Nucleic Acids Research, 2024, 1–9
https://doi.org/10.1093/nar/gkae1117
Database issue

RASP v2.0: an updated atlas for RNA structure probing data

Kunting Mu[1,2,3,†], Yuhan Fei[1,2,3,*,†], Yiran Xu[1,2,3] and Qiangfeng Cliff Zhang[1,2,3,*]

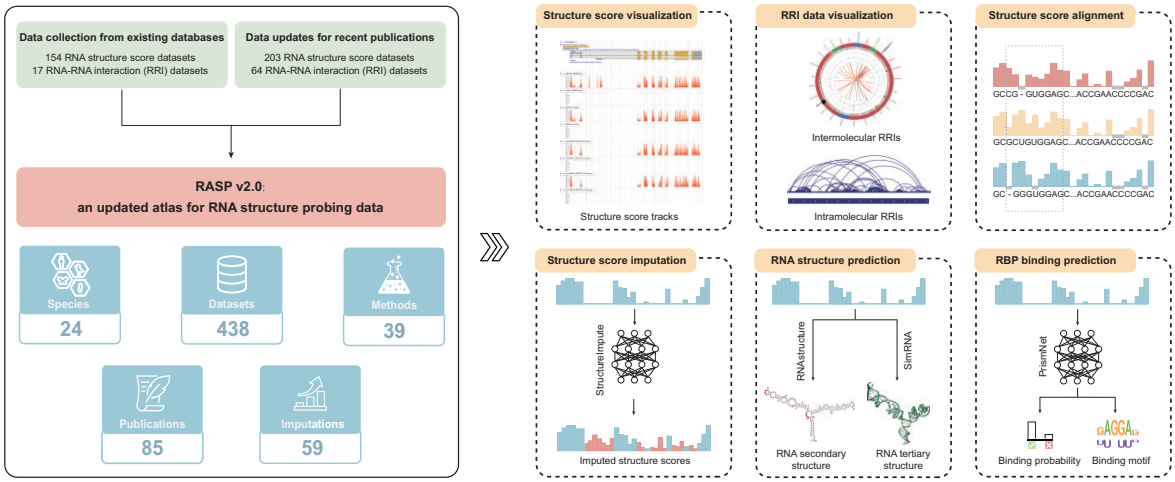
[1]MOE Key Laboratory of Bioinformatics, Center for Synthetic and Systems Biology, School of Life Sciences, Tsinghua University, Beijing 100084, China
[2]Beijing Advanced Innovation Center for Structural Biology & Frontier Research Center for Biological Structure, School of Life Sciences, Tsinghua University, Beijing 100084, China
[3]Tsinghua-Peking Center for Life Sciences, Beijing 100084, China

[*]To whom correspondence should be addressed. Tel: +86 10 627 92786; Fax: +86 10 627 92786; Email: yuhan_fei@outlook.com.
Correspondence may also be addressed to Qiangfeng Cliff Zhang. Tel: +86 10 627 95823; Fax: +86 10 627 95823; Email: qczhang@tsinghua.edu.cn
[†]The first two authors should be regarded as Joint First Authors.

Abstract

RNA molecules function in numerous biological processes by folding into intricate structures. Here we present RASP v2.0, an updated database for RNA structure probing data featuring a substantially expanded collection of datasets along with enhanced online structural analysis functionalities. Compared to the previous version, RASP v2.0 includes the following improvements: (i) the number of RNA structure datasets has increased from 156 to 438, comprising 216 transcriptome-wide RNA structure datasets, 141 target-specific RNA structure datasets, and 81 RNA–RNA interaction datasets, thereby broadening species coverage from 18 to 24, (ii) a deep learning-based model has been implemented to impute missing structural signals for 59 transcriptome-wide RNA structure datasets with low structure score coverage, significantly enhancing data quality, particularly for low-abundance RNAs, (iii) three new online analysis modules have been deployed to assist RNA structure studies, including missing structure score imputation, RNA secondary and tertiary structure prediction, and RNA binding protein (RBP) binding prediction. By providing a resource of much more comprehensive RNA structure data, RASP v2.0 is poised to facilitate the exploration of RNA structure-function relationships across diverse biological processes. RASP v2.0 is freely accessible at http://rasp2.zhanglab.net/.

Graphical abstract



Introduction

Previous studies have shown that RNA molecules participate in various biological processes, including transcription, splicing, localization, translation, and degradation(1–5). These complex cellular activities heavily depend on RNA's ability to fold into intricate structures. These structures also provide potential binding sites or pockets, allowing RNA molecules to interact with other RNAs, RNA-binding proteins (RBPs), or small molecules, thereby making RNA molecules as promising target for disease treatment(6–9). However, RNA structures are dynamic and may undergo conformational changes depending on solvent conditions, resulting in different binding states under variable cellular context (10–13). This complexity makes traditional protein structure probing methods, such as X-ray crystallography (14), nuclear magnetic resonance (NMR) spectroscopy(15), and cryo electron microscopy

Received: September 15, 2024. **Revised:** October 16, 2024. **Editorial Decision:** October 18, 2024. **Accepted:** November 14, 2024

(Cryo-EM), which are primarily used to determine a single conformation, challenging to be applied to RNA structure resolving.

