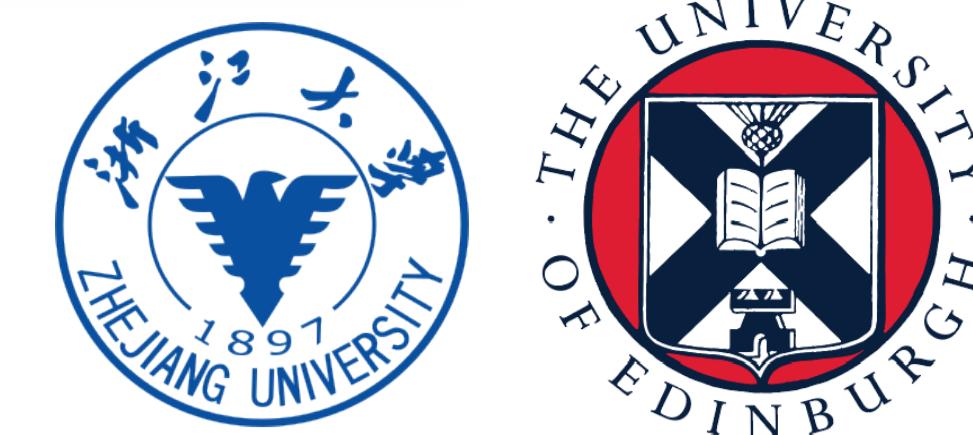


# In-silico modeling of RBP binding by deep learning disentangles directional m6A-RBP interaction dynamics

Jianche Liu<sup>1,2,4#</sup>, Xinlu Zhu<sup>1,2,4#</sup>, Yang Yin<sup>1,2,4#</sup>, Zhoutong Xu<sup>1,2,4</sup>, Jialin He<sup>3,4,5</sup>, and Xushen Xiong<sup>3,4,5\*</sup>

1. Zhejiang University-University of Edinburgh Institute (ZJU-UoE Institute), Zhejiang University School of Medicine, Zhejiang University, Haining 314400, China
2. Edinburgh Medical School: Biomedical Sciences, College of Medicine and Veterinary Medicine, The University of Edinburgh, Edinburgh, UK
3. The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 311121, China
4. Liangzhu Laboratory, School of Medicine, Zhejiang University, Hangzhou 311121, China
5. State Key Laboratory of Transvascular Implantation Devices, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 311121, China

