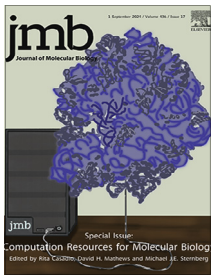

RiboVision2: A Web Server for Advanced Visualization of Ribosomal RNAs

Holly M. McCann [†] **Caeden D. Meade, Biswajit Banerjee, Petar I. Penev** [††] **Loren Dean Williams and Anton S. Petrov** [*]

NASA Center for the Origin of Life, *Georgia Institute of Technology, Atlanta, GA 30332-0400, USA*
School of Biological Sciences, *Georgia Institute of Technology, Atlanta, GA 30332, USA*

Correspondence to Anton S. Petrov: *anton.petrov@biology.gatech.edu (A.S. Petrov)*
https://doi.org/10.1016/j.jmb.2024.168556
Edited by David Mathews

Abstract

RiboVision2 is a web server designed to visualize phylogenetic, structural, and evolutionary properties of ribosomal RNAs simultaneously at the levels of primary, secondary, and three-dimensional structure and in the context of full ribosomal complexes. RiboVision2 instantly computes and displays a broad variety of data; it has no login requirements, is open-source, free for all users, and available at https://ribovision2.chemistry.gatech.edu.

Introduction

The translation of mRNA into protein, catalyzed by the ribosome,[1] set the path of biology and established the paradigm that has dominated the biological earth for over 3.8 billion years.[2,3] The first high-resolution ribosomal structures revealed the key functional regions and the major principles of its functionality. Since then, due to advances in X-ray[2,4–7] and cryo-EM techniques,[8–10] the number of ribosomal structures has grown exponentially and is currently in excess of a thousand.[11]

The information within the ribosomes can be deconstructed, compiled, represented, and visualized in multiple ways on multiple levels[12,13]: A) Sequences and multiple sequence alignments (MSAs) contain information about evolution and variability or conservation of sequences across phylogeny.[14–16] B) Secondary structures of rRNAs provide information about local interactions, structural elements and types of base pairs, and overall topological organization of RNA chains.[17,18] C) 3D structures provide detailed all-atom data about mechanism, spatial organization, and interactions within ribosomal complexes.[11,19,20]

The last decade has demonstrated substantial progress in methods that enable prediction and visualization of various elements of ribosomal structures.[21–28] Many new tools are available as online applets, enabling the robust and interactive automated processing of large ribosomal complexes. Secondary structures can now be easily predicted using template-based algorithms (R2DT),[29] and edited and visualized within web browsers. Advances in WebGL technology have led to development of powerful applets that provide 3D visualization of large macromolecular complexes with minimal hardware requirements.[30] Many of these tools have become available in collections provided by EBI (component library),[31,32] which enables easy integration across various applets.

Here we present RiboVision2, a web server for analysis and visualization of ribosomes with a focus on ribosomal RNAs (rRNAs). RiboVision2 represents a major upgrade of the original RiboVision (https://apollo.chemistry.gatech.edu/

RiboVision2/) deployed in 2013,[13] which was limited to a dozen manually processed ribosomal structures. The new version provides capabilities to process, visualize, compute, and map data for any ribosomal structure in the wwPDB,[33,34] and therefore represents a major advance of the original program. The new web server also provides the ability to upload sequences and structures of other (non-ribosomal) RNAs, even those that are not contained in the RiboVision2 database. Furthermore, it integrates modern visualization applets within a common environment.

Results

Web server description

RiboVision2 is an open-source web server (https://github.com/LDWLab/Ribovision_2.0_GT) crafted to facilitate visualization, analysis and data mapping of rRNAs on multiple levels; sequence, secondary structure and 3D structure. RiboVision2 operates in two distinct modes. In the RiboVision mode, the server features preloaded alignments of rRNAs from its database[35,36] and links that data to 2D and 3D structures available from RNAcentral[18] and wwPDB[33,34] consortia. The second, User-Upload mode, empowers users to provide proprietary alignments and 3D structures of any RNA molecules for visualization and data mapping. The essential inputs for RiboVision2 include a FASTA multiple sequence alignment (MSA) of rRNAs and a 3D structure file. However, the methods for uploading and handling data differ significantly between these two modes.

The interface of RiboVision2 consists of the Main Navigation panel, which operates across three synchronized applets located in the Visualization panel. The MSAViewer applet[37] provides interactive navigation of multiple sequence alignments. The 2D RNA Viewer (https://github.com/LDWLab/pdb-rna-viewer) provides options to display RNA secondary structures in various representations and serves as a hub for data mapping. Finally, PDBe Mol*[30,38] delivers various 3D representations of the RNA molecules and their complexes in a fast and interactive fashion. The server not only displays visualizations of ribosomal structures but also assists in calculating RNA-protein contacts and nucleotide modifications. RiboVision2 offers an optional interactive guided tour, aimed at providing a comprehensive description of each functional element within its interface.

