免疫沉淀的效率有很大的影响。阴性对照的选择对于后续判断RNA与蛋白质相互作用的强弱及可靠性也具有重要的参考作用^[8]。

RIP由于流程简单、易于操作等优点, 成为鉴定体内RNA-蛋白质相互作用最常用的方法, 为RNA结合蛋白的功能研究提供了思路。例如, 研究人员利用RIP技术验证了mRNA出核因子ALYREF与编码组蛋白的mRNA之间的相互作用, 促进了组蛋白mRNA的3'端加工过程^[10]。RIP虽然具有全基因组检测的能力, 但在实际应用中具有一定的局限性, 如RIP虽然可以鉴定RNA与蛋白质之间是否存在相互作用, 但是无法确定这种相互作用是否直接, 也难以精确识别蛋白质在RNA上的结合位点。此外, RIP特异性和分辨率欠佳, 由于在细胞裂解液中RNA-蛋白质复合物可能会形成非生理性的结合, 很容易检测到非特异性的相互作用, 不能完全准确地重现体内生理条件下发生的结合。因此, 通常需要使用其他方法进行后续验证。但随着对裂解条件和洗涤条件的优化, 可以在一定程度上得到更接近于天然水平的信息^[11]。

1.2 CLIP及其衍生方法

鉴于RIP识别特异性差、分辨率低, 并且容易识别非生理或非直接相互作用, 其应用范围十分有限。为此, 研究人员通过对其进行改良, 开发出了多种鉴定蛋白质-RNA相互作用的新方法与技术, 其中最为重要的是紫外交联免疫沉淀(CLIP)技术, 它的发展使蛋白质与RNA相互作用的研究具有较高的位置分辨率和特异性^[12]。

CLIP主要利用了蛋白质和RNA在紫外线照射下形成共价键的性质, 以捕捉蛋白质与RNA在体内的真实结合模式。通过紫外交联后, RNA与相互作用的蛋白质形成复合物, 在核酸酶消化部分游离RNA片段后, 可以利用特定的抗体将目的蛋白-RNA复合物通过免疫沉淀进行富集、纯化, 之后将纯化后的目的蛋白-RNA复合物经过解交联和RNA抽提等步骤分离目的蛋白和RNA, 最后将RNA通过逆转录合成cDNA文库后进行高通量测序, 从而在全基因组范围内研究蛋白质和RNA的相互作用网络。利用CLIP技术, 研究人员首次揭示了小鼠大脑中Nova蛋白调节神经元pri-mRNA剪接的RNA结合图谱^[13]。

与RIP技术相比, CLIP技术增加了紫外交联和

RNA酶消化步骤, 不仅保证了目标蛋白和RNA之间的结合更加紧密, 而且使得结合位点的检测更加方便灵敏。但是, 在实际应用中CLIP仍然存在一些局限性。比如, 由于该方法以紫外交联为基础, 这使得交联效率往往不高, 且需要分离交联RNA和非交联的RNA, 使得该方法对于原材料的量要求比较大。其次, 该方法步骤繁多, 对于操作人员的技术要求较高。此外, 若目的蛋白上存在多种RNA的结合, CLIP并不能识别各结合位点的比例, 因此CLIP技术的分辨率仍然有限。

为了解决CLIP技术分辨率不足、紫外交联效率低的问题, 光活性增强的CLIP (photoactivatable-ribonucleoside-enhanced CLIP, PAR-CLIP)技术应运而生。该技术核心在于利用具有光敏活性的核苷类似物代替体内正常的核苷掺入到新产生的转录本中, 如利用4-sU或6-sG等代替原本的U或G, 从而提高掺入位点的紫外交联效率^[14]。另一方面, 由于交联位点对光敏核苷具有选择偏好性, 4-sU、6-sG的加入会导致后续逆转录步骤中碱基对的变化, 进而提高了对特异性结合位点的识别和检测能力, 也通过降低背景噪音获得了更高的分辨率。利用该技术, 研究人员阐明了ASD相关的RNA结合蛋白FMRP在体内结合多种mRNA进而调控对应蛋白的表达^[15], 还发现DNA结合蛋白CTCE在体内结合RNA从而调控p53通路等^[16]。PAR-CLIP的使用局限性主要在于核苷类似物的掺入会产生细胞毒性, 亦难以大规模应用在动物水平, 且由于PCR带来的扩增偏好性仍然存在, 导致分辨率的提高依然有限。

