Genomics

https://doi.org/10.1038/s41592-023-01904-0

Identifying the structures of individual RNA isoforms inside cells

Julia C. Kenyon & Johnson Mak

Nano-DMS-MaP focuses in on the structures of individual RNA isoforms, enabling direct examination of the structural diversity of different RNAs inside cells.

A new technique by Bohn and Gribling-Burrer et al. published in this issue combines chemical probing (using dimethyl sulfate, DMS), an ultraprocessive reverse transcriptase and long-read sequencing to obtain isoform-resolved RNA structural information. This approach, termed Nano-DMS-MaP, stands to revolutionize our understanding of RNA structure and function[1].

The functions of RNA molecules are numerous: alongside their well-understood role in encoding proteins, RNAs themselves act at many regulatory levels within the cell. RNA structure, which is determined by sequence, is crucial for this regulatory activity. Even small RNA sequence differences can lead to major structural changes. The biological roles of a given RNA can therefore vary according to sequence differences resulting from transcriptional processes, such as start site selection, and post transcriptional processes, such as alternative splicing or RNA processing[2]. Such sequence variations occur in both mRNAs and noncoding RNAs. RNA structure affects how a pre-mRNA is spliced, for example, by changing the accessibility of sequences that bind splicing regulatory proteins. It can also change the functions of the mature mRNA post-splicing[2]. However, the potential roles of RNA structure before, during and after splicing have been difficult to disentangle from one another because of an absence of techniques facilitating the structural interrogation of individual RNA isoforms.

One of the most common ways to study RNA structure is chemical probing, which uses small molecules that preferentially react with nucleotides that have a particular structural conformation. This approach is particularly powerful because it can be used to measure RNA structure in cells at endogenous expression levels. It is also able to map the structures of many RNAs simultaneously, up to the genome-wide level. This is achieved by reading the positions of chemical adducts as mutations in the cDNA introduced by error-prone reverse transcription[3]. This readout method typically produces short cDNA molecules owing to the low processivity of reverse transcriptase, particularly on highly modified RNA templates or templates with high levels of structure. Because such short reads cannot be unambiguously mapped to individual RNA isoforms, most structure-probing efforts produce a hybrid signal for all isoforms of a particular RNA. This has made it difficult to differentiate the structural landscape of isoforms from one another or to see how the structures of isoforms lend specific functions to any RNA molecule. Computational algorithms to cluster the chemical probing data into ensembles have been applied to identify structural heterogeneity in sliding windows across the long RNAs, such as HIV-1 RNA, without the capacity to assign precise structure–function relationships to RNA isoforms[4]. Nano-DMS-MaP advances our ability to understand RNA structural heterogeneity by coupling the use of an ultraprocessive reverse transcriptase with long-read sequencing[1], resulting in a direct readout of the RNA structure of each different RNA isoform.

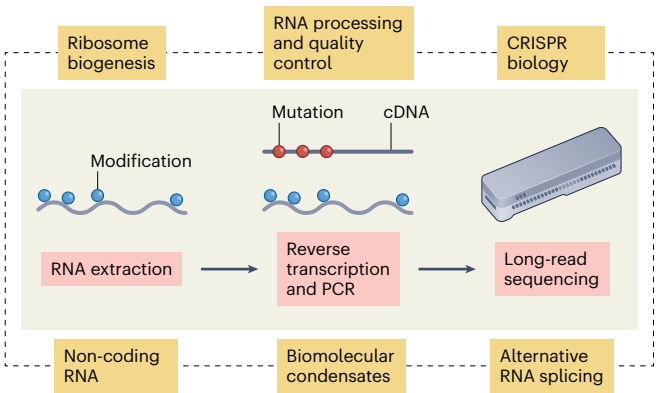

Fig. 1 | Areas standing to benefit from Nano-DMS-MaP. Areas include (but are not limited to) RNA processing and quality control, CRISPR biology, alternative RNA splicing, biomolecular condensate, non-coding RNA and ribosome biogenesis.

