Vol. 13, No. 4, 2022
ISSN 2233-4203/ e-ISSN 2093-8950
REVIEW www.msletters.org | Mass Spectrometry Letters

Combining the Power of Advanced Proteome-wide Sample Preparation Methods and Mass Spectrometry for defining the RNA-Protein Interactions

Tong Liu[1], Chaoshuang Xia[2], Xianyu Li[1,3]*, and Hongjun Yang[1]*

[1]*Beijing Key Laboratory of Traditional Chinese Medicine Basic Research on Prevention and Treatment for Major Diseases, Experimental Research Center, China Academy of Chinese Medical Sciences, Beijing, 100010, China*
[2]*Key Laboratory of Synthetic and Natural Function Molecule Chemistry of Ministry of Education, College of Chemistry and Materials Science, Northwest University, Xi'an 710069, China*
[3]*State key Laboratory of innovative Natural Medicine and TCM Injection, Jiangxi Qingfeng Pharmaceutical Co.,Ltd.*

Received December 6, 2022; Revised December 23, 2022; Accepted December 23, 2022
First published on the web December 31, 2022; DOI: 10.5478/MSL.2022.13.4.115

Abstract : Emerging evidence has shown that RNA-binding proteins (RBPs) dynamically regulate all aspects of RNA in cells and involve in major biological processes of RNA, including splicing, modification, transport, transcription and degradation. RBPs, as powerful and versatile regulatory molecule, are essential to maintain cellular homeostasis. Perturbation of RNA-protein interactions and aberration of RBPs function is associated with diverse diseases, such as cancer, autoimmune disease, and neurological disorders. Therefore, it is crucial to systematically investigate the RNA-binding proteome for understanding interactions of RNA with proteins. Thanks to the development of the mass spectrometry, a variety of proteome-wide methods have been explored to define comprehensively RNA-protein interactions in recent years and thereby contributed to speeding up the study of RNA biology. In this review, we systematically described these methods and summarized the advantages and disadvantages of each method.

Keywords : proteomics, mass spectrometry, RNA-binding proteins, enrichment, RNA-protein interactions

Introduction

As delivery intermediate of genetic information in biological systems, RNA usually interacts with proteins, which are called RBPs, to form dynamic ribonucleoprotein complexes (RNPs) for implementing crucial biological functions.[1] RBPs widely participate in various biological processes of post-transcriptional gene regulation (PTGR), and also regulate every stage of the lifetime of RNA, from transcription, translation and through to decay.[2] As one salient example, pre-mRNA is regulated by mRNA binding proteins (mRBPs) to form mature mRNA through splicing, 5′ end capping and 3′ end processing.[3] The mature mRNA is exported to cytoplasm by specific mRBPs assistance for translation.[4,5] Degradation of mRNA also requires a large number of mRBPs. Cumulative evidence has also shown that long non-coding RNA (lncRNA) may regulate biological processes such as chromatin modification, and apoptosis.[6,7] In a word, the RBPs may regulate RNA modifications, localization, translation, and stability. Reciprocally, the corresponding RNAs can also affect functions and localization of RBPs.[8] To date, a variety of RBPs have been investigated in eukaryotic cell. Individual RNA-protein interactions can form an intricate network to regulate complicated cellular processes. Of course, the broader research has uncovered that the deficiency or aberration of RBPs is closely involved in the occurrence and development of numerous diseases,[9,10] including metabolic disorders, autoimmune disease, neurological disorders and cancer.[11-14] Considering the importance of their functions, numerous studies are paying special attention to the comprehensive identification and precise quantification of RBPs. However, it is still challenging to study their interactions with RNAs due to the low stoichiometry of RBPs, the ubiquity of RNase which can easily degrade the RNA and the immense complexity of RNA-protein interactions network. Therefore, it is of great significance to design an unbiased and effectively method for large-scale identification of

*Reprint requests to Xianyu Li, Hongjun Yang
https://orcid.org/0000-0001-9625-8476
https://orcid.org/0000-0001-5501-3288
E-mail: phd_xianyuli@foxmail.com, hongjun0420@vip.sina.com

RBPs. The initial research was confined to an RNA of interest and focused on the binding proteins, or a protein of interest and RNA that associate with it.[15-17] These methods that aimed to single biomolecule of interest may omit the dynamic interaction from molecular networks. Compared with RNA-sequencing, the study of RBPs is more challenging due to low abundance and lack of amplification.[18] Benefitting from the development of mass spectrometry, recent years have seen the emergence of numerous methods aimed at investigating the RNA-binding proteome systemically. Because each method has some strengths and drawbacks, it is necessary to choose optimal methods according to relevant research question. In this review, we systematically summarized methods for studying RNA-protein interactions on proteome-wide scale based on different principles recent years and provided a selection for scientists to address a particular biological question.

The poly(A)-based technologies

Based on the principle of complementary base pairing of Watson-Crick, RNA and directly bound proteins could be captured through hybridization of the poly(A) tails of RNA and oligo(dT) beads. That opened up the systematic identification of RBPs. The first widely accepted technology was reported Castello et.al in 2012,[19] referred to as RNA-interactome capture (RIC). In this method, RNAs and proteins that directly bound were firstly "frozen" by widely used UV-crosslinking in vivo. As shown in Figure 1, cells were lysed under denaturing conditions, and the sequential RNA-protein complexes were enriched by oligo(dT) beads. The proteins were released by RNase treatment after several stringent washing. Following protease treatment, mass spectrometry-based proteomics was used to identify high confidence RBPs. Besides, RNAs from RNPs were released by proteinase K treatment and analyzed via high-throughput sequencing. Applied to HeLa cells, Castello et al.[19] discovered 860 RBPs by RIC coupled with mass spectrometry, including 300 novel RBPs. Using RIC and mass spectrometric analysis, ~800 RBPs were identified in HEK293 cells by Baltz et.al, many of which were unknown RBPs previously.[20] RIC yielded numerous RBPs from various biological origins, including several human cell lines,[21-23] mouse cell lines,[11,24,25] *Saccharomyces cerevisiae*[26,27] and *Caenorhabditis elegans*,[27] *Leishmania donovani*,[28] *Plasmodium falciparum*,[29] *Trypanosoma brucei*,[30] *Arabidopsis thaliana*,[31,32] *Drosophila*[33,34] and zebrafish.[35] Hentze et al.[36] summarized all published RNA interactomes into RBPs supersets according to different source. It greatly enriched the RBPs item. It is worth mentioning that metabolic enzymes, which play key functional roles in biological systems, were also identified as RBPs. For example, a high proportion of metabolic enzymes involved in glycolytic pathway were identified as RBPs in the human hepatocyte HuH7 cell.[27] Likewise, the RBPs that reported in the HL-1 cardiomyocyte cell included many mitochondrial metabolic enzymes involved in the tricarboxylic acid cycle, fatty acid oxidation and so on.[11] Hentze[37] laboratory presented alternative method, called enhanced RNA Interactome Capture (eRIC), which mainly used Locked Nucleic Acid (LNA)-modified capture probe instead of oligo(dT) beads. Applying the Jurkat cells, 683 and 588 RBPs were characterized from eRIC and RIC, respectively. Compared to the RIC, eRIC has notably advantages, such as excellent specificity, greater sensitively and lower contamination from rRNA and DNA. Castello et al.[38] developed a comparative RNA Interactome Capture (cRIC) by combining stable isotopic labeling by amino acids in cell clture (SILAC)-based mass spectrometry, which could study the dynamic of the RBPs upon the virus infection. As a result, 794 proteins were identified in total, 91% of which were reported to be RBPs in eukaryotic cells. Moreover, 247 RBPs changed their binding activity after infected virus. It revealed that virus infection could