Web server development

RiboVision2 follows a client–server architecture; the web server is hosted at Georgia Institute of Technology and served by Apache/RHEL8. The backend is powered by Python 3.8 and uses Django for API queries and data processing. The front-end user interface was developed using HTML5/JavaScript, leveraging the React and Vue frameworks. The data are stored in a MySQL database. RiboVision2 also relies on information provided by several external APIs and databases. Website architecture is illustrated in Figure 1.

RiboVision mode

In RiboVision mode, the multiple sequence alignment (MSA) is automatically constructed by selecting a desired subset of species from the phylogenetic browser found in the Main Navigation panel (Figure 2). This browser allows recursive selection of species from one or multiple taxonomic groups.

After loading the multiple sequence alignment, RiboVision2 calculates gap-adjusted nucleotide frequencies from the MSA. These frequencies are then utilized to compute phylogenetic conservation scores, employing methods such as Shannon Entropy or TwinCons.[39] Next, the corresponding rRNA structure is selected by the user. This selection is accomplished by either typing in or choosing a PDB ID from a list of available ribosomal structures that appears in the Main Navigation panel. This list is dynamically generated via an API from ribosome.xyz, ensuring a real-time catalog of ribosomal structures. The server then filters and displays all accessible RNA chains within the provided PDB complex. It is then the user's responsibility to select a chain that aligns with the MSA. Once the chain is selected, its structure is visualized in the 2D RNA Viewer and 3D Mol* Viewer (Figure 3).

An essential feature of RiboVision2 is its interactive visualization of secondary structures for

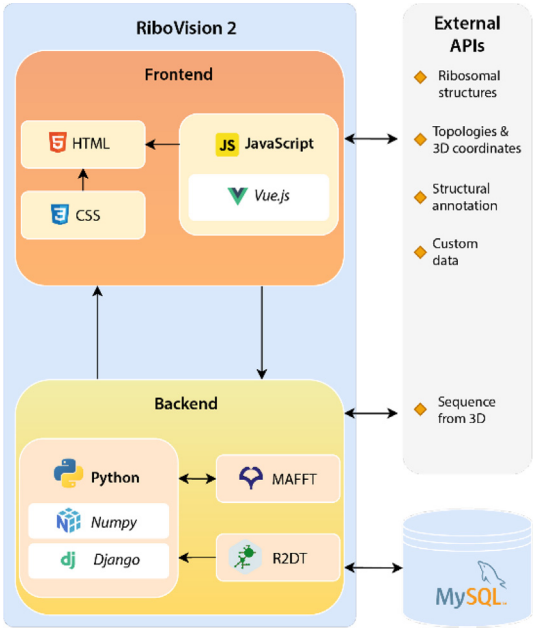

Figure 1. Schema of the RiboVision2 webserver.

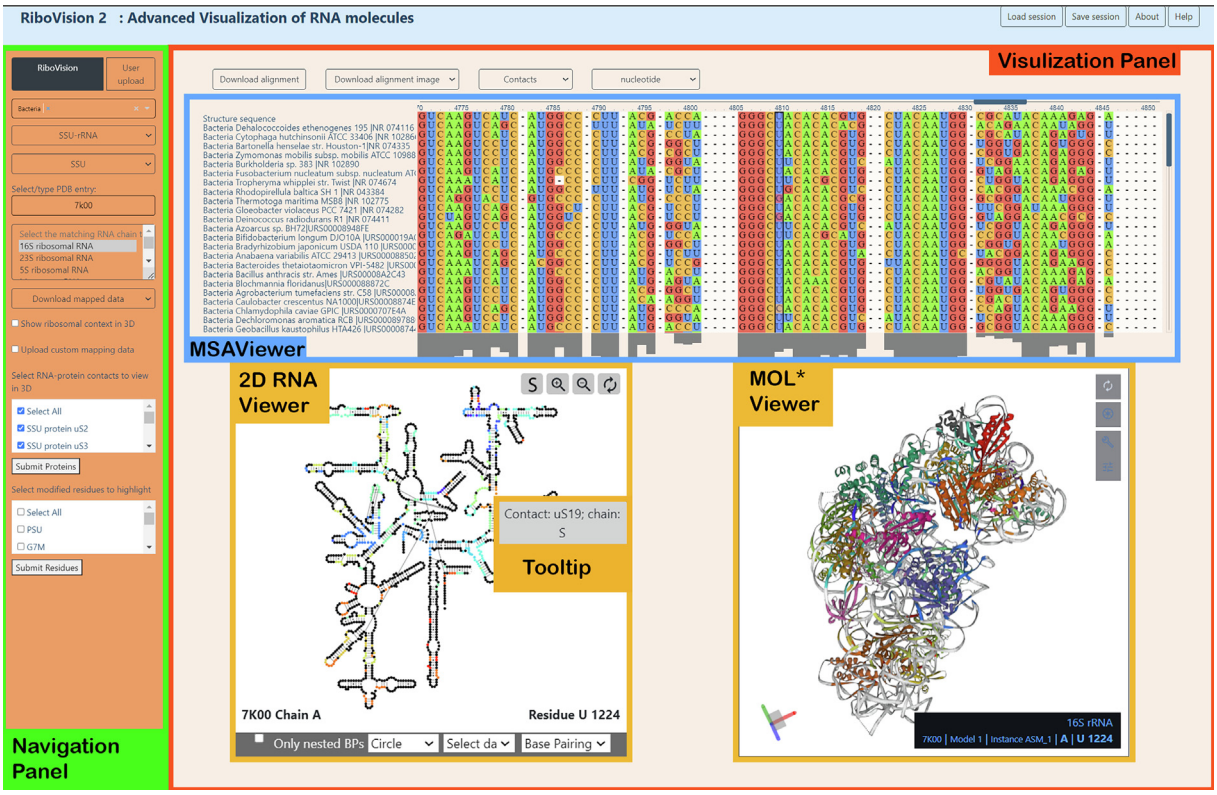

Figure 2. Frontend of RiboVision2, The Main Navigation Panel.

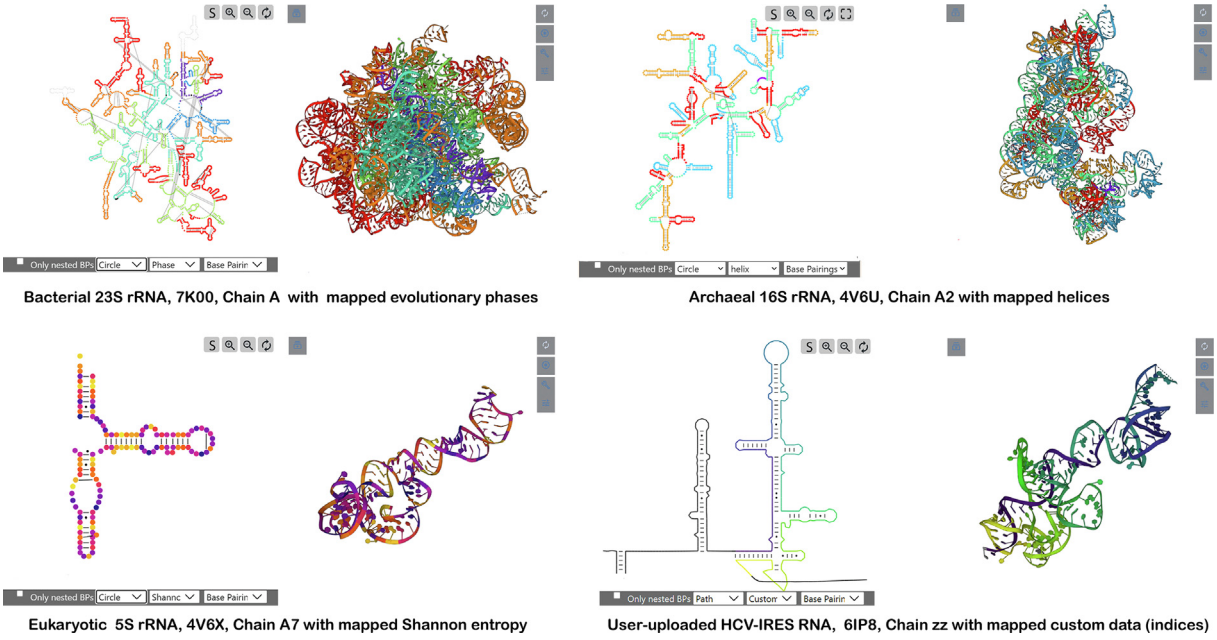

Figure 3. Demonstration RiboVision2 server capabilities.

rRNA molecules fetched from the PDBe REST API (https://www.ebi.ac.uk/pdbe/pdbe-rest-api). This visualization integrates template-derived predictions from the R2DT pipeline[29] of the RNAcentral consortium. Base pairs of secondary structures are either predicted from templates or extracted from 3D structures, and annotated according to the Leontis-Westhof notation.[40]

Finally, the 3D structures of rRNAs (alone or with ribosomal proteins) are visualized in the Mol* applet. The structural applet uses data fetched from several external databases (EMBL-EBI, ribosome.xyz, RNA 3D Hub – BGSU).

MSA

To visualize the alignment, RiboVision2 utilizes an in-house customized implementation of the robust open-source JavaScript component, MSAViewer[37] (https://github.com/plotly/react-msa-viewer/). The MSAViewer code was adapted for RiboVision2 to facilitate integration with other plug-ins. This integration allows for (i) data mapping sharing, (ii) dynamic positioning, and (iii) hover event-based highlighting of specific positions within the alignment. Furthermore, modifications were made to the MSAViewer's bar-plot feature to exhibit conservation scores and other data mappings. Each bar is color-coded based on selected alignment-derived data.