Bohn and Gribling-Burrer et al. use DMS to interrogate the conformation of adenine, cytosine and uracil residues within the RNA. DMS methylates these bases at the Watson–Crick face when they are single-stranded. Marathon reverse transcriptase[5] lengthens cDNA production beyond that of conventional enzymes, and DMS modifications are read as mutations using nanopore long-read sequencing. By optimizing the DMS concentration and incorporating data filtration steps to increase signal-to-noise ratio, Bohn, Gribling-Burrer et al. were able to produce accurate mutational readouts spanning strands of 4,000 nucleotides – long enough to assign RNA structural information to 16 RNA isoforms in HIV-1 infected cells[4]. The authors used this development to examine the structural complexity of HIV-1 RNAs inside cells and viral particles. HIV-1 makes many of its different proteins by alternative splicing, producing multiple different RNA isoforms. Much of the 5′-UTR sequence is common to the different spliced or genomic viral RNA species, and part of this sequence has been found to control viral packaging, in which viral genomic RNA is selected almost exclusively for incorporation into viral particles. As much of the packaging signal sequence is not unique to the genomic RNA, it has previously been suggested that it is the structural context of this sequence that controls packaging and enables the virus to select the full-length HIV RNA for incorporation during viral assembly. Nano-DMS-MaP has now enabled the direct visualization of isoform-specific structural

differences inside HIV-1 infected cells. The data show a structural rearrangement of the RNA upon loss of the first intron, differentiating the genomic and spliced RNAs in a clear example of how sequence differences between similar RNAs can affect their RNA structures and consequently their functions[1].

Although the technique is not unique in using nanopore-based sequencing methods to differentiate RNA structures within cells, it stands out in using DNA-based rather than RNA-based sequencing to do so. This future-proofs it to a certain extent, as direct RNA sequencing methods rely on the continual updating of RNA modification detection algorithms as nanopore chemistries evolve. Nano-DMS-MaP also enhances the efficiency and lowers the cost of conducting experiments as a result of the higher throughput of DNA across the flow cells. In examining cDNAs rather than RNAs directly, however, Nano-DMS-MaP does not differentiate structural differences resulting from natural RNA modifications. As with all RNA structural analysis techniques, Nano-DMS-MaP is also limited by the accuracy of the algorithms used for modeling. As RNA structure modeling continues to improve, so will the power of Nano-DMS-MaP. The technique could be used alongside algorithms to deconvolve RNA structural ensembles to assess the structural flexibility of individual isoforms[4].

Nano-DMS-MaP has the power to structurally interrogate the diverse RNA species made by other important viruses such as coronaviruses and influenza viruses. As cells contain a complex mixture of mRNA and non-coding RNA isoforms, Nano-DMS-MaP will permit better understanding of diverse RNA-driven processes of cellular importance. The dysregulation of RNA-driven processes has been shown to have multifarious pathological effects[6]. Some of the critical examples include the role of RNA structures in cancer biology[7], the control of cellular partitioning by liquid–liquid phase separation[8], the generation of subcellular polarity by RNA structure-controlled transport mechanisms[9] and regulation of protein expression levels via different RNA-centered processes[10] (Fig. 1).

Nano-DMS-MaP stands to enable these discoveries. In doing so, it will revolutionize our understanding of cell biology and the development of RNA therapeutics.

Julia C. Kenyon [1] ✉ **& Johnson Mak** [2] ✉

[1]Department of Pathology, Division of Virology, University of Cambridge, Addenbrookes Hospital, Cambridge, UK. [2]Institute for Glycomics, Griffith University, Gold Coast, Queensland, Australia.
✉e-mail: jck33@cam.ac.uk; j.mak@griffith.edu.au

Published online: 19 May 2023

References

1. Bohn, P., Gribling-Burrer, A.-S., Ambi, U. B. & Smyth, R. P. *Nat. Methods* https://doi.org/10.1038/s41592-023-01862-7 (2023).
2. Marasco, L. E. & Kornblihtt, A. R. *Nat. Rev. Mol. Cell Biol.* **24**, 242–254 (2023).
3. Smola, M. J. & Weeks, K. M. *Nat. Protoc.* **13**, 1181–1195 (2018).
4. Tomezsko, P. J. et al. *Nature* **582**, 438–442 (2020).
5. Guo, L. T. et al. *J. Mol. Biol.* **432**, 3338–3352 (2020).
6. Scotti, M. M. & Swanson, M. S. *Nat. Rev. Genet.* **17**, 19–32 (2016).
7. Ouyang, J. et al. *Br. J. Cancer* **126**, 1113–1124 (2022).
8. Roden, C. & Gladfelter, A. S. *Nat. Rev. Mol. Cell Biol.* **22**, 183–195 (2021).
9. Ghosh, S., Marchand, V., Gáspár, I. & Ephrussi, A. *Nat. Struct. Mol. Biol.* **19**, 441–449 (2012).
10. Mauger, D. M. et al. *Proc. Natl Acad. Sci. USA* **116**, 24075–24083 (2019).

Competing interests

The authors declare no competing interests.