RNAcanvas: interactive drawing and exploration of nucleic acid structures

Philip Z. Johnson[1] and Anne E. Simon[1,*]

[1] Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland, 20742, USA

* To whom correspondence should be addressed. Tel: +1 (301) 405-8975; Email: simona@umd.edu

ABSTRACT

Two-dimensional drawing of nucleic acid structures, particularly RNA structures, is fundamental to the communication of nucleic acids research. However, manually drawing structures is laborious and infeasible for structures thousands of nucleotides long. RNAcanvas automatically arranges residues into strictly shaped stems and loops while providing robust interactive editing features, including click-and-drag layout adjustment. Drawn elements are highly customizable in a point-and-click manner, including colours, fonts, size and shading, flexible numbering, and outlining of bases. Tertiary interactions can be drawn as draggable, curved lines. Leontis-Westhof notation for depicting non-canonical base-pairs is fully supported, as well as text labels for structural features (e.g., hairpins). RNAcanvas also has many unique features and performance optimizations for large structures that cannot be correctly predicted and require manual refinement based on the researcher's own analyses and expertise. To this end, RNAcanvas has point-and-click structure editing with real-time highlighting of complementary sequences and motif search functionality, novel features that greatly aid in the identification of putative long-range tertiary interactions, de novo analysis of local structures, and phylogenetic comparisons. For ease in producing publication quality figures, drawings can be exported in both SVG and PowerPoint formats. URL: https://rnacanvas.app

INTRODUCTION

In nucleic acids research, particularly RNA research, structures are commonly drawn in 2-dimensional (2D) format (1–3). Conventions such as the Leontis-Westhof notation (4) are also important for denoting various types of canonical and non-canonical base-pairs and interactions in RNA structures. Since manually creating 2D structure drawings is laborious for large structures (5–7), a variety of software programs have been developed to facilitate this process.

2D nucleic acid structure drawing programs can be roughly separated into two categories: 1) programs that offer flexible, force-based adjustment of drawing layouts; and 2) programs that generally apply stricter layout conventions to drawings. Programs in the first category, including forna (8) and RiboSketch (9), allow click-and-drag editing of drawings and use force-based methods (8) to help maintain the arrangement of residues into stems and loops in response to dragging. The interactivity features in this class of programs make editing drawings intuitive, and their use of force-based methods allow for great flexibility in how a drawing is arranged. These programs, however, require increased effort to (for example): 1) precisely align stems emanating from a junction or that are stacked; 2) flatten or evenly round regions of unpaired residues; and 3) align stems with the flat base of a drawing.

Programs in the second category, including VARNA (10), XRNA (http://rna.ucsc.edu/rnacenter/xrna/xrna_faq.html), PseudoViewer (11), RnaViz (12), jViz.RNA (13), RNAView (14), RNApuzzler (15), R2R (16), R2DT (17), and RNArtist (https://github.com/fjossinet/RNArtist), generally produce drawings in a more regimented, publication-friendly manner that can more efficiently communicate secondary structure and facilitate phylogenetic comparisons (2, 15–18). Importantly, programs in this category have greater functionality for precisely arranging stems and loops in a strict manner (e.g., perfectly round or flattened loops), making them preferred for the drawing of large structures, which benefit from stricter layout organization (15, 17). These programs also introduce additional methods for specifying the aspects of a drawing, such as program-specific input file formats (e.g., R2R, RNArtist) and layout templates (e.g., R2DT), and also tend to offer greater customization of individual drawing elements including bases and bond notations (e.g., Leontis-Westhof). On the other hand, features for interactivity, such as click-and-drag editing of residue layout and the shapes of tertiary interaction lines, are often lacking or limited in functionality. This category of programs may also lack point-and-click editing of base-pairs and other drawing elements, thus requiring input files or application forms for edits, where bases and other elements are specified by numeric position making editing drawings cumbersome and less intuitive.

Nucleic acid structure drawings are rarely static but change as substructures are confirmed using biochemical methods or through phylogenetic comparisons, and thus drawings should also be easily editable to reflect a researcher's current knowledge. Towards this goal, RNAcanvas was designed to combine some of the best features of both types of structure drawing programs, merging the interactivity functions of the first category (e.g., click-and-drag layout adjustments) with the orderly layouts (e.g., residues maintained in strictly shaped stems and loops) and detailed drawing customizations of the second category. Furthermore, RNAcanvas provides extensive point-and-click

editing features for individual components of the drawing, such as base letters, lines of bonds, and base outlines and numberings. RNAcanvas is also one of the few programs to fully support Leontis-Westhof notation for depicting canonical and non-canonical RNA base-pairs and interactions. RNAcanvas has been optimized to function smoothly with large structures, and contains unique features such as real-time highlighting of complementary sequences, motif search, and export of drawings in both scalable vector graphics (SVG) and PowerPoint (PPTX) formats, which greatly facilitates the creation of figures for scientific presentations. Exported SVG and PPTX drawings are also formatted such that individual objects (e.g., base letters, bond lines) can be altered in vector graphics editors such as Adobe Illustrator or PowerPoint. RNAcanvas was formerly named RNA2Drawer, which was initially available only in a download format (19).

MATERIALS AND METHODS

Versions of software tools used for RNAcanvas are currently the most recent (2023), however continuous enhancements to the program to provide additional functionality and updating of tools will continue.

Programming Languages and General Tools

RNAcanvas is a single-page web application created with HTML/CSS/JavaScript. Babel v7.16.10 is used to compile JavaScript code to the ECMAScript 2009 (ES5) standard. The application is built using the webpack module bundler v4.46.0. TypeScript v3.9.10 is used to add type annotations over JavaScript code. All JavaScript application code to be received by the user is written with TypeScript type annotations. Unit testing is performed on the Node.js runtime environment v16.18.1 and using the Jest framework v24.9.0. The uuid library v8.3.2 is used to generate universally unique identifiers (UUIDs), which are assigned to a variety of components (e.g., text elements of bases) to simplify retrieval in the application code.

Interactive Drawing

Within the application, the drawing itself is a scalable vector graphics (SVG) document, which is a type of element that web browsers can render. Web browsers provide built-in methods for detecting user interaction with SVG documents (e.g., mouse clicks, hovering and dehovering of elements within the SVG document). This allows user interaction with the drawing to be programmed in an event-driven manner. The SVG specification is mature and extensive, allowing additional customization of drawings. The SVG.js library v3.1.1 is also used to simplify code for manipulating the drawing and managing user interactions.

Peripheral UI Elements

Peripheral user interface (UI) elements, such as the top menu, bottom information bar, and right-side forms are built using the React framework v16.14.0. The react-color library v2.19.3 is used for the colour picker UI element. The react-select library v3.2.0 is used for the UI element for picking fonts.

The React Testing Library v9.5.0 and the enzyme library v3.11.0 are used for unit testing of peripheral UI elements.

Exporting Drawings

The SVG.js library v3.1.1 is used to produce SVG files of drawings. (SVG files contain the textual representation of an SVG document, such as the drawing of the application.) To produce PowerPoint files of drawings, the PptxGenJS library v3.9.0 is used. Rather than convert all the SVG elements of a drawing to their nearest PowerPoint object counterpart, many drawing elements (e.g., the curved lines of tertiary bonds) are included in PPTX files as individual SVG image elements. Because of this, exported PPTX files require PowerPoint version 2016 or later. SVG image elements in exported PPTX files can still be manipulated as though they were PowerPoint objects and can often be converted to corresponding PowerPoint objects using the PowerPoint "Convert to Shape" feature.

RESULTS

RNAcanvas offers extensive options to customize drawings of nucleic acid secondary structures and tertiary interactions based on graphical user interaction, with drawings edited using "tools" and forms. Tools control how the user interacts with the drawing itself. For example, different tools allow the user to drag elements of the drawing, select element(s) for editing, or pair and unpair bases. Forms pop-up on the right-side of the application and allow properties of selected elements (or the drawing as a whole) to be directly edited through text and numeric inputs and various picker components and toggles (e.g., fonts of bases, colours of bonds, base numbering). All structure drawing figures in this report were produced entirely within the RNAcanvas web app (unless stated otherwise).

Layout Flexibility

Using the Dragging Tool, the layout of a drawing can be adjusted by dragging with the mouse (Fig. 1). Stems can be dragged around loops and the outermost loop can be dragged to rotate the drawing with the arrangement of bases into stems and loops strictly maintained. Using the Flattening Tool, individual loops can be flattened (or reverted to unflattened). When flattened, the outermost loop is entirely flat (Fig. 1B, a-c), and inner loops with more than one child stem assume a triangular shape (Fig. 1B, d-g). Flattening inner loops with multiple child stems often helps to condense a drawing and highlight conserved branching patterns in structures. Flattening an inner loop with only one child stem (Fig. 1B, h and i) aligns the child stem with the parent stem, effectively "straightening" the two stems. Using the Flipping Tool, individual stems can be flipped across their parent loops (Fig. 1B, j-m). RNAcanvas also has a Layout form in which the rotation of the drawing can be precisely specified. The Layout form also has a field for specifying the termini gap of the drawing, which is the distance between the first and last bases when the outermost loop of a drawing is round (Fig. 1B, label n and dashed line). The termini gap can be set to zero (for a fully circular outermost loop) or to larger values (e.g., for a semicircle outermost loop).

Customization of Drawing Elements

The following customizations can be accomplished using the Editing Tool:

Base customizations. Fonts, font sizes and base colours can be personalised (Fig. 2A). Bases can be outlined, and outlines can be adjusted for size, line thickness and line and fill colours and transparencies. Chemical probing data such as SHAPE data (20) can be easily incorporated into the structure through base and/or outline colouring. A “dots” style drawing, in which only the colour-coded outlines of bases are shown, is an effective way to communicate chemical probing data (Fig. 2B) and dots can be easily suppressed to reveal the underlying bases. RNAcanvas has a dedicated form to help with styling bases according to a set of data. This form allows the user to progressively select and style bases within different data ranges (e.g., low, moderate, and high SHAPE reactivities). Base letters can be directly edited, and bases can subsequently be inserted, appended, and removed. A substructure can be specified when inserting/appending a subsequence of bases. The numbering of bases can also be specified via the offset, increment and anchor properties in the Numbering form, with text and lines customizable in terms of fonts, colour, dimensions, etc. (Fig. 2A). Individual bases can also be manually numbered regardless of what numbers are assigned to other bases in the drawing.

Bond customizations. In a drawing, consecutive bases are linked by a primary bond. Most primary bonds are not visible (by default) when a structure is initially drawn. Secondary bonds join base-pairs in the secondary structure of a drawing and affect the layout of the drawing. Tertiary bonds communicate base-pairs/interactions that do not affect the layout of the drawing. All bonds have a line that connects the two bases in the bond. Primary and secondary bonds have straight lines. Tertiary bonds have curved lines that can be dragged with the mouse to adjust their shape. The lines of bonds can be customized in terms of colour and thickness, as well as base padding, which is the distance between the end of a bond line and the base connected by the bond (Fig. 2C). Bond lines can be dashed and their ends rounded.

Elements such as rectangles, circles, triangles, and text can be “strung” onto (associated with) any bond (Fig. 2C). This attachment of elements to specific bonds allows for the full Leontis-Westhof notation (4) for canonical and non-canonical base-pairs and interactions to be depicted on secondary and tertiary bonds. As with other drawing elements, strung elements can be customized in terms of colour, transparencies, dimensions, line dashing, fonts, etc. Strung elements can be dragged with the mouse and displaced from their parent bond, which adds versatility to their usage. For example, when dragged next to nearby bases, strung elements can mark the location with additional information (Fig. 2D) or can be used to shade hairpins. Text strung elements can be used to name structural features as well as the 5′ and 3′ termini (Fig. 2D). All strung elements maintain their position and orientation relative to their parent bond when the layout of the drawing is adjusted.

Shading a series of bases and outlining structures. Highlighting a selection of bases can be accomplished by: i) setting primary bond base paddings to zero; ii) making primary bonds sufficiently thick to encompass the font size of the bases; iii) setting the line end caps of the primary bonds to round; and iv) sending the shading below the bases (Fig. 2E). Tertiary pairings can also be denoted by shading the two interacting sequences and connecting one base of each sequence with a single

tertiary bond (Fig. 2E, in orange). It is also possible to outline structures in a drawing by shading bases using this technique and then sending base letters below the shading (Fig. 2F).

"Stick-and-ball" drawings. "Stick-and-ball" drawings (Fig. 2G) can be created by making base letters invisible (e.g., by assigning them the background colour), increasing the base padding of secondary bonds to shorten them, and making secondary bonds sufficiently thick to overlap with each other (thus forming the "stick" portion). The "ball" portion is composed of circle strung elements that are dragged over the loops in the structure and sized to match the sizes of the loops.

Structure Editing and Large Structure Exploration

Structures thousands of nucleotides long, such as the structures of positive-strand RNA virus genomes, are increasingly studied as a whole (5, 6, 21). Such large structures cannot be predicted directly (5) and require manual refinement by RNA researchers based on phylogenetic comparisons, experimental results, and the researcher's own expertise. RNAcanvas has optimized performance for editing large structures, e.g., when dragging to adjust the layout.

The Pairing Tool is used to pair and unpair bases (i.e., add and remove secondary and tertiary bonds) in a point-and-click manner. The Pairing Tool also has the option of highlighting sequences that are complementary to the selected sequence (Fig. 3, selected sequence highlighted in yellow and complementary sequences highlighted in pink). The number of possible pairing partners within the structure can be designated, and options are available to exclude G-U/G-T pairs and permit partial mismatches between pairing partners. The Pairing Tool aids in refinement of output by structure prediction programs such as Mfold (22) and RNAfold (23) and helps to identify possible long-range base-pairings that cannot currently be predicted computationally (5). RNAcanvas also has a form for finding input motifs (Find Motifs form) (Fig. 3). Both complement highlighting (using the Pairing Tool) and the Find Motifs form support partial mismatching and the use of IUPAC single letter codes (24). Complement highlighting further supports a percentage cap on the G-U/G-T base-pair composition of highlighted complements. The Find Motifs form can also treat an input motif to search for as a regular expression (as defined in Computer Science).

Input and Output Formats

RNAcanvas accepts primary sequences (unstructured) as well as structures in dot-bracket notation (also called Vienna format) and in Connectivity Table (CT) files. Input structures may contain pseudoknots, which are drawn using tertiary bonds. RNAcanvas outputs drawings in SVG and PowerPoint formats. SVG files can be opened in vector graphics editors such as Adobe Illustrator and Inkscape. Drawings are exported such that all drawing elements (e.g., bases text, bond lines) are exported as individual SVG or PowerPoint objects. For instance, base letters are exported as SVG text elements or PowerPoint text boxes. This allows for further manipulation of exported drawings in vector graphics editors and PowerPoint. A scaling factor may also be specified when exporting a drawing. All elements in an exported drawing are individually scaled according to the scaling factor. Scaling elements individually on export eases the manipulation of exported drawings

compared to when exported drawings are scaled as a whole after export using a separate application such as Adobe Illustrator. Additionally, RNAcanvas can save drawings in a file format unique to RNAcanvas with ".rnacanvas" extension. Downloaded drawing files contain a complete representation of a drawing, allowing drawings to be reopened at later times exactly as they were when saved.

