W160–W167 *Nucleic Acids Research, 2023, Vol. 51, Web Server issue* *Published online 26 April 2023*
https://doi.org/10.1093/nar/gkad296

e-RNA: a collection of web-servers for the prediction and visualisation of RNA secondary structure and their functional features

Volodymyr Tsybulskyi[1,2,†], Egor Semenchenko[1,2,†] and Irmtraud M. Meyer[1,2,3,*]

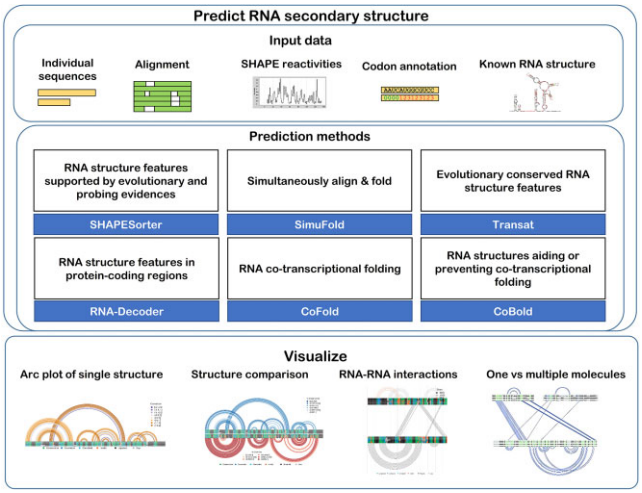
[1]Laboratory of Bioinformatics of RNA Structure and Transcriptome Regulation, Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13125 Berlin, Germany, [2]Department of Biology, Chemistry and Pharmacy, Institute of Chemistry and Biochemistry, Thielallee 63, Freie Universität Berlin, 14195 Berlin, Germany and [3]Department of Mathematics and Computer Science, Institute of Computer Science, Takustraße 9, Freie Universität Berlin, 14195 Berlin, Germany

Received January 30, 2023; Revised March 23, 2023; Editorial Decision April 07, 2023; Accepted April 19, 2023

ABSTRACT

e-RNA is a collection of web-servers for the prediction and visualisation of RNA secondary structures and their functional features, including in particular RNA–RNA interactions. In this updated version, we have added novel tools for RNA secondary structure prediction and have significantly updated the visualisation functionality. The new method CoBold can identify transient RNA structure features and their potential functional effects on a known RNA structure during co-transcriptional structure formation. New tool ShapeSorter can predict evolutionarily conserved RNA secondary structure features while simultaneously taking experimental SHAPE probing evidence into account. The web-server R-Chie which visualises RNA secondary structure information in terms of arc diagrams, can now be used to also visualise and intuitively compare RNA–RNA, RNA–DNA and DNA–DNA interactions alongside multiple sequence alignments and quantitative information. The prediction generated by any method in e-RNA can be readily visualised on the web-server. For completed tasks, users can download their results and readily visualise them later on with R-Chie without having to re-run the predictions. e-RNA can be found at http://www.e-rna.org.

GRAPHICAL ABSTRACT



INTRODUCTION

The last few years have seen a strongly increased interest in RNA biology. It is now widely recognised that the functional roles of RNA is not limited to simply transmitting protein information and enabling protein synthesis. RNA performs a much broader range of key biological tasks which we are only beginning to discover and fully understand. RNA plays an active role in the regulation of translation and alternative splicing, directs chemical modifications, and catalyses many important biochemical reactions. These functional roles are often exerted by the structure of the RNA or its trans interactions with other RNA transcripts.

The prediction and analysis of the RNA secondary structure *in vivo* is thus essential for understanding how gene expression is regulated in transcriptomes in a range of biological systems.

*To whom correspondence should be addressed. Tel: +49 30 9406 3292; Fax: +49 30 9406 3291; Email: irmtraud.meyer@cantab.net
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

In the past decades, substantial progress has been made in studying the role of RNA secondary structure features in viruses and unicellular eukaryotes (1–6). More recently, in the time of the global pandemic, new experimental structure-probing approaches have allowed to investigate the RNA secondary structure of the SARS-Cov2 virus. This helped to identify several conserved structural motifs which can potentially become major therapeutic targets. (7,8). Moreover, long-standing questions on the role of the RNA secondary structure in splicing regulation are now being investigated on a transcriptome-wide scale (9) and continue to be addressed in the relation to transcription (10) and in the context of disease (11) In addition to proving a key for the fundamental understanding of many essential biological processes, the analysis of the RNA secondary structure is also crucial for biotechnological applications that aim to target challenges in living systems. Examining the influence of secondary structure on translation efficacy (12) and molecule stability (13) has, for example, become one of the key components to successful mRNA vaccine design.

Although the structure of any RNA or transcript *in vivo* is naturally three-dimensional, it often suffices to study the so-called RNA secondary structure in order to investigate its potential functional roles. This is due to the fact that the folding process is known to be hierarchical, where the transition between primary (linear) conformation is more energy favourable and occurs on shorter timescales than the adaptation of a higher-order structure (14). We define the RNA secondary structure of a given RNA as the set of nucleotide positions that form the so-called consensus base-pairs $\{G-C, C-G, U-A, A-U, G-U, U-G\}$. Note that we explicitly include the so-called wobble base-pairs $\{G-U, U-G\}$ in all of our RNA structure predictions and RNA structure modelling.