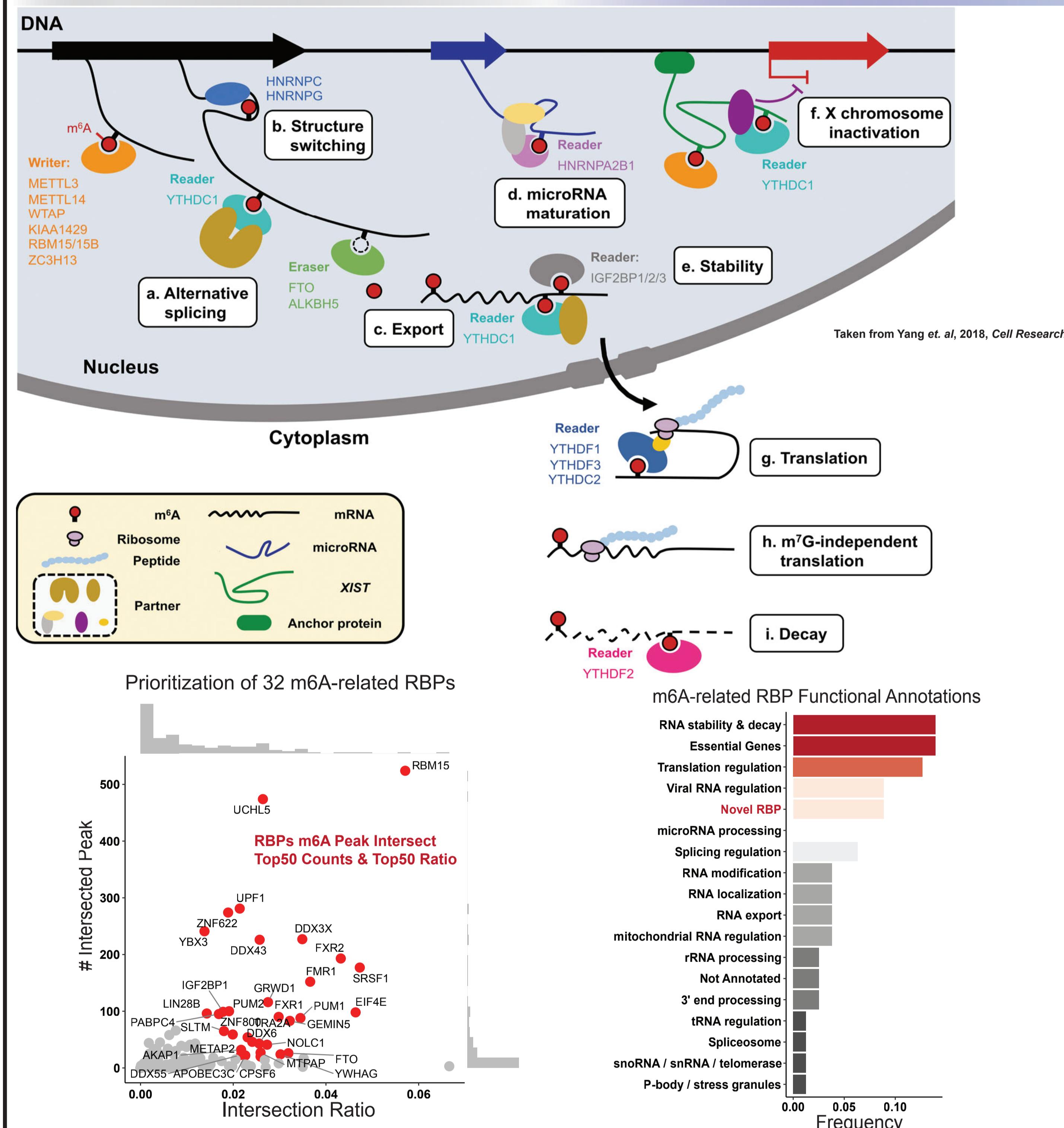


# Abstract

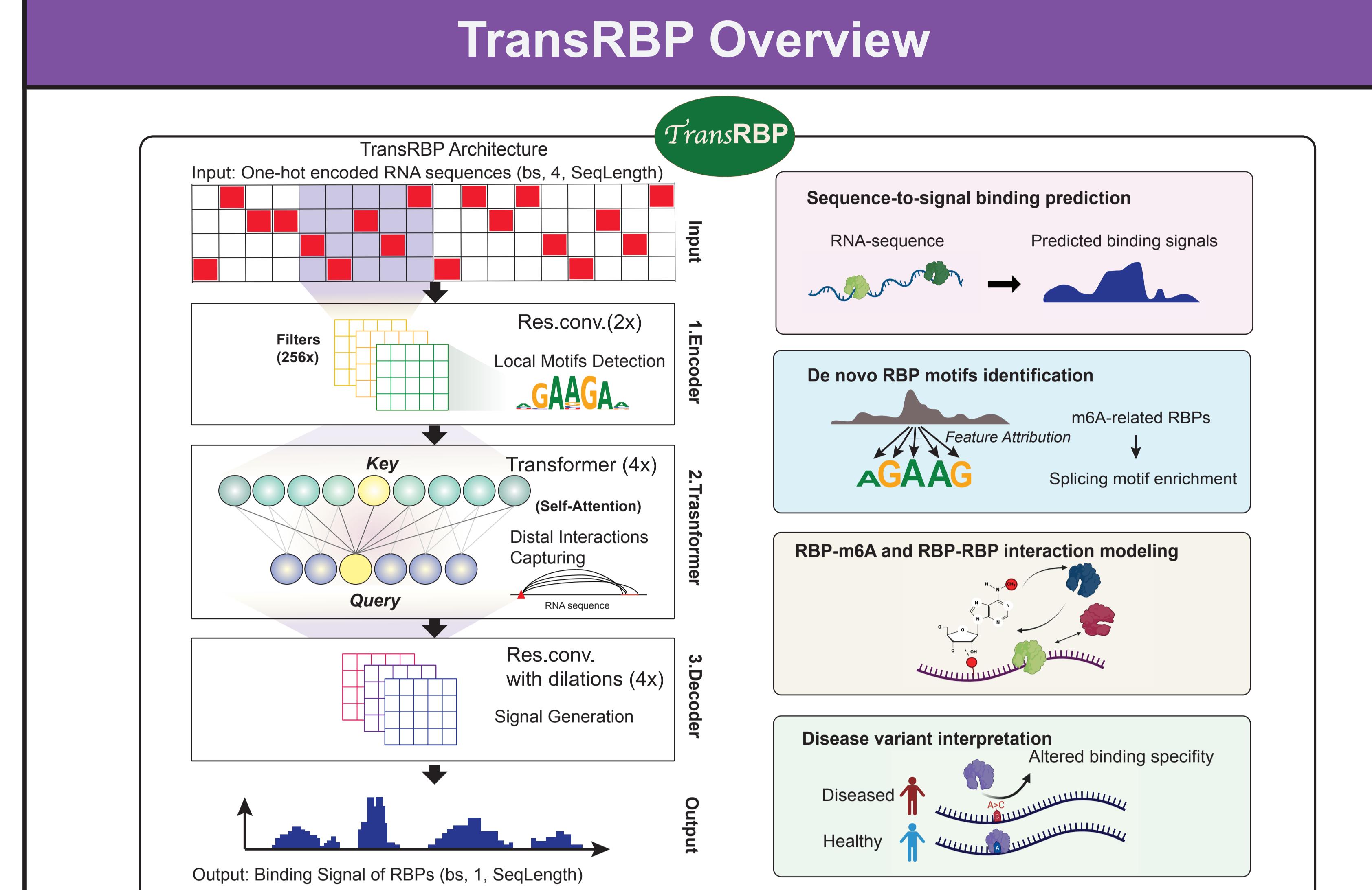
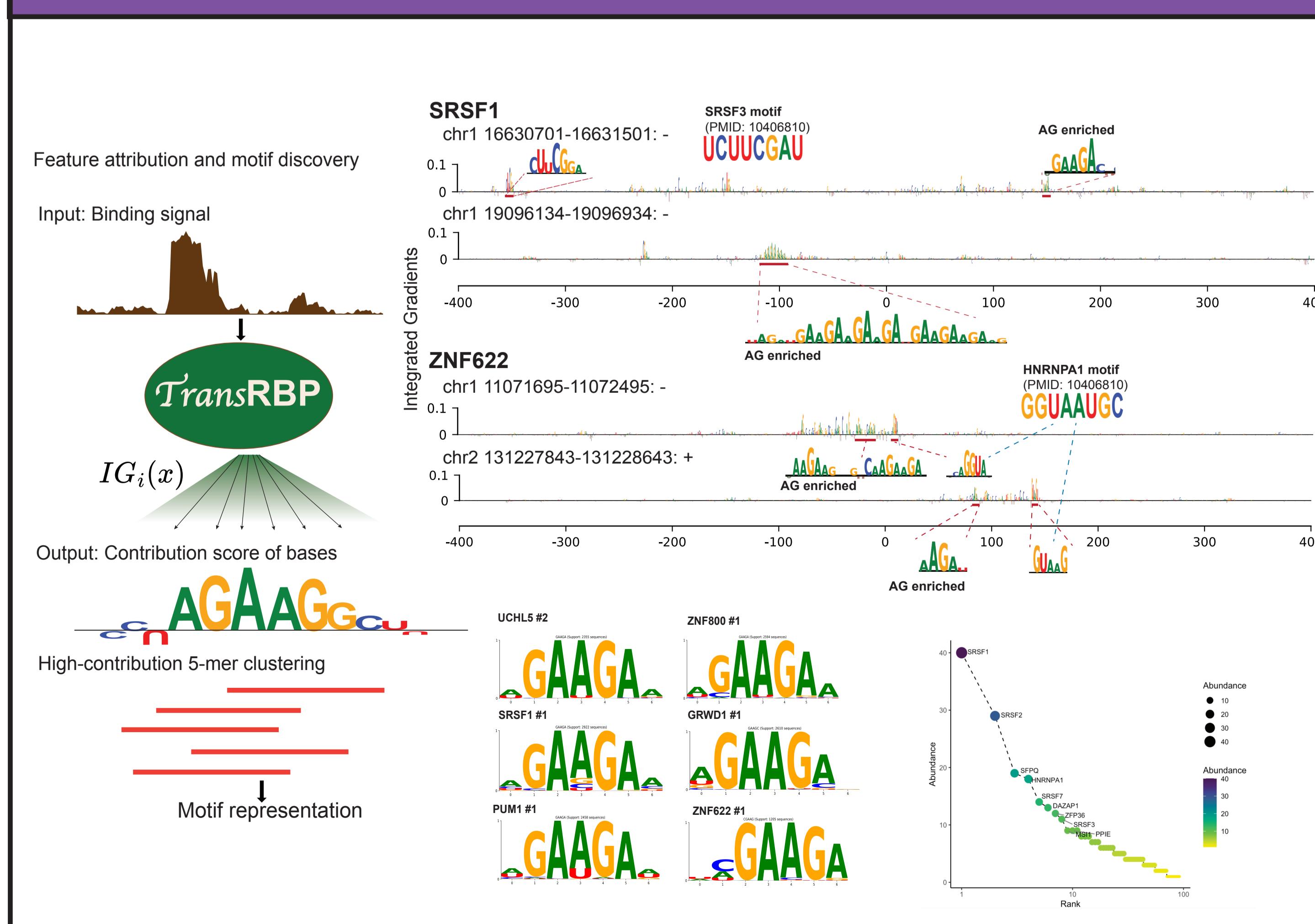
N6-methyladenosine (m6A) is the most prevalent internal mRNA modification that plays an essential role in regulating the RNA life cycle. While RNA-binding proteins (RBPs) have been revealed to interact with m6A through reading, writing, and erasing, intricate mechanisms underlying the interaction between m6A and RBPs need further clarification. In this study, we first prioritize 32 RBPs that are closely associated with m6A modification through peak intersection specificity. Surprisingly, the interaction dynamics of these RBPs with m6A remain largely unexplored. To address this issue, we developed TransRBP, a CNN-transformer-based deep learning (DL) framework, which predicts the base-resolution binding of RBPs from RNA sequences. TransRBP outperforms the state-of-the-art DL framework and approaches to the accuracy level of the eCLIP experimental replicate. Through the model interpretation by the integrated-gradient attribution, we uncover striking enrichments of RNA splicing motifs in these RBPs. Motif matching by the database strongly suggests the interaction in a subset of these RBPs with SRSF family. Additionally, by in silico saturated mutation surrounding the m6A site, we disentangle the influence of m6A on RBP binding behaviors. Specifically, our model recapitulates the repelling of PUM2 binding to m6A and systematically evaluates the repelling ability of m6A across other RBPs. Together, our findings advance the understanding of m6A-RBP interactions through in-silico modeling, as well as offering new insights into mechanistic interpretation in m6A-RBP-splicing dynamics and the underlying genetic diseases.