DISCUSSION

RNAcanvas has many enhancements over the free desktop-based RNA2Drawer predecessor (19). Being web-based, RNAcanvas is easier for users to access and is still available for use at no cost. The RNAcanvas web app also uses the SVG web standard to present drawings, which is expansive and permits wide customization of drawings. Other improvements in the web app include unlimited undo/redo functionality, greater layout flexibility (e.g., the ability to flatten and circularize loops individually), the availability of strung elements on bonds, support for Leontis-Westhof notation, ability to add text labels and base markers, enhanced flexibility in numbering bases, and the Find Motifs form.

Although RNAcanvas offers flexible options to customize the layout of a drawing (Fig. 1), the program does not currently support certain layout features. For instance, RNAcanvas cannot currently incorporate pseudoknots into the layout of a drawing, unlike some programs such as PseudoViewer (Table 1) (11). Certain types of structures such as frameshifting elements (25) are often drawn such that pseudoknot(s) are conveyed via the layout of the drawing. The initial layouts of drawings made by RNAcanvas (especially for large structures) can also contain overlapping regions that need to be manually untangled by the user. While drawings produced by many programs often contain overlapping regions (17), some programs such as RNApuzzler/RNAturtle have made algorithmic advancements in the initial layout of an RNA drawing to avoid overlaps (15). R2DT has also demonstrated the effectiveness of template-based drawing for a wide array of small and large structures (17). Programs such as PseudoViewer, RNApuzzler and R2DT, however, do not provide the expansive features of RNAcanvas for editing drawings, such as click-and-drag layout adjustment and point-and-click editing of the structure and individual drawing elements, which allow structure drawings to be moulded beyond their initial forms. RNAcanvas is also unique in its focus on the extended exploration of structures and the continued editing of drawings that occurs as the researcher's knowledge of a structure grows.

RNAcanvas is able to handle large (at least 3,000 nt) structures with only modest lag times, making it significantly faster than other drawing platforms with similar click-and-drag layout adjustment. The primary performance bottleneck of RNAcanvas is the re-rendering of base letters when the layout of a drawing is changed. Note that among Chrome, Firefox and Safari, the Chrome web browser currently has the best performance when handling large RNAcanvas structures. Many biological transcripts, such as positive-strand RNA virus genomes and long noncoding RNAs, have lengths over 3 kb and thus future performance improvements will need to include new strategies for minimizing lag

times when dragging portions of very large structures. Ultimately, RNAcanvas aspires to allow entire structures of any size biological transcript to be effectively drawn and edited on personal devices.

AVAILABILITY

RNAcanvas web app: https://rnacanvas.app

RNAcanvas GitHub repository: https://github.com/pzhaojohnson/rnacanvas

ACCESSION NUMBERS

Not applicable.

SUPPLEMENTARY DATA

Not applicable.

ACKNOWLEDGEMENTS

We would like to thank Dr. Bruce Shapiro and Dr. Wojciech Kasprzak for their invaluable help in developing RNA2Drawer, the predecessor to RNAcanvas, and all of the advice that they have provided over the years that has propelled this project forward.

FUNDING

The work was funded by National Science Foundation MCB-1818229 and United States Department of Agriculture 308291-00001 to AES. PZJ was partially supported by National Science Foundation Graduate Fellowship Award DGE-1840340.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

1. Ponty,Y. and Leclerc,F. (2015) Drawing and editing the secondary structure(s) of RNA. *Methods Mol Biol*, **1269**, 63–100.

https://doi.org/10.1007/978-1-4939-2291-8_5

http://www.ncbi.nlm.nih.gov/pubmed/25577373

2. Shabash,B. and Wiese,K.C. (2017) RNA Visualization: Relevance and the Current State-of-the-Art Focusing on Pseudoknots. *IEEE/ACM Trans Comput Biol Bioinform*, **14**, 696–712.

https://doi.org/10.1109/TCBB.2016.2522421

http://www.ncbi.nlm.nih.gov/pubmed/26915129

3. Vicens,Q. and Kieft,J.S. (2022) Thoughts on how to think (and talk) about RNA structure. *Proc Natl Acad Sci U S A*, **119**, e2112677119.

https://doi.org/10.1073/pnas.2112677119

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC9169933

4. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.

https://doi.org/10.1017/s1355838201002515

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1370104

5. Lu,Z. and Chang,H.Y. (2018) The RNA Base-Pairing Problem and Base-Pairing Solutions. *Cold Spring Harb Perspect Biol*, **10**, a034926.

https://doi.org/10.1101/cshperspect.a034926

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC6280703

6. Liu,J., Carino,E., Bera,S., Gao,F., May,J.P. and Simon,A.E. (2021) Structural Analysis and Whole Genome Mapping of a New Type of Plant Virus Subviral RNA: Umbravirus-Like Associated RNAs. *Viruses*, **13**, 646.

