

RESEARCH ARTICLE

Open Access



Pairwise visual comparison of small RNA secondary structures with base pair probabilities

Serge Léger^{2†}, Maria Beatriz Walter Costa^{4†} and Dan Tulpan^{1,2,3*†} 

Abstract

Background: Predicted RNA secondary structures are typically visualized using dot-plots for base pair binding probabilities and planar graphs for unique structures, such as the minimum free energy structure. These are however difficult to analyze simultaneously.

Results: This work introduces a compact unified view of the most stable conformation of an RNA secondary structure and its base pair probabilities, which is called the **Circular Secondary Structure Base Pairs Probabilities Plot (CS²BP²-Plot)**. Along with our design we provide access to a web server implementation of our solution that facilitates pairwise comparison of short RNA (and DNA) sequences up to 200 base pairs. The web server first calculates the minimum free energy secondary structure and the base pair probabilities for up to 10 RNA or DNA sequences using RNAfold and then provides a two panel comparative view that includes CS²BP²-Plots along with the traditional graph, planar and circular diagrams obtained with VARNA. The CS²BP²-Plots include highlighting of the nucleotide differences between two selected sequences using ClustalW local alignments. We also provide descriptive statistics, dot-bracket secondary structure representations and ClustalW local alignments for compared sequences.

Conclusions: Using circular diagrams and colour and weight-coded arcs, we demonstrate how a single image can replace the state-of-the-art dual representations (dot-plots and minimum free energy structures) for base-pair probabilities of RNA secondary structures while allowing efficient exploration and comparison of different RNA conformations via a web server front end. With that, we provide the community, especially the biologically oriented, with an intuitive tool for ncRNA visualization.

Web-server: <https://cs2bp2plot.cluster.gctools.nrc.ca/>

Background

Visual analysis of biological sequences is a crucial step in bioinformatics and computational biology and contributes substantially to biological data interpretation and algorithm development. Traditional representations of RNA secondary structures, for example (such as linear arc diagrams, circular diagrams and planar graphs), present a single state view of a structural conformation of a nucleic acid, which provide fundamental insights into the cellular function of

both coding and non-coding RNAs [1]. In reality, an RNA molecule transitions among a variety of energy states due to thermodynamic variations in its environment [2]. Therefore, more advanced and accurate visualization approaches are needed to characterize the whole ensemble of secondary structure conformations for an RNA sequence.

Visualization approaches [3–6] based on predicted RNA secondary structures [7] have been developed to separately address these needs. Single state RNA conformations can be depicted with *planar graphs*, *linear arc diagrams* and *circular diagrams*, based on different ways of representing the RNA backbone.

In linear arc diagrams (Fig. 1a), the backbone is represented by a straight line. The bases are consecutively placed along the line and paired bases are connected with arcs. To save space and accommodate longer RNA

* Correspondence: dtulpan@uoguelph.ca

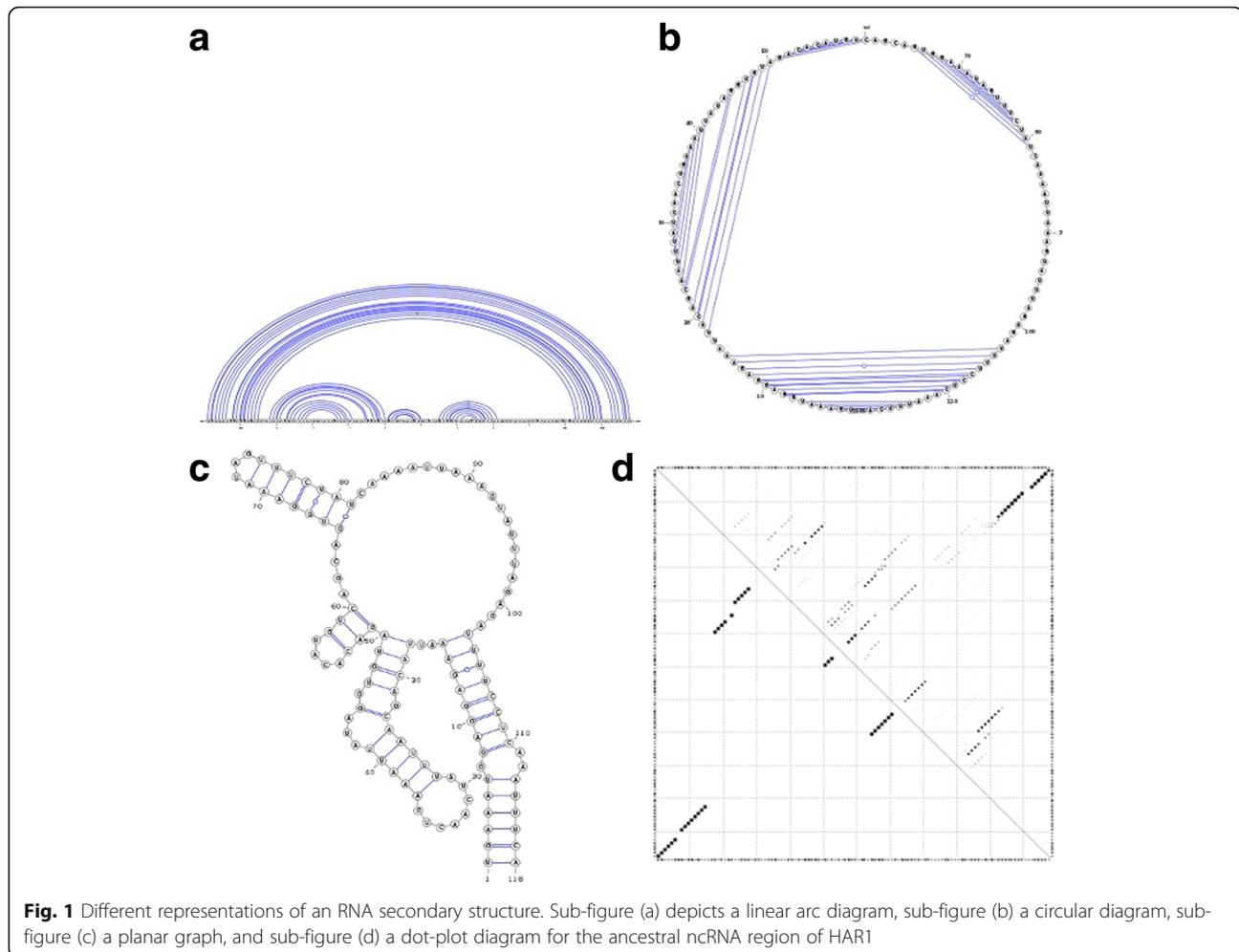
[†]Serge Léger, Maria Beatriz Walter Costa and Dan Tulpan contributed equally to this work.

¹Department of Animal Biosciences, Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, Ontario, Canada

²Digital Technologies Research Center, National Research Council Canada, 100 des Aboiteaux St, Moncton, NB E1A7R1, Canada

