Peer-Review Report

Peer Review of “Exploring the Accuracy of Ab Initio Prediction Methods for Viral Pseudoknotted RNA Structures (Preprint)”

Daniela Saderi[1]; Vaishnavi Nagesh[2]; Randa Salah Gomaa Mahmoud[3]; Toba Olatoye[4]; Femi Qudus Arogundade[5]

[1]PREreview, Portland, OR, United States

[2]Mammoth Biosciences, Brisbane, CA, United States

[3]Zagazig University, Zagazig, Egypt

[4]Kwara State Teaching Service Commission, Ilorin, Nigeria

[5]ASAPBIO, Nigeria

Related Articles:
Preprint (bioRxiv): https://www.biorxiv.org/content/10.1101/2024.03.21.586060v1
Authors' Response to Peer-Review Reports: https://bio.jmirx.org/2024/1/e67586/
Published Article: https://bio.jmirx.org/2024/1/e58899/

(JMIRx Bio 2024;2:e65154) doi: 10.2196/65154

This is a peer-review report submitted for the preprint “Exploring the Accuracy of Ab Initio Prediction Methods for Viral Pseudoknotted RNA Structures.”

This review is the result of a virtual collaborative live review discussion organized and hosted by PREreview and JMIR Publications on June 20, 2024. The discussion was joined by 11 people: 2 facilitators, 2 members of the JMIR Publications team, 1 author, and 6 live review participants, including 2 who agreed to be named: Mike Chang and Heba Abdullah Mohammed Ali. The authors of this review have dedicated additional asynchronous time over the course of 2 weeks to help compose this final report using the notes from the live review. We thank all participants who contributed to the discussion and made it possible for us to provide feedback on this preprint.

Summary

The study [1] examines the performance of 5 RNA-folding engines for predicting complex viral pseudoknotted RNA structures. This research fills a critical gap in the field by comparing the efficiency of minimal free energy (MFE) and maximum expected accuracy (MEA) using a curated dataset of 26 viral RNA sequences with known secondary structures. Contrary to prevailing assumptions favoring MEA models, their findings reveal that pKiss, an MFE-folding engine, outperforms Vsfold 5 in terms of the sensitivity, positive predictive value (PPV), and F_1-score, while laying emphasis on the importance of the PPV and sensitivity parameters in understanding and determining the superior accuracy of pKiss to predict correct base pairs and minimize incorrect predictions. The authors also point out that the engine still needed additional data to achieve high accuracy as well as a better understanding of thermodynamics at the intracellular level.

The statistical analyses used to evaluate the results were 2-way ANOVA and Tukey multiple comparisons test, which provided robust insights into the performance differences among the tested engines. The research integrates bioinformatics with statistics and advanced data science methodologies to promote our understanding of computational RNA biology. The study provides important insights into the relative advantages and disadvantages of both approaches in predicting pseudoknotted RNA structures by contrasting MFE models and MEA models. It also highlights avenues for future research to focus on the development of more sophisticated energy models and MFE engines, like pKiss, to enhance prediction capabilities, especially in the context of viral replication and gene regulation, which may lead to a better understanding of the functional roles of pseudoknotted RNA structures. Overall, this research contributes significantly to the field of computational and molecular biology.

Below, we list major and minor concerns that were discussed by participants of the live review, and where possible, we provide suggestions on how to address those issues.

List of Major Concerns and Feedback

- It would be helpful to provide more context on why percent error was chosen as the primary metric for evaluating different engines, considering alternatives like mean absolute error (MAE) and mean squared error could enhance the analysis. For instance, MAE is robust against outliers, making it a valuable metric, especially when outlier removal is part of the process. Although MAE is less sensitive to extreme values, it can offer a useful qualitative check on the models. On the other hand, the mean squared error’s sensitivity to outliers can be advantageous when the spread of the forecast is important. Including these metrics could provide a more comprehensive evaluation.
- The authors have conducted a comprehensive and insightful study, revealing important differences in prediction accuracy between Vsfold 5 and pKiss. One area that could further enhance the manuscript is the exploration of how auxiliary parameters (eg, Mg2+ binding, dangling end options, H-type penalties) are managed across the various RNA-folding