In recent years, numerous sequencing-based methods have been developed to probe RNA structures in a high-throughput fashion, many of which can be applied in cells (16). These structure probing methods can be categorized into footprinting-based and proximity ligation-based methods (6,17). Footprinting-based methods utilize reagents or enzymes to selectively modify or digest RNA molecules based on their structure, generating structural signals that represent the probabilities for a nucleotide is in single-stranded conformation. DMS-seq (18), icSHAPE(19), SHAPE-MaP (20), Structure-seq (21) and RL-seq (22), are representative examples of these methods. Proximity ligation-based methods employ reagents or proteins to capture spatially proximate RNA molecules, directly identifying intramolecular RNA–RNA interactions (RRIs) (within the same RNA molecule) or intermolecular RRIs (between different RNA molecules). PARIS (23), SPLASH (24), RIC-seq (25) and KARR-seq (26) are representative examples of these methods. These high-throughput methods have produced extensive structural data both *in vitro* and *in vivo* across different species, providing invaluable resources for studying RNA structure.

A few databases, including RMDB (27), RSVdb (28) and RASP (29), have been developed as repositories for footprinting-based structure probing data. Among which, RMDB primarily focuses on RNA structures identified by low-throughput methods, while RSVdb specializes in RNA structures detected using the DMS reagent. RASP stands out as one of the first databases to include RNA structure probing data generated from multiple high-throughput methods. In addition, other databases such as RISE (30), RAID v2.0 (31) and NPInter v5.0 (32) have been created to curate RRIs based on proximity ligation-based structure probing data. However, these databases primarily focus on intermolecular RRIs, often overlooking intramolecular interactions. Overall, the existing RNA structure databases primarily focus on a single type of structure probing data, and usually provide very limited capacities for online structural analysis.

Here we present RASP v2.0, a substantially updated version of the RASP database featuring a much larger collection of datasets along with enhanced online structural analysis functionalities. Specifically, RASP v2.0 contains 438 RNA structure datasets derived from 85 publications, and integrates RNA structure data generated from 39 experimental methods, encompassing both footprinting-based and proximity ligation-based techniques. RASP v2.0 improved the structure score coverage of 59 transcriptome-wide RNA structure datasets using a deep learning-based model, which facilitate the study of low-abundance RNAs. RASP v2.0 also offers three online analysis modules for structure-related analyses. In summary, RASP v2.0 is a powerful database that provides comprehensive RNA structure probing data and supports various online structural analysis functionalities.

Materials and methods

Data collection

To develop a more comprehensive RNA structure database, we integrated RNA structure datasets from our previously released databases, i.e. RASP(29) and RISE(30). In addition, we collected and curated all published RNA structure datasets, covering research up to March 2024. As shown in Figure 1, RASP v2.0 incorporates RNA structure data derived from both footprinting-based and proximity ligation-based structure probing methods. Specifically, we searched the NCBI PubMed database (33) using 'RNA structure' and 'RNA–RNA interaction' as keywords and manually identified all publications with downloadable RNA structure datasets. Processed data from the NCBI GEO database (34) or supplementary files in these publications are incorporated into RASP v2.0 following the approach used in RASP and RISE. Finally, we curated information from these datasets, including experimental methods, reagents, species, cell lines, experimental conditions, DOI number, and other relevant details to build RASP v2.0. Detailed information is available in Supplementary Table S1 and S2.

Structure score imputation

RNA structural profiles from transcriptome-wide RNA structure data usually contain missing signals for low-abundance transcripts. To improve data quality, particularly in datasets with low structure score coverage, we used a deep learning-based model, StructureImpute (35), to impute the missing signals. Here we focused on transcriptome-wide RNA structure datasets with scores ranging from 0 to 1 that fulfill the StructureImpute requirements. We categorized the datasets into A/C-only data and A/U/T/C/G data based on the nucleotides detected by different methods, and we only focused on signals for the corresponding positions.

To enable imputation across different cellular context, we fine-tuned the meta model using specific datasets corresponding to each condition. The meta model was trained on a diverse mixture of icSHAPE datasets, including HEK293 whole cell (*in vivo* and *in vitro*), HEK293 chromatin (*in vivo*), HEK293 nucleosome (*in vivo*), and HEK293 cytoplasmic (*in vivo*), and can be downloaded from https://github.com/Tsinghua-gongjing/StructureImpute. Specifically, we sliced RNA sequences into 100 nt fragments and collected those with 100% structure score coverage for model finetuning. Fragments were then clustered using BLASTn (36) with an E-value equal to 10, and split into training and validation sets at a 7:3 ratio. During the training and validation process, we randomly masked 30% nucleotides and used the flanking region for missing signal imputation as described in StructureImpute. Next, we used the default parameters to finetune the meta model, with a batch size of 800 and a learning rate of 1E-5 for up to 100 epochs, stopping when no improvement was observed after 20 epochs. Finally, transcripts with missing structure scores are iteratively imputed following the strategy used in StructureImpute. In each iteration, transcripts were segmented using a window size of 100 nt and a step size of 10 nt and those fragments with coverage >50% would be imputed. The imputed structure scores from each iteration would be used as input for the next round until the data coverage reaches 80% for at most eight iterations. To evaluate the performance of our model, we calculated the Pearson correlation coefficient between the true and imputed structure scores for all positions with missing signals in the validation set, following the strategy mentioned in StructureImpute.

