Software Review

Mfold®: RNA modeling program

Authors: Prof: Michael Zuker, Professor of Mathematical Sciences;
License: Freeware, Copyright © Washington University.
Reviewed By: Dr. Ahmed Mansour Alzohairy
Genetic Department, Faculty of Agriculture, Zagazig University, Egypt.

Mfold at a glance

RNA structure is very important in many biological processes, including translation regulation in messenger RNA and replication of single-stranded RNA viruses. Much of Michel Zuker’s work is on studying RNA structure. He developed Mfold program as a tool for predicting the secondary structure of RNA, mainly by using thermodynamic methods (the Gibbs free energy). That sophisticated RNA modeling program takes into account many parameters, (e.g. pH, temperature, and the local composition bias of RNA), that affect the RNA folding. Using thermodynamics in the prediction algorithm of RNA structure is not common with Single sequence secondary structure prediction software. Up till now, only three programs, (Mfold, RNA fold, RNAshapes), is known to use this feature (Table 1). Recently, Mfold has been replaced by UNAFold, a software package that is much easier to install and run and it offers many more types of computations.

Mfold History

Santa Lucia (1998) have published his work on the unified view of polymer, dumbbell, and oligonucleotide DNA thermodynamics in PNAS journal. That paper was the base for M. Zuker (1999) and his colleagues to study the expanded sequence dependence of thermodynamic parameters for the prediction of RNA secondary Structure. Since then, Zuker has worked extensively to develop mfold, a program for predicting the secondary structure of RNA and DNA, mainly by using thermodynamic methods, which was published at Nucleic Acids Research journal in 2003 (Chain, 2005).

Mfold contribution to research

Many researchers in the field of Molecular biology have used using thermodynamics based Mfold modules during their original research for predicting the secondary structure of RNA and DNA. With many citations in peer reviewed journals, Mfold become very important bioinformatics software in molecular biology for prediction of secondary RNA structure. The freeware license and its efficient up to date modules beside its quick ability to produce results make it one of the most popular programs predicting secondary structure of RNA of nowadays.

Advantages

- Mfold is very easy to use and free of charge online.
- Mfold is internet based program that runs on (almost) any computer that has access to the Internet.
- Mfold has many versions and updated regularly.

Disadvantages:

- The energetic model that Mfold uses ignores 3-D interactions or protein-RNA interactions that could stabilize a suboptimal fold.
- Prediction from Mfold does not always correspond to the true biological fold of the RNA. So that, the researcher should be able to make his own decision.
- The researcher should be able to judge the stability of the optimal fold when Mfold returns suboptimal models.
- You cannot force pseudo-knots in Mfold.
- The time Mfold needs to analyze sequences increases very rapidly with their length.

Software Design

Mfold can predict the energetically optimal secondary structure of an RNA molecule by using of Zuker algorithm. Versions 3.xxx of Mfold for Unix and Linux operating systems are available via a free license for academic and nonprofit use only; commercial use is available for a fee.

*Corresponding author: alzohairy@yahoo.com
<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Links</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CentroidFold</td>
<td>Secondary structure prediction based on generalized centroid estimator</td>
<td><a href="http://www.ncrna.org/centroidfold/">http://www.ncrna.org/centroidfold/</a></td>
<td>Hamada et al. (2009)</td>
</tr>
<tr>
<td>CONTRAfold</td>
<td>Secondary structure prediction method based on conditional log-linear models (CLLMs), a flexible class of probabilistic models which generalize upon SCFGs by using discriminative training and feature-rich scoring.</td>
<td><a href="http://contra.stanford.edu/contrafold/server.html">http://contra.stanford.edu/contrafold/server.html</a></td>
<td>Do CB et al. (2006).</td>
</tr>
<tr>
<td>CyloFold</td>
<td>Secondary structure prediction method based on placement of helices allowing complex pseudoknots.</td>
<td><a href="http://cylofold.abcc.nicifcrf.gov/">http://cylofold.abcc.nicifcrf.gov/</a></td>
<td>Bindewald et al., 2010</td>
</tr>
<tr>
<td>PknotsRG</td>
<td>A dynamic programming algorithm for the prediction of a restricted class of RNA pseudoknots.</td>
<td><a href="http://bibiserv.techfak.uni-bielefeld.de/pknotsrg/submission.html">http://bibiserv.techfak.uni-bielefeld.de/pknotsrg/submission.html</a></td>
<td>Reeder et al., 2007</td>
</tr>
<tr>
<td>RNAsizes</td>
<td>MFE RNA structure prediction based on abstract shapes. Shape abstraction retains adjacency and nesting of structural features, but disregards helix lengths, thus reduces the number of suboptimal solutions without losing significant information. Furthermore, shapes represent classes of structures for which probabilities based on Boltzmann-weighted energies can be computed.</td>
<td><a href="http://bibiserv.techfak.uni-bielefeld.de/rnashapes/">http://bibiserv.techfak.uni-bielefeld.de/rnashapes/</a></td>
<td>Giegerich et al., 2004, Voß et al., 2006</td>
</tr>
<tr>
<td>RNAstructure</td>
<td>A program to predict lowest free energy structures and base pair probabilities for RNA or DNA sequences. Structure prediction can be constrained using experimental data, including SHAPE, enzymatic cleavage, and chemical modification accessibility. Graphical user interfaces are available for Windows and for Mac OS-X/Linux. Programs are also available for use with Unix-style text interfaces. Additionally, a C++ class library is available.</td>
<td><a href="http://rna.urmc.rochester.edu/RNAstructure.html">http://rna.urmc.rochester.edu/RNAstructure.html</a></td>
<td>Mathews et al. 2004, Mathews 2004</td>
</tr>
<tr>
<td>UNAFold</td>
<td>The UNAFold software package is an integrated collection of programs that simulate folding, hybridization, and melting pathways for one or two single-stranded nucleic acid sequences.</td>
<td><a href="http://www.bioinfo.rpi.edu/applications/hybrid/download.php">http://www.bioinfo.rpi.edu/applications/hybrid/download.php</a></td>
<td>Markham, Zuker, 2008</td>
</tr>
</tbody>
</table>

*Table: 1. Comparative table for other online for RNA Single sequence secondary structure prediction.*
Limitation

The current limit of the server is 3,000 bases. This is convenient for most non-coding RNAs except the largest ribosomal ones.

Mfold availability online

Servers in USA:
- www.bioinfo.rpi.edu/applications/mfold/
- http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi
- http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi

European mirrors of mfold:
- biobi serv.techfak.uni-bielefeld.de/cgi-bin/mfold_submit
- bioweb.pasteur.fr/seqanal/interfaces/mfold-simple.html

Australian counterpart:
- mfold.burnet.edu.au

References


www.gerfbb.com