engines utilized. For example, Vsfold 5, although being an MEA model, may encounter challenges if its handling of Mg2+ binding or dangling ends significantly diverges from what is optimal for the studied RNAs. The authors' observation in section 3.1 that "the low percent error exhibited by pKiss could be the result of the pseudoknot 'enforce' constraint, but it is more likely that this outcome was multivariable, equating to the Turner energy model used, and the sensitive auxiliary parameters enforced by the program" is particularly insightful. This highlights the complexity of RNA structure prediction algorithms. To build on these findings, a structured comparative analysis of parameter handling across different software tools could be highly beneficial. This analysis would not only clarify why certain engines performed better than others but also help in identifying best practices or potential biases in prediction methodologies. Such an addition would significantly strengthen the study's conclusions and provide valuable guidance for future research in RNA structure prediction.

- To build on these findings, a structured comparative analysis of parameter handling across different software tools could be highly beneficial. This analysis would not only clarify why certain engines performed better than others but also help in identifying best practices or potential biases in prediction methodologies. Such an addition would significantly strengthen the study's conclusions and provide valuable guidance for future research in RNA structure prediction.
- In section 3.1 of the manuscript, no significant difference in percent error was identified. However, it does not specify the statistical test employed nor the method used for adjusting *P* values, which are essential details for validating the results. Additionally, the term "Vij" is introduced early in the manuscript but is not contextualized until page 13. Providing this context earlier would enhance the reader's understanding.
- It would be beneficial if "false positive" and "false negative" were more clearly defined, particularly in the context of mRNA detection. To improve clarity, the authors might consider specifying that sensitivity is the appropriate measure for detecting mRNA among known positives, while specificity is the appropriate measure for detecting mRNA among known negatives, where the probability of false positives is 1 – specificity. Additionally, using the Youden index (J), which is defined as sensitivity + specificity – 1, could provide a helpful summary of detection accuracy. This index ranges from –1 (indicating 100% incorrect detection) to 1 (indicating 100% correct detection), offering a clear metric for assessing performance [2].
- Providing the link to the dataset will allow better compliance with open science practices. Please add the link to the dataset as it appears to be missing from the reviewed version of the manuscript. When sharing the dataset, it would be important to also include the associated metadata and appropriate documentation that matches the methods described in the manuscript. For guidelines on how to share data so that it's as reusable as it can be, authors may refer to the Findability, Accessibility, Interoperability, and Reuse (FAIR) principles of data sharing [3].
- Figure 5B displays the PPV as three distinct blocks rather than continuous values, with varying sensitivity within these blocks. This nonrandom binning of PPV suggests the need for further investigation to understand the underlying causes.
- In the Discussion section, the authors stated "We have provided evidence suggesting that MEA software is not always the optimal method of topological prediction when applied to short viral pseudoknotted RNA." This is a significant claim and would benefit greatly from specific references to support the evidence provided in the study. Citing the relevant figures and results that support this claim would significantly enhance comprehension and readability. For example, "As demonstrated in Figure 4, the MEA software Vsfold 5 exhibited higher percent errors in predicting knotted base pairs compared to MFE software like pKiss." Additionally, referencing previous studies that have reported similar findings or that discuss the limitations of MEA methods in RNA structure prediction in the Discussion section would strengthen the credibility of the authors' claims by showing that similar limitations have been observed by other researchers. This helps readers understand that the study is building upon existing knowledge. For instance, "Previous studies have also highlighted the limitations of MEA methods in RNA folding predictions, particularly for pseudoknotted structures (in-text citations)."

List of Minor Concerns and Feedback

Overall, the reviewers really appreciated how clearly the figures and results were presented. Below are some minor suggested improvements.