Typically, these Watson−Crick and non-canonical base pairs of the secondary structures form stronger bonds than those that are involved in forming higher-order structures. Therefore, tertiary structures are often viewed as an ensemble of structures descendant from the secondary structure (15).

RNA molecules in any biological environment *in vivo* can readily form more than a unique secondary structure. For example, during RNA synthesis shortly after the 5′- end is synthesised, a partially synthesised molecule starts folding into transient structures ('co-transcriptional folding') before settling into one (or more) functional RNA structures (16,17). In addition, these transient structures may play their own functional roles, e.g. to influence the transcription kinetics or to facilitate splicing (10). The function of an RNA *in vivo* may also be determined by its trans interaction partners in that particular cellular environment, e.g. direct trans interactions with other RNAs (18) or with RNA-binding proteins (19,20). Also, individual nucleotide modifications (21) and changes in the physiological conditions such as temperature and pH (22) may change the nominal RNA structure formation or the RNA's nominal trans interactions *in vivo* and thereby alter the functional roles of the transcript that are thereby expressed in a cell-specific way (23).

The most widely-used methods for predicting RNA secondary structure to date are so-called thermodynamic methods which assume that the thermodynamically most stable RNA structures (in terms of Gibbs free energy) are most likely to play a functional role, even in *in vivo* conditions where the conditions of thermodynamic equilibrium are typically not given and where trans interaction partners may play a decisive role (24). These methods employ dynamic programming to find the (typically pseudo-knot free) most stable RNA secondary structure that corresponds to the minimum free energy (MFE) (25). This is why these methods are typically referred to as MFE methods and the RNA secondary structures they predict as MFE structures.

Multiple thermodynamic models with different sets of free parameters, reflecting the energy gains of a possible base-pairing, were proposed and fine-tuned over the last four decades (26). MFE methods typically achieve a high prediction accuracy in predicting the RNA secondary structure of short molecules shorter than 200 nucleotides. Their prediction accuracy, however, significantly decreases with increasing sequence length (27).

A conceptually different approach to RNA secondary structure employs probabilistic approaches. These methods usually rely on the assumption that base-pairs of functionally important RNA secondary structure features and RNA–RNA interaction are evolutionary conserved or co-evolved to maintain their base-pairing potential (28). These observations from carefully compiled multiple-sequence alignments (MSAs) can be extracted and used to train the free parameters of underlying prediction models based on a likelihood-ratio test (LRT) (29), stochastic context-free grammars (SCFG) (30) or Markov Chain Monte Carlo (MCMC) models (31). A major challenge in applying these probabilistic methods is their dependence on the quality of the input MSA. The MSA defines the amount of co-variation information that the prediction tool can utilise. Since primary sequence conservation does not necessarily imply the presence of a functional RNA secondary structure, any variation within a given MSA has to be carefully and quantitatively evaluated before drawing biological conclusions (32,33). Importantly, any MSA can be viewed as an imprint of the overall constraints of the corresponding cellular *in vivo* environment on the transcripts within the MSA, as it reflects patterns of compensatory mutations and conservation that are compatible with the functional RNA secondary structure as well as any functionally important trans interaction partners. To conclude, probabilistic methods for RNA secondary structure prediction are conceptually well suited for capturing the overall evolutionary signals and constraints that are encoded in RNA transcripts, yet they are typically more challenging to employ as they often (but not always (31)) require an input MSA.

The emergence of experimental structure probing methods such as SHAPE (34) and DMS-seq (35) and their further developments (36–38) was a major experimental advancement that also impacted computational RNA secondary structure prediction. It is now possible to experimentally probe the pairing status (i.e. paired versus unpaired) of individual nucleotides in a transcript *in vivo* on a transcriptome-wide scale. The corresponding raw data,

Table 1. Comparison of the e-RNA prediction programs with other publicly available web-servers

e-RNA program	Closest analog	Common features	Differences
CoFold	none	Predicts most stable RNA structure arising in co-transcriptional folding.	—
ShapeSorter	ViennaRNA RNAprobing (45)	Integrate SHAPE reactivity profiles into RNA structure prediction.	RNAprobing predicts a single MFE structure. ShapeSorter predicts all high-confidence helices supported by evolutionary and experimental evidence.
SimulFold	Freiburg RNA-tools LocARNA (46)	Co-estimate an MSA and a common RNA structure.	SimulFold employs a Bayesian MCMC, can also handle pseudo-knotted RNA structures and can also co-estimate evolutionary trees. LocARNA uses an MFE-based RNA structure prediction concept.
Transat	Rivas laboratory R-scape (32) and CaCoFold (47)	Predict high-confidence RNA structure features based on an input MSA.	R-scape predicts high-confidence base-pairs, CaCoFold predicts entire RNA structures. Transat predicts high-confidence helices.
RNA-Decoder	none	Predicts conserved RNA structures overlapping protein-coding regions for an input MSA with known protein-coding regions.	—
CoBold	none	Predicts transient RNA structure features and their likely impact on a given reference RNA structure.	—
R-Chie	none	Visualise and compare cis- and trans-RNA–RNA, RNA–DNA and DNA–DNA interactions including quantitative information.	—