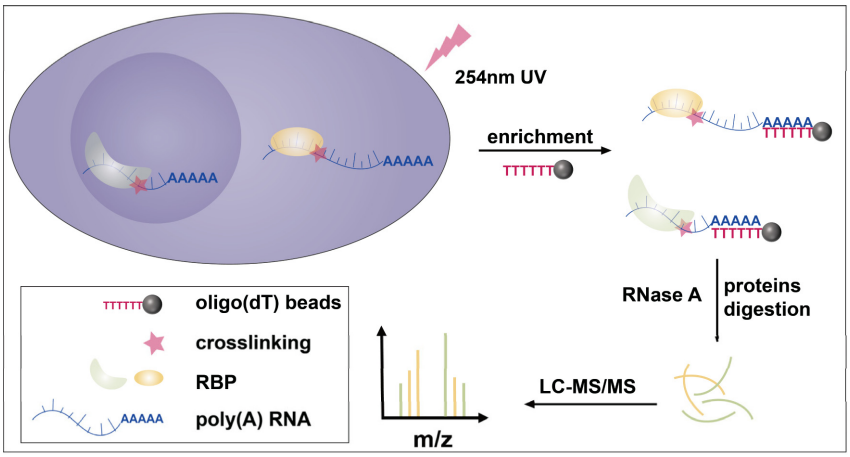

Figure 1. Schematic representation of the RNA-interactome capture. Only the significant steps are shown.

rearrange the cellular RBPs to regulate the life cycle of virus. They also used cRIC tandem quantitative mass spectrometry to systematically and comprehensively discovered the cellular RBPs that responded to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection.[39] In their work, they uncovered that the cellular RBPs could be remodeled upon the SARS-CoV-2 infection. The RNA metabolic pathways were profoundly affected. They also described viral RNA interactome capture (vRIC) and revealed dozens of cellular RBPs and six viral RBPs that directly interacted with viral RNA. Furthermore, available antiviral drugs targeting host RBPs were identified as being capability of inhibiting infection against the SARS-CoV-2. It demonstrated that the RBPs have enormous potential for antiviral therapies.

Notably, most RBPs comprise one or multiple well-known RNA-binding domains (RBDs), where they interact with target RNA.[40] With the in-depth study, however, researchers noticed that many novel RBPs are lacking canonical RBDs. Meanwhile, the mutations of RBDs could be the cause of numerous monogenic diseases.[41] Consequently, much attention was devoted to identifying the regions of RBPs interaction with RNA. As shown in Figure 2, in order to depict an atlas of RNA-binding sites in a proteome-wide manner, Castello et al.[42,43] developed RBDmap-based approach, which combined UV cross-linking, oligo(dT) capture, controlled proteolysis and mass spectrometry to discover high-confidence RNA-binding sites on a proteome-wide scale. In brief, the RNA–protein complexes were primarily purified with the aforementioned RIC. After elution, the RBPs were partial digested and separated by a second round of oligo(dT) purification yielding two classes of peptides: (1) peptides released into the supernatant and (2) peptides remain bound to RNA. Subsequently, peptides remain bound to RNA were further subjected to proteolysis generating two types of peptides, including the peptides still covalently bound to the RNA, referred to X-link, and its neighboring peptides (N-link). It was difficult to characterize X-link due to various mass shift resulting from residual nucleotides, while N-link could be easily identified by mass spectrometry and search algorithms. Most importantly, the RNA-binding region of RBPs, namely RBDpep, composed of X-link peptides and N-link peptides. The RNA-binding sites were detected through analysis of the RNA-bound and released fractions by mass spectrometry-based quantitative proteomics. Ultimately, they identified 1174 RNA-binding sites derived from 529 RBPs in HeLa cells. Plus, RBDmap has also proven its utility in HL-1 cardiomyocytes, for which RBDmap dissected the 568 RNA-binding regions of 368 murine proteins by mass spectrometry.

Overall, with the development and application of mass spectrometry, the advent of RNA interactome capture not only opened up the opportunities to system-wide identification of RBPs but also supported for the comprehensive study of RNA-protein interaction. Moreover, the discovery of novel RBPs with lack of classical RBD but closely relate to biological function implicated that the potential roles of RBPs need to be investigated in depth urgently.[36] Nevertheless, the poly(A)-based technologies are only limited to RBPs that bound poly(A) RNA. It is not suitable for unearthing the nonpoly(A) RNAs, including most ncRNAs and organisms without poly(A) RNA such as bacteria and many archaea.[44] Where proteins interacting with mRNA are of interest, the poly(A)-based technologies are recommendable.

The nucleotide analogs-based methods

Further efforts to extend into the nonpoly(A) RNA bound proteome had focused on methods that were used to nucleotide analogs tandem with click chemistry to capture proteins bound to nascent RNA, regardless of the polyadenylation state of RNA. Two more promising methods were published in different cells and systems. Bao et al.[45] presented an RNA interactome using click chemistry (RICK) method, which enabled to identify a wide range of RNA bound proteome. As illustrated in

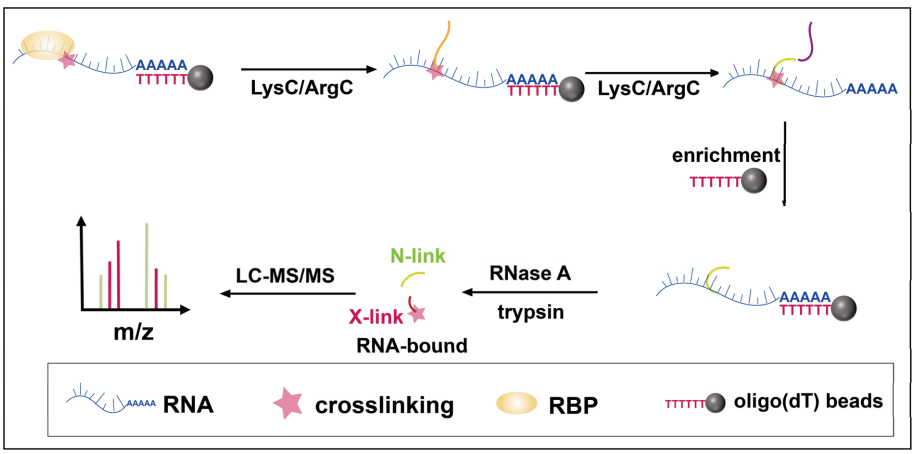

Figure 2. Schematic diagram for RBDmap. Only the significant steps are shown.