- In the Abstract section: Please identify the abbreviation PPV as positive predictive value.
- Page 3, first paragraph after Figure 1: Definitions of pseudoknot should be referenced.
- Page 3, second paragraph after Figure 1: Please identify the NMR abbreviation as nuclear magnetic resonance.
- Page 7: The manuscript acknowledges the skewness in the data and provides a rationale for its presence. It's noted that this skewness impacts the training and testing phases, often contributing to false positives and false negatives. It would be beneficial if the authors could elaborate on how they addressed data imbalance, particularly in relation to reducing false positives and false negatives. This additional detail would enhance the understanding of the methods used to manage data skewness and improve model performance.
- Page 8, second paragraph: Mathews et al. 2019 should be corrected to Mathews, 2019 [4].
- Page 8, equation 1: Add a "%" next to *100, giving the output of x%.
- Page 10, Figure 4: In the title, "accurcy" should be corrected to "accuracy."
- Page 10, Figure 4: The bar of the SD of Vienna (knotted) is not presented.

- Page 10, Figure 4: The bars of the SD seem to be widely large, indicating significant variability in the results, so a test of the normality of data distribution should be performed before comparisons. This is also observed for the kinefold results in Figures 5 and 6.
- Page 12, Figure 6B: The color bar on the heat maps is missing.

Concluding Remarks

One of the authors of the manuscript (VM) was present during the call and provided some additional information regarding the source code, which we would like to report here as an additional resource for the reader.

Author's note: Although an original source code was not implemented within this investigation, several well-established web servers were used to generate the data present within this investigation. The link to each web server is provided below.

- VSfold5 [5]
- pKiss [6]
- Kinefold [7]
- NUPACK 3.0 [8]
- RNAfold [9]

Acknowledgments

PREreview and JMIR Publications thank the authors of the preprint for posting their work openly for feedback. We also thank all participants of the live review call for their time and for engaging in the lively discussion that generated this review.

Conflicts of Interest

DS was a facilitator of this call and one of the organizers. No other competing interests were declared by the reviewers.

References

1. Medeiros V, Pearl JM, Carboni M, Er E, Zafeiri S. Exploring the accuracy of ab initio prediction methods for viral pseudoknotted RNA structures. Preprint posted online on February 23, 2023. [FREE Full text] [doi: 10.1101/2024.03.21.586060]
2. Youden index. ScienceDirect. 2010. URL: https://www.sciencedirect.com/topics/medicine-and-dentistry/youden-index [accessed 2024-08-26]
3. FAIR principles. GO FAIR. URL: https://www.go-fair.org/fair-principles/ [accessed 2024-08-26]
4. Mathews DH. How to benchmark RNA secondary structure prediction accuracy. Methods. Jun 01, 2019;162-163:60-67. [FREE Full text] [doi: 10.1016/j.ymeth.2019.04.003] [Medline: 30951834]
5. vsfold5: RNA pseudoknot prediction server. Server for Structural Biology. URL: http://www.rna.it-chiba.ac.jp/~vsfold/vsfold5/ [accessed 2024-08-26]
6. pKiss. Bielefeld BioInformatics Service. URL: https://bibiserv.cebitec.uni-bielefeld.de/pkiss [accessed 2024-08-26]
7. KineFold request form. KineFold Web Server. URL: http://kinefold.curie.fr/cgi-bin/form.pl [accessed 2024-08-26]
8. Analysis: input. NUPACK. URL: https://nupack.org/analysis/input [accessed 2024-08-26]
9. RNAfold web server. Institute for Theoretical Chemistry: VennaRNA Web Services. URL: http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi [accessed 2024-08-26]

Abbreviations

FAIR: Findability, Accessibility, Interoperability, and Reuse
MAE: mean absolute error
MEA: maximum expected accuracy
MFE: minimal free energy
PPV: positive predictive value

Edited by G Eysenbach; this is a non–peer-reviewed article. Submitted 06.08.24; accepted 06.08.24; published 11.09.24.

Please cite as:
Saderi D, Nagesh V, Mahmoud RSG, Olatoye T, Arogundade FQ
Peer Review of "Exploring the Accuracy of Ab Initio Prediction Methods for Viral Pseudoknotted RNA Structures (Preprint)"
JMIRx Bio 2024;2:e65154
URL: https://bio.jmirx.org/2024/1/e65154
doi: 10.2196/65154
PMID: