Structural bioinformatics

PseudoViewer3: generating planar drawings of large-scale RNA structures with pseudoknots

Yanga Byun and Kyungsook Han*
Department of Computer Science and Engineering, Inha University, Incheon 402-751, South Korea.

Associate Editor: Prof. Ivo Hofacker

ABSTRACT

Motivation: Pseudoknots in RNA structures make visualization of RNA structures difficult. Even if a pseudoknot itself is represented without a crossing, visualization of the entire RNA structure with a pseudoknot often results in a drawing with crossings between the pseudoknot and other structural elements, and requires additional intervention by the user to ensure that the structure graph is overlap-free. Many programs such as web services prefer to obtain an overlap-free graph in one-shot rather than get a graph with overlaps to be edited. There are few programs for visualizing RNA pseudoknots, and PseudoViewer has been the almost only program that automatically draws RNA secondary structures with pseudoknots. The previous version of PseudoViewer visualizes all the known types of RNA pseudoknots as planar drawings, but visualizes some hypothetical pseudoknots as nonplanar drawings.

Results: We developed a new version of PseudoViewer for efficiently visualizing large RNA structures with any types of pseudoknots, both known and hypothetical, as planar drawings in one-shot. It is about 10 times faster than the previous algorithm, and produces a more compact and aesthetic structure drawing. PseudoViewer3 supports both web services and web applications.


Contact: khan@inha.ac.kr

Supplementary information: available at Bioinformatics online.

1 INTRODUCTION

Many programs have been developed for visualizing RNA secondary structures (De Rijk et al., 2003; Hofacker, 2003; Wiese and Glen, 2006), but there are few programs for visualizing pseudoknotted RNA structures automatically. From the standpoint of graph theory, a drawing of an RNA secondary structure is a tree, whereas a drawing of an RNA secondary structure with a pseudoknot is a graph with inner cycles within the pseudoknot, and possibly outer cycles formed between the pseudoknot and other structural elements. Thus, visualizing RNA pseudoknot structures is computationally more difficult than depicting RNA secondary structures.

A drawing of an RNA structure is planar if no two structural elements overlap or intersect. Visualizing a pseudoknotted RNA structure as a planar drawing is complicated, because it requires both the pseudoknots and the entire structure enclosing the pseudoknots to be planar. Overlaps in structure drawings may be removed by a program’s editing facility. But, some systems such as web services cannot use the editing facility, and thus prefer to obtain an overlap-free structure drawing in one-shot.

We previously developed a representation and algorithm for visualizing RNA pseudoknots as planar drawings. PseudoViewer1 is the first program that visualizes RNA secondary structures with H-type pseudoknots (Han et al., 2002). PseudoViewer2 (Han and Byun, 2003) visualizes RNA secondary structures with all types of pseudoknots in PseudoBase (van Batenburg et al., 2001). Previous versions of PseudoViewer focused on known RNA pseudoknot structures, and did not consider hypothetical pseudoknots that are topologically possible. Hence, they visualize some hypothetical pseudoknots as nonplanar drawings. A large number of hypothetical pseudoknots are often considered by prediction algorithms to find a good RNA pseudoknot structure for a given RNA sequence.

As an extension of PseudoViewer2, we developed PseudoViewer3 for visualizing RNA structures with all types of pseudoknots, both known and hypothetical, as planar drawings. PseudoViewer3 generates a compact drawing by representing a loop as a path of circles and line segments. PseudoViewer3 is about 10 times faster than PseudoViewer2.

2 METHODS

In the schematic diagram of RNA structures, stems and loops are simplified as line segments and circles, respectively. The space taken by non-paired bases of a pseudoknot is the convex hull of the structural elements of the pseudoknot. The convex hull is not necessarily a polygon since it consists of line segments and arcs. In the simplified structure we examine three types of overlaps: between line segments, between a line segment and a circle, and between circles.

For a planar drawing, we developed a heuristic method of partitioning stems into groups of non-crossing stems. Each group consists of stems that do not cross each other. Any two groups of stems can be drawn without a crossing, and we choose two largest groups for a planar drawing. The stems in the remaining groups, if any, are represented as dotted lines. Layout of a pseudoknot starts with an arc representation of the pseudoknot. The arcs representing the stems are put into two halfplanes separated by the backbone (Figure 1A-i). When visualizing the RNA structure, the backbone direction (either up or down) at base i is opposite to that of base j for each base pair (i, j). Note that in the final drawing of a pseudoknot (Figure 1A-iv), the backbone direction changes from up to down for the stems in the lower halfplane, whereas the backbone direction changes from down to up.
for the stems in the upper halfplane. Figure 1B shows a skeletal view of a hypothetical pseudoknot with 30 stems, visualized by PseudoViewer3.