2D RNA Viewer

The 2D RNA Viewer from PDBe's component library[32,33] has been adapted to facilitate data mapping and interactivity with the other components. This viewer was enhanced by adding several drop-down menus for data visualization. The first allows users to select between displaying nucleotides as individual letters (default), smooth paths, or circles. The second allows users to choose structural, evolutionary, or custom-uploaded data to map as colors onto the nucleotides. When data is mapped from this menu or from selections of protein contacts or modified residues, tooltips are enabled which display the mapping values for each nucleotide upon hovering. In the third menu, users can choose types of base pairs to be displayed based on the Leontis-Westhof definitions.[40] Base pairs are represented as lines accompanied by Leontis-Westhof symbols illustrating their specific types. Hovering over a base pair reveals its type and identifies the nucleotides involved in the pairing. In RiboVision mode, data for drawing the secondary structure diagram are obtained from the PDBe REST API based on the selected PDB structure. Canonical and non-canonical base pairs are obtained from the FR3D API.[41] In User-Upload mode, secondary structures are dynamically computed using R2DT and canonical base pairs are predicted based on template structures.[42]

Mol* Viewer

3D representations of selected or uploaded RNA structures are displayed as ribbon diagrams using an enhanced version of the PDBe Mol* Viewer[30] (available at https://github.com/LDWLab/pdbe-molstar-GT). Coloring of individual nucleotides on large structures was significantly accelerated by creating a color wrapper that incorporates a set of color themes within the Mol* Viewer.[30] This enables the efficient mapping of selected data onto nucleotides utilizing precomputed color palettes. Additionally, the structures of proteins in contact with the RNA chain can be visualized along with the chain, and the nucleotides they are in contact with shown in the same color as the protein.

Applet synchronization

Data visualizations are synchronized across the MSAViewer, 2D RNA Viewer, and Mol* Viewer. When a nucleotide is hovered on in any of the three viewers, the same nucleotide is highlighted in each of the other viewers, and its index and type are displayed in the 2D and Mol* viewers, creating a clear visual connection of the components of the RNA structures across three levels. When a user selects data to map as colors onto the 2D viewer, the same data is simultaneously mapped onto the Mol* Viewer and displayed as color bars beneath the MSAViewer. Modified residues or protein contacts selected are shown in the same colors in both the 2D and 3D structures.

Data saving

Users can map calculated or user-supplied data onto each level of sequence/structure representation. The alignment displayed in the MSAViewer can be saved as a FASTA (sequence) or PNG (figure) file. The secondary structure diagram, along with the current data layer, can be saved as an SVG file using the 2D RNA Viewer. The current visual representation of the 3D RNA complex can be saved directly from the Mol* Viewer.

RiboVision2 also offers the option to save all relevant data—conservation scores, structural and evolutionary information, protein contacts, nucleotide modifications, as well as custom data—as a CSV table or PyMol script. These options become accessible in the Main Navigation panel after structures are selected or uploaded and the data are computed. Data saving functionality is supported in both RiboVision and User-Upload modes.

RiboVision2 data

MSAs. RiboVision2 contains a pre-uploaded set of alignments for rRNAs. These alignments have been derived for 150 bacterial, archaeal, and eukaryotic species, following the procedure described in Bernier et al.[35] This procedure incorporates information from available 3D structures. Due to the independent treatment of RNA chains by

PDBe and R2DT, alignments for 5.8S rRNA and 28S rRNA are provided as separate objects.

Ribosomal structures

RiboVision2 supports visualization of all structures that are currently deposited into wwPDB[34] via a set of APIs:

The secondary structures are obtained from EBI using the following URL pattern: "https://www.ebi.ac.uk/pdbe/static/entry/{pdb_id}_{entity_id}_{chain_id}.json", for example https://www.ebi.ac.uk/pdbe/static/entry/7k00_22_a.json.

The 3D structures are fetched from the RCSB coordinate server using the following template: "https://models.rcsb.org/v1/{pdb_id}/atoms?label_entity_id={entity_id}&encoding = bcif".

The desired 3D structure can be selected by typing the corresponding PDB ID or by choosing from the automatically generated pre-filtered list of available ribosomal structures, supplied by ribosome.xyz[11] using the API template: "https://api.ribosome.xyz/neo4j/get_rna_class/?rna_class={rRNA}&format = json", where, "rRNA" represents the RNA class of a selected molecule, for example, 23SrRNA.

Shannon entropy

The Shannon entropy (the conservation score) ranges from 0 (full conservation) to 2 (fully random) for a 4-letter alphabet. The Shannon entropy is calculated based on the gap-adjusted nucleotide frequencies derived from the MSA. Gap frequencies are randomly assigned to one of the four RNA nucleotides using a probability of 0.25.

TwinCons

When two groups are selected in the phylogenic browser of RiboVision2, an additional option, TwinCons, becomes available for computing additional conservation metrics. TwinCons[39] operates using a composite MSA containing two predefined groups. The distribution of nucleotides at a given position of each group is transformed to a vector of gap-adjusted nucleotide frequencies.