Full list of author information is available at the end of the article



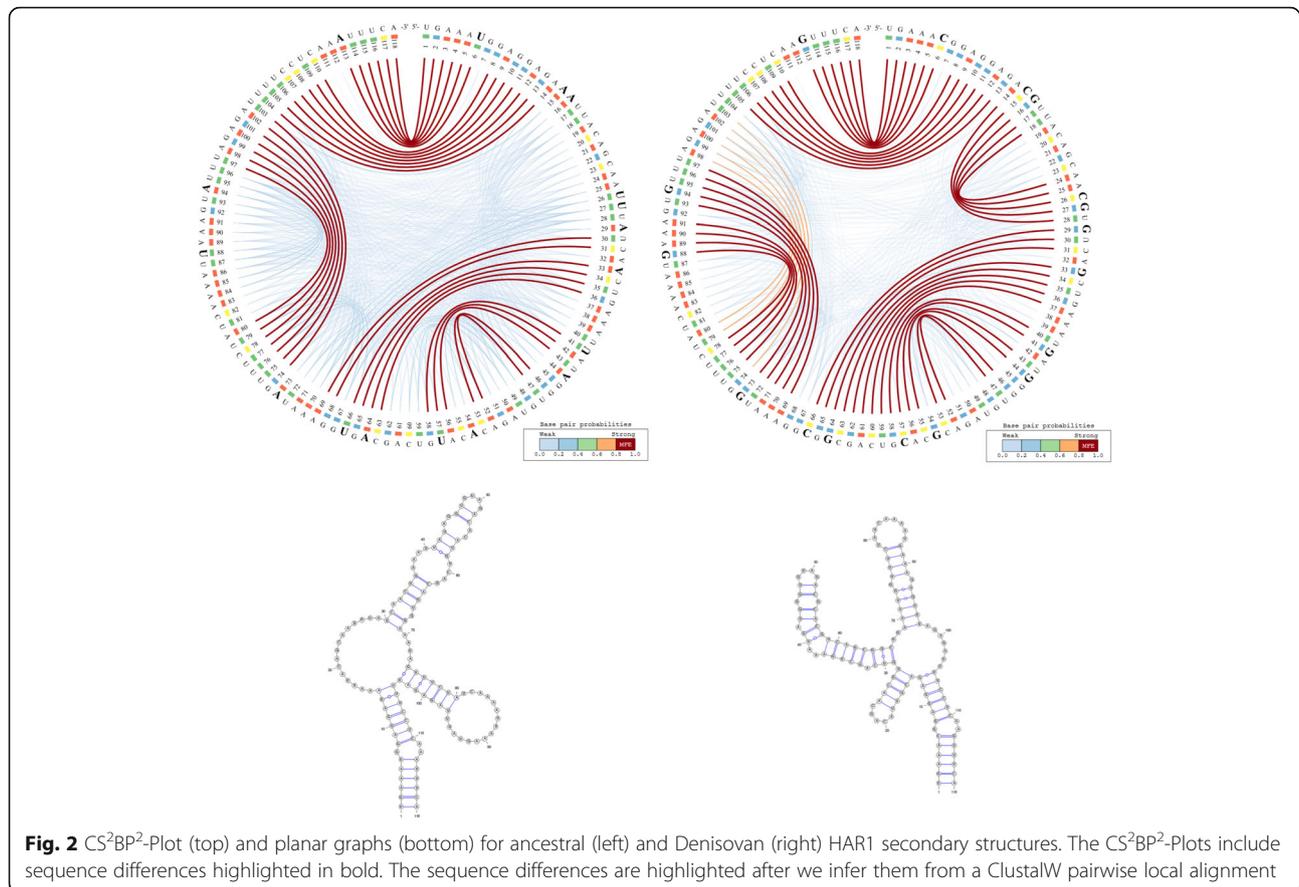


sequences, the backbone can be displayed in a circular format as originally proposed by Nussinov et al. in 1978 [8]. A circular diagram (Fig. 1b) will therefore consist of a circular backbone with bases lined up consecutively along the perimeter of the circle, while chords will connect the paired bases. A planar graph (Fig. 1c) is by far the most used visualization method to date, due to its flexibility and simplicity. It represents the backbone as a convoluted planar curve with predefined distances between neighbouring paired bases and a minimum number of overlaps.

Base-pair probabilities are typically represented with *dot-plots* (Fig. 1d) – bi-dimensional graph representations with positions and bases of the sequence being present on both x and y axis and dots being placed at positions where base pairs occur. In this case, the dot sizes are proportional to the base pair probabilities. An extensive presentation of RNA dot-plots and their usage is explored by Churkin and Barash [9]. Other RNA secondary structure alignment and pairwise comparison methods have been published in the past, most notably RNAforester [10], R-chie [5] and BEAGLE [11]. RNAforester calculates RNA secondary

structures alignments using dot-parentheses representations as input and applies a tree alignment model on all sequence secondary structures in a progressive fashion. R-chie focuses on highlighting structure and primary sequence conservation and variation in multiple RNA secondary structures. BEAGLE exploits a new encoding for RNA secondary structure and a substitution matrix of RNA structural elements to perform RNA structural alignments and easily identify structural similarities between RNAs.

In 2013, Aalberts and Jannen [4] introduced the RNAbows diagrams, the first successful attempt to combine two representations of RNA secondary structures with base pair probabilities. It captures both the minimum free energy (MFE) structure and the base pair probabilities into a single image. While their method is versatile and allows for comparison of pairs of RNA secondary structures, their choice of linear arc diagrams places serious constraints on the maximum length of RNA sequences that can be represented (less than 100 bases in a legible format). Moreover, pairwise (bottom-up) comparisons require sequences of equal length that can be horizontally aligned.



Other bioinformatics tools have been recently proposed for visualization or comparison of RNA secondary structures, such as RNA-TVCurve [12] for secondary structure comparisons and TRAVeLer [13] for visualization of structures in the presence of a template. While interesting, they solve different problems and require additional constraints that limit their applicability for both comparison and visualization of RNA secondary structures with base-pair probabilities, such as the need of external structure templates, as is the case of TRAVeLer.

To alleviate these limitations and to propose an alternative tool that is informative and easy to use for biologists, we present the Circular Secondary Structure Base Pairs Probabilities Plot (CS²BP²-Plot) – an intuitive visual representation of an RNA secondary structure that includes all possible base pair probabilities. The CS²BP²-Plot uses a chord diagram layout comprised of two concentric graphical layers representing (i) the RNA sequence and corresponding positions for each base pair (outer layer), and, (ii) the base pair probabilities and MFE base pairings (inner layer).

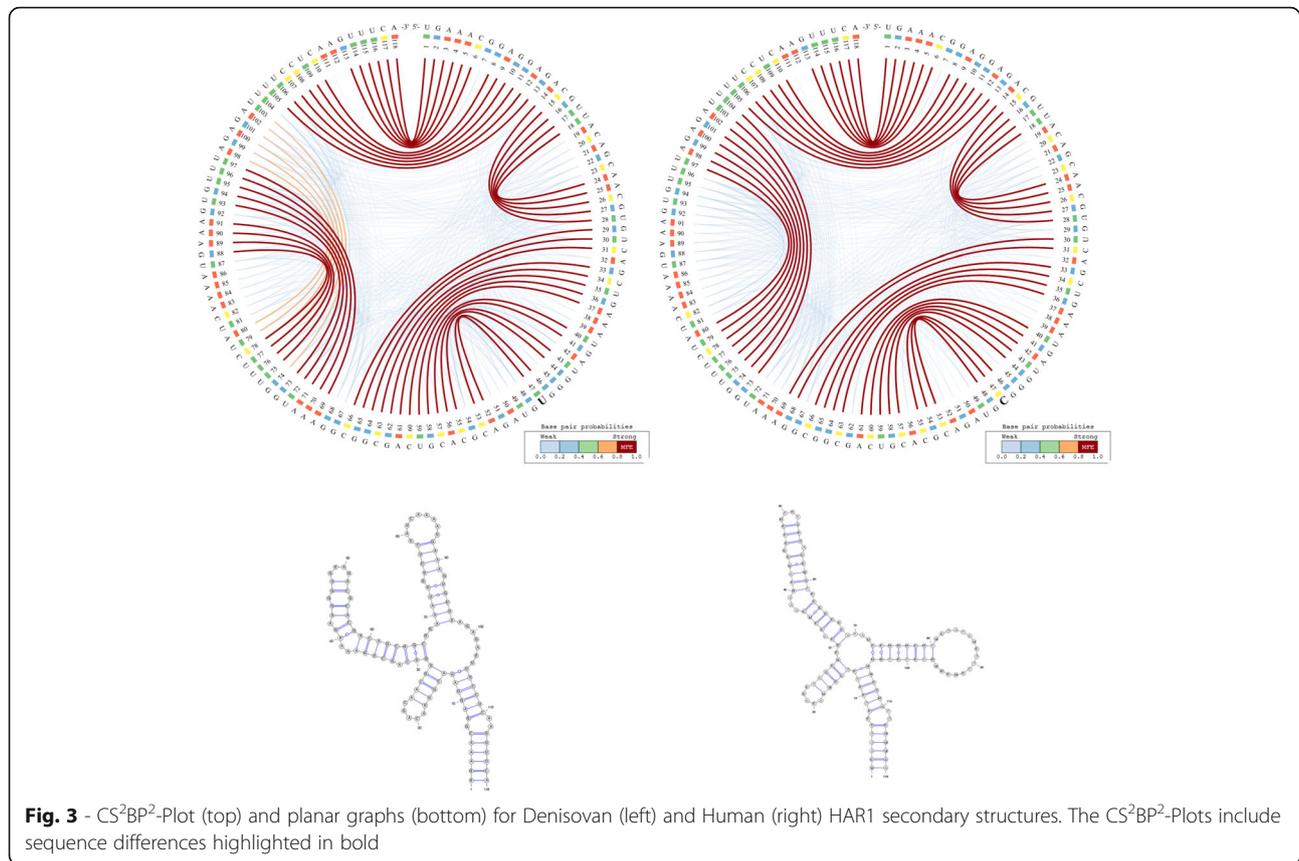
Results and discussion

We consider five use cases in this study to further show how the CS²BP²-Plot can be successfully used as a

comparative tool to discover important characteristics of RNA molecules and shed light on pertinent biological questions. The applications we show are very diverse and exemplify how our new tool can be used in studies of evolution, function and engineering of ncRNAs.