https://doi.org/10.3390/v13040646

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC8068935

7. Chkuaseli,T. and White,K.A. (2020) Activation of viral transcription by stepwise largescale folding of an RNA virus genome. *Nucleic Acids Res*, **48**, 9285–9300.

https://doi.org/10.1093/nar/gkaa675

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC7498350

8. 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.

https://doi.org/10.1093/bioinformatics/btv372

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4595900

9. Lu,J.S., Bindewald,E., Kasprzak,W.K. and Shapiro,B.A. (2018) RiboSketch: versatile visualization of multi-stranded RNA and DNA secondary structure. *Bioinformatics*, **34**, 4297–4299.

https://doi.org/10.1093/bioinformatics/bty468

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC6289134

10. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.

https://doi.org/10.1093/bioinformatics/btp250

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2712331

11. Byun,Y. and Han,K. (2009) PseudoViewer3: generating planar drawings of large-scale RNA structures with pseudoknots. *Bioinformatics*, **25**, 1435–1437.

https://doi.org/10.1093/bioinformatics/btp252

http://www.ncbi.nlm.nih.gov/pubmed/19369500

12. De Rijk,P., Wuyts,J. and De Wachter,R. (2003) RnaViz 2: an improved representation of RNA secondary structure. *Bioinformatics*, **19**, 299–300.

https://doi.org/10.1093/bioinformatics/19.2.299

http://www.ncbi.nlm.nih.gov/pubmed/12538259

13. Shabash,B. and Wiese,K.C. (2019) jViz.RNA 4.0-Visualizing pseudoknots and RNA editing employing compressed tree graphs. *PLoS One*, **14**, e0210281.

https://doi.org/10.1371/journal.pone.0210281

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC6502502

14. Yang,H., Jossinet,F., Leontis,N., Chen,L., Westbrook,J., Berman,H. and Westhof,E. (2003) Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res*, **31**, 3450–3460.

https://doi.org/10.1093/nar/gkg529

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC168936

15. Wiegreffe,D., Alexander,D., Stadler,P.F. and Zeckzer,D. (2019) RNApuzzler: efficient outerplanar drawing of RNA-secondary structures. *Bioinformatics*, **35**, 1342–1349.

https://doi.org/10.1093/bioinformatics/bty817

http://www.ncbi.nlm.nih.gov/pubmed/30239566

16. Weinberg,Z. and Breaker,R.R. (2011) R2R--software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics*, **12**, 3.

https://doi.org/10.1186/1471-2105-12-3

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3023696

17. Sweeney,B.A., Hoksza,D., Nawrocki,E.P., Ribas,C.E., Madeira,F., Cannone,J.J., Gutell,R., Maddala,A., Meade,C.D., Williams,L.D., *et al.* (2021) R2DT is a framework for predicting and visualising RNA secondary structure using templates. *Nat Commun*, **12**, 3494.

https://doi.org/10.1038/s41467-021-23555-5

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC8190129

18. Muller,G., Gaspin,C., Etienne,A. and Westhof,E. (1993) Automatic display of RNA secondary structures. *Comput Appl Biosci*, **9**, 551–561.

https://doi.org/10.1093/bioinformatics/9.5.551

http://www.ncbi.nlm.nih.gov/pubmed/7507400

19. Johnson,P.Z., Kasprzak,W.K., Shapiro,B.A. and 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.

https://doi.org/10.1080/15476286.2019.1659081

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC6844559

20. Wilkinson,K.A., Merino,E.J. and Weeks,K.M. (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc*, **1**, 1610–1616.

https://doi.org/10.1038/nprot.2006.249

http://www.ncbi.nlm.nih.gov/pubmed/17406453

21. Ziv,O., Price,J., Shalamova,L., Kamenova,T., Goodfellow,I., Weber,F. and Miska,E.A. (2020) The Short- and Long-Range RNA-RNA Interactome of SARS-CoV-2. *Mol Cell*, **80**, 1067-1077.e5.

https://doi.org/10.1016/j.molcel.2020.11.004

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC7643667

22. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, **31**, 3406–3415.

https://doi.org/10.1093/nar/gkg595

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC169194

23. Lorenz,R., Bernhart,S.H., Höner Zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol Biol*, **6**, 26.

https://doi.org/10.1186/1748-7188-6-26

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3319429

24. Cornish-Bowden,A. (1985) Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Res*, **13**, 3021–3030.

https://doi.org/10.1093/nar/13.9.3021

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC341218

25. Firth,A.E. and Brierley,I. (2012) Non-canonical translation in RNA viruses. *J Gen Virol*, **93**, 1385–1409.

https://doi.org/10.1099/vir.0.042499-0

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3542737

TABLE AND FIGURE LEGENDS

Figure 1. Various layout options for drawings in RNAcanvas. (A) The same input structure as in (B, middle) before any layout adjustments save for being rotated 90°. (B) Labels a-n and dashed line (lower left) added in PowerPoint. (a) Round outermost loop. (b) Flat outermost loop. (c) Rotated flat outermost loop. (d, e) Round inner loops. (f, g) Flattened inner loops with multiple child stems. (h)

Flattened inner loop with only one child stem that has been condensed. (i) Flattened inner loop with only one child stem that has been stretched. (j-m) Flipped stems. (n) Termini gap of the round outermost loop (indicated by dashed line).