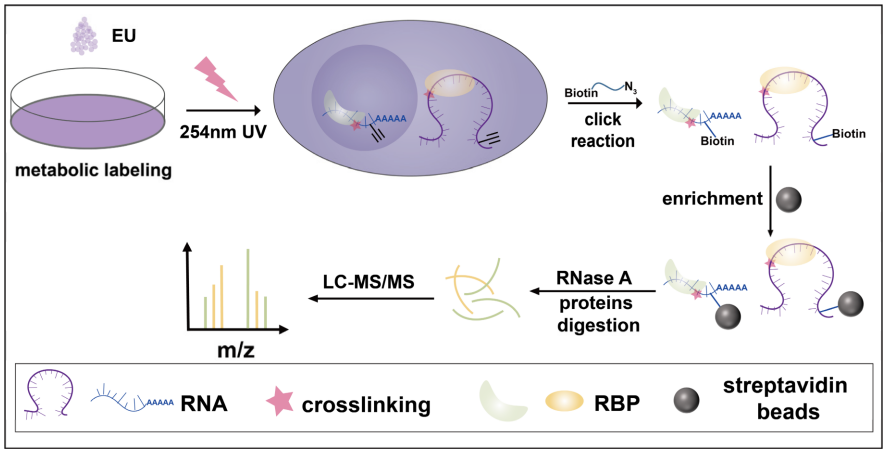

Figure 3. Schematic illustration of the nucleotide analogs-based methods. Only the significant steps are shown.

Figure 3, they exploited an alkynyl uridine analog, 5-ethynyluridine (EU), to label the newly transcribed RNA in living cells, followed by in vivo fixed the RNA-proteins with 254 nm UV light. Combined with click chemistry, the alkynyl could react with the azide-biotin to gain biotinylated RNA. Subsequent biotinylated RNA and covalently cross-linked proteins were captured via streptavidin beads. The resulting isolated RBPs were characterized and quantified by mass spectrometry. Consequently, a high-confidence RNA-binding proteomes containing 720 proteins and 518 proteins were identified from HeLa cells and mouse embryonic stem cells (mESCs), respectively. Contemporary, Chen's lab[46] introduced the same procedure, but used a photoactivatable uridine analog 4-thiouridine (4SU) by 365-nm UV light to cross-link RNA-proteins instead of 254nm light directly. This unbiased method was termed click chemistry-assisted RNA interactome capture (CARIC) and yielded 597 proteins constituting HeLa cells RNA interactome with the help of mass spectrometry. These two methods that were designed to profile RBPs by mass spectrometer rely on the alkynyl uridine analog-assisted click chemistry.

Overall, the combination of mass spectrometric methods and nucleotide analogs-based methods overcomes the limitations of the poly(A)-based technologies. CARIC and RICK were not restricted to the analysis of RBPs that associated with polyadenylated RNA, but could extend to the proteomic profiling interacting with all types of RNA. In addition, it might facilitate an in-depth interrogation of the complex regulatory mechanisms of the ncRNA bound proteins. Unfortunately, the nucleotide analogs-based methods embed artificially introduced nucleotide analogues into newly synthesized RNA through transcription. Therefore, the method is aimed at the analysis of newly transcribed RNA-bound proteins during process of cell culture, which mainly have the following two defects. Firstly, metabolic labeling may interfere with cellular physiological processes, such as the inhibition of rRNA synthesis, nucleolar stress response,[47] and decreased cell viability.[44,48] Secondly, bias may be introduced due to ignoring RNA that is already present. For studies targeting at nascent RNA-binding proteins, the nucleotide analogs-based methods may be a better choice.

The phase separation-based strategies

Phase separation has advanced the field by offering strategy that can be used for isolating RNA and proteins. The rationale of strategy is that RNA can repartition to the upper aqueous phase in the acidic guanidinium thiocyanate/phenol/chloroform (also called Trizol), while proteins remain in the lower organic phase.[49] Owing to the opposing physicochemical properties of the RNA and proteins, RNA-protein complexes would be concentrated at the interphase. Exploiting the phase separation offered a promising novel direction for the field of comprehensive identification of RNA–protein interactions. Therefore, three teams successionally published effective strategies, named Orthogonal Organic Phase Separation (OOPS)[50] and Protein-Crosslinked RNA Extraction (XRNAX)[51] as well as Phenol Toluol Extraction (PTex)[52], that repurposed phase separation to enrich the cellular RNA-binding proteomes from the interface layer and catalog them with mass spectrometric analysis (Figure 4). Indeed, the crude RNA-protein complexes were obtained at the interface using Trizol in OOPS and XRNAX. The OOPS strategy undergone three consecutive rounds of phase separation to purify RNA-protein complexes. The RBPs migrated from the interface to the organic phase followed by RNase treatment. Benefit of high-throughput mass spectrometry, OOPS resulted in 1267 RNA-binding proteomes for U2OS cells, 1410 RNA-binding proteomes for HEK293 cells, and 1165 RNA-binding proteomes for MCF10A cells by mass spectrometry-based proteomics analysis, of which 759 proteins were shared by three human cell lines. Notably,

OOPS was used to obtain 364 RBPs of *E. coli,* including new discovered sRNA binding protein (ProQ)[53]. It was the first systematically retrieved the RBPs from prokaryotes. By performing a single Trizol phase separation, XRNAX was able to produce crude RNA-protein complexes that were then subjected to limited protease digestion, resulting in RNA crosslinked to peptides. Reliable RBPs separation was achieved through an additional silica-based columns purification prior to mass spectrometric analysis. Applied to MCF7, HeLa and HEK293 cell lines, XRNAX yielded 1207, 1239 and 1357 high-confidence RBPs from three cell lines respectively. They discovered 858 RBPs shared among the three cell lines. Furthermore, the dynamics of RNA-proteins interactions was explored during arsenite stress using XRNAX incorporated an additional silica enrichment and mass spectrometric analysis. In terms of the results, most RBPs did not change the interaction with RNA, whereas several proteins represented remarkable decline their RNA-binding capacity under arsenite stress. Coincidentally, Hentze et al.[54] introduced a rapid strategy to enrich RNA-proteins, known as complex capture (2C). The principle on the basis of the strategy is that the nucleic acid can selectively retain to silica matrix columns. Based on this, the cross-linked RNA-proteins can also retain to the silica columns compared with non-cross-linked proteins[55]. The following year Shchepachev et al.[56] described total RNA-associated protein purification (TRAPP) strategy which mainly based on recovery of all cross-linked RNA-proteins on silica beads. The described strategy entailed the extraction of cross-linked RNA-proteins under acidic condition and in-gel trypsin digestion of RBPs, followed by liquid chromatography-mass spectrometric analysis. They used yeast and *E. coli* for proof-of-principle experiments and discovered more novel RBPs. Although this strategy has broad applicability for RBPome characterization, the silica-based methods suffer from low recovery from silica gel columns. PTex described an alternative strategy that utilized a neutral mixture of phenol: toluol to shift RNA, proteins and cross-linked RNPs into the aqueous phase, while the DNA and lipids were distributed to the interface. The aqueous phase was subsequently recovered and extracted twice using acidic phenol. On this condition, the RNPs were migrated to the interface, away from RNA in the upper aqueous layer and proteins in the lower organic layer. Eventually, the interface was subjected to precipitate using ethanol to highly enrich the cross-linked RNPs. The purified RNPs could be protease digestion and directly analysis by liquid chromatography-mass spectrometry. By applying this strategy, 3037 RBPs were significantly enriched from HEK293 cells.