为了进一步提高CLIP技术的分辨率, 研究人员开发了单核苷酸分辨率CLIP (individual nucleotide resolution CLIP, iCLIP)技术。iCLIP技术利用交联位置能够抑制逆转录进行的特点, 定位cDNA截断位置上游的第一个氨基酸位置即是交联位点, 从而可以在单核苷酸水平上精准定位蛋白质和RNA的结合位点。由iCLIP产生的cDNA往往不具有5'端接头, 因此需要通过额外的环化和线性化步骤捕获截断的cDNA进行测序^[17]。iCLIP技术主要用于精确分析RNA结合蛋白的识别位点, 应用十分广泛, 在可变剪接、可变加尾、RNA甲基化、mRNA出核转运等多种基因表达过程的功能研究中均有涉及, 例如: 研究人员利用iCLIP方法在基因组范围内鉴定了重要出核因子ALYREF的结合位点, 发现其主要结合在成熟mRNA的5'和3'区域^[18]。

iCLIP技术的产生确实极大地提高了RNA与蛋

白质结合位点的分辨率,而且随机标签的加入消除了PCR扩增所带来的偏好性,但其不足之处在于反应步骤比较繁琐,环化过程也带来了效率问题,导致建库成功率低,因此迫切需要一种新的技术能够在提高分辨率和简化反应流程中取得一个平衡,由此催生了优化的CLIP (enhanced CLIP, eCLIP)技术。eCLIP技术的分辨率虽不及iCLIP,但也比普通的CLIP提高了数倍,操作过程与iCLIP相比也简便了许多。该方法与iCLIP的不同之处在于省略了其中的环化步骤,并分两步为RNA加上接头,减少了由于接头连接效率不高导致的RNA的损失,提高了建库的效率,显著降低了后续的测序成本。利用eCLIP方法,科学家们揭示了eIF4A3在转录后基因调控中的作用^[19]。由于其位点分辨率比CLIP更高,操作流程比iCLIP更简便,eCLIP逐渐成为了研究RNA与蛋白质相互作用的重要工具,被广泛应用于各类RNA结合蛋白研究。除了上述几种重要的体内鉴定RNA结合蛋白与其特异性RNA底物的方法外,研究人员还开发出了多种在体外鉴定RNA结合蛋白底物的方法。

1.3 EMSA

凝胶迁移(EMSA)是体外检测RNA与特定目标蛋白结合的有效方法。EMSA基于在电场均匀的凝胶电泳中,核酸-蛋白质复合物比游离探针迁移速率慢的原理,通过比较信号在聚丙烯酰胺凝胶上的位置,对蛋白质与RNA之间的相互作用进行定性或定量分析。

EMSA检测需要将细胞裂解液或蛋白质的纯化样品与放射性同位素³²P标记的RNA探针一起孵育,以形成RNA-蛋白质复合物,并加入肝素排除蛋白质与探针的非特异结合,随后将RNA-蛋白质复合物在非变性聚丙烯酰胺凝胶上进行电泳。电泳完成后,干燥凝胶,用放射自显影法指示RNA-蛋白质复合物和游离探针的位置。游离探针迁移速度快,而RNA-蛋白质复合物的迁移速度较慢,因此该过程也被称为“凝胶阻滞”或“带移”。由于EMSA简单、灵敏的特点,近年来被广泛应用于DNA/RNA结合蛋白的鉴定、RNA结合蛋白与特定RNA序列的相互作用检测等研究中。例如,Chi等^[20]通过EMSA实验证明出核因子ZC3H18能够特异性地与乙肝病毒mRNA上的顺式原件SEP1结合,进而揭示了SEP1促进无内含子mRNA出核的作用机制。

传统的EMSA检测使用的是放射性同位素标记的探针,不可避免地需要更严格的实验条件以规避

对人体或环境造成的伤害与污染。近年来,多种安全有效的荧光染料或荧光素已逐步取代了放射性标记探针作为新的检测系统。相较于传统系统而言,这些新型的荧光染料或荧光素更加安全与精确,并且产生的背景噪音也更少。除此以外,为了标准化实验流程,各类EMSA商用试剂盒也逐步被开发,如化学发光EMSA或SYBR绿色核酸&红色蛋白质染色EMSA等,与传统EMSA试剂盒相比具有更加简便、快速的优势。

1.4 MITOMI

Martin等^[7]于2012年开发出了基于机械诱导的分子互作捕获(mechanically induced trapping of molecular interactions, MITOMI)技术,用于体外高通量分析RNA与蛋白质的相互作用。该技术基于微流体系统平台,利用配体亲和原理,能对RNA文库进行整体功能分析。由于其更高的检测通量以及精准性,除了用于研究蛋白质与RNA的相互作用之外,该技术还可用于测量转录因子的结合亲和力、并行获取数百种蛋白质-配体相互作用的亲和常数和动力学速率、绘制蛋白质相互作用网络图、鉴定药理学抑制剂并进行高通量低成本分子诊断等^[21-22]。