Figure 2. Various customizations of drawing elements. (A) Fonts, font sizes and colours of bases text are all customizable. Bases can be outlined with circles and numbered. Base outlines and numberings are also customizable in terms of colours, dimensions, fonts, etc. This drawing has a base numbering offset of -121 nt. (B) Drawing presented in "dots" style in which only colour-coded outlines of bases are shown. Dots style is an effective way to communicate chemical probing data, such as the example SHAPE data shown in this drawing. Dots colour legend added in PowerPoint. (C) Various customizations of bond lines and "strung" elements. Entire Leontis-Westhof notation for depicting canonical and non-canonical base-pairs and interactions is supported. (D) Text labels and base markers. Hairpin H1, stem-loop SL2, linker region L3, the 5′end and the 3′OH were all labelled within the program. Bases also can have various markers (i.e., coloured triangles, squares, circles) associated with them. These text labels and base markers are "strung" elements of nearby primary and secondary bonds that have been dragged to their current positions. Due to being strung elements, these text labels and base markers maintain both their positions and orientations relative to their parent bonds when the layout of the drawing is adjusted. (E) Shading sequences of bases. Tertiary pairing connected by a single tertiary bond is shown (orange). (F) Drawing with portions outlined. (G) Drawing with portions in the "stick-and-ball" style.

Figure 3. Screenshot showing highlighting of complementary sequences with the Pairing Tool and the Find Motifs form. The 5′CAACC sequence highlighted in yellow is currently selected. Complementary sequences 5′GGUUG found by the program are highlighted in pink. (Right) Find Motifs form is open. Currently the options to match Us and Ts and to use IUPAC single letter codes are toggled. The motif 5′CYRCVA (containing the IUPAC single letter codes Y, R and V) was searched for and six matching motifs were found. The second matching motif from the bottom is being hovered with the mouse cursor (highlighted in grey). Clicking on a matching motif will centre the matching motif on the screen and flash its base letters.

Table 1. Comparison of programs for 2D drawing of nucleic acid structures. Solid dots indicate that a feature is supported by a program, though the robustness of feature implementations may vary between different programs. VARNA and jViz.RNA have hollow dots for web-based use since they require a Java plugin for use in a web browser. Point-and-click structure editing refers to the ability to add and remove base-pairs and interactions in the secondary and tertiary structures of a drawing graphically using the mouse. Motif finding functionality refers to features similar to the Find Motifs form in RNAcanvas. Point-and-click element editing is the ability to select individual or groups of elements to edit by clicking on them and/or dragging a selecting box over them, rather than being required to select elements by less direct means (e.g., by inputting numeric positions). Text labels for structural features refer to the ability to place pieces of text next to structural features such as hairpins and linker regions that name them. Flexible base numbering is the ability to assign any number to any base regardless of what numbers are assigned to other bases. This is useful when a drawing

contains two discontinuous portions of a larger sequence and the numbering for one portion must be offset from the other. Granular output scaling refers to the ability to scale drawings at the level of individual elements when exported, as opposed to scaling an exported drawing as a whole after export. Drawing using input files and automated drawing are grouped together since the use of input files can lend itself to automated drawing via the automated creation of input files. Automated drawing also includes the ability to generate large numbers of structure drawings in an automated way in general, such as by allowing other software written by the user to interface with the structure drawing software, such as with VARNA.