Database implementation

RASP v2.0 was developed using Django for the back end and using HTML, CSS and JavaScript for the front end. RNA

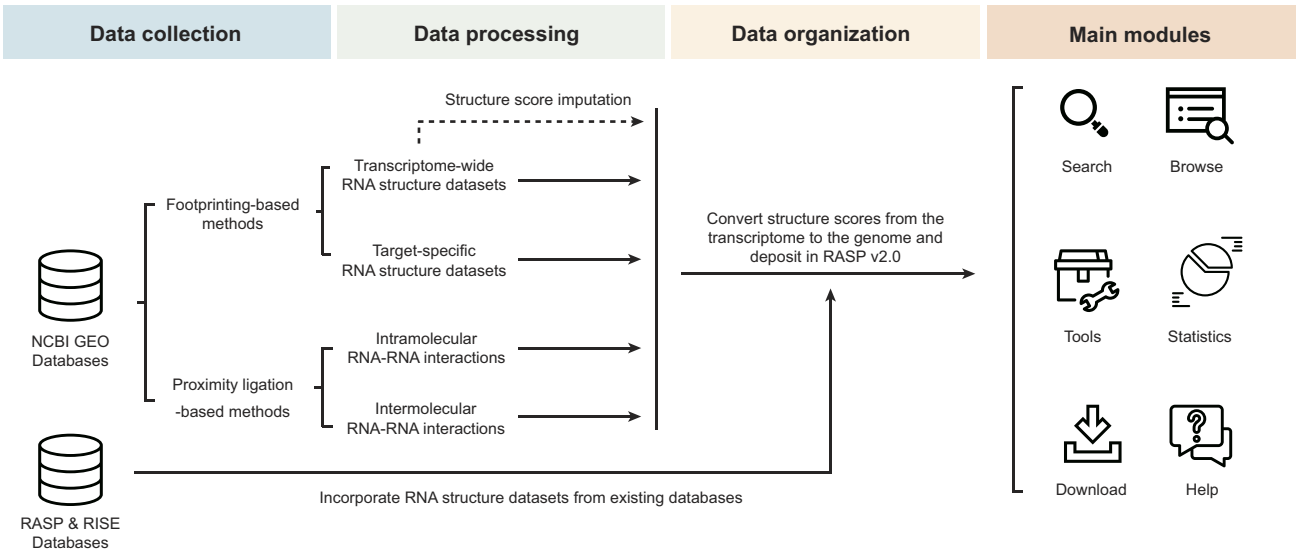


Figure 1. Flowchart of RASP v2.0.

structure data are stored in a MySQL database. Visualizations of RNA structure data, including structure scores, intermolecular RRIs, and intramolecular RRIs, are generated using JBrowser2 (37), IGV-web (38) and Circos plot (39), respectively. VARNA (40), forna (41) and Molstar (42) were utilized to visualize RNA secondary and tertiary structures. StructureImpute(35), RNAstructure (43), SimRNA (44) and PrismNet (45) were utilized to perform structure score imputation, RNA secondary structure, RNA tertiary structure, and RBP binding prediction. RASP v2.0 is installed on a workstation with four CPUs and three 1080 Ti GPUs, publicly accessible at http://rasp2.zhanglab.net/ with no login credentials required.

Results

Expanded structure probing data

We have collected 216 transcriptome-wide RNA structure datasets, 141 target-specific RNA structure datasets, and 81 proximity ligation-based probing datasets. This collection increases both the quantity and variety of RNA structure probing data compared to the previous version (Figure 2A). Specifically, we updated the transcriptome-wide RNA structure datasets from RASP (Figure 2B) and added target-specific RNA datasets that were not previously included (Figure 2C). We have also incorporated 81 proximity ligation-based probing datasets (containing 45 974 567 intramolecular RRIs and 4 628 511 intermolecular RRIs) (Figure 2D). In contrast to the diverse structure data obtained from footprinting-based structure probing methods (Figure 2E), the RRI data are predominantly derived from *Homo sapiens*, *Mus musculus* and different virus (Figure 2F). Further analysis shown that most of the intermolecular RRI data provided in RASP v2.0 are mRNA-related (Figure 2G). More detailed information is available in Supplementary Table S3.

Improved coverage for low coverage structure probing data

Some of the transcriptome-wide structure probing datasets exhibit low data coverage, limiting analyses of RNA structure and functions. Here we used a deep learning-based model, StructureImpute (35), to obtain insights of the missing signals. As shown in Figure 3A, we filtered 59 valid datasets from 218 transcriptome-wide structure datasets that were suitable for missing structure score imputation. These datasets exhibited <80% structure score coverage and included >500 transcripts. Next, we individually fine-tuned the meta model from StructureImpute on these 59 datasets. The results demonstrated an averaged Pearson correlation of 0.725 on the validation set for the 59 datasets, compared to 0.498 without fine-tuning, indicating that the imputed structure scores accurately represent the experimental probing data (Figure 3B). Finally, we applied these fine-tuned models for missing structure score imputation, increasing the average structure score coverage for these datasets from 48.7% to 71.0% (Figure 3C). We also present the improvements of structure score coverage for each dataset (Figure 3D).

As shown in the example regions within signal recognition particle (SRP) RNA from the icSHAPE_HEK293 dataset, we observed a 0.762 Pearson correlation coefficient between the ground truth and the imputed structure scores for all nucleotides with missing signals (Figure 3E). By building structural models with/without SHAPE reactivity score constraints, we found that the structure generated with imputed SHAPE reactivity score constraints is consistent with the one generated with original SHAPE reactivity score constraints (Figure 3G, H), and is similar to the known RNA secondary structure of SRP RNA from RNAcentral (46). However, these structures were entirely different from those based on missing SHAPE reactivity score constraints and minimum free energy (Figure 3I, J). This highlights that imputed structure scores can guide structural modeling when experimental probing data is unavailable.

Database improved user interfaces and visualization

We redesigned the user interface in RASP, enhancing both user-friendliness and stability. We expanded the search and visualization capabilities to incorporate proximity ligation-

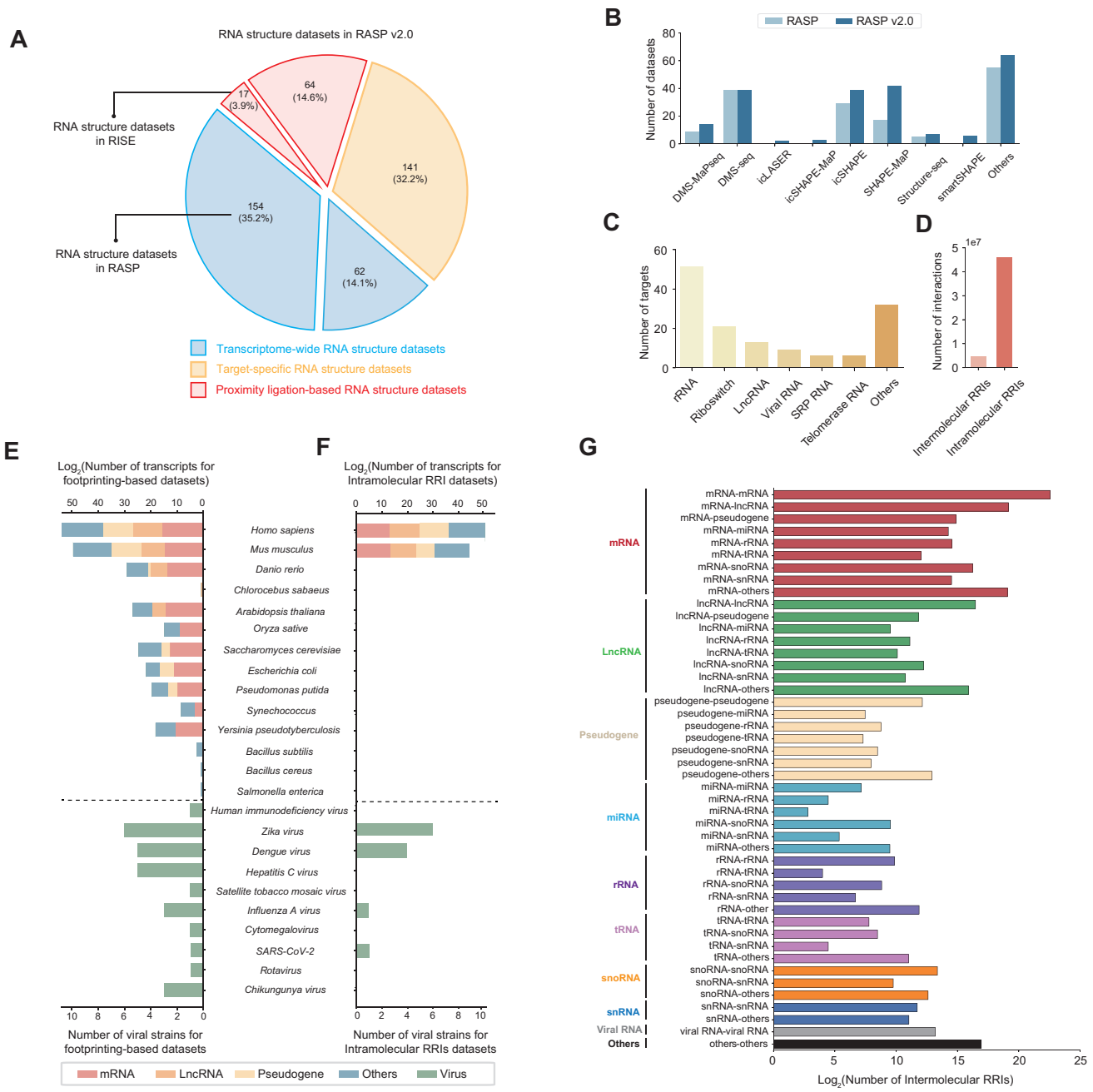


Figure 2. Data statistics of RASP v2.0. (**A**) Composition of RNA structure datasets in RASP v2.0. (**B**) Comparison of dataset numbers for transcriptome-wide structure probing methods in RASP and RASP v2.0. (**C**) Target numbers of target-specific structure probing data in RASP v2.0, categorized by RNA types. (**D**) Number of intramolecular and intermolecular RRIs detected by proximity ligation-based methods in RASP v2.0. (**E**) Number of transcripts or strains identified by footprinting-based structure probing methods across RNA types in different species. Dashed lines are used to distinguish viruses from other species. (**F**) Number of transcripts or strains involving intramolecular RRIs across RNA types in different species. Dashed lines are used to distinguish viruses from other species. (**G**) Statistical information on different types of intermolecular RRIs identified by proximity ligation-based structure probing methods.

based structure data. Moreover, the search module in RASP v2.0 was upgraded to include three query modes: 'Search Gene', 'Search Sequence', and 'Search Genomic Coordinate' (Figure 4A). The 'Search Gene' mode allows users to query interested transcripts using a specific gene symbol or transcript ID, while the 'Search Sequence' mode and the 'Search Genomic Coordinate' mode enable users to search gene candidates with a user-defined sequence or genomic region. After searching through one of the three modes above, a match list will be provided, including gene symbol, transcript link list, structure score browser link, intermolecular browser link, intramolecular browser link, and other important information (Figure 4A). Users can visualize structure scores and RRIs through these links. Additionally, we updated JBrowse (47) to JBrowse 2 (37) for a more stable display of structure score tracks (Figure 4B), and introduced IGV-web (38) (Figure 4C) and Circos plot(39) (Figure 4D) to visualize intramolecular and intermolecular RRIs.