![Diagram](image)

Fig. 1. (A) Example of converting an arc representation of a pseudoknot into a planar drawing. (i) Arc representation of a pseudoknot with stems both in the upper and lower halfplanes. (ii) Layout of the stems in the lower halfplane. (iii) Layout of the stems in the upper halfplane. (iv) Final visualization of the pseudoknot. (B) Skeletal view of a hypothetical pseudoknot with 30 stems. The bracket view of the pseudoknot is [[(([(())]([()])])][((()[]))]]

3 RESULTS AND DISCUSSION

Visualizing a pseudoknotted RNA structure as a planar drawing requires a large amount of time for overlap detection. While PseudoViewer2 examined possible overlaps between almost every structural element, PseudoViewer3 examines overlaps between structural elements that are connected to a same loop, which significantly reduces the time for overlap detection especially for large RNA structures. The execution time of PseudoViewer3 depends on the complexity of crossing relation of the stems in the RNA structure rather than on the RNA sequence length. A pseudoknotted RNA structure with more than 4,000 bases is visualized within a second. Table 1 shows the execution times for visualizing several RNA structures with 1,000 or more bases. PseudoViewer3 is about 10 times faster than PseudoViewer2.

Figure 2 shows two drawings of *E. coli* 16S rRNA, visualized by PseudoViewer2 and PseudoViewer3, respectively. In both drawings, the red box shows the pseudoknot part. By showing a loop region as a path of circles and line segments, PseudoViewer3 generates a more compact drawing than PseudoViewer2. Supplementary Figure 1 shows a planar drawing of *Aquifex pyrophilus* 16 rRNA visualized by PseudoViewer3. It includes three H-type pseudoknots and a big pseudoknot that includes the H-type pseudoknots. Neither PseudoViewer1 nor PseudoViewer2 can generate a planar drawing of the structure.

The structure drawing visualized by PseudoViewer3 can be stored in a file of several types, which include PNG, GIF, EPS, and Scalable Vector Graphics (SVG). SVG is an XML-based vector format that produces high-quality, interactive web graphics. The structure drawing in SVG format supports mouse-over events, which are useful for showing information on structural elements, such as base numbers and base pairs. PseudoViewer3 also generates the bracket view of the structure with a small number of bracket types, in which the stems in a group of non-crossing stems are represented with a same bracket type. Supplementary Table 1 shows the pseudoknot patterns and the topological relations of the stems, which were generated by PseudoViewer3 for the 856 RNA structures of the CRW database (http://www.rna.icmb.utexas.edu).

In summary, PseudoViewer3 efficiently generates an overlap-free drawing of RNA structures with pseudoknots. PseudoViewer3 visualizes RNA structures with any type of pseudoknots, and its web service can be used with any system that requires RNA structure drawings. PseudoViewer3 can be used either as a web application or web service. To facilitate the use of the web service, we provide sample web service clients in C# and Java.

<table>
<thead>
<tr>
<th>RNA</th>
<th>#bases</th>
<th>Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PseudoViewer2</td>
<td>PseudoViewer3</td>
</tr>
<tr>
<td><em>Escherichia coli</em> 16S</td>
<td>1542</td>
<td>420</td>
</tr>
<tr>
<td><em>Escherichia coli</em> 23 5P</td>
<td>1656</td>
<td>840</td>
</tr>
<tr>
<td><em>Escherichia coli</em> 23 3P</td>
<td>1266</td>
<td>810</td>
</tr>
<tr>
<td><em>Escherichia coli</em> 23</td>
<td>2904</td>
<td>4080</td>
</tr>
<tr>
<td><em>Thermus thermophilus</em> 23S 5′</td>
<td>1683</td>
<td>640</td>
</tr>
<tr>
<td><em>Thermus thermophilus</em> 23S 3′</td>
<td>1230</td>
<td>810</td>
</tr>
<tr>
<td><em>Thermus thermophilus</em> 23S</td>
<td>2893</td>
<td>3840</td>
</tr>
<tr>
<td>Hypothetical RNA sequence</td>
<td>4094</td>
<td>14220</td>
</tr>
</tbody>
</table>

Table 1. Execution times of PseudoViewer2 and PseudoViewer3 on an Intel Core 2 1.83 GHz processor with 2GB RAM. All the RNA data except the last one are from XRNA (http://rna.ucsc.edu/rnacenter/xrna/xrna.html).

**ACKNOWLEDGEMENTS**

This work was supported by the KOSEF grant (No. R01-2008-000-10197-0) funded by the Korea government (MEST).

**REFERENCES**