Proximity labeling (PL) is a widely applicable tool for deciphering of molecular interactions in subcellular organelles and compartments of interest with nanometer-scale spatial resolution.[57] It would be an important step to a better separating subcellular region in complex cells. PL also integrated with RNA-proteins enrichment strategy to profile RNA-protein interactions in different subcellular regions. Qin et al.[58] have extensive experience in subcellular molecular interactions using PL. In particular, they turned attention to decipher of the RNA-protein interactions in spatially and temporally resolution. They developed a method that peroxidase-catalyzed PL combined with phase separation (APEX-PS) to categorize RBPs from specific subcellular regions (Figure 5a). By combining APEX-PS with mass spectrometry, they generated several RBPs datasets for nuclear, nucleolar, and outer mitochondrial membrane (OMM). Moreover, they mined novel RBPs function from outer mitochondrial membrane and confirmed the localization of OMM SYNJ2BP after puromycin treatment. Similarity, Chen's group[59] reported a strategy that used subcellular fractionation, acidic guanidinium-thiocyanate-phenol-chloroform biphasic extraction, and quantitative mass spectrometry to enrich RBPs from specific subcellular organelles (Figure 5b). 1775 and 2245 RBPs were uncovered by this strategy from the cell

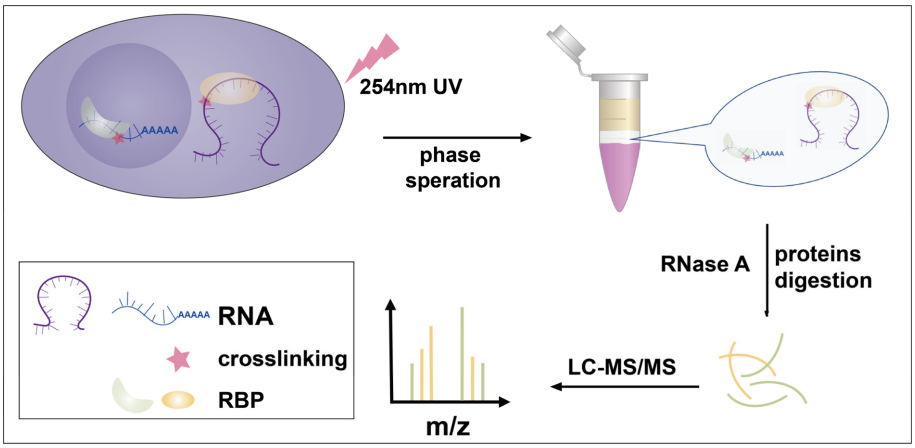

Figure 4. Schematic overview of the phase separation-based strategies. Only the significant steps are shown.

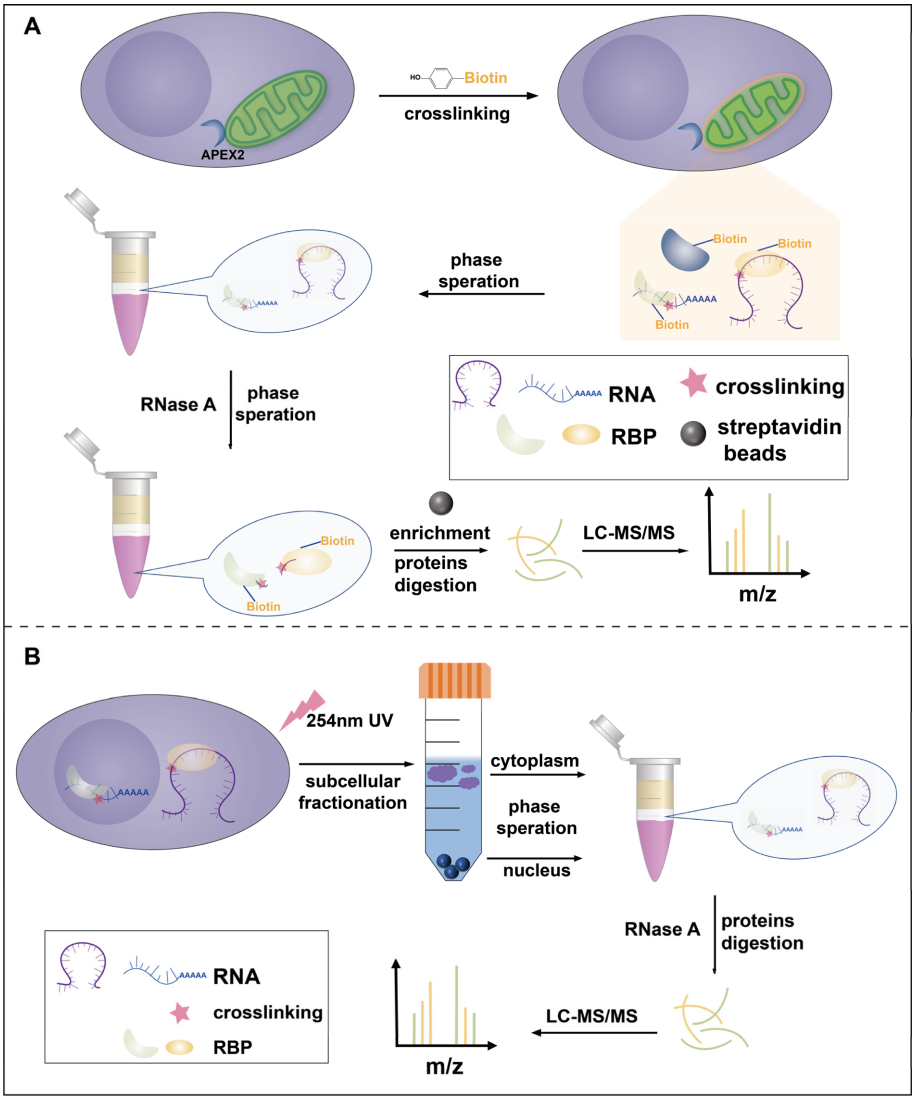

Figure 5. a. Schematic of APEX-PS. Only the significant steps are shown; b. the workflow for enrichment of nucleus and cytoplasm RNA-binding proteins. Only the significant steps are shown.

nucleus and cytoplasm, respectively. Further analysis discovered 403 unique RBPs from cell nucleus and 873 unique RBPs from cell cytoplasm. Among the all RBPs datasets not only contained a large number of known RBPs, but also revealed 614 RBPs that have never been reported previously. Systematically classified the RBPs subcellular localizations and provided an extra information about latent biological functions. In order to verify the novel identified RBPs, they characterized 2157 RNA binding interfaces originate from 892 RBPs using the modified RBDmap method. It could elucidate complex binding structural regarding the RNA-protein interactions.