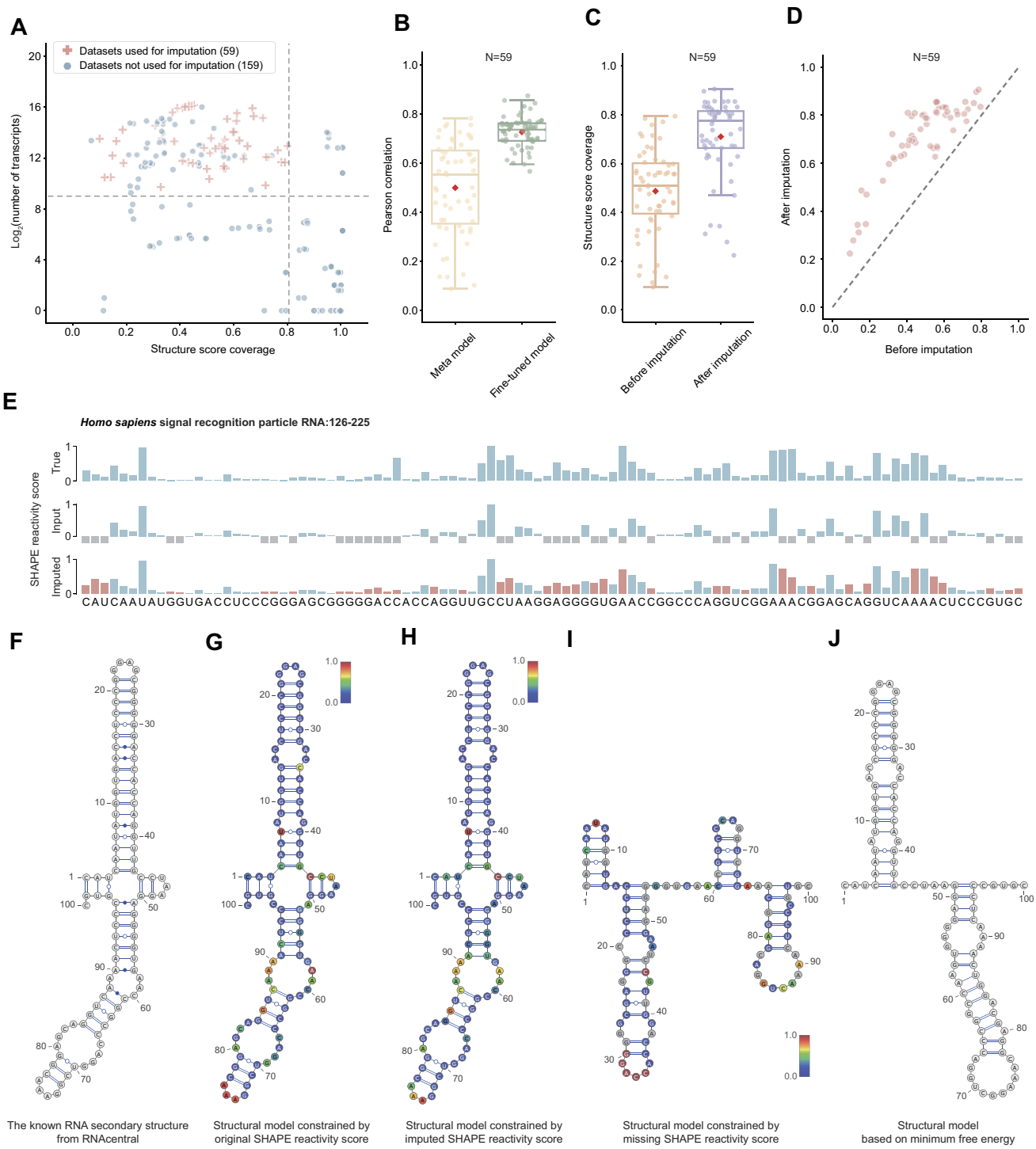


Figure 3. Imputation results for transcriptome-wide RNA structure datasets. (**A**) Distribution of average structure score coverage and transcript numbers for 218 transcriptome-wide RNA structure datasets. The red cross marks the selected datasets for imputation, the blue dot marks transcriptome-wide structure probing datasets, and the dashed line marks the imputation cutoff. (**B**) Performance comparison between the meta model and fine-tuned model based on the 59 selected transcriptome-wide RNA structure datasets. (**C**) Structure score coverage before and after imputation for the 59 selected transcriptome-wide RNA structure datasets. (**D**) Improvements of structure score coverage for the 59 selected transcriptome-wide RNA structure datasets. (**E**) An example showing the true and imputed structural scores. The upper track shows the true structure scores, the middle track shows the randomly masked structure scores to simulate the missing structure scores, and the bottom track shows the imputed structure scores. Blue, red and grey bars represent true structure score, imputed structure score and missing structure score, respectively. The height of a bar represents the value of the SHAPE reactivity score. (**F**) The known RNA secondary structure of signal recognition particle RNA from RNAcentral, including non-canonical base pairs. (**G**) An example of signal recognition particle RNA showing the structural model with original SHAPE reactivity score constrains. (**H**) An example of signal recognition particle RNA showing the structural model with imputed SHAPE reactivity score constrains. (**I**) An example of signal recognition particle RNA showing the structural model with missing SHAPE reactivity score constrains. (**J**) An example of signal recognition particle RNA showing the structural model generated by minimum free energy without SHAPE reactivity score constrains.

Figure 4. Updated Search and Browse modules. (**A**) Three types of search approaches in the search module, and the corresponding match lists. (**B**) Visualization of structure score tracks using JBrowse 2. (**C**) Visualization of intramolecular RRIs using IGV-Web. (**D**) Visualization of intermolecular RRIs plotted using Circos Plot.