Case study 1: visualization of evolutionary changes in non-coding RNAs

Human accelerated regions (HARs) are sequences in the DNA that have an accumulation of human specific changes [14], which were caused by accelerated rates of nucleotide substitutions in humans, when compared to other vertebrates. Located at the end of chromosome 20 in humans, the HAR1 in particular, shows the highest substitution rate among HARs. The HAR1 is 118 bp long and has 18 substitutions that are specific to the human species. The HAR1 is part of two overlapping long ncRNAs, HAR1F and HAR1R, both of which are expressed very specifically in the brain in a developmental time point that is fundamental for brain formation [14]. Different studies of experimental biology have shown that the 118 bp HAR1 [14–18] folds into a stable secondary structure, which differs between the human species and the other vertebrates. If we consider the structure of



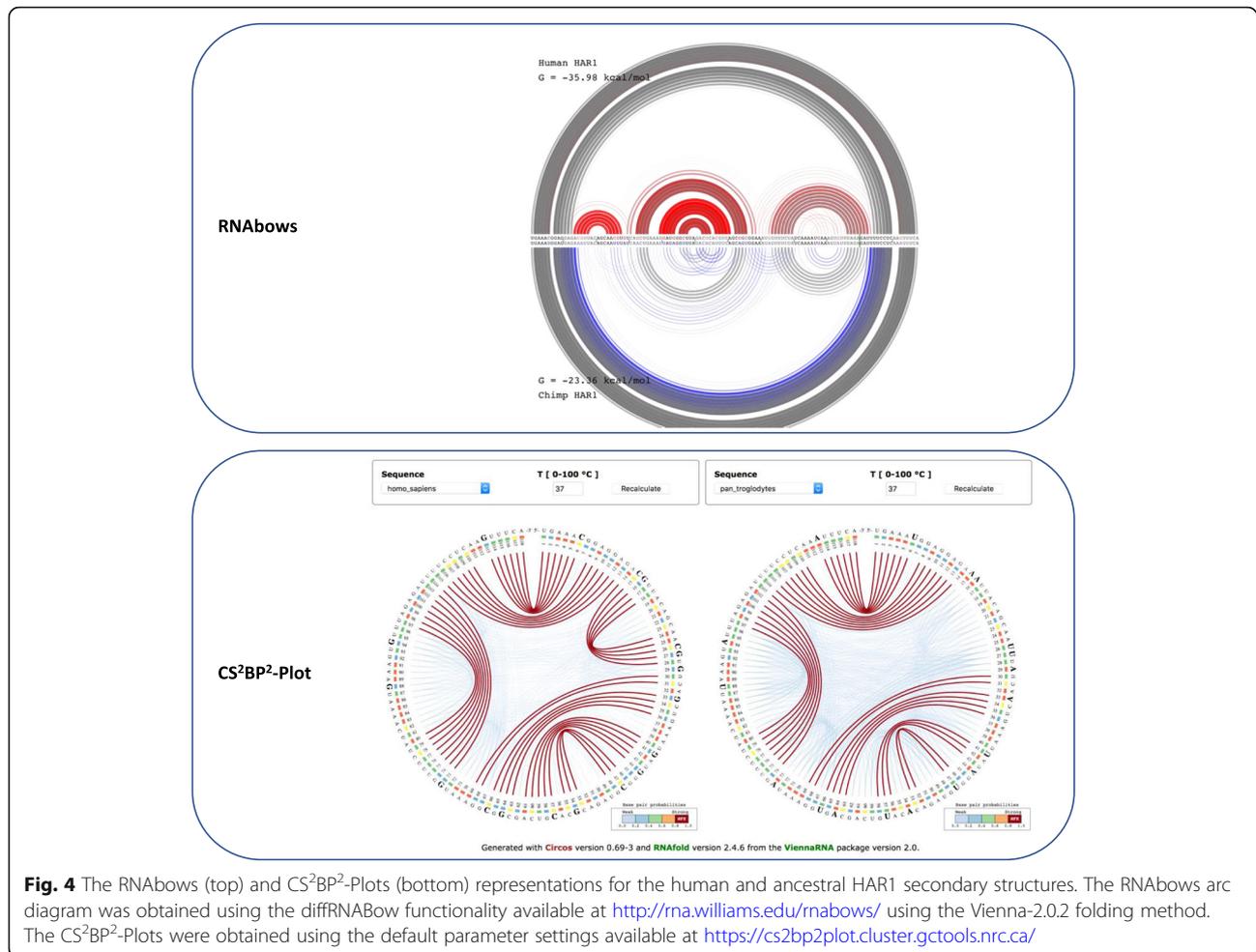
non-human species as the ancestral structure, that means that the HAR1 has been kept conserved in non-human species since the last common ancestor and changed only in the human lineage. This is a likely indicator of adaptive evolution acting on the human version of HAR1. To confirm this hypothesis, it is of great importance to compare structures. In this way, we can better understand how the HAR1 evolved in humans. Since the HAR1 is extremely similar in non-human vertebrates, we considered the chimpanzee sequence as a proxy of the ancestral sequence.

We use the CS²BP²-Plot to highlight the major similarities and differences among the predicted secondary structures for the ancestral, an archaic human (Denisovan) and the modern human HAR1. For convenience and comparison purposes, the planar graph representation of each secondary structure is also included.

The information represented in Fig. 2 can be interpreted based on the predominance and the location of various visual cues related to colour and position. For example, the CS²BP²-Plot of the ancestral HAR1 structure (Fig. 2 – left side images) displays two internal sequences with no base pairings between positions 15 and 29 and between positions 81 and 96. By comparison, the corresponding internal sequence in the Denisovan (Fig. 2 – right side images and Fig. 3 – left side images) HAR1 structure contains two solid

4 base-pair stems formed due to 2 base pair mutations at positions 88 and 94 replacing the A and U bases in ancestral with the more stable Gs in Denisovan.

A total of 17 single and consecutive base pair mutations (positions 6, 15–16, 26–27, 29, 33, 41, 44, 54, 57, 64, 66, 73, 88, 94 and 113) can be identified between the ancestral and the Denisovan HAR1 structures, while only one mutation (U replaced by C) at position 47 occurred between the Denisovan and the modern human structures. The large number of mutations that distinguish the Denisovan from the ancestral sequences was apparently caused not only by the emerging of 2 new 4-base pair stems in Denisovan, but also caused by the breakdown of multiple existing stems in the ancestral, as well as the creation of new stems in Denisovan. The orange arcs in the Denisovan HAR1 CS²BP²-Plot suggest the existence of a powerful energetic pressure to create a stem between positions 74–80 and 98–103. This stem appears in the human HAR1 structure due to changes in the stem structures between positions 69–78 and 88–97. Importantly, the single mutation that differentiates the sequences of modern and archaic humans caused a stabilization of a stem that reverted back to the ancestral state. This observation is biologically very relevant, since it also corroborates the hypothesis that the HAR1 structure adapted in the human



lineage towards an increase in stability. This is a novel piece of information that could not be acquired with the visual comparison of the classical dot-plots.

In a previous work on the HAR1 [19], we constructed a model showing that the most likely last step in turning the ancestral state into the modern human state was exactly the substitution at position 47 that distinguishes the modern human sequence from that of the Denisovan, which provides an independent support for the conclusions we drew from our visualizations.