Compared with poly(A)-based technologies, the phase separation-based strategies eliminated the focus on poly(A) RNA and based entirely on inherent physicochemical properties of RNA and proteins. It mainly lied in the difference of the solubility between RNA, proteins and RNA-protein complexes. Meanwhile, phase separation-based strategies held enormous potential towards characterizing the prokaryotes RBPs from a system-wide perspective and thereby facilitated a better mining of RBPs biology functions with RNA-proteins interactions in prokaryotes. This strategy requires less than 1% of the cells needed by previous RBPs-enrich methods. OOPS and XRNAX are valuable approaches that manipulated simply and rapidly. Problematically, since glycoproteins and RNA-protein complexes have the same solubility, glycoproteins cannot be distinguished using phase separation-based strategies, creating a problem of glycoprotein contamination.

The chemical labeling-based approaches

Intense research have unearthed numerous RNA crosslinkers that used to connect interacting of RNA to each other, such as methylene blue and psoralen.[60] Among them, psoralen is a class of planar and ternary heterocyclic compounds that can form stable covalent bonds with pyrimidine bases (especially thymine and uracil) through cycloaddition reaction.[61] The reasoning is that psoralen first intercalated into the pyrimidine bases after enter cells. Then the cycloaddition reaction was occurred by long wavelength ultraviolet irradiation (320 – 410 nm) to crosslink RNA and psoralen with covalent bonds.[62] This photoactivation reaction could be used to efficient and unbiased capture of RNA, study of RNA interactions could be addressed, opening new opportunities to understand RNA biological functions on a genome-wide scale. Exploiting the high reactivity of psoralen with uracil under 365 nm UV irradiation, Qin's lab synthesized a psoralen probe (PP) and developed a new PP-based RBPs identifying approach.[63] (As shown in Figure 6a) The PP could tag the nucleic acid and capture the nucleic acid-proteins complex in cells, followed by eluting enriched RBPs by treatment with RNase A. Subsequently, the product was digested with trypsin and characterized by quantitative mass spectrometry after stable isotopic dimethyl labelling. Applying this approach, the authors revealed a total of 2986 RBPs in HeLa cells, which covered ~70% of RBPs originated from HeLa cells reported by previous works. It also included 782 low abundant candidate RBPs. Meanwhile, 178 metabolic enzymes widely involved in metabolic pathways, which were discovered in candidate RBPs. Furthermore, the

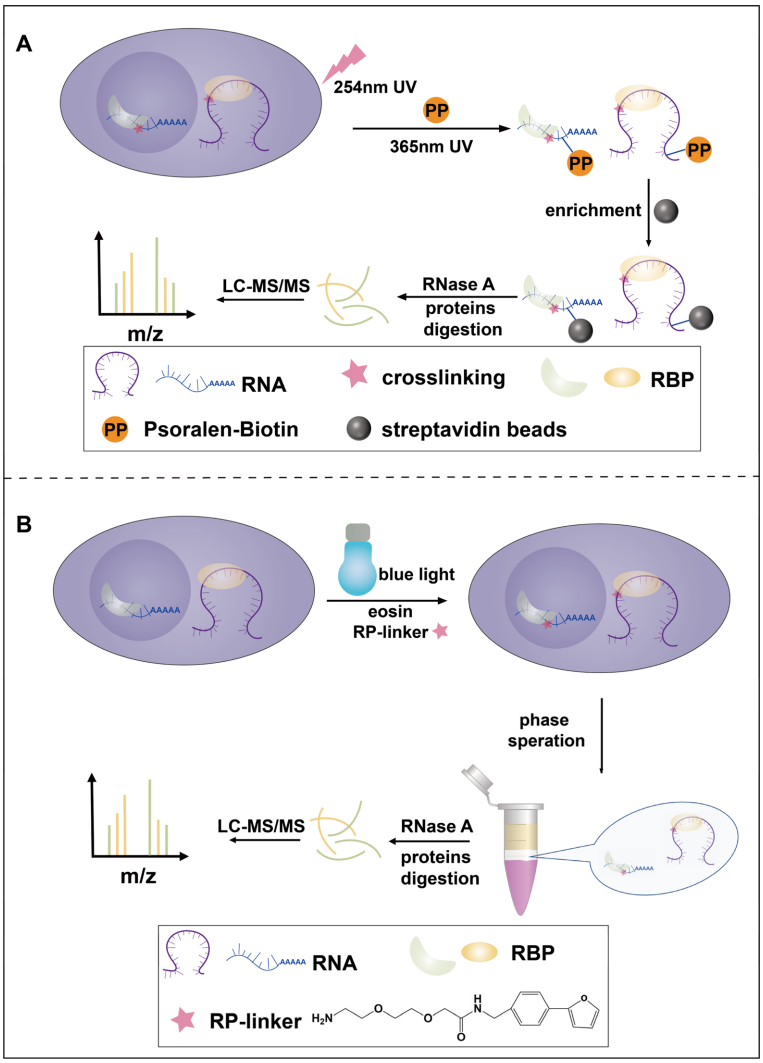

Figure 6. a. Schematic representation of the PP-based approach. Only the significant steps are shown; b. Experimental scheme of the PhotoCAX-MS. Only the significant steps are shown.

authors employed PP-based approach to investigate a large-scale dynamic of RNA-proteins interactions upon Actinomycin D-stimulation. As a result, the distribution map of RBPs with different decreasing/decay rates in RNPs was obtained. More importantly, it also provided a new high-throughput way to evaluate candidate RBPs. Additionally, PP-based approach can be combined with other methods to achieve more comprehensive and extensive coverage of RNA-binding proteome.

Zou et al.[64] designed a dual-functional photocatalytic RNA-protein crosslinker (RP-linker) which could be adopted for in-situ labelling and efficient crosslinking of RNA-protein complexes under blue light-triggered photocatalyst. Besides, they reported a comprehensive profiling of RNA-protein interaction approach by integrating photocatalytic crosslinking with phase separation and mass spectrometry, termed as "PhotoCAX-MS", as shown in Figure 6b. The RP-linker was composed of a furan group and amine group for labelling RNAs and connecting RBPs, respectively. Thus, RNA-protein complexes were covalent linked by RP-linker upon photocatalytic conditions in cultured human HEK293 cells. And then the crosslinked RNA-protein complexes were enriched by phase separation for subsequent mass spectrometric analysis. With this approach, they identified 2044 RBPs from HEK293 cells. The PhotoCAX-MS could further yield novel insights concerning the dynamics remodeling of RNA-protein interactions in macrophage cells upon LPS-assisted and SILAC labelling. 1926 proteins were identified by liquid chromatography-tandem mass spectrometry, 1299 of which were shared in three independent biological replicates. Further analysis revealed 11 up-regulated RNA-binding proteins and 12 down-regulated RNA-binding proteins. Noticeably, the Photo CAX-MS was also applied to dig the proteins directly binding with the 5' untranslated regions of SARS-CoV-2 RNA. A total of 193 RBPs were discovered by liquid chromatography-tandem mass spectrometry. Further analysis emerged the potential biological function of host RBPs in the SARS-CoV-2 infections.