however, needs to first be heavily interpreted computationally in order for the raw data to be converted into actual RNA secondary structures for individual transcripts. Most commonly, this raw data is used as additional input information to MFE methods for RNA secondary structure prediction in terms of so-called reactivity values which are then converted into pseudo-energy terms inside the algorithm. These pseudo-energy terms slightly alter the nominal pairing energies of nucleotides within an input RNA in a sequence-position specific way and thereby alter the resulting, predicted MFE structure. To date, only a few probabilistic prediction methods can incorporate experimental probing data (39,40). One recent tool ShapeSorter by us which combines maximum-likelihood estimation with PGMs (Probabilistic Graphical Models) has recently been published and is now available on at e-RNA web server (41).

In light of the improvements and opportunities that transcriptome-wide RNA structure probing has brought to the field, there is a corresponding need for visualising RNA structure predictions alongside (1) the corresponding experimental structure probing evidence and (2) the evolutionary evidence in terms of an MSA (42). This is the mandate that our e-RNA web-server aims to address since its inception in 2012 (43). e-RNA provides a convenient way of generating predictions using a range of comparative and non-comparative methods and then visualising the resulting predictions alongside various sources of evidence (e.g. experimental and/or evolutionary). For this, we deploy an updated R-Chievisualisation suite (44) and the most recent computational prediction methods developed by the Meyer lab.

COMPARISON OF E-RNA TO OTHER WEB-SERVERS

Other well-known web-servers for RNA structure prediction include e.g. the ViennaRNA web-server and a range of others, for an overview see Table 1. Our prediction methods are almost exclusively based on probabilistic concepts (the only exception being CoFold). They all detect potential functional RNA structure features based on evolutionarily evidence and – in case of ShapeSorter – additional experimental evidence, e.g. in form of SHAPE probing reactivities. The RNA structure prediction programs of the ViennaRNA web-server, however, assume that (i) functional RNA structures correspond to the most stable RNA structures (in terms of Gibbs free energy) and (ii) that these structures form in an environment of thermodynamic equilibrium. As a glance at Table 1 shows, several of our methods are unique.

PROGRAMS

Our updated version of e-RNA includes web-servers for the novel RNA secondary structure prediction tools CoBold and ShapeSorter as well as a significantly updated version of our R-Chie visualisation suite. In this section, we introduce the different methods available at e-RNA.

Transat

Transat is a fully probabilistic method for predicting evolutionary conserved RNA secondary structure features based on a multiple sequence input alignment (MSA) (48),

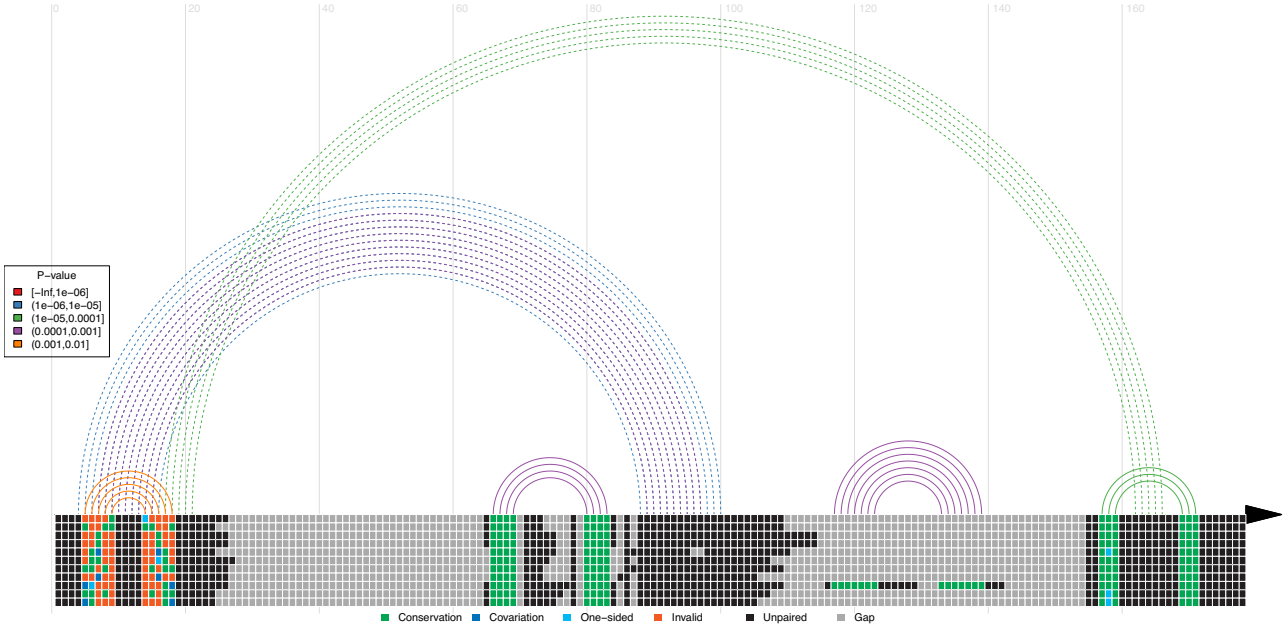


Figure 1. Experimentally confirmed case of co-transcriptional folding. We can observe three possible RNA secondary structures formations coloured differently and multiple sequence alignment annotation for one of the structures.