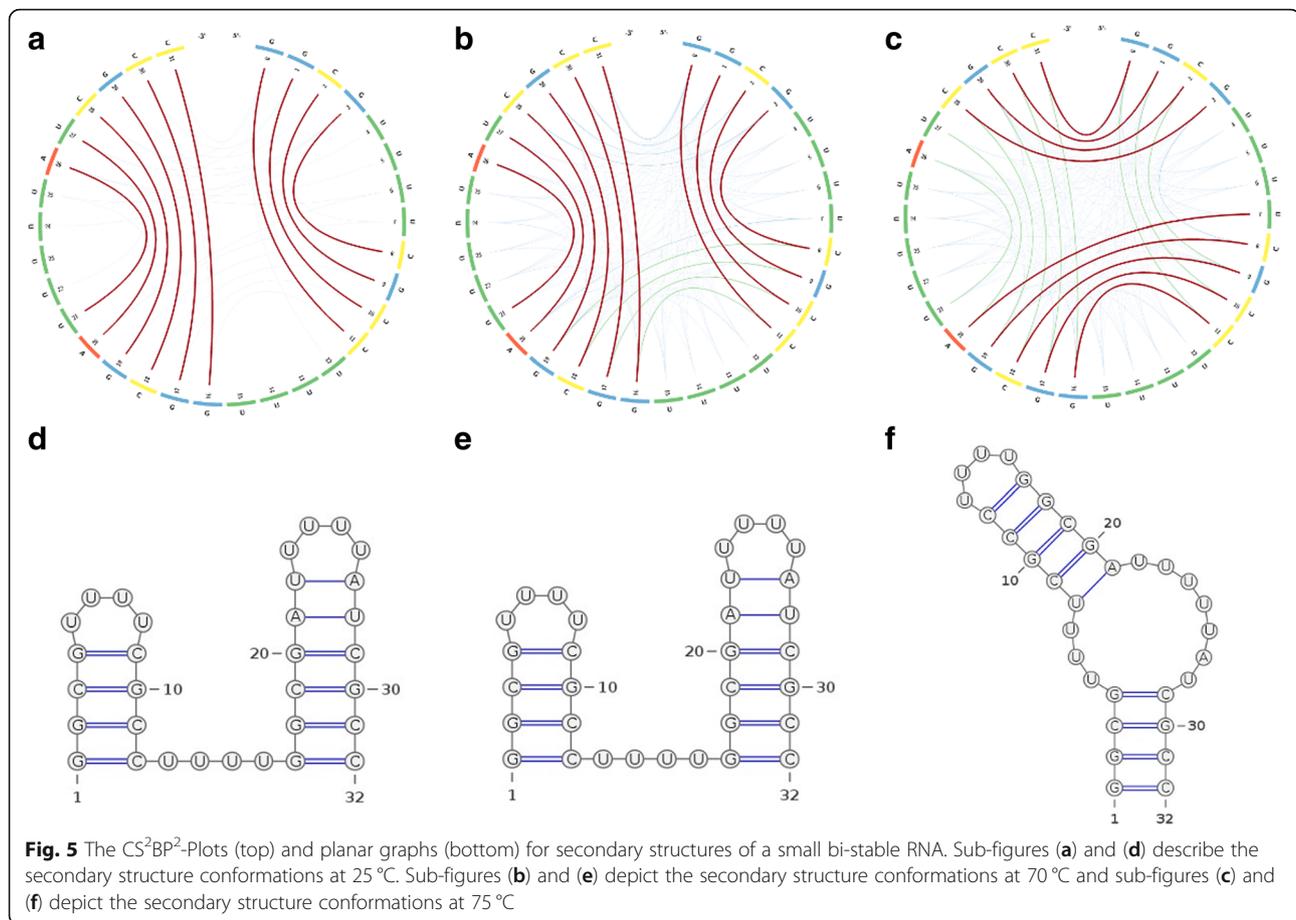
We can also notice in the CS²BP²-Plot that the human secondary structure is by far the most stable with a total $MFE_{\text{human}} = -33.5 \text{ Kcal/mol}$ out of all three HAR1 structures ($MFE_{\text{ancestral}} = -20.0 \text{ kcal/mol}$; $MFE_{\text{Denisovan}} = -32.1 \text{ Kcal/mol}$). This is likely due to the emergence of stronger base pairings in the modern human structure (low number of blue arcs), which could have occurred in the human lineage as a result of adaptive pressures.

We also use this use case to show how the CS²BP²-Plot representation of human versus ancestral HAR1 ncRNAs compares with the RNAbows representations (Fig. 4). The stem structure differences between the two secondary

structures of human and ancestral are explained in detail above and are much easier to identify when using the CS²BP²-Plot versus the RNAbows representation.

Case study 2: visualization of energetic transitions between RNA secondary structures

The RNA folding process leading to a native fold with a minimum free energy is typically perceived as a succession of transitions between intermediate states or conformations that represent local minima in the free energy landscape. While the free energy barrier between two consecutive states of an RNA structure is typically small allowing for smooth and fast transitions, large energy barriers are sometimes encountered. These barriers act as folding traps and slow down the folding process, leading to bi-stable RNA structure. The first bi-stable RNA structure had around 150 bp in length and represented a ribozyme, which was reported in 2000 [20]. Since then, other studies proposed shorter RNA sequences (20–40 bp) with computationally proven bi-stable secondary structures [21, 22]. Nevertheless their experimental validation, based on UV-melting and gel assays, was not possible at the time due to their significantly



fast timescale folding process. In 2003, Hobartner and Micura [23] successfully used ¹H NMR spectroscopy to study a series of bi-stable RNAs of 18–20 bp in length, each RNA comprising of two competing stem-loop motifs.

The CS²BP²-Plot can be used to visualize the energetic transitions between two stable secondary structures of the RNA sequence 4 proposed by Hobartner and Micura [23].

Images (a) and (d) in Fig. 5 depict the secondary structure formed at 25 °C with an MFE = -17.85 Kcal/mol. The two stems formed at positions 0–3/8–11 and 16–21/26–31 are in agreement with the reported 85–15 equilibrium by imino proton NMR spectroscopy. We can also notice that virtually no competing base pairs (no green or orange arcs) exist at this temperature. When the temperature is increased to 70 °C (images (b) and (e) in Fig. 5), two competing stems depicted by green arcs appear at positions 0–3/28–31 and 7–11/16–20 and the MFE of the secondary structure decreases to -4.06 Kcal/mol. With further increase in temperature, the new stems will break the existing ones, thus forming a new secondary structure for the RNA sequence 4. Images (c) and (f) in Fig. 5 depict the second stable conformation of the RNA structure formed at 75 °C with MFE = -2.63 Kcal/mol, which is

supported by the UV-melting analysis reported in Hobartner and Micura [23].

Case study 3: gRNAs for CRISPR/Cas systems

Guide RNAs (gRNAs) are typically used in CRISPR/Cas systems to direct sequence-specific DNA cleavage at desired locations along the target sequence. Nevertheless, gRNAs have a wide spectrum of cleavage effectivity. While there is a large amount of active research in the field, the mechanisms and factors governing their activities are still poorly understood.

Thyme et al., 2016 [24] suggest that there are two potential mechanisms that decrease gRNA performance: (i) weak gRNA sequence content that do not form active Cas9-gRNA complexes, and (ii) gRNAs with in vivo refractory target sites.

Here, we use CS²BP²-Plot to: (i) provide a visual interpretation that supports the hypothesis that gRNA secondary structure plays a significant role in the modulation of Cas9 cleavage efficiency, and (ii) suggest a third potential mechanism that might contribute to better gRNA performance, which is the need to design gRNAs with minimal self-folding structures.