In conclusion, the chemical labeling-based approaches greatly improved the specificity of RBPs enrichment and provided an ideal platform for in-depth RNA-protein interaction research. Of course, these approaches are still unbiassed which can map all kind of RNA bound proteins. Although tremendous progress has been made for chemical labelling-based approaches to map all kind of RNA bound proteins, these approaches are still far from reaching the level of integration sufficient to address bioanalytical challenges. For example, PP-based methods tend to favor more abundant non-coding RNA-binding proteins and thus may inhibit the identification of mRNA-binding proteins. Due to the use of phase separation, RBPs with glycosylation may not be successfully recognized by PhotoCAX-MS.

Conclusions and perspectives

Recent efforts opened up new opportunities to study RNA-protein interactions and expanded the methods to characterize RNA-binding proteome with high-throughput mass spectrometry. We reviewed here experimental large-scale methods for RNA-protein complexes enrichment, with the characteristics of each method. Researchers can design a suitable method for specific issue. Considering different enrichment mechanisms, these methods can complement each other in some aspects. Taking advantage of these methods, it is promising to achieve more comprehensive and extensive coverage of RNA binding proteome, and thereby laying the foundation for the in-depth study of RNA-protein interactions. Nevertheless, given the fact that UV light that used to covalent crosslink RNA-protein complexes has poor penetrability, the application of these methods is confined to cultured cells (generally monolayer cells), fly and *Arabidopsis thaliana*, as well as bacteria. It is also worth noting that the current methods are not compatible with the identification of RNA-binding proteome from mammal tissue. To overcome this limitation, it is necessary to develop more crosslinking methods, such as new chemical crosslinkers, for replacing UV crosslinking in the fix of RNA-proteins complexes in future. In addition, proper control conditions and stringent screening cut-off RBPs should be carefully considered. The specific and precision of RNA-binding proteome can provide more confident data to support further biological study. The field of mass spectrometry-based proteomics is rapidly expanding and creating more convenience for comprehensive studying biological functions. It can be expected to open a new way to investigate RNA-binding proteome for understanding interaction of RNA with proteins by combining the advantages of proteome-wide sample preparation methods and mass spectrometry.

Acknowledgements

We are grateful for financial support from the Fundamental Research Funds for the Central Public Welfare Research Institutes (ZZ13-YQ-080)?Scientific and technological innovation project of China Academy of Chinese Medical Sciences (CI2021B017, CI2021A05032, XTCX2021001).

Conflict of interest statement

There are no conflicts no declare.

Reference

1. Gerstberger, S.; Hafner, M.; Tuschl, T. *Nature Reviews Genetics* **2014**, 15, 829. DOI: 10.1038/nrg3813.
2. Nair, L.; Chung, H.; Basu, U. *Nature Reviews Molecular Cell Biology* **2020**, 21, 123. DOI: 10.1038/s41580-019-

