

Uniwersytet Jagielloński
Wydział Matematyki i Informatyki
INSTYTUT INFORMATYKI I MATEMATYKI KOMPUTEROWEJ
Studia dzienne

Nr indeksu: 1104482

Marta Szynczewska

Zbiór i ocena jakości i aktualności narzędzi do przewidywania struktury drugorzędowej RNA

Opracowano zgodnie z Ustawą o prawie autorskim i prawach pokrewnych z dnia 4 lutego 1994 r. (Dz.U. 1994 nr 24 poz. 83) wraz z nowelizacją z dnia 25 lipca 2003 r. (Dz.U. 2003 nr 166 poz. 1610) oraz z dnia 1 kwietnia 2004 r. (Dz.U. 2004 nr 91 poz. 869)

Kraków 2016

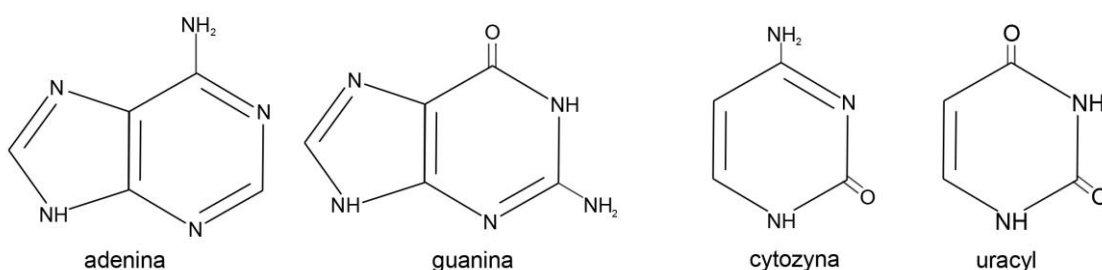
Spis treści

Wstęp teoretyczny	3
RNA – definicja	3
Struktury RNA	3
Metody	6
Afold	6
Carnac	7
CentroidFold	8
CentroidHomfold	10
ContraFold	12
DotKnot	13
HotKnots	14
TurboKnot	15
CMfinder	16
CRWrnafold	17
CyloFold	18
Mastr	19
MaxExpect	20
PPfold	23
RNAwolf	24
Sfold	24
taveRNA	24
TurboFold	24
Opcje narzędzi	25
Podsumowanie	29
Bibliografia	30

Wstęp teoretyczny

RNA – definicja

RNA, zwane inaczej kwasem rybonukleinowym, to długa liniowa makrocząsteczka polinukleotydowa. Nukleotydy połączone są wiązaniami fosfodiesterowymi 3' → 5'. Każdy nukleotyd zbudowany jest z jednostki cukrowej, przynajmniej jednej reszty fosforanowej oraz zasady azotowej. W przypadku RNA resztą cukrową jest ryboza oraz wyróżniamy cztery podstawowe zasady azotowe: adeninę oznaczaną literą A, guaninę (G), cytozynę (C) i uracyl (U).



Cząsteczki RNA występują głównie w postaci pojedynczej nici, jednakże łańcuch RNA może się zwinąć i stworzyć strukturę *spinki do włosów* o budowie dwuniciowej helisy. W tych strukturach adenina tworzy pary z uracylem, a guanina z cytozyną.

Struktury RNA

a) Pierwszorzędowa

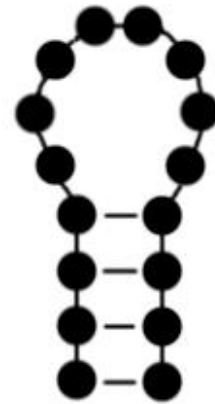
Strukturę pierwszorzędową RNA definiujemy jako ciąg nukleotydów ułożonych jeden po drugim.



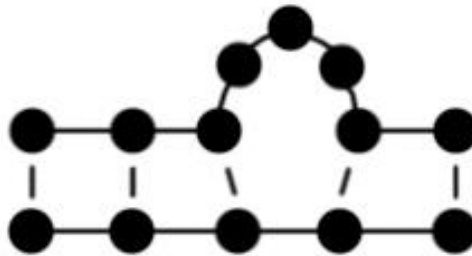
b) Drugorzędowa

Strukturę drugorzędową RNA definiujemy jako położenie ciągu nukleotydów na płaszczyźnie. Strukturami, jakie możemy zauważyć są m.in. spinki do włosów, wybrzuszenia, pętle wewnętrzne, węzły, pseudowęzły oraz odcinki dwuniciowe.