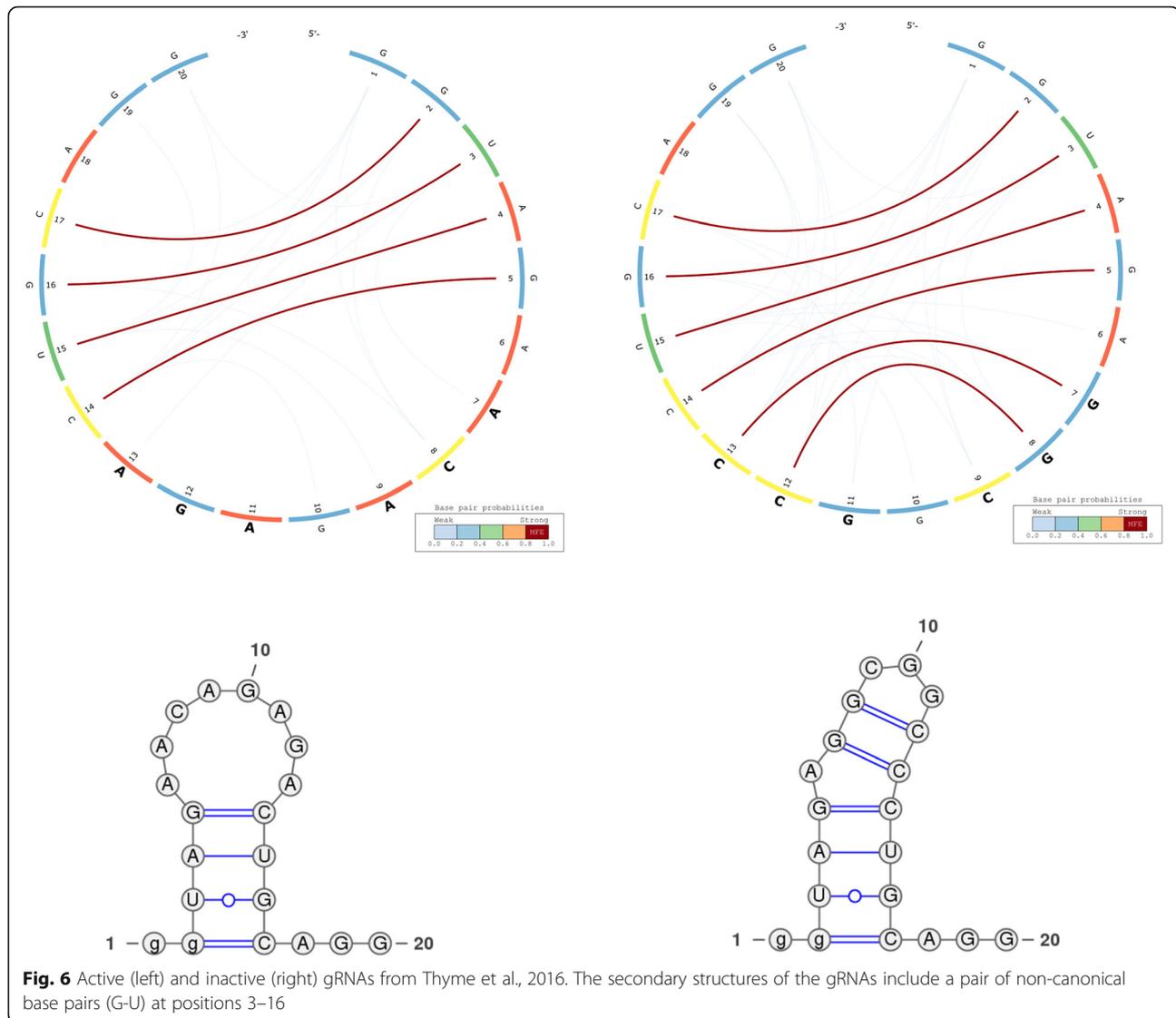


Figure 6 depicts the predicted secondary structures of an active (left) and an inactive (right) gRNA sequence from Thyme et al. (2016). The inactive gRNA has a significantly more stable secondary structure (MFE = -4.40 kcal/mol) compared to the active gRNA (MFE = -2.6 kcal/mol). This suggests that strong hairpin formation hampers the ability of gRNAs to interact with the desired target, an effect that can be observed on all gRNA sequences and their mutated variants from Thyme et al. (2016).

Case study 4: RiboSNitches

A RiboSNitch is a regulatory RNA in which a specific Single Nucleotide Polymorphism (SNP) has a structural consequence that results in a local or global conformational change in the secondary structure, which could lead to a disease phenotype.

Unlike a riboswitch, a RiboSNitch results in a permanent change in regulation and can thus lead to disease phenotypes. RiboSNitches represent a novel diagnose and therapeutic target, since small molecules can repartition the RNA structural ensemble [25].

Here we use CS²BP²-Plots to computationally validate and visualize one of the riboSNitches presented by Corley et al. [26], which represents a sequence flanking an SNV in the 3' untranslated region of the activated RNA polymerase II transcriptional co-activator p15 (SUB1). A single base mutation at position 51 has a major effect not only on the MFE structure, but also on the entire structural landscape of the RNA secondary structure as depicted in Fig. 7, which can be more clearly visualized with our new representation. In a similar manner, the CS²BP²-Plot could be used to investigate other putative RiboSNitches, providing valuable and intuitive visual

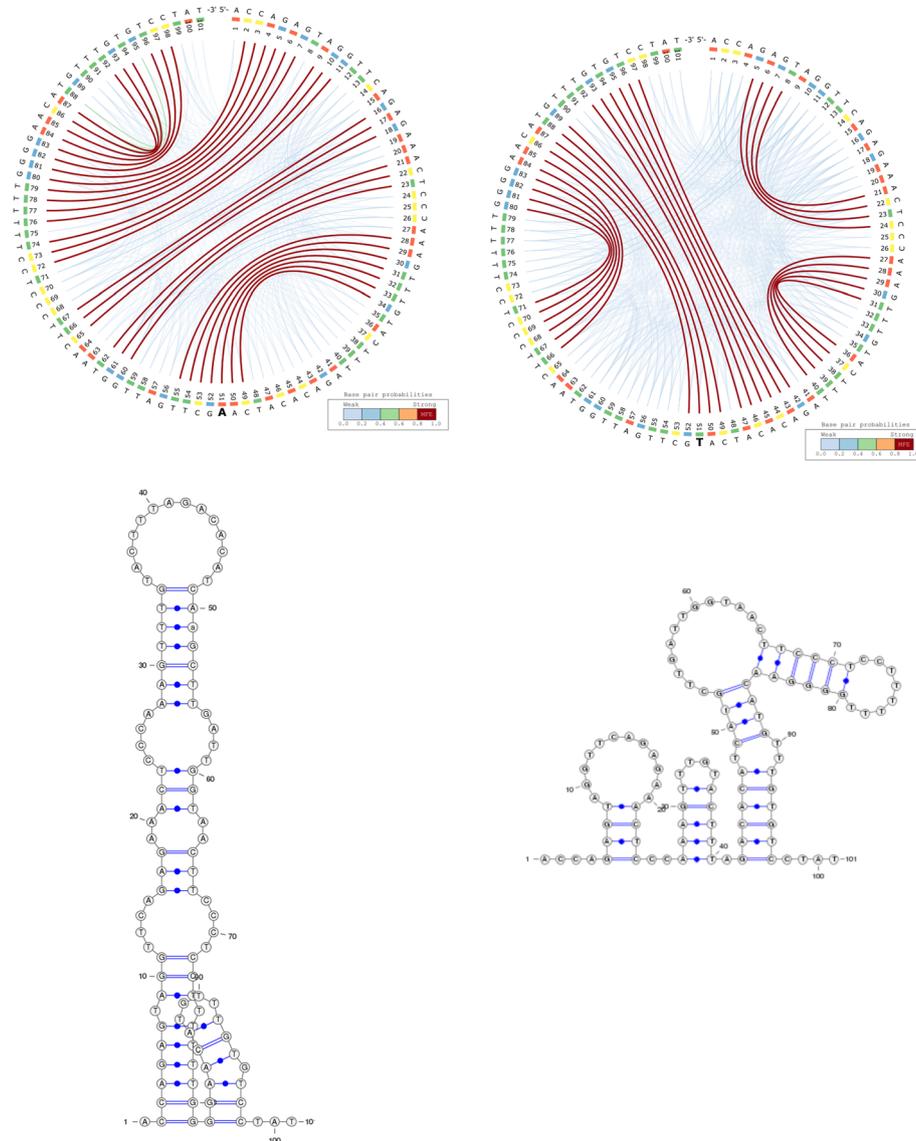


Fig. 7 The CS^2BP^2 -Plots and planar graph diagrams of the sequence flanking an SNV in the 3'UTR of SUB1. The activated RNA polymerase II transcriptional co-activator p15 or SUB1 homolog was previously identified as a riboSNitch by Corley et al. [24]