0209-0.
3. Dreyfuss, G.; Kim, V.N.; Kataoka, N. *Nature Reviews Molecular Cell Biology* **2002**, 3, 195. DOI: 10.1038/nrm760.
4. Koehler, A.; Hurt, E. *Nature Reviews Molecular Cell Biology* **2007**, 8, 761. DOI: 10.1038/nrm2255.
5. Sonenberg, N.; Hinnebusch, A.G. *Cell* **2009**, 136, 731. DOI: 10.1016/j.cell.2009.01.042.
6. Chen, R.; Liu, Y.; Zhuang, H.; Yang, B.; Hei, K.; Xiao, M.; Hou, C.; Gao, H.; Zhang, X.; Jia, C.; Li, L.; Li, Y.; Zhang, N. *Nucleic Acids Res.* **2017**, 45, 9947. DOI: 10.1093/nar/gkx600.
7. Cooper, T.A.; Wan, L.; Dreyfuss, G. *Cell* **2009**, 136, 777. DOI: 10.1016/j.cell.2009.02.011.
8. Huang, A.; Zheng, H.; Wu, Z.; Chen, M.; Huang, Y. *Theranostics* **2020**, 10, 3503. DOI: 10.7150/thno.42174.
9. Jankowsky, E.; Harris, M.E. *Nature Reviews Molecular Cell Biology* **2015**, 16, 533. DOI: 10.1038/nrm4032.
10. Gebauer, F.; Schwarzl, T.; Valcarcel, J.; Hentze, M.W. *Nature Reviews Genetics* **2021**, 22, 185. DOI: 10.1038/s41576-020-00302-y.
11. Liao, Y.; Castello, A.; Fischer, B.; Leicht, S.; Foeehr, S.; Frese, C.K.; Ragan, C.; Kurscheid, S.; Pagler, E.; Yang, H.; Krijgsveld, J.; Hentze, M.W.; Preiss, T. *Cell Reports* **2016**, 16, 1456. DOI: 10.1016/j.celrep.2016.06.084.
12. Jazurek, M.; Ciesiolka, A.; Starega-Roslan, J.; Bilinska, K.; Krzyzosiak, W.J. *Nucleic Acids Res.* **2016**, 44, 9050. DOI: 10.1093/nar/gkw803.
13. Pereira, B.; Billaud, M.; Almeida, R. *Trends In Cancer* **2017**, 3, 506. DOI: 10.1016/j.trecan.2017.05.003.
14. Castello, A.; Fischer, B.; Hentze, M.W.; Preiss, T. *Trends Genet.* **2013**, 29, 318. DOI: 10.1016/j.tig.2013.01.004.
15. Ramanathan, M.; Porter, D.F.; Khavari, P.A. *Nat. Methods* **2019**, 16, 225. DOI: 10.1038/s41592-019-0330-1.
16. McHugh, C.A.; Russell, P.; Guttman, M. *Genome Biol.* **2014**, 15, 203. DOI: 10.1186/gb4152.
17. Koenig, J.; Zarnack, K.; Luscombe, N.M.; Ule, J. *Nature Reviews Genetics* **2012**, 13, 77. DOI: 10.1038/nrg3141.
18. Nechay, M.; Kleiner, R.E. *Curr. Opin. Chem. Biol.* **2020**, 54, 37. DOI: 10.1016/j.cbpa.2019.11.002.
19. Castello, A.; Fischer, B.; Eichelbaum, K.; Horos, R.; Beckmann, B.M.; Strein, C.; Davey, N.E.; Humphreys, D.T.; Preiss, T.; Steinmetz, L.M.; Krijgsveld, J.; Hentze, M.W. *Cell* **2012**, 149, 1393. DOI: 10.1016/j.cell.2012.04.031.
20. Baltz, A.G.; Munschauer, M.; Schwanhaeusser, B.; Vasile, A.; Murakawa, Y.; Schueler, M.; Youngs, N.; Penfold-Brown, D.; Drew, K.; Milek, M.; Wyler, E.; Bonneau, R.; Selbach, M.; Dieterich, C.; Landthaler, M. *Mol. Cell* **2012**, 46, 674. DOI: 10.1016/j.molcel.2012.05.021.
21. Beckmann, B.M.; Horos, R.; Fischer, B.; Castello, A.; Eichelbaum, K.; Alleaume, A.-M.; Schwarzl, T.; Curk, T.; Foehr, S.; Huber, W.; Krijgsveld, J.; Hentze, M.W. *Nature Communications* **2015**, 6, 1. DOI: 10.1038/ncomms10127.
22. Conrad, T.; Albrecht, A.-S.; Costa, V.R.d.M.; Sauer, S.; Meierhofer, D.; Orom, U.A. *Nature Communications* **2016**, 7, 1. DOI: 10.1038/ncomms11212.
23. Kramer, K.; Sachsenberg, T.; Beckmann, B.M.; Qamar, S.; Boon, K.-L.; Hentze, M.W.; Kohlbacher, O.; Urlaub, H. *Nat. Methods* **2014**, 11, 1064. DOI: 10.1038/nmeth.3092.
24. Kwon, S.C.; Yi, H.; Eichelbaum, K.; Foehr, S.; Fischer, B.; You, K.T.; Castello, A.; Krijgsveld, J.; Hentze, M.W.; Kim, V.N. *Nat. Struct. Mol. Biol.* **2013**, 20, 1122. DOI: 10.1038/nsmb.2638.
25. Liepelt, A.; Naarmann-de Vries, I.S.; Simons, N.; Eichelbaum, K.; Foehr, S.; Archer, S.K.; Castello, A.; Usadel, B.; Krijgsveld, J.; Preiss, T.; Marx, G.; Hentze, M.W.; Ostareck, D.H.; Ostareck-Lederer, A. *Mol. Cell. Proteomics* **2016**, 15, 2699. DOI: 10.1074/mcp.M115.056564.
26. Mitchell, S.F.; Jain, S.; She, M.; Parker, R. *Nat. Struct. Mol. Biol.* **2013**, 20, 127. DOI: 10.1038/nsmb.2468.
27. Matia-Gonzalez, A.M.; Laing, E.E.; Gerber, A.P. *Nat. Struct. Mol. Biol.* **2015**, 22, 1027. DOI: 10.1038/nsmb.3128.
28. Nandan, D.; Thomas, S.A.; Nguyen, A.; Moon, K.-M.; Foster, L.J.; Reiner, N.E. *PLoS One* **2017**, 12, e0170068. DOI: 10.1371/journal.pone.0170068.
29. Bunnik, E.M.; Batugedara, G.; Saraf, A.; Prudhomme, J.; Florens, L.; Le Roch, K.G. *Genome Biol.* **2016**, 17, 1. DOI: 10.1186/s13059-016-1014-0.
30. Lueong, S.; Merce, C.; Fischer, B.; Hoheisel, J.D.; Erben, E.D. *Mol. Microbiol.* **2016**, 100, 457. DOI: 10.1111/mmi.13328.
31. Reichel, M.; Liao, Y.; Rettel, M.; Ragan, C.; Evers, M.; Alleaume, A.-M.; Horos, R.; Hentze, M.W.; Preiss, T.; Millar, A.A. *Plant Cell* **2016**, 28, 2435. DOI: 10.1105/tpc.16.00562.
32. Marondedze, C.; Thomas, L.; Serrano, N.L.; Lilley, K.S.; Gehring, C. *Sci. Rep.* **2016**, 6, 1. DOI: 10.1038/srep29766.
33. Sysoev, V.O.; Fischer, B.; Frese, C.K.; Gupta, I.; Krijgsveld, J.; Hentze, M.W.; Castello, A.; Ephrussi, A. *Nature Communications* **2016**, 7, 1. DOI: 10.1038/ncomms12128.
34. Wessels, H.-H.; Imami, K.; Baltz, A.G.; Kolinski, M.; Beldovskaya, A.; Selbach, M.; Small, S.; Ohler, U.; Landthaler, M. *Genome Res.* **2016**, 26, 1000. DOI: 10.1101/gr.200386.115.
35. Despic, V.; Dejung, M.; Gu, M.; Krishnan, J.; Zhang, J.; Herzel, L.; Straube, K.; Gerstein, M.B.; Butter, F.; Neugebauer, K.M. *Genome Res.* **2017**, 27, 1184. DOI: 10.1101/gr.215954.116.
36. Hentze, M.W.; Castello, A.; Schwarzl, T.; Preiss, T. *Nature Reviews Molecular Cell Biology* **2018**, 19, 327. DOI: 10.1038/nrm.2017.130.
37. Perez-Perri, J.I.; Rogell, B.; Schwarzl, T.; Stein, F.; Zhou, Y.; Rettel, M.; Brosig, A.; Hentze, M.W. *Nature Communications* **2018**, 9, 1. DOI: 10.1038/s41467-018-