i.e. TRANSAT is a comparative tool. It predicts so-called helices, i.e. consecutive stretches of base-pairs. To detect these, the method employs different probabilistic models of evolution for detecting paired and unpaired alignment columns within the input MSA. In the first step, the algorithm identifies helices within the individual, un-gapped sequences of the input MSA, before mapping them back onto the input MSA and treating the resulting (potentially gapped) helices as candidate helices. Each candidate helix then gets assigned a log-likelihood value which quantitatively compares the two competing hypotheses (paired versus unpaired), before estimating a p-value for the log-likelihood value. To estimate p-values, TRANSAT generates a null distribution of log-likelihood values for helices in carefully randomised versions of the original input MSA. The overall output of TRANSAT is a set of conserved helices within the input MSA that can be readily ranked by their corresponding p-values. On the e-RNA web-server, the user needs to provide an input MSA and – optionally – a phylogenetic tree linking the sequences in the input MSA. The current web-server of TRANSAT is an updated version of the program described in the original publication. The new version 2.0 includes a memory optimisation which noticeably reduced the computation time, especially for alignments of >100 sequences.

COFOLD

COFOLD is a non-comparative, thermodynamic RNA secondary structure method that—unlike the commonly used MFE methods—also captures one overall effect that co-transcriptional folding has on RNA structure formation, thereby resulting in a substantially increased prediction accuracy, especially for long input sequences (49).

COBOLD

COBOLD (50) is a set of two methods—one comparative and probabilistic, one non-comparative and MFE driven—that both (i) identify potential transient RNA structure features for a given input MSA or individual RNA sequence, respectively and (ii) judge the likely effect of these transient features on the co-transcriptional formation of a given reference RNA secondary structure (positive, neutral, negative). The non-comparative mode of COBOLD uses a slightly modified version of the previously described COFOLD method to detect *all* potentially relevant RNA transient structures in the sequence of interest (49). The comparative mode of COFOLD employs TRANSAT to identify potential transient helices, see above. E-RNA provides a special web-server for readily visualising the predicted transient features and how they relate exactly to the features of the given RNA secondary structure.

SHAPESORTER

SHAPESORTER (41) is a comparative method and a natural extension of the probabilistic method TRANSAT described above that also takes experimental SHAPE probing evidence in terms of a SHAPE reactivity profile into account. Both methods share the probabilistic models of evolution. SHAPESORTER integrates experimental SHAPE evidence via dedicated PGMs (Probabilistic Graphical Models) that are integrated into the calculation of the log-likelihood values for candidate helices in the input MSA. Similarly to TRANSAT, SHAPESORTER also estimates *P*-values for its predicted helices, thereby allowing the user to readily rank and prioritise its predictions. The user input to SHAPESORTER consists of an input MSA as well as SHAPE reactivity profile for the reference sequence inside the MSA (top sequence).