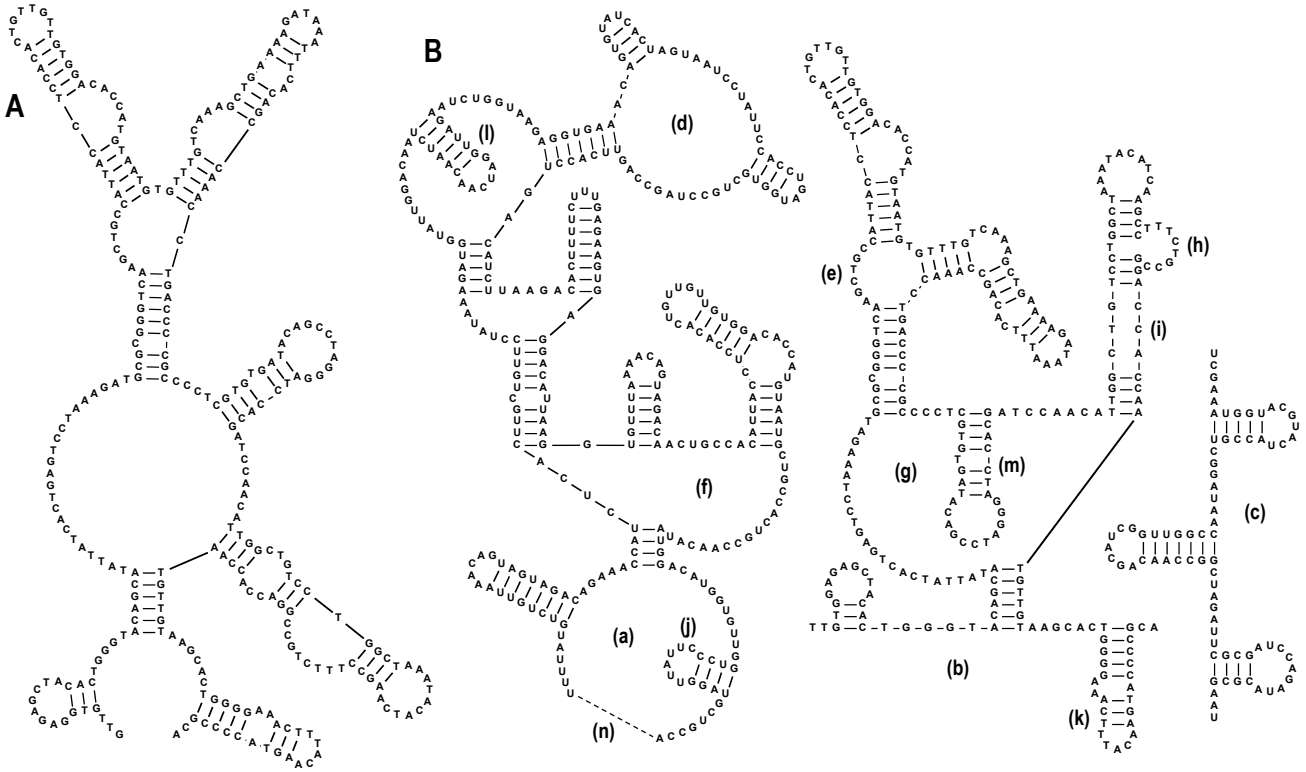

Figure 1

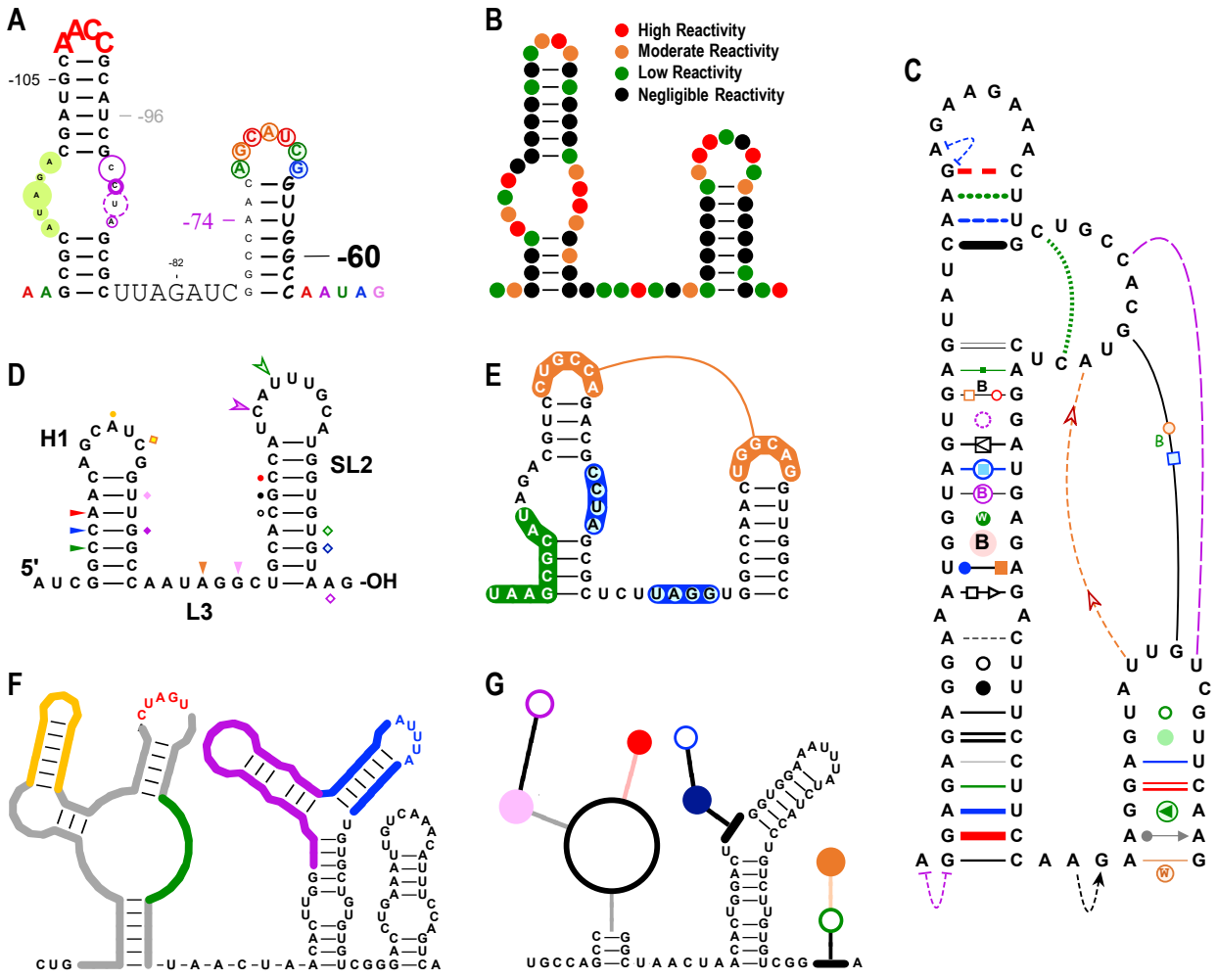
A
-105
-96
-74
-82
-60
AAG—CUUAGAUC
B
High Reactivity
Moderate Reactivity
Low Reactivity
Negligible Reactivity
C
D
H1
SL2
5'
-OH
L3
E
F
CUG
G
UGCCAG—CUAACUAA—UCGG—A