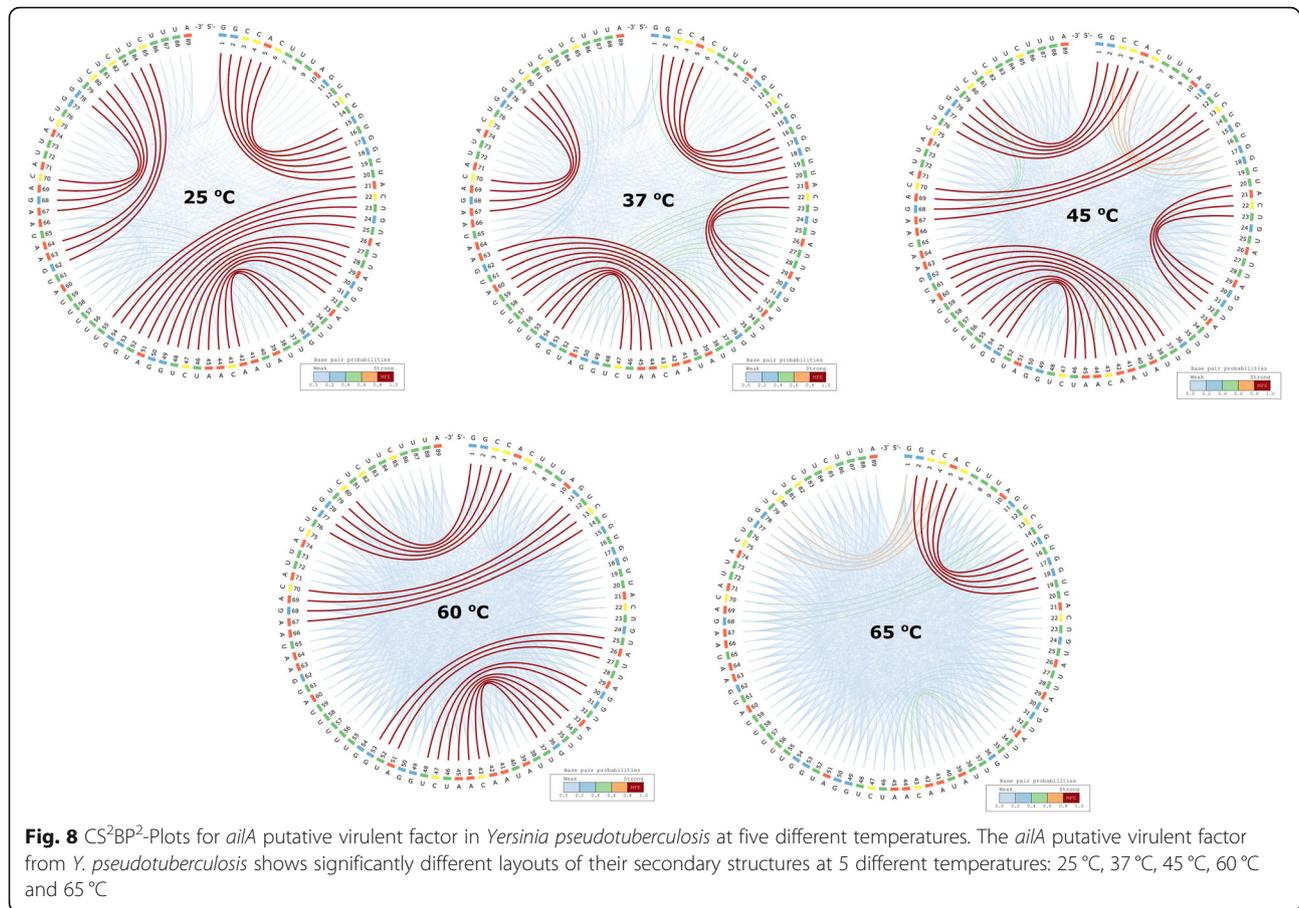
clues about the impact of SNPs on the entire landscape of the structure.

Case study 5: RNA thermometers

The RNA world is very diverse and their role in living organisms spans a large spectrum. Some RNAs are able to control gene expression in a temperature-dependent manner and are called “RNA thermometers”, such as those identified in the structurome of *Yersinia pseudotuberculosis* [27]. Righetti et al. (2016) identified two candidate virulence factors for this pathogen: *ailA* (attachment invasion locus protein) and *cnfY* (cytotoxic necrotizing factor). They reported the genome-scale landscape of RNA structures of the human pathogen *Y. pseudotuberculosis* at

three physiologically relevant temperatures reflecting environmental (25 °C), host body (37 °C), and heat shock (42 °C) conditions at single-nucleotide resolution.

The CS^2BP^2 -Plots and planar graph secondary structure representations from Fig. 8 are in agreement with the findings of Righetti et al. (2016), suggesting that the expression of virulence-relevant functions in *Y. pseudotuberculosis* and reprogramming of its metabolism in response to temperature is associated with a restructuring of some of its RNAs. We also performed *in-silico* experiments for higher temperatures (60 °C and 65 °C) and we observed a changing structural landscape for *ailA* by losing existing stems while forming new ones, which could qualify this molecule as a 5-phase riboswitch. Although experimental



evidence is needed to further confirm these observations, our plots allow for a straightforward comparison of the landscape change when the temperature varies.

Conclusions

Our results demonstrate that a single image consisting of circular diagrams combined with color and weight-coded arcs (the CS²BP²-Plot) can efficiently represent both the minimum free energy secondary structure and the base pair probabilities for an RNA structure. The CS²BP²-Plot significantly simplifies the interpretation, analyses and comparisons of RNA secondary structures, thereby making evolutionary and energetic transition results more easily understandable.

Methods

The CS²BP²-Plot currently use Circos [28] as a graphical library and Go Language (<https://golang.org/>) scripts to prepare the input files for Circos. Nevertheless, other more versatile approaches could be considered in the future, such as the JavaScript D3 library, for enhanced interactivity and user experience. Our current input consists of up to 10 RNA sequences in FASTA format. In addition the RNA sequences were computationally folded with the

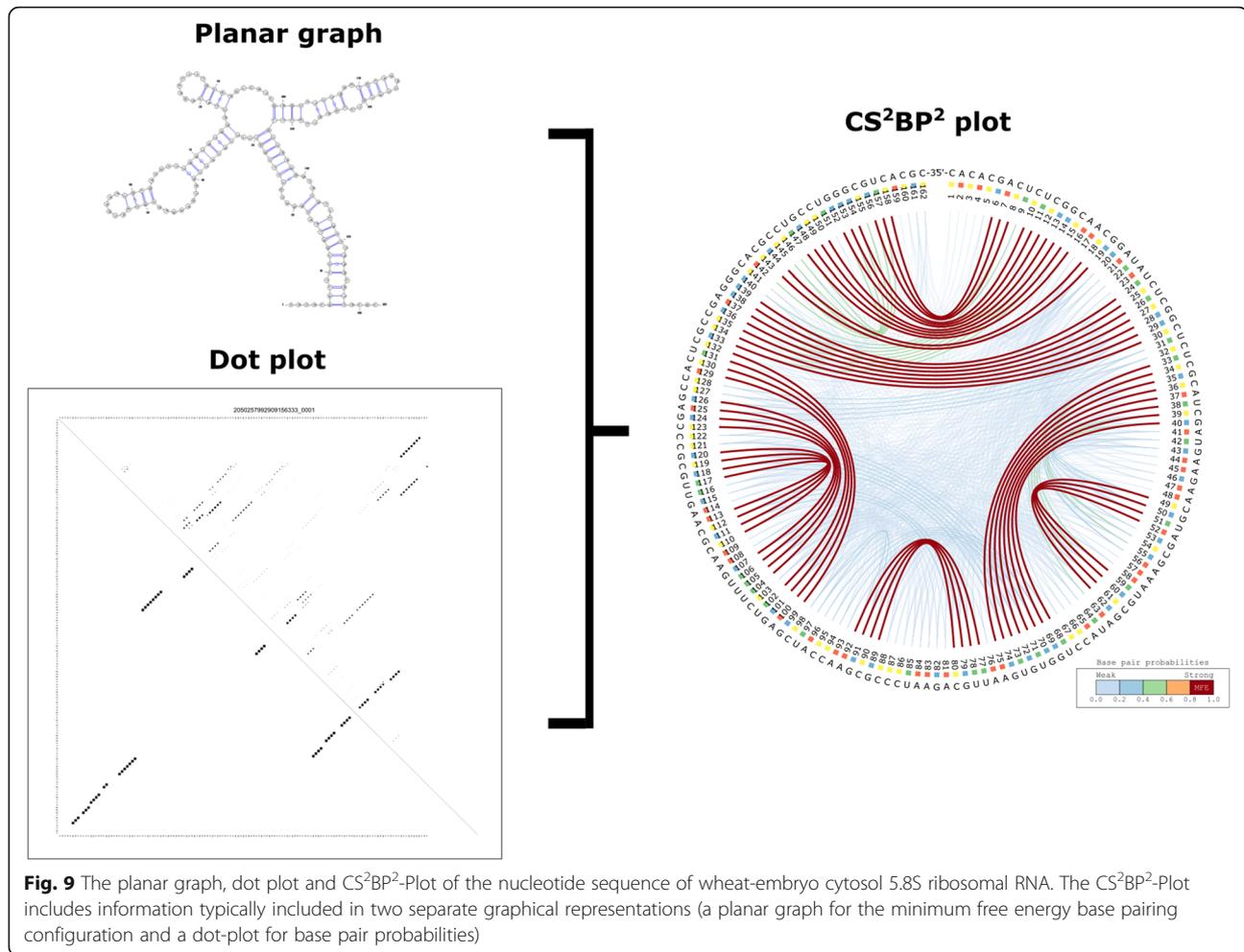
aid of RNAfold from the ViennaRNA Package [6] version 2.0 at different temperatures. Nevertheless other RNA folding algorithms can also be used, such as Mfold [29], RNAsoft [30] and RNAstructure [31]. Default RNAfold parameters (`--partfunc = 1`, `--temp = 37`, `--dangles = 2`) were used, with the exception of the bi-stable RNA secondary structures calculated at non-standard temperatures (25° C, 70° C and 75° C).