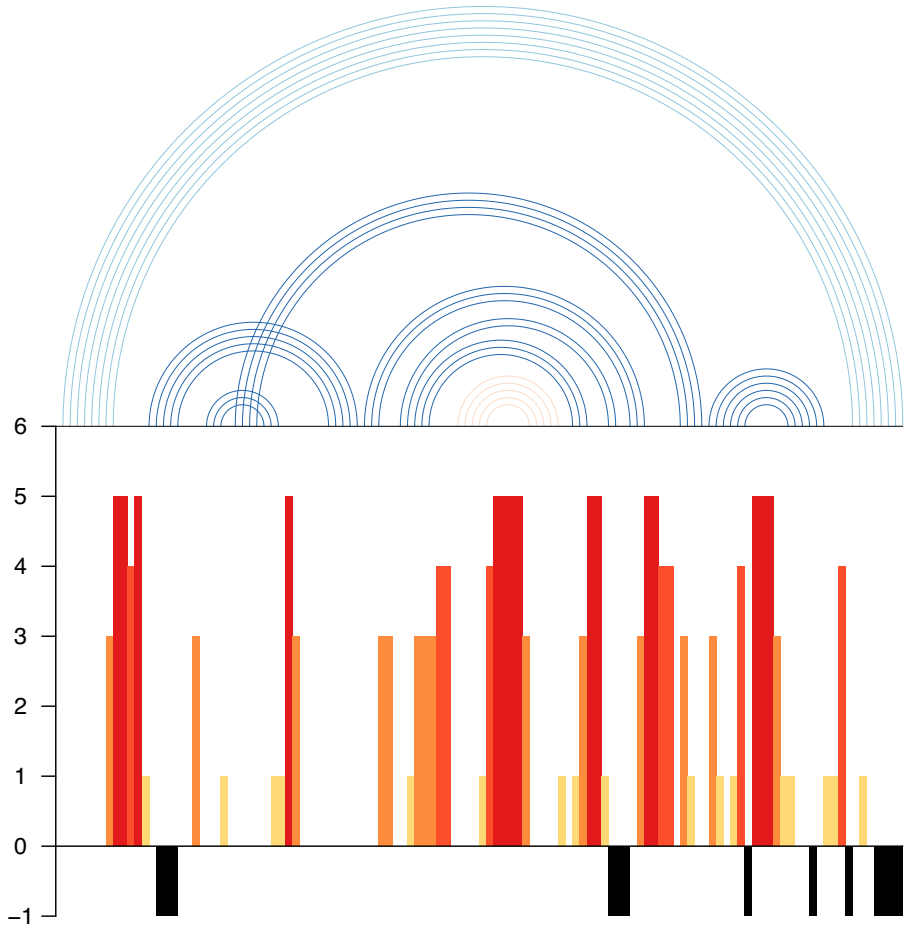


Figure 2. Example of ShapeSorter prediction for SAMI riboswitch of T.Tengcongensis organism. Arcs are coloured by p-value of prediction, and the graph with a line underneath arcs shows binned SHAPE reactivity values, obtained by SHAPESORTER pre-processing. Maximum is 5 and -1 stands for the absence of SHAPE value for nucleotide position, bars are coloured based on the value of reactivity (yellow is 0 and red is 5, black is the absence of SHAPE reactivity value).

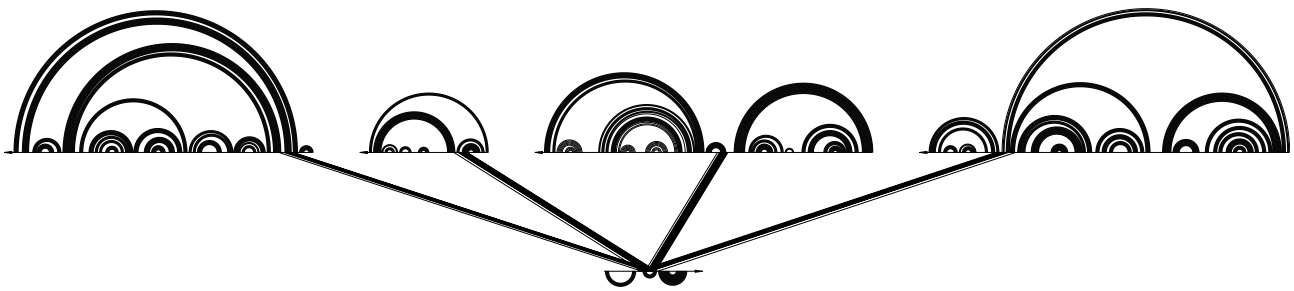


Figure 3. Example of RNA–RNA interactions. Here we can observe 'CyaR' sRNA at the bottom of the figure, and its known target mRNAs (luxS,nadE,ompX,yqaE) at the top of the figure. Arcs are showing cis interactions while lines show trans interactions (54).

SIMULFOLD

SIMULFOLD (31) is a comparative, probabilistic method that employs a Bayesian MCMC. Unlike almost all comparative methods for RNA secondary structure prediction, it does not require an input MSA, but only a set of homologous RNAs. SIMULFOLD is able to simultaneously predict an RNA secondary structure (including pseudo-knots), a corresponding MSA as well as a corresponding evolutionary tree or network. In addition to the set of un-aligned input sequences, the user can specify an initial MSA (which is used as a starting point for sampling MSAs within the Bayesian MCMC). The user can also opt to employ the method for co-estimating only one two or one of the three potential output features (RNA structure, MSA, phylogenetic tree). SIMULFOLD employs three different priors for evaluating potential secondary structures, multiple sequence alignments and phylogenetic trees. Optimal running time is achieved by sampling from the posterior distribution

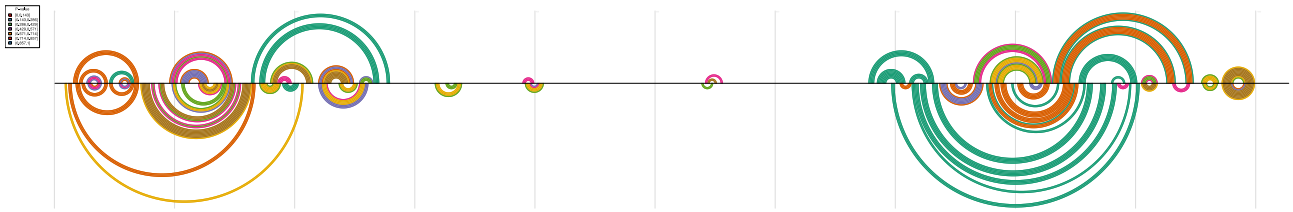


Figure 4. Short RNA structure feature overlapping one 3' splice site in M segment of influenza A. Arcs are coloured based on prediction p-value. Top arcs are related to the Human secondary structure and the bottom is to Avian secondary structure (53).

using the Bayesian MCMC method instead of analytically calculating the posterior distribution.