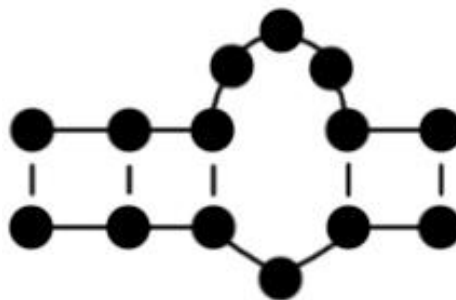
Spinka do włosów (hairpin) - struktura składająca się z części dwuniciowej oraz pętli zewnętrznej.



Wybrzuszenie (bulge loop) - struktura składająca się z przynajmniej dwóch niesparowanych nukleotydów tylko na jednej z nici.



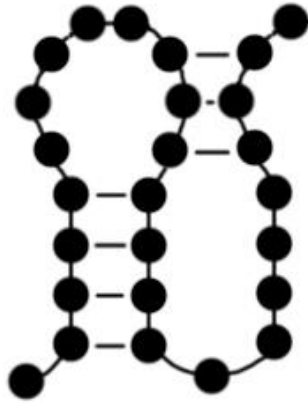
Pętla wewnętrzna (internal loop) - struktura składająca się z niesparowanych nukleotydów na obu niciach pomiędzy dwoma odcinkami podwójnej helisy.



Węzeł (junction) - rozgałęzienie przynajmniej trzech odcinków podwójnej helisy.

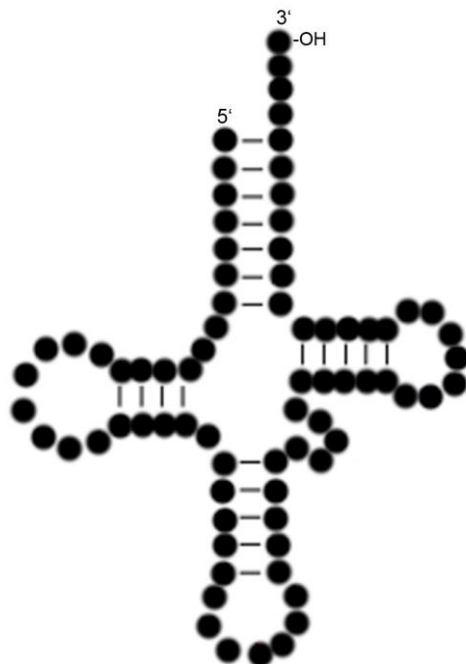


Pseudowęzeł (pseudoknot) - struktura opierająca się na oddziaływaniach pomiędzy nukleotydami wchodzącymi w skład innej struktury, np. spinki do włosów, a innymi nukleotydami.



Często zdarza się, że zasady nie tworzą idealnych par typu Watsona-Cricka. Uracyl może tworzyć parę z guaniną, ponieważ występują wiązania wodorowe pomiędzy N3 uracylu i C6 guaniny oraz C2 uracylu i N1 guaniny. Jednakże para guaniny i cytozyny jest od niej silniejsza.

Najbardziej znanym przykładem struktury drugorzędowej RNA jest tRNA, które wyglądem przypomina liść koniczyny.



Metody

Afold

Metoda Afold powstała między 2003 a 2006 rokiem. Autorem narzędzia jest Aleksey Ogurtsov. Służy ona głównie do przewidywania struktury drugorzędowej pętli wewnętrznych RNA dla sekwencji o długości nie większej niż 28 000 nukleotydów. Przetwarzanie tak długiej sekwencji przy użyciu tego algorytmu trwa ok. 28 godzin na komputerze z 2GB RAM.

W algorytmie wykorzystywany jest model najbliższego sąsiada NNM (nearest-neighbor model) z założeniem, że kara dla pętli wewnętrznej zależy od dwóch zmiennych:

$$F(s,d) = f_{\text{Len}}(s) + f_{\text{Diff}}(d),$$

gdzie s jest ilością niesparowanych nukleotydów w pętli, a d jest różnicą w długości pomiędzy dwoma niesparowanymi regionami tworzącymi pętlę. W celu oceny poprawności pętli, algorytm korzysta z listy kandydatów.