Figure 2

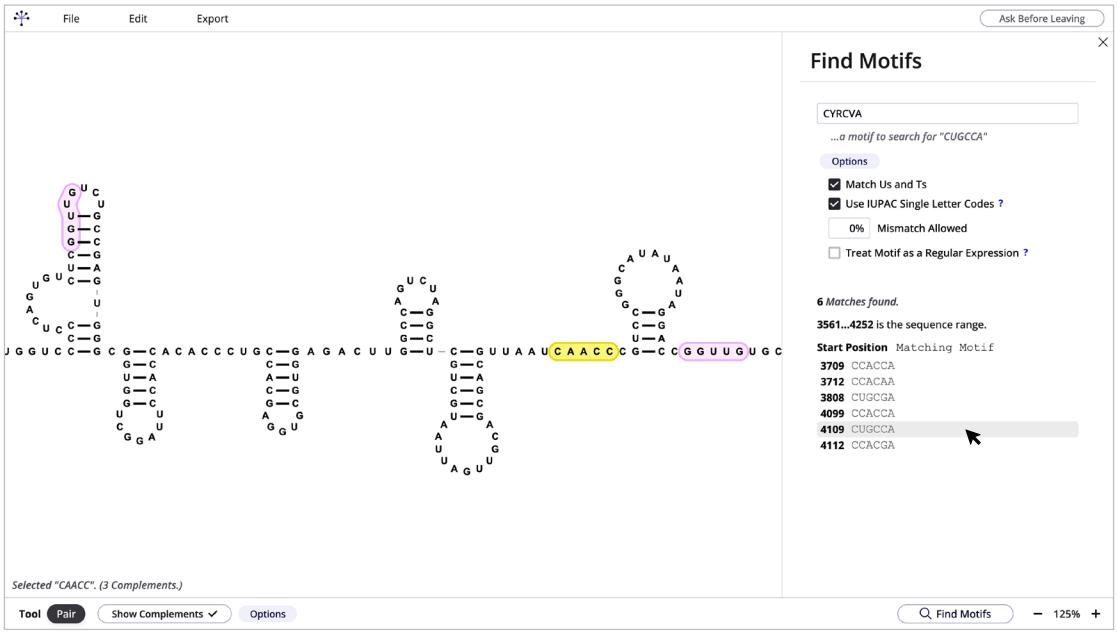
File
Edit
Export
Ask Before Leaving
Find Motifs
CYRCVA
...a motif to search for "CUGCCA"
Options
Match Us and Ts
Use IUPAC Single Letter Codes ?
0%
Mismatch Allowed
Treat Motif as a Regular Expression ?
6 Matches found.
3561...4252 is the sequence range.
Start Position Matching Motif
3709 CCACCA
3712 CCACAA
3808 CUGCGA
4099 CCACCA
4109 CUGCCA
4112 CCACGA
Selected "CAACC". (3 Complements.)
Tool
Pair
Show Complements
Options
Find Motifs
125%

Figure 3

	RNAcanvas	VARNA	XRNA	PseudoViewer	RnaViz	jViz.RNA	forna	RiboSketch	RNAView	R2R	R2DT	RNArtist
Web-Based	●	○	●	●		○	●	●	●		●	
Click-and-Drag Layout Adjustment	●	●	●		●	●	●	●				
Force-Based Drawing Layouts						●	●	●				
Draggable, Curved Tertiary Bonds	●											
Incorporation of Pseudoknots into the Layout			●	●	●	●	●	●		●	●	
Point-and-Click Structure Editing	●	●			●	●	●	●				●
Real-Time Highlighting of Complementary Sequences	●											
Motif Finding Functionality	●											
Point-and-Click Element Editing	●	●	●		●	●						●
Leontis-Westhof Notation	●	●			●			●	●			●
Text Labels for Structural Features	●				●					●		
Flexible Base Numbering	●											
SVG Output	●	●	●	●			●	●		●	●	●
PPTX Output	●											
Granular Output Scaling	●											
Raster Image Output (e.g., PNG, TIFF, JPEG)		●	●	●	●	●	●	●	●		●	●
Template-Based Drawing					●						●	
Drawing Using Input Files and Automated Drawing		●	●						●	●	●	●

Table 1