# Background

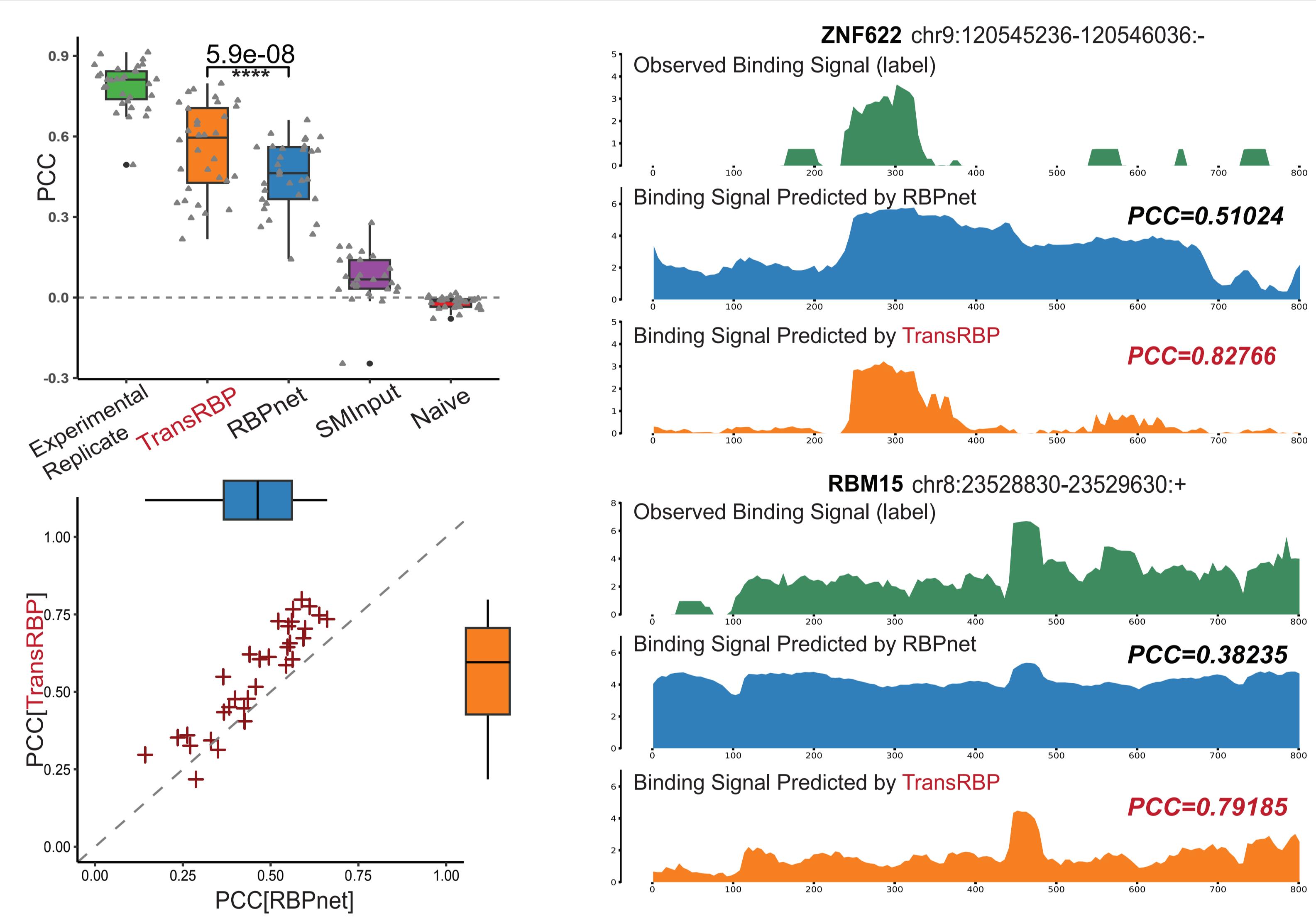
**RBPs have associated with m6A dynamics, but the mechanism remain largely elusive.**