Using a pair of vectors as input at every position of the composite MSA, TwinCons calculates joint conservation using the substitution matrix blastn[43] (available at https://github.com/LDWLab/TwinCons). The resulting score signifies:

a) Joint conservation in both groups (high positive values, reaching a maximum of 6.75); b) Divergence or signature positions, where nucleotides are highly conserved within each group but differ between the groups (resulting in low negative values, up to −2.25); c) Random values observed in one or both groups (ranging between 0 and 2).

Quantitatively, TwinCons represents the transformation cost between the two vector columns via the substitution matrix blastn. This approach facilitates the simultaneous detection and visualization of jointly conserved, divergent (or signature), and highly variable positions within rRNAs.

Protein contacts

RiboVision2 automatically computes RNA-protein interactions for every ribosomal complex analyzed in the RiboVision mode. The computation of RNA-protein contacts utilizes the NeighborSearch module of BioPython,[44] employing the KD Tree algorithm[45] with a cutoff distance set at 3.5 Å. Once this computation is performed, a list of proteins within the threshold from the selected RNA chain becomes accessible in the Main Navigation panel.

Users can select one or several proteins from this list. Upon selection, corresponding RNA nucleotides are colored in the 2D RNA Viewer. Moreover, detailed information about specific contacts can be explored through tooltips. A tooltip will appear upon hovering over a nucleotide in the 2D applet; it will reveal the protein name and chain ID within the ribosomal complex. In addition, RNA nucleotides in contact with selected rProteins are highlighted in the 3D applet using the same color scheme as in the 2D representation.

Chemical modifications

RiboVision2 offers the ability to map and explore modifications of nucleotides within ribosomal structures. When a selected RNA chain contains modified nucleotides (as specified in the "_entity_poly.pdbx_seq_one_letter_code" section of a chosen mmCIF file), the abbreviated modification (e.g., 1MG, PSU, 5MU, etc.) will be displayed in the "List of Modified Nucleotides" section. Users can then select one or several modified residue types from this list, leading to the simultaneous highlighting of the respective groups of modified nucleotides in both the 2D and 3D applets.

Users can interactively access detailed information about a particular nucleotide's modifications by hovering over it in the 2D RNA Viewer.

Structural and evolutionary data (helices, expansion segments, and phases)

RiboVision2 offers visualization of structural and evolutionary data related to rRNAs, including helices, expansion segments, and ancestral expansion segments (Figure 3). To depict helices, we used definitions from Petrov,[46] where we ensure each nucleotide belongs uniquely to one helix, containing contiguous base-paired or stacked bases. Definitions of Ancestral Expansion Segments and evolutionary phases were taken from the evolutionary accretion model.[47,48]

The evolutionary data were deposited into RiboVision2's database for a limited number of anchor structures. When users select a ribosomal sequence, RiboVision2 maps pre-existing evolutionary data onto the chosen sequence. This data can then be visualized in both 2D and 3D applets. For detailed information at a nucleotide level, interactive exploration is enabled through tooltips.

User-Upload data and mapping

RiboVision2 also contains a User-Upload mode, which allows users to upload a custom alignment (or single sequences) along with custom 3D RNA structures (as mmCIF or PDB files), which have not been previously uploaded to structural databases. Thus, the User-Upload mode is not limited to ribosomal structures that have been already deposited into the PDB. Furthermore, it is not limited to rRNAs, but supports processing and visualization of any RNA molecule, given that a user can upload an MSA (or a single RNA sequence) and the 3D structure. From the supplied information, the User-Upload mode will then attempt to compute the 2D diagram. Generation of the 2D layout is achieved by the R2DT algorithm, which is automatically executed at the backend of the server. The R2DT algorithm[29] provides visualization of 2D RNA structures in standardized layouts using template derived algorithms.

MSAs

In the User-Upload mode of RiboVision2, the initial step involves uploading an MSA or a single sequence of the desired RNA molecule in the fasta format. The fasta header format is generally flexible. However, TwinCons[39] calculation is supported only when the headers can be distinctly partitioned into two groups based on their prefixes within a composite MSA (*e.g.* "Archaea_species1"/, "Archaea_species2", "Bacteria_species1", "Bacteria_species2"). RiboVision2 automatically detects and partitions the sequences into two groups (according to the headers), then computes the TwinCons score.

Once the custom alignment is uploaded, RiboVision2 proceeds to compute the conservation statistics. These numerical values are graphically depicted via a bar plot situated beneath the Multiple Sequence Alignment (MSA).