RNA-DECODER

RNA-DECODER is a unique, comparative and fully probabilistic method for detecting RNA secondary structure within an input MSA that is known to be partly or fully protein-coding such as mRNAs, RNA or DNA genomes and pre-mRNAs. Compared to all other existing methods for RNA secondary structure prediction, RNA-DECODER is unique in the sense that it explicitly takes the known protein-coding context of the input MSA into account when detecting evolutionarily conserved RNA structures. This feature is key as any protein-coding regions correspond to an additional layer of evolutionary constraints that needs to be carefully distinguished (and disentangled) from any evolutionary constraint due to RNA secondary structure. The method employs an SCFG with multiple specifically designed and carefully parameterised probabilistic models of evolution that are able to capture one or two potentially overlapping and conceptually quite different evolutionary constraints. It is key to note that the grammar underlying RNA-DECODER was designed to specifically detect multiple, adjacent RNA structures within the same input MSA that may be separated by potentially long regions that are devoid of any conserved RNA structure features. RNA-DECODER requires as input an MSA where any known protein-coding regions are annotated. RNA-DECODER produces as output either the best RNA secondary structure annotation (RNA structure prediction mode of RNA-DECODER) or – alternatively – the base-pairing probabilities for each position in the input MSA (scanning mode of RNA-DECODER) (3,51–53).

R-CHIE and R4RNA

R-CHIE and R4RNA were originally developed for visualising the results of all the programs above. They allow the visualisation of cis-interactions at the nucleotide level resolution in RNAs of interest (43). As an additional functionality, they can also now be used to visualise RNA–RNA, RNA–DNA and DNA–DNA interactions (44) alongside the corresponding evolutionary evidence (i.e. MSA) and additional quantitative evidence (e.g. SHAPE reactivities, p-values) that can be assigned to individual sequence positions or individual base-pairs. Good example cases are shown in Figures 1–3 and 4. Figure 1 shows the co-transcriptional structural features of the ZTP riboswitch. Here, we show multiple conflicting transient structures alongside corresponding evolutionary evidence for the different organisms within multiple sequence alignments. As another example, we show the visualisation of the SHAPE-SORTER prediction using a histogram of the SHAPE probing data for the predicted RNA secondary structure (see Figure 2) (41). An additional key feature of the new R-CHIE is its ability to show multiple RNA entities and their cis- and trans-interactions, see for example Figure 3 in which an sRNA targets several mRNAs within the transcriptome which clash with known RNA secondary structure features (54). R-CHIE is particularly well suited to readily compute and predict two alternative RNA structure annotations for the same RNA, e.g. one predicted and one reference one, see for example Figure 4 which shows the similarity of the RNA secondary structure features of the M segment in influenza A, once for an alignment of Human and once for an alignment of Avian sequences.

WEB-SERVER SPECIFICATION

The maximum input size that e-RNA servers can accommodate is capped at 200 000 characters, e.g. an alignment of 20 sequences of around 1000 nt length or similar. The down-loadable software that users can execute locally on their own machines has no such input limitations.

e-RNA web runs on a RedHat operating system using Apache server software with an HTML front end and combined Perl and JavaScript back end. Each user query is submitted via PBS queue system.

CONCLUSION

Here, we have described e-RNA, a collection of web-servers for the prediction and visualisation of RNA secondary structure and their functional features. There are many prediction tools presented on web-server that adapt to the available user-provided data evidence and the specific type of RNA secondary structure features that users want to define. All web servers are available as one-click applications, with self-described examples. We also provide post-launch compressed results and a command to run on user's own computer or server. In case of technical problems with the web servers, we have provided a contact email address. Additionally, with links to GitHub, we provide the source code in a separate sub-tab, and some programs are also presented as singularity containers.

DATA AVAILABILITY

E-RNA is freely accessible at: https://e-rna.org/.

FUNDING

Helmholtz-Gemeinschaft. Funding for open access charge: Helmholtz Association.
Conflict of interest statement. None declared.