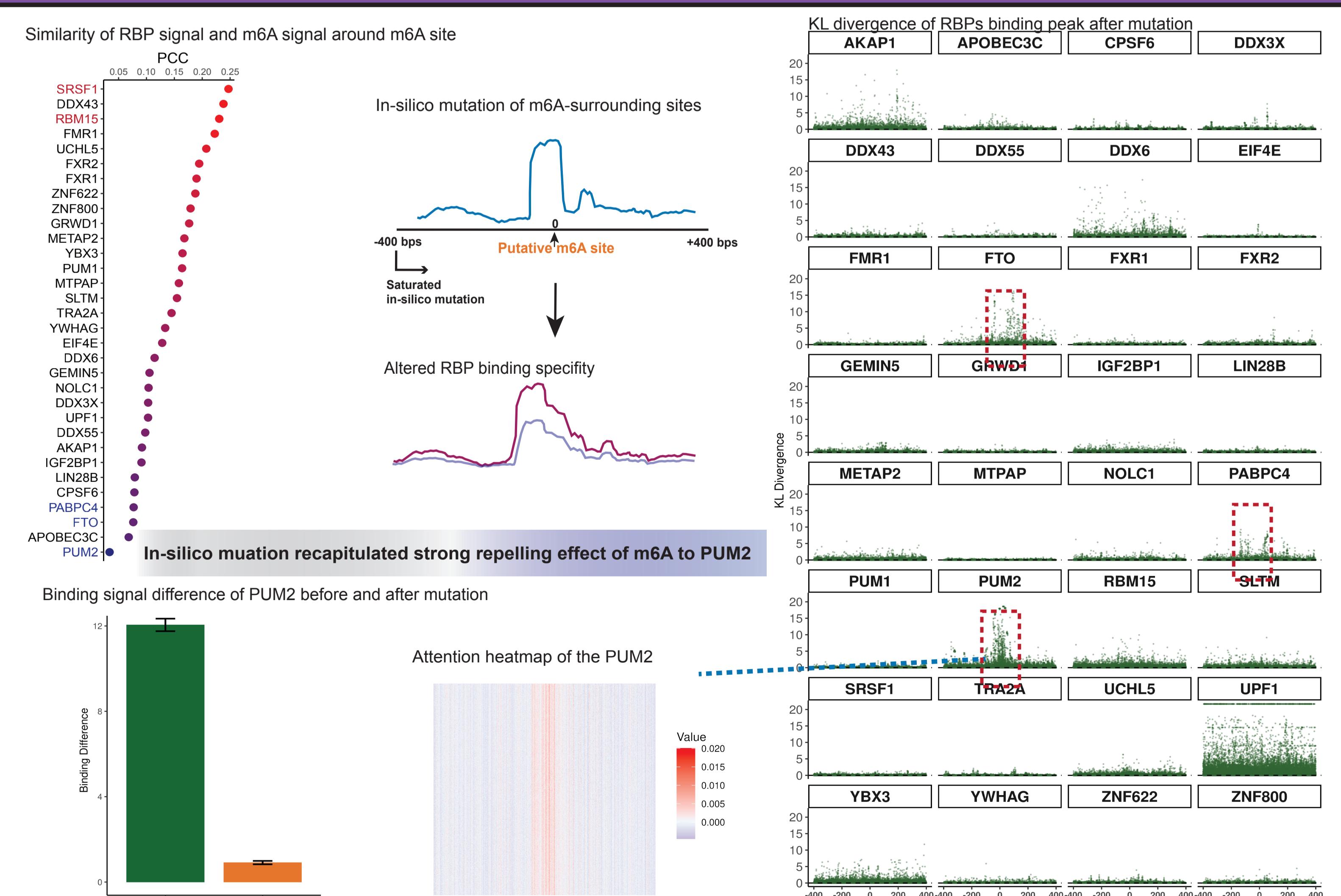
# De novo RBP Motif Discovery enriched m6A-related RBPs in splicing motifs



# Model performance evaluation



# In-silico mutation disentangled RBP-m6A interactions



# Conclusion

**Our sequence-to-signal model for RBP binding prediction outperform the state-of-art model.  
Feature attribution has enriched the m6A-related RBP to splicing motifs.  
In-silico mutation analysis recapitulated the repelling effect of the m6A to PUM2.**