PDB structure upload

RiboVision2 offers two distinct options for uploading custom RNA structures. The first option involves uploading the structure in PDB format, which comes with specific requirements:

A) The PDB file must contain only a single RNA chain. If the initial RNA complex comprises multiple chains (including protein chains), the desired RNA chain needs to be extracted as a separate PDB file before being uploaded into RiboVision2. B) Alongside the PDB structure, the complete RNA sequence must be provided to generate the comprehensive 2D layout of the RNA molecule. PDB files might skip nucleotides, especially in flexible regions. To ensure the generation of the 2D structure from the entire RNA sequence, a separate upload of the full RNA sequence is necessary in PDB mode.

Upon uploading both the structure and its corresponding full RNA sequence, RiboVision2 maps both sequences onto the MSA objects. The full RNA sequence is required to generate the 2D layout via the R2DT algorithm. This double mapping procedure ensures synchronization in nucleotides between the MSA and the 2D as well as 3D structures. Synchronization enables all derived or uploaded properties to be simultaneously mapped and visualized within the three representations of RNA. Moreover, this mapping facilitates synchronized interactivity and nucleotide highlighting across the three templates.

mmCIF structure upload

Secondly, RiboVision2 also supports the upload of 3D structures in mmCIF format, accommodating files with multiple chains. However, specific requirements must be met for proper processing: A) The mmCIF file should include attributes such as auth_seq_id, auth_comp_id, label_entity_id, and auth_asym_id. These attributes are essential for identification and processing of the selected RNA chain. B) The mmCIF file must contain the full RNA sequence within its pdbx_seq_one_letter_code record. C) The mmCIF file should include a pdbx_poly_seq_scheme section providing the mapping between nucleotide IDs in the full sequence and those resolved in the 3D structure.

If any of these elements are absent or incomplete in the mmCIF file, users are encouraged to save the desired RNA chain in PDB format and obtain the full RNA sequence for upload using the PDB structure upload option. Upon supplying the mmCIF file, users must also provide the entity ID corresponding to the desired RNA chain, matching the RNA molecule type in the supplied MSA.

Generation of RNA secondary structures

Unlike the RiboVision mode that utilizes pre-generated 2D maps, the User-Upload mode utilizes the R2DT pipeline at the server's backend. The R2DT pipeline generates secondary structure diagrams of uploaded RNA molecules from templates.

Once the MSA and the 3D structure are uploaded, the server automatically processes the inputs. It extracts the full sequence and initiates

the 2D layout generation twice: first, using default options, and if unsuccessful, it retries with the --skip_ribovore_filters option.

Users are encouraged to consult the relevant documentation on R2DT (accessible at https://docs.r2dt.bio/en/latest/usage.html) and exercise patience, as R2DT alone may take several minutes to complete.

Custom data mapping

Once the MSA and 3D structure have been uploaded in the User-Upload mode and the secondary structure has been generated by R2DT, users can map and visualize custom data in all three representations of the uploaded RNA molecule. RiboVision2 supports the upload of custom data in CSV format, containing two columns (Nucleotide_id and Data_value). Upon upload, RiboVision2 processes the data, converting it into the viridis color scheme and color-coding the corresponding nucleotides within each of the three applets based on the supplied data.

Discussion

In the current work we described RiboVision2 – a web server designed for visualization of rRNAs. RiboVision2 is an upgrade of a previous version of RiboVision and provides a set of integrated tools and applets to explore rRNAs across phylogeny, and to visualize any ribosomal complex in the PDB. RiboVision2 computes evolutionary and structural data and allows for instant mapping of the data to any structure selected by users, overcoming the major limitations of the previous version, which supported only a handful of preloaded ribosomal structures.

RiboVision2 also includes the User-Upload mode, offering visualization capabilities for ribosomal structures not in the PDB. Additionally, when coupled with R2DT at the backend, it provides the ability to visualize other classes of RNA molecules (e.g., RNase P or miRNAs) for which a 3D structure is currently available. Both RiboVision and User-Upload modes offer the capability to save data and visualizations in various formats.

The current limitations of RiboVision2 include its dependency on RNA 2D templates currently available within R2DT. Thus, currently the server provides very limited support for mito-ribosomal RNAs. Additionally, at present, the User-Upload mode derives base pairing from predicted 2D structures rather than directly from uploaded 3D structures. Moreover, RiboVision2 currently supports data mapping only using several pre-defined color schemes; the coloring repertoire will be extended in future developments by incorporating additional color palettes, and providing users with the option to select a desired color schema.

RiboVision2 shares visualization philosophy with the previously established server ProteoVision[36] with emphasis on rRNAs and complements it in terms of the ribosomal data. RiboVision2 provides an API service (https://ribovision2.chemistry.gatech.edu/desire-api/) for rRNA nomenclature, sequences, alignments, and annotations.

Conclusion

Here, we presented RiboVision2, an upgrade of the previous release with modern visualization applets integrated into a single web server. This web server enables the exploration of structural features of ribosomal complexes as well as custom RNA molecules across the tree of life, mapping phylogenetic, structural, or evolutionary properties, and incorporating custom data. Additionally, users can save resulting visualizations in various formats suitable for publication-quality images. We hope that RiboVision2 will meet the needs of structural and evolutionary biologists (including those who work in the origins of life field) and can be of interest to a broad audience that enjoys molecular visualizations.