MITOMI平台结合了两种不同的技术,即微阵列和微流体检测。MITOMI微阵列由一个平面基板(通常是涂有环氧的载玻片)组成,其上具有由标准DNA微阵列机器人印制的放置微量生物溶液(约1~0 nL)的微腔。MITOMI微阵列可包含成千上万个生物溶液斑点,包括DNA^[23-25]、RNA^[7]、小分子、细胞培养基^[26]、人血清、抗体和活细胞^[27]等。微流体检测器件由两层组成:与基板接触并包含微流体通道网络以执行生物学测定的流动层,以及控制流体流动的控制层。

MITOMI技术首先需要以DNA寡核苷酸作为RNA文库的转录模板,通过生物素和荧光素标记的polyT单链DNA探针对转录完成的RNA进行捕获,从而使RNA固定在微腔中;然后,将目标蛋白与偶联德克萨斯红染料的抗体一起孵育,使孵育物流过捕获的RNA库;最后,使用微阵列扫描仪对每个微腔中与RNA结合的蛋白质进行量化分析。Martin等^[7]利用MITOMI技术研究了RNA突变体文库与茎环结合蛋白SLBP的相互作用,精确鉴定了一系列与该蛋白具有高亲和力的RNA的序列和结构特征。

由于MITOMI技术需要使用DNA探针对RNA进行捕获,不可避免地引入了技术误差。Tome等^[28]于2014年开发了HiTS-RAP (high-throughput sequencing

RNA affinity profiling)技术。HiTs-RAP除了具有高通量的优势外, 还通过大肠杆菌复制终止蛋白对RNA进行原位捕获, 不需要引入DNA探针, 进一步提高了分析的精度与准确度。除此以外, 近年来许多以蛋白质为中心的新兴分析技术也被开发出来, 如基于RNA编辑原理的TRIBE (targets of RNA-binding proteins identified by editing)技术^[29]、能有效排除蛋白质复合物干扰的cDNAs交联和分析(crosslinking and analysis of cDNAs, CRAC)技术^[30]、基于SpyTag的CLIP (SpyTag-based CLIP, SpyCLIP)技术^[31]等等。这些不断被开发的分析技术为科研人员提供了有力的研究工具, 推动了RNA-蛋白质互作调控网络研究的飞速发展。

2 以RNA为中心鉴定其结合蛋白质的分析技术

RNA在多种生理和病理条件下发挥重要的功能, RNA结合蛋白对于RNA在细胞内发挥生物学功能必不可少^[32-33]。因此, 鉴定RNA-蛋白质复合物中的蛋白质种类对理解RNA的生理功能极为重要。以RNA为中心鉴定RNA结合蛋白的技术旨在以感兴趣的RNA为研究对象, 鉴定与其互作的蛋白质。现有的技术大多是利用体外合成的RNA或寡核苷酸探针作为诱饵来捕获与RNA互作的蛋白质, 这些RNA或寡核苷酸探针通过掺入生物素、荧光染料等修饰的核糖核苷酸进行化学标记, 或者掺入天然或人工合成的寡核苷酸配体(如S1、D8、MS2发夹结构等)加以识别。标记好的RNA或探针可以进一步用于在体内或体外鉴定与RNA互作的蛋白质^[34], 并通过质谱对这些蛋白质加以分析。

在体内鉴定RNA结合蛋白的技术主要包括MS2-BioTRAP (MS2 *in vivo* biotin tagged RNA affinity purification)、ChIRP (chromatin isolation by RNA purification)、CHART (capture hybridization analysis of RNA targets)、RAP-MS (RNA antisense purification)和IC (interactome capture)等。以上这些技术不依赖于RNA的二级结构, 所得结果能很好地反应体内RNA与蛋白质的相互作用。

体外的互作实验可以作为体内鉴定的重要补充手段, 为验证RNA与蛋白质的相互作用提供更直接的证据。其中应用最为广泛的包括RNA沉降(RNA pull-down)、蛋白质微阵列技术等。但体内外技术各有优缺点, 需要根据实验的实际需求进行选择^[35-40]。

2.1 MS2-BioTRAP

MS2-BioTRAP是Tsai等^[35]基于MS2和生物素亲

和纯化技术开发的一种捕获体内RNA-蛋白质复合物的方法。它的核心原理是MS2噬菌体来源的MCP蛋白(MS2 coat protein)能高亲和性地结合特定RNA茎环结构上的单链环区域(MCP binding sites, MBS)^[41]。该方法首先需要构建一个能稳定表达MCP-HB融合蛋白的细胞系, HB标签(histidine and biotin tag)包含两个6×His序列、一个烟草蚀纹病毒(tobacco etch virus, TEV)切割位点以及一段内源生物素化信号序列^[42], 其中的生物素化信号序列能够被内源生物素转移酶催化带上生物素标记, 便于后续亲和纯化。此外, 还需要构建一个在靶RNA末端带有串联重复的MBS茎环序列的表达质粒, 将其转染进上述稳转细胞系中。MCP与MBS茎环序列之间的紧密结合使得基于HB标签的亲和纯化成为可能, 只需通过链霉亲和素亲和纯化就能得到与靶RNA上特异结合的蛋白质。最后, 再利用质谱、蛋白质免疫印迹等方法对所得蛋白质进行后续分析^[35]。

MS2-BioTRAP能够鉴定与某一感兴趣RNA特异结合的蛋白质, 利用此方法可以分析在RNA转录、加工、定位、翻译及降解相关过程中与特定RNA相互作用的蛋白质因子。比如, 有研究利用该方法鉴定了能与淋巴组织增强因子-1 (lymphoid enhancer factor-1, LEF1)内部核糖体进入位点(IRES)结合的蛋白质, 该研究拓展了我们对IRES介导翻译的理解, 同时也验证了该方法鉴定RNA结合蛋白的可行性^[35]。

MS2-BioTRAP的一大优势在于能够反映RNA和蛋白质在体内的真实结合情况, 避免了体外非特异性的结合。此外, 该方法可以应用于任意一条感兴趣的RNA, 并且和其他方法相比更加简单快速, 只需将MBS茎环序列添加到目的序列上即可, 不需要针对每一独立的RNA去单独设计纯化探针。但是, 由于目的RNA及MCP都是在细胞内过量表达, 因此会造成假阳性率高的问题, 并且当大量MCP结合到RNA上后可能会影响RNA二级结构形成、相关蛋白质与RNA的亲合性以及RNA降解等过程^[43]。

2.2 ChIRP、CHART和RAP-MS

MS2-BioTRAP技术的出现解决了RNA-蛋白质体外组装导致的假阳性问题, 但此方法步骤繁多, 并且由于目的RNA及MCP过量表达也不可避免地会产生一些假阳性的结果。近年来, 研究人员开发出了一系列基于寡核苷酸探针纯化鉴定RNA结合蛋白的技术。ChIRP、CHART和RAP-MS是其中三项重要的技术, 越来越多地被用于在体内鉴定与RNA互

作的蛋白质。这三项技术基于相同的理念,即将与目的RNA互补的反义DNA寡核苷酸制成生物素标记的探针,然后将其与细胞裂解液共同孵育,之后再利用链霉亲和素偶联的磁珠分离得到DNA-RNA-蛋白质复合物。通过二代测序或质谱分析,鉴定与RNA互作的蛋白质,并确定发生这些互作的基因组区域^[37-40]。虽然这三项技术都是使用反义DNA寡核苷酸从交联细胞中分离获取特异的DNA-RNA-蛋白质复合物,但它们在探针设计和交联方法上有所不同。对于ChIRP和RAP-MS这两种方法,使用的是均匀覆盖整个目的RNA所有可能的杂交位点上的探针,确保目的RNA的所有片段都能被探针捕获到^[37-38,40]。而CHART基于RNase H消化DNA-RNA杂合双链中RNA的原理,预先筛选目的RNA上的可能的DNA结合区域,并以此为依据设计探针,降低了探针的使用量从而提高信噪比^[39]。在交联方法上,ChIRP和CHART使用甲醛或戊二醛,在交联RNA-蛋白质的同时也会导致蛋白质-DNA和蛋白质-蛋白质的交联;而RAP-MS利用紫外交联,只在具有直接相互作用的RNA和蛋白质之间建立共价键。因此,RAP-MS可以特异性地富集与RNA存在直接相互作用的蛋白质^[37-40]。

上述三种方法解决了当今RNA生物学中的关键需求,研究者利用这些技术鉴定了多种lncRNA(long noncoding RNA)的互作蛋白。利用ChIRP技术,Chu等^[44]在体内鉴定了81个与Xist RNA互作的蛋白质。2020年,Ninomiya等^[45]利用ChIRP技术发现在nSBs(nuclear stress bodies)中富集了包括SRSFs在内的多种剪接因子,进一步揭示了nSBs在基因表达调控中的重要作用。Gandhi等^[46]利用RAP技术发现了YBX1蛋白与lincNMR具有直接相互作用,并进一步揭示lincNMR通过YBX1-RRM2-TYMS-TK1通路调控核苷酸代谢,从而调控肿瘤细胞增殖的作用机制。

尽管这些技术能够在体内高度特异地鉴定目的RNA及其互作蛋白,并能确定发生互作的基因组区域,但目前这些技术只适用于表达丰度较高的RNA。如何提高这些技术的效率,从而应用于低表达的RNA,是亟待解决的问题。

2.3 RNA相互作用组捕获(RIC)及进阶捕获方法(serIC和eRIC)

MS2-BioTRAP、ChIRP、CHART和RAP-MS都可用于鉴定特定RNA的互作蛋白,这些基于RNA的特性开发出来的鉴定某一类RNA的互作蛋白的

技术一直备受领域内研究人员的关注。近几十年,材料技术的日益革新极大地推动了RNA研究领域的发展^[47],研究人员开发出了一系列精确筛选RNA与蛋白质相互作用的高分辨率、高准确性的工具和方法。