06557-8.
38. Garcia-Moreno, M.; Noerenberg, M.; Ni, S.; Jarvelin, A.I.; Gonzalez-Almela, E.; Lenz, C.E.; Bach-Pages, M.; Cox, V.; Avolio, R.; Davis, T.; Hester, S.; Sohier, T.J.M.; Li, B.; Heikel, G.; Michlewski, G.; Sanz, M.A.; Carrasco, L.; Ricci, E.P.; Pelechano, V.; Davis, I.; Fischer, B.; Mohammed, S.; Castello, A. *Mol. Cell* **2019**, 74, 196. DOI: 10.1016/j.molcel.2019.01.017.
39. Kamel, W.; Noerenberg, M.; Cerikan, B.; Chen, H.; Ja, A.I.; Kammoun, M.; Lee, J.Y.; Shuai, N.; Garcia-Moreno, M.; Andrejeva, A.; Deery, M.J.; Johnson, N.; Neufeldt, C.J.; Cortese, M.; Knight, M.L.; Lilley, K.S.; Martinez, J.; Davis, I.; Bartenschlager, R.; Mohammed, S.; Castello, A. *Mol. Cell* **2021**, 81, 2851. DOI: 10.1016/j.molcel.2021.05.023.
40. Benhalevy, D.; Anastasakis, D.G.; Hafner, M. *Nat. Methods* **2018**, 15, 1074. DOI: 10.1038/s41592-018-0220-y.
41. Sun, W.; Wang, N.; Liu, H.; Yu, B.; Jin, L.; Ren, X.; Shen, Y.; Wang, L. *Nat. Chem.* **2022**. DOI: 10.1038/s41557-022-01038-4.
42. Castello, A.; Frese, C.K.; Fischer, B.; Jarvelin, A.I.; Horos, R.; Alleaume, A.-M.; Foehr, S.; Curk, T.; Krijgsveld, J.; Hentze, M.W. *Nat. Protoc.* **2017**, 12, 2447. DOI: 10.1038/nprot.2017.106.
43. Castello, A.; Fischer, B.; Frese, C.K.; Horos, R.; Alleaume, A.-M.; Foehr, S.; Curk, T.; Krijgsveld, J.; Hentze, M.W. *Mol. Cell* **2016**, 63, 696. DOI: 10.1016/j.molcel.2016.06.029.
44. Smith, T.; Villanueva, E.; Queiroz, R.M.L.; Dawson, C.S.; Elzek, M.; Urdaneta, E.C.; Willis, A.E.; Beckmann, B.M.; Krijgsveld, J.; Lilley, K.S. *Curr. Opin. Chem. Biol.* **2020**, 54, 70. DOI: 10.1016/j.cbpa.2020.01.009.
45. Bao, X.; Guo, X.; Yin, M.; Tariq, M.; Lai, Y.; Kanwal, S.; Zhou, J.; Li, N.; Lv, Y.; Pulido-Quetglas, C.; Wang, X.; Ji, L.; Khan, M.J.; Zhu, X.; Luo, Z.; Shao, C.; Lim, D.-H.; Liu, X.; Li, N.; Wang, W.; He, M.; Liu, Y.-L.; Ward, C.; Wang, T.; Zhang, G.; Wang, D.; Yang, J.; Chen, Y.; Zhang, C.; Jauch, R.; Yang, Y.-G.; Wang, Y.; Qin, B.; Anko, M.-L.; Hutchins, A.P.; Sun, H.; Wang, H.; Fu, X.-D.; Zhang, B.; Esteban, M.A. *Nat. Methods* **2018**, 15, 213. DOI: 10.1038/nmeth.4595.
46. Huang, R.; Han, M.; Meng, L.; Chen, X. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, 115, E3879. DOI: 10.1073/pnas.1718406115.
47. Burger, K.; Muehl, B.; Kellner, M.; Rohrmoser, M.; Gruber-Eber, A.; Windhager, L.; Friedel, C.C.; Doelken, L.; Eick, D. *RNA Biol.* **2013**, 10, 1623. DOI: 10.4161/rna.26214.
48. Tani, H.; Akimitsu, N. *RNA Biol.* **2012**, 9, 1233. DOI: 10.4161/rna.22036.
49. Chomczynski, P.; Sacchi, N. *Nat. Protoc.* **2006**, 1, 581. DOI: 10.1038/nprot.2006.83.
50. Queiroz, R.M.L.; Smith, T.; Villanueva, E.; Marti-Solano, M.; Monti, M.; Pizzinga, M.; Mirea, D.-M.; Ramakrishna, M.; Harvey, R.F.; Dezi, V.; Thomas, G.H.; Willis, A.E.; Lilley, K.S. *Nat. Biotechnol.* **2019**, 37, 169. DOI: 10.1038/s41587-018-0001-2.
51. Trendel, J.; Schwarzl, T.; Horos, R.; Prakash, A.; Bateman, A.; Hentze, M.W.; Krijgsveld, J. *Cell* **2019**, 176, 391. DOI: 10.1016/j.cell.2018.11.004.
52. Urdaneta, E.C.; Vieira-Vieira, C.H.; Hick, T.; Wessels, H.-H.; Figini, D.; Moschall, R.; Medenbach, J.; Ohler, U.; Granneman, S.; Selbach, M.; Beckmann, B.M. *Nature Communications* **2019**, 10, 1. DOI: 10.1038/s41467-019-08942-3.
53. Smirnov, A.; Foerstner, K.U.; Holmqvist, E.; Otto, A.; Guenster, R.; Becher, D.; Reinhardt, R.; Vogel, J. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, 113, 11591. DOI: 10.1073/pnas.1609981113.
54. Asencio, C.; Chatterjee, A.; Hentze, M.W. *Life Science Alliance* **2018**, 1, e201800088. DOI: 10.26508/lsa.201800088.
55. Koo, K.; Foegeding, P.M.; Swaisgood, H.E. *Biotechnol. Tech.* **1998**, 12, 549. DOI: 10.1023/a:1008859632378.
56. Shchepachev, V.; Bresson, S.; Spanos, C.; Petfalski, E.; Fischer, L.; Rappsilber, J.; Tollervey, D. *Mol. Syst. Biol.* **2019**, 15, e8689. DOI: 10.15252/msb.20188689.
57. Qin, W.; Cho, K.F.; Cavanagh, P.E.; Ting, A.Y. *Nat. Methods* **2021**, 18, 133. DOI: 10.1038/s41592-020-01010-5.
58. Qin, W.; Myers, S.A.; Carey, D.K.; Carr, S.A.; Ting, A.Y. *Nature Communications* **2021**, 12, 1. DOI: 10.1038/s41467-021-25259-2.
59. Yan, S.; Zhao, D.; Wang, C.; Wang, H.; Guan, X.; Gao, Y.; Zhang, X.; Zhang, N.; Chen, R. *Anal. Chim. Acta* **2021**, 1168, 338609. DOI: 10.1016/j.aca.2021.338609.
60. Harris, M.E.; Christian, E.L. 2009. RNA Crosslinking methods. In *Methods In Enzymology, Vol 468: Biophysical, Chemical, And Functional Probes Of Rna Structure, Interactions And Folding, Pt A*. D. Herschalag, editor. 127.
61. Zhang, Y.; Chan, P.P.Y.; Herr, A.E. *Angewandte Chemie-International Edition* **2018**, 57, 2357. DOI: 10.1002/anie.201711441.
62. Aw, J.G.A.; Shen, Y.; Wilm, A.; Sun, M.; Lim, X.N.; Boon, K.-L.; Tapsin, S.; Chan, Y.-S.; Tan, C.-P.; Sim, A.Y.L.; Zhang, T.; Susanto, T.T.; Fu, Z.; Nagarajan, N.; Wan, Y. *Mol. Cell* **2016**, 62, 603. DOI: 10.1016/j.molcel.2016.04.028.
63. Zhang, Z.; Liu, T.; Dong, H.; Li, J.; Sun, H.; Qian, X.; Qin, W. *Nucleic Acids Res.* **2021**, 49, e65. DOI: 10.1093/nar/gkab156.
64. Luo, H.; Tang, W.; Liu, H.; Zeng, X.; Ngai, W.S.C.; Gao, R.; Li, H.; Li, R.; Zheng, H.; Guo, J.; Qin, F.; Wang, G.; Li, K.; Fan, X.; Zou, P.; Chen, P.R. *Angewandte Chemie-International Edition* **2022**, 61, e202202008. DOI: 10.1002/anie.202202008.