REFERENCES

1. Mauger,D.M., Golden,M., Yamane,D., Williford,S., Lemon,S.M., Martin,D.P. and Weeks,K.M. (2015) Functionally conserved architecture of hepatitis C virus RNA genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3692–3697.
2. Zeng,Z., Shi,J., Guo,X., Mo,L., Hu,N., Sun,J., Wu,M., Zhou,H. and Hu,Y. (2018) Full-length genome and molecular characterization of dengue virus serotype 2 isolated from an imported patient from Myanmar. *Virol. J.*, **15**, 131.
3. Watts,J.M., Dang,K.K., Gorelick,R.J., Leonard,C.W., Bess Jr,J.W., Swanstrom,R., Burch,C.L. and Weeks,K.M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, **460**, 711–716.
4. Lavender,C.A., Gorelick,R.J. and Weeks,K.M. (2015) Structure-based alignment and consensus secondary structures for three HIV-related RNA genomes. *PLoS Comput. Biol.*, **11**, e1004230.
5. Simon,L.M., Morandi,E., Luganini,A., Gribaudo,G., Martinez-Sobrido,L., Turner,D.H., Oliviero,S. and Incarnato,D. (2019) In vivo analysis of influenza A mRNA secondary structures identifies critical regulatory motifs. *Nucleic Acids Res.*, **47**, 7003–7017.
6. Qi,Y., Zhang,Y., Zheng,G., Chen,B., Zhang,M., Li,J., Peng,T., Huang,J. and Wang,X. (2021) In vivo and in vitro genome-wide profiling of RNA secondary structures reveals key regulatory features in *Plasmodium falciparum*. *Front. Cell. Inf. Microbiol.*, **11**, 673966.
7. Manfredonia,I., Nithin,C., Ponce-Salvatierra,A., Ghosh,P., Wirecki,T.K., Marinus,T., Ogando,N.S., Snijder,E.J., van Hemert,M.J., Bujnicki,J.M. *et al.* (2020) Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Res.*, **48**, 12436–12452.
8. Sun,L., Li,P., Ju,X., Rao,J., Huang,W., Ren,L., Zhang,S., Xiong,T., Xu,K., Zhou,X. *et al.* (2021) In vivo structural characterization of the SARS-CoV-2 RNA genome identifies host proteins vulnerable to repurposed drugs. *Cell*, **184**, 1865–1883.
9. Kalmykova,S., Kalinina,M., Denisov,S., Mironov,A., Skvortsov,D., Guigó,R. and Pervouchine,D. (2021) Conserved long-range base pairings are associated with pre-mRNA processing of human genes. *Nat. Commun.*, **12**, 2300.
10. Herzel,L., Ottoz,D.S., Alpert,T. and Neugebauer,K.M. (2017) Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat. Rev. Mol. Cell Biol.*, **18**, 637–650.
11. Singh,N.N., Lee,B.M. and Singh,R.N. (2015) Splicing regulation in spinal muscular atrophy by an RNA structure formed by long-distance interactions. *Ann. New York Acad. Sci.*, **1341**, 176–187.
12. Mauger,D.M., Cabral,B.J., Presnyak,V., Su,S.V., Reid,D.W., Goodman,B., Link,K., Khatwani,N., Reynders,J., Moore,M.J. *et al.* (2019) mRNA structure regulates protein expression through changes in functional half-life. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 24075–24083.
13. Wayment-Steele,H.K., Kim,D.S., Choe,C.A., Nicol,J.J., Wellington-Oguri,R., Watkins,A.M., Parra Sperberg,R.A., Huang,P.-S., Participants,E. and Das,R. (2021) Theoretical basis for stabilizing messenger RNA through secondary structure design. *Nucleic Acids Res.*, **49**, 10604–10617.
14. Celander,D.W. and Cech,T.R. (1991) Visualizing the higher order folding of a catalytic RNA molecule. *Science*, **251**, 401–407.
15. Mustoe,A.M., Brooks,C.L. and Al-Hashimi,H.M. (2014) Hierarchy of RNA functional dynamics. *Ann. Rev. Biochem.*, **83**, 441–466.
16. Lai,D., Proctor,J.R. and Meyer,I.M. (2013) On the importance of cotranscriptional RNA structure formation. *Rna*, **19**, 1461–1473.
17. Watters,K.E., Strobel,E.J., Yu,A.M., Lis,J.T. and Lucks,J.B. (2016) Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat. Struct. Mol. Biol.*, **23**, 1124–1131.
18. Villamizar,O., Chambers,C.B., Riberdy,J.M., Persons,D.A. and Wilber,A. (2016) Long noncoding RNA Saf and splicing factor 45 increase soluble Fas and resistance to apoptosis. *Oncotarget*, **7**, 13810–13826.
19. Flores,J.K. and Ataide,S.F. (2018) Structural changes of RNA in complex with proteins in the SRP. *Front. Mol. Biosci.*, **5**, 7.
20. Sun,L., Fazal,F.M., Li,P., Broughton,J.P., Lee,B., Tang,L., Huang,W., Kool,E.T., Chang,H.Y. and Zhang,Q.C. (2019) RNA structure maps across mammalian cellular compartments. *Nat. Struct. Mol. Biol.*, **26**, 322–330.
21. Mazloomian,A. and Meyer,I.M. (2015) Genome-wide identification and characterization of tissue-specific RNA editing events in D. melanogaster and their potential role in regulating alternative splicing. *RNA Biol.*, **12**, 1391–1401.
22. Wu,M. T.-P. and D'Souza,V. (2020) Alternate RNA structures. *Cold Spring Harbor Perspect. Biol.*, **12**, a032425.
23. Meyer,I.M. (2017) In silico methods for co-transcriptional RNA secondary structure prediction and for investigating alternative RNA structure expression. *Methods*, **120**, 3–16.
24. 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.
25. Zuker,M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48–52.
26. 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.
27. Higgs,P.G. (2000) RNA secondary structure: physical and computational aspects. *Quart. Rev. Biophys.*, **33**, 199–253.
28. Eddy,S.R. and Durbin,R. (1994) RNA sequence analysis using covariance models. *Nucleic Acids Res.*, **22**, 2079–2088.
29. Muse,S.V. (1995) Evolutionary analyses of DNA sequences subject to constraints of secondary structure. *Genetics*, **139**, 1429–1439.
30. Dowell,R.D. and Eddy,S.R. (2004) Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction. *BMC Bioinformatics*, **5**, 71.
31. Meyer,I.M. and Miklós,I. (2007) SimulFold: simultaneously inferring RNA structures including pseudoknots, alignments, and trees using a Bayesian MCMC framework. *PLoS Comput. Biol.*, **3**, e149.
32. Rivas,E., Clements,J. and Eddy,S.R. (2017) A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods*, **14**, 45–48.
33. Rivas,E., Clements,J. and Eddy,S.R. (2020) Estimating the power of sequence covariation for detecting conserved RNA structure. *Bioinformatics*, **36**, 3072–3076.
34. 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.
35. Tijerina,P., Mohr,S. and Russell,R. (2007) DMS footprinting of structured RNAs and RNA–protein complexes. *Nat. Protoc.*, **2**, 2608–2623.
36. Zubradt,M., Gupta,P., Persad,S., Lambowitz,A.M., Weissman,J.S. and Rouskin,S. (2017) DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods*, **14**, 75–82.
37. Smola,M.J. and Weeks,K.M. (2018) In-cell RNA structure probing with SHAPE-MaP. *Nat. Protoc.*, **13**, 1181–1195.
38. Flynn,R.A., Zhang,Q.C., Spitale,R.C., Lee,B., Mumbach,M.R. and Chang,H.Y. (2016) Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat. Protoc.*, **11**, 273–290.
39. Sükösd,Z., Knudsen,B., Kjems,J. and Pedersen,C.N. (2012) PPfold 3.0: fast RNA secondary structure prediction using phylogeny and auxiliary data. *Bioinformatics*, **28**, 2691–2692.
40. Sahoo,S., Świtnicki,M.P. and Pedersen,J.S. (2016) ProbFold: a probabilistic method for integration of probing data in RNA secondary structure prediction. *Bioinformatics*, **32**, 2626–2635.
41. Tsybulskyi,V. and Meyer,I.M. (2022) ShapeSorter: a fully probabilistic method for detecting conserved RNA structure features supported by SHAPE evidence. *Nucleic Acids Res.*, **00**, gkac405.
42. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
43. Lai,D., Proctor,J.R., Zhu,J. Y.A. and Meyer,I.M. (2012) R-CHIE: a web server and R package for visualizing RNA secondary structures. *Nucleic Acids Res.*, **40**, e95.
44. Tsybulskyi,V., Mounir,M. and Meyer,I.M. (2020) R-chie: A web server and R package for visualizing cis and trans RNA–RNA,