CRediT authorship contribution statement

Holly M. McCann: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Methodology, Formal analysis, Data curation. **Caeden D. Meade:** Visualization, Validation, Software, Data curation. **Biswajit Banerjee:** Resources, Data curation, Software, Writing – review & editing. **Petar I. Penev:** Conceptualization, Software, Visualization. **Loren Dean Williams:** Conceptualization, Funding acquisition, Project administration, Writing – review & editing. **Anton S. Petrov:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by NASA grant 80NSSC24K0344 to L.D.W., A.S.P. The authors

are thankful to Prof. Khanh Dao Duc, Mr. Artem Kushner (ribosome.xyz), Mr. Arsh Suri, Dr. Anton I. Petrov (R2DT), Prof. Craig Zirbel, and Dr. Alex Rose for their support and contributions.

Received 31 January 2024;
Accepted 25 March 2024;
Available online 30 March 2024

† Current address: Department of Biomedical Data Science, Stanford, CA 94305, USA.

†† Current address: Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA 94720, USA.

References

1. Agmon, I., Bashan, A., Yonath, A., (2006). On ribosome conservation and evolution. *Isr. J. Ecol. Evol.* **52**, 359–374.
2. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., et al., (2001). High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* **107**, 679–688.
3. Woese, C.R., Fox, G.E., (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5088–5090.
4. Wimberly, B.T., Brodersen, D.E., Clemons Jr., W.M., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., et al., (2000). Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339.
5. Selmer, M., Dunham, C.M., Murphy IV, F.V., Weixlbaumer, A., Petry, S., Kelley, A.C., et al., (2006). Structure of the 70 S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942.
6. Ramakrishnan, V., (2002). Ribosome structure and the mechanism of translation. *Cell* **108**, 557–572.
7. Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905–920.
8. Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A., et al., (1995). A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome. *Nature* **376**, 441–444.
9. Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., et al., (2013). Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**, 1108–1119.
10. Anger, A.M., Armache, J.-P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., et al., (2013). Structures of the human and Drosophila 80S ribosome. *Nature* **497**, 80–85.
11. Kushner, A., Petrov, A.S., Dao Duc, K., (2022). RiboXYZ: a comprehensive database for visualizing and analyzing ribosome structures. *Nucleic Acids Res.* **51**, D509–D516.
12. Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J. R., D'Souza, L.M., Du, Y., et al., (2002). The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformat.* **3**, 2.
13. Bernier, C.R., Petrov, A.S., Waterbury, C.C., Jett, J., Li, F., Freil, L.E., et al., (2014). RiboVision suite for visualization and analysis of ribosomes. *Faraday Discuss.* **169**, 195–207.
14. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al., (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596.
15. Nawrocki, E.P., Burge, S.W., Bateman, A., Daub, J., Eberhardt, R.Y., Eddy, S.R., et al., (2014). Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* **43**, D130–D137.
16. Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., Eddy, S.R., (2003). Rfam: an RNA family database. *Nucleic Acids Res.* **31**, 439–441.
17. RNAcentral Consortium, (2016). RNAcentral: a comprehensive database of non-coding RNA sequences. *Nucleic Acids Res.* **45**, D128–D134.
18. RNAcentral Consortium, (2020). RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Res.* **49**, D212–D220.
19. Petrov, A.I., Zirbel, C.L., Leontis, N.B., (2013). Automated classification of RNA 3D motifs and the RNA 3D Motif Atlas. *RNA* **19**, 1327–1340.
20. Lawson, C.L., Berman, H.M., Chen, L., Vallat, B., Zirbel, C. L., (2023). The Nucleic Acid Knowledgebase: a new portal for 3D structural information about nucleic acids. *Nucleic Acids Res.*.
21. Yang, H., Jossinet, F., Leontis, N., Chen, L., Westbrook, J., Berman, H., et al., (2003). Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res.* **31**, 3450–3460.
22. Shabash, B., Wiese, K.C., (2019). jViz.RNA 4.0-Visualizing pseudoknots and RNA editing employing compressed tree graphs. *PLoS One* **14**, e0210281.
23. Johnson, P.Z., Kasprzak, W.K., Shapiro, B.A., Simon, A.E., (2019). RNA2Drawer: geometrically strict drawing of nucleic acid structures with graphical structure editing and highlighting of complementary subsequences. *RNA Biol.* **16**, 1667–1671.
24. Johnson, P.Z., Simon, A.E., (2023). RNAcanvas: interactive drawing and exploration of nucleic acid structures. *Nucleic Acids Res.* **51**, W501–W508.
25. Das, R., Watkins, A.M., (2021). RiboDraw: semiautomated two-dimensional drawing of RNA tertiary structure diagrams. *NAR Genom. Bioinformat.* **3**
26. Mathews, D.H., (2006). RNA secondary structure analysis using RNAstructure. *Curr. Protoc. Bioinformat.*. Chapter 12, Unit 12.16.
27. Ponty, Y., Leclerc, F., (2015). Drawing and editing the secondary structure(s) of RNA. *Methods Mol. Biol.* **1269**, 63–100.
28. Eddy, S. R. & Nawrocki, E. P. (2009). Structural rna homology search and alignment using covariance models.
29. Sweeney, B.A., Hoksza, D., Nawrocki, E.P., Ribas, C.E., Madeira, F., Cannone, J.J., et al., (2021). R2DT is a framework for predicting and visualising RNA secondary structure using templates. *Nature Commun.* **12**, 3494.
30. Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Berka, K., Bazgier, V., et al., (2021). Mol* Viewer: modern web app for 3D visualization and analysis of large

biomolecular structures. *Nucleic Acids Res.* **49**, W431–W437.
31. Velankar, S., van Ginkel, G., Alhroub, Y., Battle, G.M., Berrisford, J.M., Conroy, M.J., et al., (2015). PDBe: improved accessibility of macromolecular structure data from PDB and EMDB. *Nucleic Acids Res.* **44**, D385–D395.
32. Mir, S., Alhroub, Y., Anyango, S., Armstrong, D.R., Berrisford, J.M., Clark, A.R., et al., (2017). PDBe: towards reusable data delivery infrastructure at protein data bank in Europe. *Nucleic Acids Res.* **46**, D486–D492.
33. Armstrong, D.R., Berrisford, J.M., Conroy, M.J., Gutmanas, A., Anyango, S., Choudhary, P., et al., (2019). PDBe: improved findability of macromolecular structure data in the PDB. *Nucleic Acids Res.* **48**, D335–D343.
34. wwPDB consortium, (2018). Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **47**, D520–D528.
35. Bernier, C.R., Petrov, A.S., Kovacs, N.A., Penev, P.I., Williams, L.D., (2018). Translation: the universal structural core of life. *Mol. Biol. Evol.* **35**, 2065–2076.
36. Penev, P.I., McCann, H.M., Meade, C.D., Alvarez-Carreño, C., Maddala, A., Bernier, C.R., et al., (2021). ProteoVision: web server for advanced visualization of ribosomal proteins. *Nucleic Acids Res.* **49**, W578–W588.
37. Yachdav, G., Wilzbach, S., Rauscher, B., Sheridan, R., Sillitoe, I., Procter, J., et al., (2016). MSAViewer: interactive JavaScript visualization of multiple sequence alignments. *Bioinformatics.* **32**, 3501–3503.
38. Sehnal, D., Rose, A.S., Koča, J., Burley, S.K., Velankar, S., (2018). Mol*: towards a common library and tools for web molecular graphics. *In: Proc. Wkshp. Mol. Graph. Vis. Anal. Mol. Graph*. Eurographics Association, Brno, Czech Republic, pp. 29–33.
39. Penev, P.I., Alvarez-Carreño, C., Smith, E., Petrov, A.S., Williams, L.D., (2021). TwinCons: conservation score for uncovering deep sequence similarity and divergence. *PLoS Comp. Biol.* **17**, e1009541.
40. Leontis, N.B., Westhof, E., (2001). Geometric nomenclature and classification of RNA base pairs. *RNA.* **7**, 499–512.
41. Petrov, A.I., Zirbel, C.L., Leontis, N.B., (2011). WebFR3D—a server for finding, aligning and analyzing recurrent RNA 3D motifs. *Nucleic Acids Res.* **39**, W50–W55.
42. Elias, R., Hoksza, D., (2017). TRAVeLer: a tool for template-based RNA secondary structure visualization. *BMC Bioinformat.* **18**, 487.
43. States, D.J., Gish, W., Altschul, S.F., (1991). Improved sensitivity of nucleic acid database searches using application-specific scoring matrices. *Methods* **3**, 66–70.
44. Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., et al., (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423.
45. Berg, M., Cheon, O., Kreveld, M., Overmars, M., (2008). Orthogonal Range Searching. *In: Computational Geometry: Algorithms and Applications*. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 95–120.
46. Petrov, A.S., Bernier, C.R., Gulen, B., Waterbury, C.C., Hershkovits, E., Hsiao, C., et al., (2014). Secondary structures of rRNAs from all three domains of life. *PLoS One* **9**, e88222.
47. Petrov, A.S., Gulen, B., Norris, A.M., Kovacs, N.A., Bernier, C.R., Lanier, K.A., et al., (2015). History of the ribosome and the origin of translation. *Proc. Natl. Acad. Sci. USA* **112**, 15396–15401.
48. Petrov, A.S., Bernier, C.R., Hsiao, C., Norris, A.M., Kovacs, N.A., Waterbury, C.C., et al., (2014). Evolution of the ribosome at atomic resolution. *Proc. Natl. Acad. Sci. USA* **111**, 10251–10256.