Database added new analysis functions

Basic usage of online analysis functions

The input data types for the analysis modules in RASP v2.0 are uniform, with Figure 5A as an example. The left box is for RNA sequence input, while the right box is for RNA structure input. Some relevant parameters are required to be set before submitting the task. For each analysis module, we also provide one or two examples to help users quickly understand the functionality of the module.

Structure score imputation

RASP v2.0 supports missing structure score imputation using StructureImpute (35), which predicts missing structure scores using flanking sequences. Currently, this module provides 59 fine-tuned models derived from corresponding transcriptome-wide RNA structure datasets. As shown in Figure 5B, after entering the sequence and structure scores (where 'Null' represents missing structure scores), the results page displays structure scores after imputation, with green showing raw structure scores and orange showing imputed structure scores. Each input RNA sequence must be at least 100 nt in length.

RNA structure prediction

In addition to the existing RNA secondary structure prediction module in RASP, RASP v2.0 introduced an RNA tertiary structure prediction module using SimRNA (44,48). SimRNA applies Monte Carlo simulations to explore 3D RNA conformations, allowing the secondary structure as a constraint for RNA tertiary structure. As shown in Figure 5C, users can directly predict RNA secondary structure based on structure scores as in RASP, and then use these predictions to constrain tertiary structure predictions. Predicted RNA secondary structure can be visualized locally via VARNA (40) or online via forna (41). Predicted RNA tertiary structures can be visualized online using Molstar (42), and pdb/cif files can be downloaded for molecular docking.

RBP binding prediction

RASP v2.0 features an RBP binding prediction module based on PrismNet (45,49), a deep learning-based approach that

Figure 5. Applications of online analysis modules. (**A**) Input box example. 1 denotes the input window for the RNA sequence, 2 denotes the input window for the RNA structure, 3 denotes the parameter selection window related to this module, and 4 denotes the submit button. (**B**) Results page for the structure score imputation module. Green bars represent raw structure scores, orange bars represent imputed structure scores, and grey bars represent positions with 'Null' values. Users can download the '.txt' format of the imputed structure score by clicking the download icons. (**C**) Results page for the structure prediction module. Users can visualize the predicted RNA secondary and tertiary structure or copy VARNA commands for local visualization by clicking the 'Go' buttons. (**D**) Results page for the RBP binding prediction module. Users can visualize high attention region of saliency map and potential motifs by clicking 'View/Hide' buttons.

predicts potential RBP binding sites using RNA sequences and their structure scores. This module currently supports 168 RBP binding predictions across various cell lines. As shown in Figure 5D, the results page displays basic information and inferred results of input sequences, including binding probability, RBP binding sites, high attention region of saliency map, and potential RBP motifs. Additionally, we have collected known RBP motifs from the CISBP-RNA Database (50), users can compare the similarity between predicted motifs and known motifs. Each input RNA sequence or structure must be at least 101 nt in length.

Discussion

With the advent of high-throughput RNA structure probing methods, large-scale RNA structure datasets are accumulating rapidly. These structure probing data are essential for modeling higher-order RNA structures and elucidating the relationships between structure and function of RNA molecule. Through extensive data collection, we developed RASP v2.0, which includes 438 RNA structure datasets from 24 species across 39 experimental methods. We also improved the data quality of 59 transcriptome-wide RNA structure datasets through data imputation. Nevertheless, we would like to note those that the imputed data should be used with caution. Finally, RASP v2.0 introduces three online analysis modules to assist researchers for studying RNA structure.

Currently, the analysis of footprinting-based and proximity ligation-based structure probing data remains relatively independent. Therefore, integrating these datasets through joint analysis is crucial for a more comprehensive understanding of higher-order RNA structures. This integration will be invaluable for elucidating molecular mechanisms and advancing treatments for human diseases. Furthermore, developing a unified computational pipeline to standardize experimental data processing is highly desirable, as it would ensure the comparability of results across different technologies and studies. As more data are generated, we will continue to update RASP v2.0 as a repository for RNA structure information and annotations.

Data availability

The RASP v2.0 database is freely accessible at http://rasp2.zhanglab.net/. The RNA structure data can be downloaded from http://rasp2.zhanglab.net/download/.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We thank Jiasheng Zhang for the help in stress testing and network configuration for database. We thank Dr. Kui Xu for the help with StructureImpute usage and RISE database maintenance, and Dr. Pan Li for the help with transcriptome-wide RNA structure data analysis.

Funding

National Key Research and Development Project of China [2022YFF1203100 to Q.C.Z.]; National Natural Science Foundation of China [32 230 018 and 32 125 007 to Q.C.Z., 32 100 504 to Y.F.]; Beijing Natural Science Foundation [M23011 to Y.F.]; Postdoctoral Foundation of Tsinghua-Peking Center for Life Sciences (to Y.F.); Beijing Advanced Innovation Center for Structural Biology, and the Tsinghua-Peking Joint Center for Life Sciences. Funding for open access charge: National Natural Science Foundation of China; National Key Research and Development Project of China; Beijing Natural Science Foundation; Postdoctoral Foundation of Tsinghua-Peking Center for Life Sciences.

Conflict of interest statement

None declared.

References

1. Statello,L., Guo,C.J., Chen,L.L. and Huarte,M. (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.*, **22**, 96–118.
2. Welsh,S.A. and Gardini,A. (2023) Genomic regulation of transcription and RNA processing by the multitasking Integrator complex. *Nat. Rev. Mol. Cell Biol.*, **24**, 204–220.
3. Mattick,J.S., Amaral,P.P., Carninci,P., Carpenter,S., Chang,H.Y., Chen,L.L., Chen,R., Dean,C., Dinger,M.E., Fitzgerald,K.A., *et al.* (2023) Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.*, **24**, 430–447.
4. Houseley,J. and Tollervey,D. (2009) The many pathways of RNA degradation. *Cell*, **136**, 763–776.
5. Roden,C. and Gladfelter,A.S. (2021) RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.*, **22**, 183–195.
6. Wang,X.W., Liu,C.X., Chen,L.L. and Zhang,Q.C. (2021) RNA structure probing uncovers RNA structure-dependent biological functions. *Nat. Chem. Biol.*, **17**, 755–766.
7. Van Nostrand,E.L., Freese,P., Pratt,G.A., Wang,X., Wei,X., Xiao,R., Blue,S.M., Chen,J.Y., Cody,N.A.L., Dominguez,D., *et al.* (2020) A large-scale binding and functional map of human RNA-binding proteins. *Nature*, **583**, 711–719.
8. Childs-Disney,J.L., Yang,X., Gibaut,Q.M.R., Tong,Y., Batey,R.T. and Disney,M.D. (2022) Targeting RNA structures with small molecules. *Nat. Rev. Drug Discov.*, **21**, 736–762.
9. Zhu,Y., Zhu,L., Wang,X. and Jin,H. (2022) RNA-based therapeutics: an overview and prospectus. *Cell Death. Dis.*, **13**, 644.
10. Ganser,L.R., Kelly,M.L., Herschlag,D. and Al-Hashimi,H.M. (2019) The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.*, **20**, 474–489.
11. Wan,Y., Kertesz,M., Spitale,R.C., Segal,E. and Chang,H.Y. (2011) Understanding the transcriptome through RNA structure. *Nat. Rev. Genet.*, **12**, 641–655.
12. Cruz,J.A. and Westhof,E. (2009) The dynamic landscapes of RNA architecture. *Cell*, **136**, 604–609.
13. Spitale,R.C. and Incarnato,D. (2023) Probing the dynamic RNA structurome and its functions. *Nat. Rev. Genet.*, **24**, 178–196.
14. Rasmussen,S.G.F., DeVree,B.T., Zou,Y., Kruse,A.C., Chung,K.Y., Kobilka,T.S., Thian,F.S., Chae,P.S., Pardon,E., Calinski,D., *et al.* (2011) Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature*, **477**, 549–555.
15. Riek,R., Hornemann,S., Wider,G., Billeter,M., Glockshuber,R. and Wüthrich,K. (1996) NMR structure of the mouse prion protein domain PrP(121–231). *Nature*, **382**, 180–182.
16. Strobel,E.J., Yu,A.M. and Lucks,J.B. (2018) High-throughput determination of RNA structures. *Nat. Rev. Genet.*, **19**, 615–634.
17. Zhang,J., Fei,Y., Sun,L. and Zhang,Q.C. (2022) Advances and opportunities in RNA structure experimental determination and computational modeling. *Nat. Methods*, **19**, 1193–1207.
18. Rouskin,S., Zubradt,M., Washietl,S., Kellis,M. and Weissman,J.S. (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature*, **505**, 701–705.
19. Spitale,R.C., Flynn,R.A., Zhang,Q.C., Crisalli,P., Lee,B., Jung,J.W., Kuchelmeister,H.Y., Batista,P.J., Torre,E.A., Kool,E.T., *et al.* (2015) Structural imprints in vivo decode RNA regulatory mechanisms. *Nature*, **519**, 486–490.
20. Siegfried,N.A., Busan,S., Rice,G.M., Nelson,J.A. and Weeks,K.M. (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods*, **11**, 959–965.
21. Ding,Y., Tang,Y., Kwok,C.K., Zhang,Y., Bevilacqua,P.C. and Assmann,S.M. (2014) In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*, **505**, 696–700.
22. Solayman,M., Litfin,T., Zhou,Y. and Zhan,J. (2022) High-throughput mapping of RNA solvent accessibility at the single-nucleotide resolution by RtcB ligation between a fixed 5'-OH-end linker and unique 3'-P-end fragments from hydroxyl radical cleavage. *RNA Biol.*, **19**, 1179–1189.
23. Lu,Z., Zhang,Q.C., Lee,B., Flynn,R.A., Smith,M.A., Robinson,J.T., Davidovich,C., Gooding,A.R., Goodrich,K.J., Mattick,J.S., *et al.* (2016) RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell*, **165**, 1267–1279.
24. Aw,J.G., Shen,Y., Wilm,A., Sun,M., Lim,X.N., Boon,K.L., Tapsin,S., Chan,Y.S., Tan,C.P., Sim,A.Y., *et al.* (2016) In vivo mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell*, **62**, 603–617.
25. Cai,Z.K., Cao,C.C., Ji,L., Ye,R., Wang,D., Xia,C., Wang,S., Du,Z.C., Hu,N.J., Yu,X.H., *et al.* (2020) RIC-seq for global in situ profiling of RNA–RNA spatial interactions. *Nature*, **582**, 432–+.
26. Wu,T., Cheng,A.Y., Zhang,Y., Xu,J., Wu,J., Wen,L., Li,X., Liu,B., Dou,X., Wang,P., *et al.* (2024) KARR-seq reveals cellular higher-order RNA structures and RNA–RNA interactions. *Nat. Biotechnol.*, https://doi.org/10.1038/s41587-023-02109-8.
27. Cordero,P., Lucks,J.B. and Das,R. (2012) An RNA mapping DataBase for curating RNA structure mapping experiments. *Bioinformatics*, **28**, 3006–3008.
28. Yu,H., Zhang,Y., Sun,Q., Gao,H. and Tao,S. (2021) RSVdb: a comprehensive database of transcriptome RNA structure. *Brief Bioinform*, **22**, bbaa071.
29. Li,P., Zhou,X., Xu,K. and Zhang,Q.C. (2021) RASP: an atlas of transcriptome-wide RNA secondary structure probing data. *Nucleic Acids Res.*, **49**, D183–D191.

30. Gong,J., Shao,D., Xu,K., Lu,Z., Lu,Z.J., Yang,Y.T. and Zhang,Q.C. (2018) RISE: a database of RNA interactome from sequencing experiments. *Nucleic Acids Res.*, **46**, D194–D201.
31. Yi,Y., Zhao,Y., Li,C., Zhang,L., Huang,H., Li,Y., Liu,L., Hou,P., Cui,T., Tan,P., *et al.* (2017) RAID v2.0: an updated resource of RNA-associated interactions across organisms. *Nucleic Acids Res.*, **45**, D115–D118.
32. Zheng,Y., Luo,H., Teng,X., Hao,X., Yan,X., Tang,Y., Zhang,W., Wang,Y., Zhang,P., Li,Y., *et al.* (2023) NPInter v5.0: ncRNA interaction database in a new era. *Nucleic Acids Res.*, **51**, D232–D239.
33. Sayers,E.W., Bolton,E.E., Brister,J.R., Canese,K., Chan,J., Comeau,D.C., Connor,R., Funk,K., Kelly,C., Kim,S., *et al.* (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **50**, D20–D26.
34. Clough,E., Barrett,T., Wilhite,S.E., Ledoux,P., Evangelista,C., Kim,I.F., Tomashevsky,M., Marshall,K.A., Phillippy,K.H., Sherman,P.M., *et al.* (2024) NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Res.*, **52**, D138–D144.
35. Gong,J., Xu,K., Ma,Z., Lu,Z.J. and Zhang,Q.C. (2021) A deep learning method for recovering missing signals in transcriptome-wide RNA structure profiles from probing experiments. *Nat. Mach. Intell.*, **3**, 995–1006.
36. Morgulis,A., Coulouris,G., Raytselis,Y., Madden,T.L., Agarwala,R. and Schaffer,A.A. (2008) Database indexing for production MegaBLAST searches. *Bioinformatics*, **24**, 1757–1764.
37. Diesh,C., Stevens,G.J., Xie,P., De,J., Martinez,T., Hershberg,E.A., Leung,A., Guo,E., Dider,S., Zhang,J., *et al.* (2023) JBrowse 2: a modular genome browser with views of synteny and structural variation. *Genome Biol.*, **24**, 74.
38. Robinson,J.T., Thorvaldsdottir,H., Turner,D. and Mesirov,J.P. (2023) igv.Js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics*, **39**, btac830.
39. Krzywinski,M., Schein,J., Birol,I., Connors,J., Gascoyne,R., Horsman,D., Jones,S.J. and Marra,M.A. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.*, **19**, 1639–1645.
40. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
41. Kerpedjiev,P., Hammer,S. and Hofacker,I.L. (2015) Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. *Bioinformatics*, **31**, 3377–3379.
42. Sehnal,D., Bittrich,S., Deshpande,M., Svobodova,R., Berka,K., Bazgier,V., Velankar,S., Burley,S.K., Koca,J. and Rose,A.S. (2021) Mol* viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.*, **49**, W431–W437.
43. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinf.*, **11**, 129.
44. Boniecki,M.J., Lach,G., Dawson,W.K., Tomala,K., Lukasz,P., Soltysinski,T., Rother,K.M. and Bujnicki,J.M. (2016) SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res.*, **44**, e63.
45. Sun,L., Xu,K., Huang,W., Yang,Y.T., Li,P., Tang,L., Xiong,T. and Zhang,Q.C. (2021) Predicting dynamic cellular protein-RNA interactions by deep learning using in vivo RNA structures. *Cell Res.*, **31**, 495–516.
46. RNAcentral Consortium (2021) RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Res.*, **49**, D212–D220.
47. Buels,R., Yao,E., Diesh,C.M., Hayes,R.D., Munoz-Torres,M., Helt,G., Goodstein,D.M., Elsik,C.G., Lewis,S.E., Stein,L., *et al.* (2016) JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.*, **17**, 66.
48. Moafinejad,S.N., de Aquino,B.R.H., Boniecki,M.J., Jeyeram,I., Nikolaev,G., Magnus,M., Farsani,M.A., Badepally,N.G., Wirecki,T.K., Stefaniak,F., *et al.* (2024) SimRNAweb v2.0: a web server for RNA folding simulations and 3D structure modeling, with optional restraints and enhanced analysis of folding trajectories. *Nucleic Acids Res.*, **52**, W368–W373.
49. Xu,Y., Zhu,J., Huang,W., Xu,K., Yang,R., Zhang,Q.C. and Sun,L. (2023) PrismNet: predicting protein-RNA interaction using in vivo RNA structural information. *Nucleic Acids Res.*, **51**, W468–W477.
50. Ray,D., Kazan,H., Cook,K.B., Weirauch,M.T., Najafabadi,H.S., Li,X., Gueroussov,S., Albu,M., Zheng,H., Yang,A., *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, **499**, 172–177.

Received: September 15, 2024. **Revised:** October 16, 2024. **Editorial Decision:** October 18, 2024. **Accepted:** November 14, 2024