Linear arc diagrams and planar graphs were generated with VARNA [3] version 3–93 using corresponding parameter settings (`-algorithm line`, `-algorithm circular`), while dot-plots were calculated with RNAfold version 2.4.6.

The circular secondary structure base pair probabilities plot (CS²BP²-plot)

The CS²BP²-Plot challenges the classical ways of representing RNA secondary structures and combines base pairings with dot-plot values in a single graphical representation (Fig. 9) capable of assisting biologists in quickly spotting similarities and differences among a large number of secondary structures and their corresponding RNA sequences.

The use of a circular diagram instead of a linear one is justified by the equivalent of π times savings in horizontal



spread, where π (3.14 ...) equals the circle length (i.e. the length of the RNA sequence) divided by the circle diameter (i.e. the width of the graphical representation).

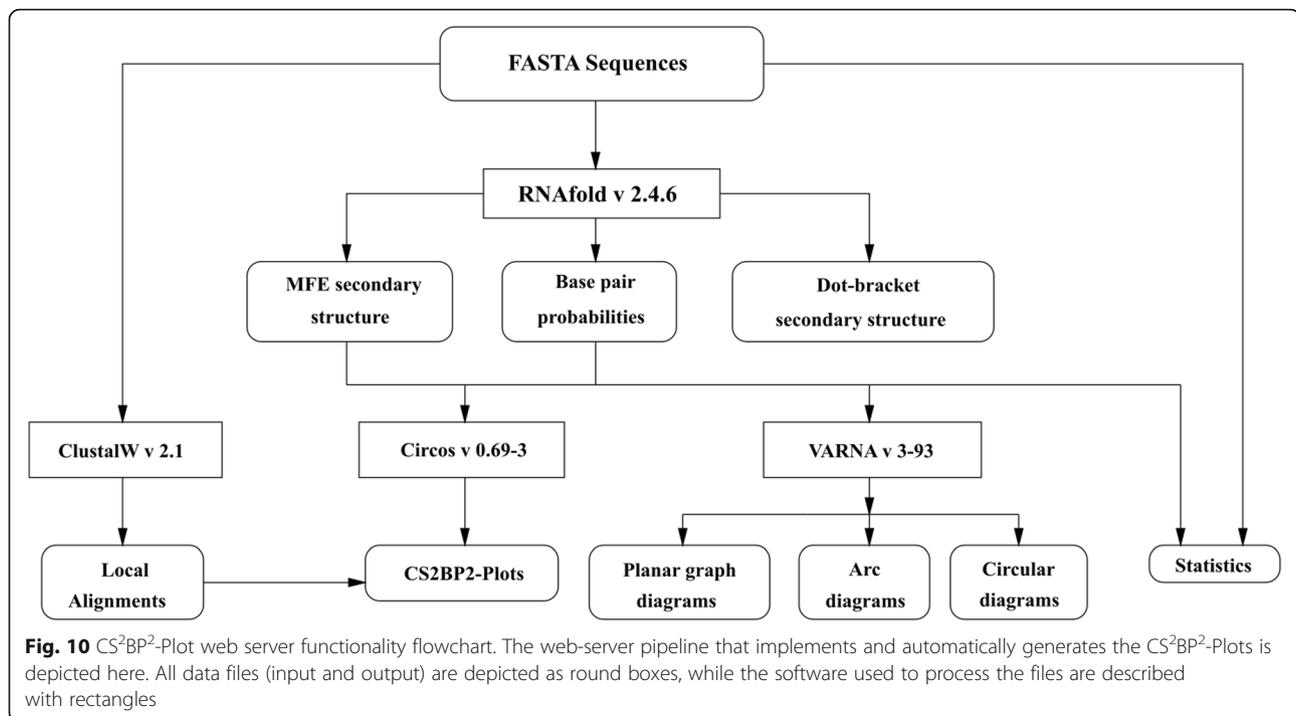
The outer layer (the RNA sequence) consists of 4 types of equally spaced blocks colored and labeled corresponding to each base (A, C, G and U).

The inner layer consists of a set of colored arcs that connect base pairs on the RNA sequence showing hydrogen bonds depicted by short segments in a typical linear graph RNA secondary structure representation. The arc colours, currently using a 5-color palette assigned values within 0.2 unit intervals spanning the interval [0,1], are assigned based on the corresponding base pair probabilities, ranging from blue (less stable) to dark orange (more stable) with green representing base pairs with medium energetic stability. The thickness of each arc is also proportional with the uncertainty probabilities using 5 incremental sizes ranging from 1 (less stable) to 9 (more stable). The dark red arcs represent the *MFE* base pairing corresponding to the most stable interactions. Each plot includes a legend summarizing the probabilistic range significance for each color.

Originally implemented in Perl and currently converted to the Go programming language and using Circos version 0.69–3 [28] as graphical library, the CS²BP²-Plot can be applied to represent and compare secondary structures of various types of RNAs, such as small (< 200 bp) non-coding (ncRNAs) and bi-stable small RNAs.

The CS²BP²-plot web server

We provide a web server implementation that allows pairwise comparison of two predicted RNA secondary structures. As depicted in Fig. 10, the web server accepts as input up to 10 RNA sequences, each no longer than 200 bases and uses RNAfold version 2.4.6 from the ViennaRNA package version 2.0 [6] to predict their secondary structures and corresponding base pair probabilities. The user can adjust the temperature and 5 other parameters such as the linearity of the RNA molecule (linear or circular), G-quadruplex formation incorporation, forbidden lonely pairs, and forbidden GU pairs along the whole sequence or at the ends of helices.



Once the secondary structures and base pair probabilities are calculated for all sequences, two sequences are selected for visualization and their corresponding CS²BP²-Plots are generated using Circos version 0.69–3 [28] and a customized template design that allows for highlighting of base pair differences between two sequences based on a local alignment performed using ClustalW version 2.1. This provides a significant advantage over other secondary structure comparative tools such as diffRNABow [4], which work only for two sequences of the same length. In addition to this, we use VARNA version 3–93 to generate the corresponding traditional planar graph, arc and circular diagram plots for the convenience of the users that are more accustomed with this type of visualization.