RNA相互作用组捕获(RNA interactome capture, RIC)作为最早的体内鉴别RNA-蛋白质互作组的技术,通过紫外照射活细胞交联RNA-蛋白质互作组。裂解细胞后,使用oligo dT磁珠富集含polyA尾的mRNA-蛋白质组,经过非变性条件漂洗后,最终洗脱的蛋白质经质谱分析被一一鉴别。在此技术的基础上,衍生出两种新型的RIC方法,即serIC(serial RNA interactome capture)^[48]和eRIC(enhanced RNA interactome capture)^[49]。serIC引入了RNA结合蛋白重复严格的纯化过程和DNA酶消化过程,从而获得高度特异性的RNA-蛋白质互作组。eRIC通过利用锁核酸(locked nucleic acid, LNA)修饰探针,提升RNA的识别能力和亲和力,高特异性富集RNA结合蛋白组。与传统的RIC相比,serIC和eRIC鉴定的RNA-蛋白质互作组范围更小,显著降低了非特异性或不相关蛋白组分的比例,因此鉴定的RNA-蛋白质互作组结果更加可靠。serIC除能够鉴定RNA-蛋白质互作组以外,同时也能鉴别DNA-RNA-蛋白质互作组。

Ørom研究组利用serIC鉴定出了一些参与DNA修复过程的新型RNA-蛋白质互作组分,例如新鉴定的RNA互作蛋白BCLAF1通过与磷酸化BRCA1蛋白互作,偶联DNA修复和可变剪接过程^[48]。Hentze实验室通过eRIC技术鉴别出了61种与甲基化RNA互作相关的蛋白质,其中70%为新的甲基化RNA互作蛋白^[49]。

2.4 RNA沉降(RNA pull-down)

用内源方法鉴定感兴趣的RNA结合蛋白可以较为真实地反映生物体内的互作情况,但是由于生物体内环境的复杂,内源的纯化分析方法得到的结果很难判断是来源于直接还是间接相互作用。因此,通常需要体外的互作实验来验证RNA与蛋白质直接的相互作用。其中,RNA沉降(RNA pull-down)^[50]是迄今为止最常用的体外验证方法。

RNA沉降首先利用线性化的质粒或者PCR产物,通过体外转录得到感兴趣的RNA,再将得到的RNA交联到固体基质,如琼脂糖微球(agarose beads)或磁珠(magnetic beads)上,并将其与全细胞裂解液共孵育,之后通过变性洗脱,便可将与目标

RNA有相互作用的蛋白质富集下来, 然后利用层析、质谱等方法对组分进行分离鉴定。该方法的关键点在于如何高效地将RNA与固体基质结合以及如何特异性地将RBPs洗脱下来。目前已有多种手段实现了目标RNA与固体基质的有效交联。例如, 将高碘酸氧化的双链RNA交联到酰肼修饰的固体基质上可以提高其结合蛋白的特异性^[51]。与此同时, 生物素由于分子量小, 且能与链霉亲和素可逆地结合, 也已经成为一种在RNA与固体基质交联过程中被广泛应用的修饰基团^[52]。

但是, 另一方面, 如何特异性地将RBPs洗脱下来这一问题却一直没有被很好地解决。传统方法利用盐浓度梯度对目标蛋白进行洗脱, 通过不同的盐浓度实现洗杂和洗脱的目的。这个过程不可避免地会导致信噪比低的问题, 使得鉴定潜在的RNA互作蛋白变得困难。造成此问题的原因主要是某些蛋白质的表达量相对较低或RNA结合能力较弱, 难以对其进行纯化检测; 同时, 非特异性蛋白会结合在固体基质上, 在洗脱时被一并洗脱下来, 造成对目标蛋白信号采集的干扰。因此, 根据实验需求采用合适的洗脱方法变得非常重要。

为了克服此困难, Cáceres实验室在RNA沉降的基础上开发了利用RNase而非基于盐浓度梯度洗脱RBPs的新方法^[53]。他们利用RNase A/T1处理消化固体基质上的探针RNA, 然后对洗脱后的蛋白质进行了分析, 鉴定了与pri-let-7a-1和pri-miR-101-1这两种不同miRNA相互作用的蛋白质组。这两种miRNA分别与hnRNPL和PTB有相互作用, 该结果与此前利用盐浓度梯度洗脱的结果一致^[54], 而且由于背景信号的显著降低, 他们得以鉴定到几种与pri-let-7a-1和pri-miR-101-1结合的新蛋白^[53]。

这种基于RNase洗脱的方法得到了广泛使用, 如Alam等^[55]利用该方法发现G3BPs通过结合IFITM1、IFITM2及IFITM3的转录本, 影响它们的翻译过程, 进而调控癌细胞的发生与迁移过程; Frye实验室则借此方法阐明SRSF2通过结合非5mC甲基化形态的VTRNA1.1转录本, 抑制其加工, 从而影响细胞分化^[56]。

RNA沉降的方法在应用过程中需要注意以下问题: 首先, 由于RNA二级结构对其功能的发挥至关重要, 因此特定的标签需要优先考虑加在RNA的末端。其次, 在固体基质与全细胞裂解液孵育过程中, 必须要保证RNA的稳定性, 因此构建一个无核酸酶的环境对实验的成功非常重要; 另外, 体外纯

化的RNA在结构上与内源的RNA可能会有所差别, 且体外细胞裂解液的环境与细胞内的真实环境也会有差别, 因此实验结果可能无法准确地反映体内的真实情况。

2.5 蛋白质微阵列

由于RNA沉降的灵敏程度有限, 可能无法有效地检测到某些丰度较低的蛋白质, 同时其通量太低导致在大规模筛选实验中具有一定的局限性, 因此Siprashvili等^[57]开发了一种基于蛋白质微阵列技术(protein microarray)的方法, 用于鉴定特定目标RNA的结合蛋白。蛋白质微阵列技术主要利用一个包括有大约9 400个人类重组蛋白的微阵列, 对目标RNA的结合蛋白进行快速高通量地鉴定。简单来说, 首先将目标RNA转录出来后制备成荧光探针, 再将荧光探针与蛋白质微阵列进行孵育, 洗脱后检测微阵列孔内的荧光情况, 以判断孔内蛋白质与目标RNA间是否存在相互作用。通过将目标RNA带上不同颜色的荧光基团, 可同时对几种感兴趣的RNA进行蛋白质互作组鉴定。蛋白质微阵列技术除了具有通量高、操作简便和出结果快等优点外, 其适用范围也非常广泛, 可以覆盖任何体外转录出来的编码或非编码RNA。Siprashvili等^[57]利用蛋白质微阵列技术鉴定了几种编码和非编码RNA的结合蛋白。后续, 他们还利用此项技术发现非编码RNA SNORD50A和SNORD50B结合促癌蛋白K-Ras^[58]。

蛋白质微阵列方法同样具有一些不可避免的缺点。一方面, 该方法所能鉴定到的蛋白质完全受限于蛋白质微阵列中包含的蛋白质种类, 而且微阵列上的蛋白质可能缺乏体内甲基化、磷酸化等多种蛋白质修饰, 使得蛋白质的高级结构难以重现, 从而降低了结果的可靠性; 另一方面, 由于体内的RNA往往具有高级结构, 众多蛋白质-RNA的相互作用都依赖于RNA的高级结构, 而体外转录的RNA可能不会形成体内条件下的高级结构; 除此之外, 体外转录的RNA偶联的荧光基团如果过少, 可能导致阳性信号不易被检测到, 而偶联过多的荧光基团则可能破坏RNA本身的结构, 所以需要综合考虑RNA上偶联的荧光基团数量和间距。

由于mRNA具有polyA尾的特殊结构, 我们可以比较容易地对其进行富集纯化和蛋白质互作组分析, 但非polyA尾RNA的蛋白质互作组鉴定仍缺少有效手段。随着技术的更新换代和RNA-蛋白质互作网络的不断完善, 将会有更多的RNA结合蛋白被挖掘出来, 一些早期鉴定的经典RNA结合蛋白的功

能也可能被重新定义。

越来越多的研究表明,许多RNA发挥顺式调控作用,这种作用的发挥多数情况下依赖其互作蛋白。因此,以RNA为中心鉴定RNA结合蛋白,揭示RNA-蛋白质互作网络是目前的研究热点之一。揭示RNA发挥作用的相关分子机制,需要RNA-蛋白质互作鉴定技术的支持,这种需求必然会促进RNA-蛋白质相互作用预测和实验鉴定技术的进步。同样,相关技术的进一步突破和完善,必将推动RNA研究领域的发展。

3 总结与展望

以往, RNA与蛋白质互作系统的研究很大程度上是受制于技术手段的匮乏,难以进行纯化分析,但随着技术手段的发展与革新, RNA与蛋白质互作研究已进入了百花齐放的阶段,不断有各种新的技术被发展应用到具体研究中。随着研究的进一步深入,这些技术手段不仅仅被应用在简单的RNA与蛋白质结合分析中,甚至被用于细胞内复杂RNA-蛋白质机器及无膜结构的纯化与分析中。相信随着科学技术的进步, RNA与蛋白质互作系统的研究将逐步进入到更为深入和精细的层面,这将有助于在更深层次上理解RNA与蛋白质在细胞活动中所担任的角色,解释其中涉及的重要分子机制及其生物学意义。

[参 考 文 献]

- [1] Wang F, Tidei JJ, Polich ED, et al. Positive feedback between RNA-binding protein HuD and transcription factor SATB1 promotes neurogenesis. *Proc Natl Acad Sci USA*, 2015, 112: E4995-5004
- [2] Klattenhoff CA, Scheuermann JC, Surface LE, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell*, 2013, 152: 570-83
- [3] Ryder JJ, Garrison K, Song F, et al. Genetic associations in peripheral joint osteoarthritis and spinal degenerative disease: a systematic review. *Ann Rheum Dis*, 2008, 67: 584-91
- [4] Salim NN, Feig AL. Isothermal titration calorimetry of RNA. *Methods*, 2009, 47: 198-205
- [5] Wong I, Lohman TM. A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc Natl Acad Sci USA*, 1993, 90: 5428-32
- [6] Campbell ZT, Bhimsaria D, Valley CT, et al. Cooperativity in RNA-protein interactions: global analysis of RNA binding specificity. *Cell Rep*, 2012, 1: 570-81
- [7] Martin L, Meier M, Lyons SM, et al. Systematic reconstruction of RNA functional motifs with high-throughput microfluidics. *Nat Methods*, 2012, 9: 1192-4
- [8] Gagliardi M, Matarazzo MR. RIP: RNA immunoprecipitation. *Methods Mol Biol*, 2016, 1480: 73-86
- [9] 熊冰钰, 叶珂璠, 崔凌云, 等. 蛋白质-RNA相互作用鉴定技术研究进展. *生物技术进展*, 2020, 10: 217-25
- [10] Fan J, Wang K, Du X, et al. ALYREF links 3'-end processing to nuclear export of non-polyadenylated mRNAs. *EMBO J*, 2019, 38: e99910
- [11] Popova VV, Kurshakova MM, Kopytova DV. Methods to study the RNA-protein interactions. *Mol Biol (Mosk)*, 2015, 49: 472-81
- [12] 邢伶俐, 张彦, 黄静, 等. CLIP相关技术的研究进展. *军事医学*, 2013, 37: 867-9
- [13] Ule J, Jensen KB, Ruggiu M, et al. CLIP identifies Nova-regulated RNA networks in the brain. *Science*, 2003, 302: 1212-5
- [14] 陈伟, 许馨, 孙绍光. RNA结合蛋白与RNA相互作用鉴定技术. *中国生物化学与分子生物学报*, 2017, 33: 103-7
- [15] Ascano M Jr, Mukherjee N, Bandaru P, et al. FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature*, 2012, 492: 382-6
- [16] Saldana-Meyer R, Gonzalez-Buendia E, Guerrero G, et al. CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev*, 2014, 28: 723-34
- [17] Konig J, Zarnack K, Rot G, et al. iCLIP -- transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. *J Vis Exp*, 2011, 50: 2638
- [18] Shi M, Zhang H, Wu X, et al. ALYREF mainly binds to the 5' and the 3' regions of the mRNA *in vivo*. *Nucleic Acids Res*, 2017, 45: 9640-53
- [19] Hocq R, Paternina J, Alasseur Q, et al. Monitored eCLIP: high accuracy mapping of RNA-protein interactions. *Nucleic Acids Res*, 2018, 46: 11553-65
- [20] Chi B, Wang K, Du Y, et al. A Sub-Element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. *Nucleic Acids Res*, 2014, 42: 7305-18
- [21] Garcia-Cordero JL, Maerkl SJ. Mechanically induced trapping of molecular interactions and its applications. *J Lab Autom*, 2016, 21: 356-67
- [22] Karr JR, Sanghvi JC, Macklin DN, et al. A whole-cell computational model predicts phenotype from genotype. *Cell*, 2012, 150: 389-401
- [23] Geertz M, Shore D, Maerkl SJ. Massively parallel measurements of molecular interaction kinetics on a microfluidic platform. *Proc Natl Acad Sci USA*, 2012, 109: 16540-5
- [24] Maerkl SJ, Quake SR. A systems approach to measuring the binding energy landscapes of transcription factors. *Science*, 2007, 315: 233-7
- [25] Rockel S, Geertz M, Hens K, et al. iSLIM: a comprehensive approach to mapping and characterizing gene regulatory networks. *Nucleic Acids Res*, 2013, 41: e52
- [26] Garcia-Cordero JL, Nembrini C, Stano A, et al. A high-throughput nanoimmunoassay chip applied to large-scale vaccine adjuvant screening. *Integr Biol (Camb)*, 2013, 5:

- 650-8
- [27] Denervaud N, Becker J, Delgado-Gonzalo R, et al. A chemostat array enables the spatio-temporal analysis of the yeast proteome. *Proc Natl Acad Sci USA*, 2013, 110: 15842-7
- [28] Tome JM, Ozer A, Pagano JM, et al. Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat Methods*, 2014, 11: 683-8
- [29] McMahon AC, Rahman R, Jin H, et al. TRIBE: hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins. *Cell*, 2016, 165: 742-53
- [30] Granneman S, Kudla G, Petfalski E, et al. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci USA*, 2009, 106: 9613-8
- [31] Zhao Y, Zhang Y, Teng Y, et al. SpyCLIP: an easy-to-use and high-throughput compatible CLIP platform for the characterization of protein-RNA interactions with high accuracy. *Nucleic Acids Res*, 2019, 47: e33
- [32] Holoch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet*, 2015, 16: 71-84
- [33] Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet*, 2014, 15: 7-21
- [34] Marchese D, de Groot NS, Lorenzo Gotor N, et al. Advances in the characterization of RNA-binding proteins. *Wiley Interdiscip Rev RNA*, 2016, 7: 793-810
- [35] Tsai BP, Wang X, Huang L, et al. Quantitative profiling of *in vivo*-assembled RNA-protein complexes using a novel integrated proteomic approach. *Mol Cell Proteomics*, 2011, 10: M110 007385
- [36] Castello A, Horos R, Strein C, et al. System-wide identification of RNA-binding proteins by interactome capture. *Nat Protoc*, 2013, 8: 491-500
- [37] Chu C, Chang HY. ChIRP-MS: RNA-directed proteomic discovery. *Methods Mol Biol*, 2018, 1861: 37-45
- [38] Chu C, Quinn J, Chang HY. Chromatin isolation by RNA purification (ChIRP). *J Vis Exp*, 2012, 61: 3912
- [39] Simon MD. Capture hybridization analysis of RNA targets (CHART). *Curr Protoc Mol Biol*, 2013, Chapter 21: Unit 21.25
- [40] McHugh CA, Guttman M. RAP-MS: a method to identify proteins that interact directly with a specific RNA molecule in cells. *Methods Mol Biol*, 2018, 1649: 473-88
- [41] Keryer-Bibens C, Barreau C, Osborne HB. Tethering of proteins to RNAs by bacteriophage proteins. *Biol Cell*, 2008, 100: 125-38
- [42] Wang X, Chen CF, Baker PR, et al. Mass spectrometric characterization of the affinity-purified human 26S proteasome complex. *Biochemistry*, 2007, 46: 3553-65
- [43] Tutucci E, Vera M, Biswas J, et al. An improved MS2 system for accurate reporting of the mRNA life cycle. *Nat Methods*, 2018, 15: 81-9
- [44] Chu C, Zhang QC, da Rocha ST, et al. Systematic discovery of Xist RNA binding proteins. *Cell*, 2015, 161: 404-16
- [45] Ninomiya K, Adachi S, Natsume T, et al. LncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J*, 2020, 39: e102729
- [46] Gandhi M, Gross M, Holler JM, et al. The lincRNA lincNMR regulates nucleotide metabolism via a YBX1-RRM2 axis in cancer. *Nat Commun*, 2020, 11: 3214
- [47] Morris KV, Mattick JS. The rise of regulatory RNA. *Nat Rev Genet*, 2014, 15: 423-37
- [48] Conrad T, Albrecht AS, de Melo Costa VR, et al. Serial interactome capture of the human cell nucleus. *Nat Commun*, 2016, 7: 11212
- [49] Perez-Perri JI, Rogell B, Schwarzl T, et al. Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat Commun*, 2018, 9: 4408
- [50] Slobodin B, Gerst JE. A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. *RNA*, 2010, 16: 2277-90
- [51] Langland JO, Pettiford SM, Jacobs BL. Nucleic acid affinity chromatography: preparation and characterization of double-stranded RNA agarose. *Protein Expr Purif*, 1995, 6: 25-32
- [52] Ruby SW, Goelz SE, Hostomsky Z, et al. Affinity chromatography with biotinylated RNAs. *Methods Enzymol*, 1990, 181: 97-121
- [53] Michlewski G, Caceres JF. RNase-assisted RNA chromatography. *RNA*, 2010, 16: 1673-8
- [54] Michlewski G, Guil S, Semple CA, et al. Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol Cell*, 2008, 32: 383-93
- [55] Alam U, Kennedy D. G3BP1 and G3BP2 regulate translation of interferon-stimulated genes: IFITM1, IFITM2 and IFITM3 in the cancer cell line MCF7. *Mol Cell Biochem*, 2019, 459: 189-204
- [56] Sajini AA, Choudhury NR, Wagner RE, et al. Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation. *Nat Commun*, 2019, 10: 2550
- [57] Siphshvili Z, Webster DE, Kretz M, et al. Identification of proteins binding coding and non-coding human RNAs using protein microarrays. *BMC Genomics*, 2012, 13: 633
- [58] Siphshvili Z, Webster DE, Johnston D, et al. The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nat Genet*, 2016, 48: 53-8