RNA–DNA and DNA–DNA interactions. *Nucleic Acids Res.*, **48**, e105.
45. Gruber,A.R., Bernhart,S.H. and Lorenz,R. (2015) The ViennaRNA web services. *RNA Bioinformatics*, 307–326.
46. Smith,C., Heyne,S., Richter,A.S., Will,S. and Backofen,R. (2010) Freiburg RNA Tools: a web server integrating INTARNA, EXPARNA and LOCARNA. *Nucleic Acids Res.*, **38**, W373–W377.
47. Rivas,E. (2020) RNA structure prediction using positive and negative evolutionary information. *PLoS Comput. Biol.*, **16**, e1008387.
48. Wiebe,N.J. and Meyer,I.M. (2010) Transat—a method for detecting the conserved helices of functional RNA structures, including transient, pseudo-knotted and alternative structures. *PLoS Comput. Biol.*, **6**, e1000823.
49. Proctor,J.R. and Meyer,I.M. (2013) CoFold: an RNA secondary structure prediction method that takes co-transcriptional folding into account. *Nucleic Acids Res.*, **41**, e102–e102.
50. Martín,A.L., Mounir,M. and Meyer,I.M. (2021) CoBold: a method for identifying different functional classes of transient RNA structure features that can impact RNA structure formation in vivo. *Nucleic Acids Res.*, **49**, e19.
51. Pedersen,J.S., Meyer,I.M., Forsberg,R., Simmonds,P. and Hein,J. (2004) A comparative method for finding and folding RNA secondary structures within protein-coding regions. *Nucleic Acids Res.*, **32**, 4925–4936.
52. Pedersen,J.S., Forsberg,R., Meyer,I.M. and Hein,J. (2004) An evolutionary model for protein-coding regions with conserved RNA structure. *Mol. Biol. Evol.*, **21**, 1913–1922.
53. Bogdanow,B., Wang,X., Eichelbaum,K., Sadewasser,A., Husic,I., Paki,K., Budt,M., Hergeselle,M., Vetter,B., Hou,J. *et al.* (2019) The dynamic proteome of influenza A virus infection identifies M segment splicing as a host range determinant. *Nat. Commun.*, **10**, 5518.
54. Reißer,S. and Meyer,I.M. (2021) Investigating the concept of accessibility for predicting novel RNA–RNA interactions. bioRxiv doi: https://doi.org/10.1101/2021.06.03.446902, 03 June 2021, preprint: not peer reviewed.