The webserver also provides information regarding the sequence length and base content (single bases, AT/AU and GC), the MFE of the predicted structure, the number of MFE base pairs, the number of similar and different base pairs between the two sequences, and the dot-bracket representation of the MFE secondary structure. The information is structured so that each section can be expanded or contracted allowing the user to focus on the relevant sections. In addition, all the information generated by the web-server is downloadable as a ZIP archive.

Acknowledgements

We thank the editorial team and anonymous reviewers for excellent comments and suggestions.

Funding

The work was supported by the University of Guelph (Food from Thought Initiative) and National Research Council of Canada.

Availability of data and materials

The authors provide free access to the web server that implements the new visualization approach presented in the manuscript: <https://nrcmonsrv01.nrc.ca/cs2bp2plot>

Authors' contributions

DT and MBWC edited the paper. SL implemented the web server application. DT identified and developed the use cases. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Animal Biosciences, Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, Ontario, Canada. ²Digital Technologies Research Center, National Research Council Canada, 100 des Aboiteaux St, Moncton, NB E1A7R1, Canada. ³School of Computer Science, University of Guelph, Guelph, Ontario, Canada. ⁴Department of Computer Science, TFome Research Group, Bioinformatics Group, Interdisciplinary Center of Bioinformatics, University of Leipzig, Härtelstrasse 16-18, D-04107 Leipzig, Germany.

Received: 6 February 2019 Accepted: 14 May 2019

Published online: 29 May 2019

References

- Mortimer SA, Kidwell MA, Doudna JA. Insights into RNA structure and function from genome-wide studies. *Nat Rev Genet.* 2014;15:469–79. <https://doi.org/10.1038/nrg3681>.
- Cruz JA, Westhof E. The dynamic landscapes of RNA architecture. *Cell.* 2009;136:604–9. <https://doi.org/10.1016/j.cell.2009.02.003>.
- Darty K, Denise A, Ponty Y. VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics.* 2009;25:1974–5. <https://doi.org/10.1093/bioinformatics/btp250>.
- Aalberts DP, Jannen WK. Visualizing RNA base-pairing probabilities with RNAbow diagrams. *RNA.* 2013;19:475–8. <https://doi.org/10.1261/ma.033365.112>.
- Lai D, Proctor JR, Zhu JYA, Meyer IM. R-CHIE: a web server and R package for visualizing RNA secondary structures. *Nucleic Acids Res.* 2012;40:e95. <https://doi.org/10.1093/nar/gks241>.
- Lorenz R, Bernhart SH, Höner Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. *Algorithms Mol Biol.* 2011;6:26. <https://doi.org/10.1186/1748-7188-6-26>.
- Seetin MG, Mathews DH. RNA structure prediction: an overview of methods. *Methods Mol Biol.* 2012;905:99–122. https://doi.org/10.1007/978-1-61779-949-5_8.
- Nussinov R, Piecznik G, Griggs JR, Kleitman DJ. Algorithms for loop matchings. *SIAM J Appl Math.* 1978;35:68–82. <https://doi.org/10.1137/0135006>.
- Churkin A, Barash D. RNA dot plots: an image representation for RNA secondary structure analysis and manipulations. *Wiley Interdiscip Rev RNA.* 2013;4:205–16. <https://doi.org/10.1002/wrna.1154>.
- Höchsmann M, Töller T, Giegerich R, Kurtz S. Local similarity in RNA secondary structures. *Proceedings IEEE Comput Soc Bioinform Conf.* 2003;2:159–68. <http://www.ncbi.nlm.nih.gov/pubmed/16452790>. Accessed 11 Apr 2019.
- Mattei E, Pietrosanto M, Ferrè F, Helmer-Citterich M. Web-Beagle: a web server for the alignment of RNA secondary structures. *Nucleic Acids Res.* 2015;43:W493–7. <https://doi.org/10.1093/nar/gkv489>.
- Li Y, Shi X, Liang Y, Xie J, Zhang Y, Ma Q. RNA-TVcurve: a web server for RNA secondary structure comparison based on a multi-scale similarity of its triple vector curve representation. *BMC Bioinformatics.* 2017;18:51. <https://doi.org/10.1186/s12859-017-1481-7>.
- Elias R, Hoksza D. TRAVeLer: a tool for template-based RNA secondary structure visualization. *BMC Bioinformatics.* 2017;18:487. <https://doi.org/10.1186/s12859-017-1885-4>.
- Pollard KS, Salama SR, Lambert N, Lambot M-A, Coppens S, Pedersen JS, et al. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature.* 2006;443:167–72. <https://doi.org/10.1038/nature05113>.
- Ziegeler M, Cevc M, Richter C, Schwalbe H. NMR studies of HARI RNA secondary structures reveal conformational dynamics in the human RNA. *ChemBiochem.* 2012;13:2100–12. <https://doi.org/10.1002/cbic.201200401>.
- Prabhakar S, Visel A, Akiyama JA, Shoukry M, Lewis KD, Holt A, et al. Human-specific gain of function in a developmental enhancer. *Science.* 2008;321:1346–50. <https://doi.org/10.1126/science.1159974>.
- Beniaminov A, Westhof E, Krol A. Distinctive structures between chimpanzee and human in a brain noncoding RNA. *RNA.* 2008;14:1270–5. <https://doi.org/10.1261/rna.1054608>.
- Burbano HA, Green RE, Marčić T, Lalueza-Fox C, de la Rasilla M, Rosas A, et al. Analysis of human accelerated DNA regions using archaic hominin genomes. *PLoS One.* 2012;7:e32877. <https://doi.org/10.1371/journal.pone.0032877>.
- Walter Costa MB, Höner zu Siederdisen C, Tulpan D, Stadler PF, Nowick K. Temporal ordering of substitutions in RNA evolution: uncovering the structural evolution of the human accelerated region 1. *J Theor Biol.* 2018;438:143–50.
- Schultes EA, Bartel DP. One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science.* 2000;289:448–452. <http://www.ncbi.nlm.nih.gov/pubmed/10903205>. Accessed 8 Jan 2016.
- Flamm C, Hofacker IL, Maurer-Stroh S, Stadler PF, Zehl M. Design of multistable RNA molecules. *RNA.* 2001;7:254–65. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1370083&tool=pmcentrez&rendertype=abstract>. Accessed 8 Jan 2016.
- Tøstesen E, Chen S-J, Dill KA. RNA folding transitions and cooperativity. *J Phys Chem B.* 2001;105:1618–30. <https://doi.org/10.1021/jp002877q>.
- Höbartner C, Micura R. Bistable secondary structures of small RNAs and their structural probing by comparative imino proton NMR spectroscopy. *J Mol Biol.* 2003;325:421–31. [https://doi.org/10.1016/S0022-2836\(02\)01243-3](https://doi.org/10.1016/S0022-2836(02)01243-3).
- Thyme SB, Akhmetova L, Montague TG, Valen E, Schier AF. Internal guide RNA interactions interfere with Cas9-mediated cleavage. *Nat Commun.* 2016;7:11750. <https://doi.org/10.1038/ncomms11750>.
- Halvorsen M, Martin JS, Broadaway S, Laederach A. Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genet.* 2010;6:e1001074. <https://doi.org/10.1371/journal.pgen.1001074>.
- Corley M, Solem A, Qu K, Chang HY, Laederach A. Detecting riboSNitches with RNA folding algorithms: a genome-wide benchmark. *Nucleic Acids Res.* 2015;43:1859–68. <https://doi.org/10.1093/nar/gkv010>.
- Righetti F, Nuss AM, Twittenhoff C, Beele S, Urban K, Will S, et al. Temperature-responsive in vitro RNA structure of Yersinia pseudotuberculosis. *Proc Natl Acad Sci U S A.* 2016;113:7237–42. <https://doi.org/10.1073/pnas.1523004113>.
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* 2009;19:1639–45. <https://doi.org/10.1101/gr.092759.109>.
- Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003;31:3406–15. <https://doi.org/10.1093/nar/gkg595>.
- Andronescu M, Aguirre-Hernández R, Condon A, Hoos HH. RNAsoft: a suite of RNA secondary structure prediction and design software tools. *Nucleic Acids Res.* 2003;31:3416–22. <https://doi.org/10.1093/nar/gkg612>.
- Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics.* 2010;11:129. <https://doi.org/10.1186/1471-2105-11-129>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