Złożoność czasowa algorytmu jest równa $O(M \cdot \log^2 L)$, gdzie L oznacza długość sekwencji. W porównaniu z Mfold oraz ZUCKER, Afold wypada znacznie lepiej. Przede wszystkim jest szybszy już w etapie wypełniania macierzy oraz nie posiada ograniczenia ilości nukleotydów w pętli.[4]

Afold jest dostępny za darmo do pobrania z serwera ftp: <ftp://ftp.ncbi.nlm.nih.gov/pub/ogurtsov/Afold/>. Po przejściu na serwer otrzymujemy widok jak przedstawiony obok.

Mamy do wyboru pliki instalacyjne na system Windows i Linux. Jednakże w obecnej chwili pliki dedykowane dla Windows są niekompletne.



The screenshot shows a web browser window displaying the FTP directory listing for the Afold directory. The address bar shows the URL: ftp://ftp.ncbi.nlm.nih.gov/pub/ogurtsov/Afold/. The page title is "Indeks – /pub/ogurtsov/Afold/". The table lists the following files and directories:

Nazwa	Rozmiar	Data modyfikacji
[katalog główny]		
Linux/		28.01.2006, 00:00:00
RNA_Intern_Loops-Suppl.doc	516 kB	11.04.2006, 00:00:00
RNA_Intern_Loops-Suppl.pdf	360 kB	11.04.2006, 00:00:00
ReadMe.txt	3.4 kB	11.04.2006, 00:00:00
Windows/		04.01.2007, 00:00:00

Ryc.1. Zawartość serwera ftp.

Carnac

Głównym celem metody Carnac jest przewidywanie konserwatywnych elementów struktury drugorzędowej rodzin homologicznych niekodującego RNA.[38] W metodzie tej nie jest wymagane wcześniejsze dopasowanie sekwencji, dlatego idealnie nadaje się do danych posiadających niskie podobieństwo struktur pierwszorzędowych. Trzy podstawowe komponenty metody to: termodynamiczna minimalizacja energii, dopasowanie filogenetyczne oraz konserwacja sekwencji. Dodatkowo wykluczamy wyizolowane pary i pseudowęzły.

Algorytm składa się z trzech kroków:

Krok pierwszy to wyłonienie wszystkich potencjalnych łodyg o niskim poziomie energii swobodnej dla wszystkich sekwencji. Poszukiwanie jest realizowane za pomocą programowania dynamicznego.

Krok drugi polega na analizie wszystkich możliwych par sekwencji i tworzeniu struktury złożonej właśnie z danej pary. Aby dwie sekwencje uznać za parę muszą one spełnić warunki. W obu sekwencjach powinny znajdować się analogiczne dobrze zakonserwowane regiony oraz muszą zawierać co najmniej jedną mutację kompensacyjną.

Krok trzeci to tworzenie grafu macierzystego złożonego na podstawie relacji między najbardziej wiarygodnymi sekwencjami.

Istnieją dwa sposoby na użycie Carnac: pobranie i instalacja oprogramowania na własnym komputerze lub online na stronie: <http://bioinfo.lifl.fr/RNA/carnac/>. Obie wersje są darmowe, bez rejestracji. Dane wejściowe muszą być podane w formacie FASTA, natomiast wynik otrzymujemy w kilku formatach: Connest file (CT), PostScript oraz w formie graficznej. Każdemu procesowi zostaje nadany unikalny numer ID, pod którym zapisywane są operacje. Czas oczekiwania zależy od długości sekwencji. Jeśli długość jest mniejsza niż 300 nukleotydów to zajmuje to kilka sekund, dla większych danych czas wydłuża się do kilku minut.[6]

The screenshot shows the Carnac web interface. At the top, there's a navigation bar with links: home, web server, help, examples, and retrieve result with an ID. Below this is a sidebar with a list of supported data types: DNA (YASS, Magnolia, mrep, ProCARS), HTS (SortMeRNA, CRAC, Vidjil, StORM), RNA (Carnac, RNAfamily, Gardemia, Regliss, CG-seq, RNAspace), Proteins (Path, Protea, ReBLOSUM), and TFM (TFM-Explorer, TFM-Scan, TFM-Pvalue, TFM-CUDA). The main area contains a form with the following fields: 'Enter a name for the sequences (optional):', 'Paste your RNA sequences in FASTA format [?]', 'upload a file [Wybierz plik] Nie wybrano pliku', 'Example', 'Disable threshold correction according to GC% [?]', and 'Enter your E-mail address (optional):'. At the bottom of the form are 'Reset' and 'Run CARNAC' buttons. A footer note says 'For questions about carnac or for bug reports, please contact carnac' and the year '2013 - Bonsai bioinformatics'.

Ryc.2. Strona domyślna Carnac.

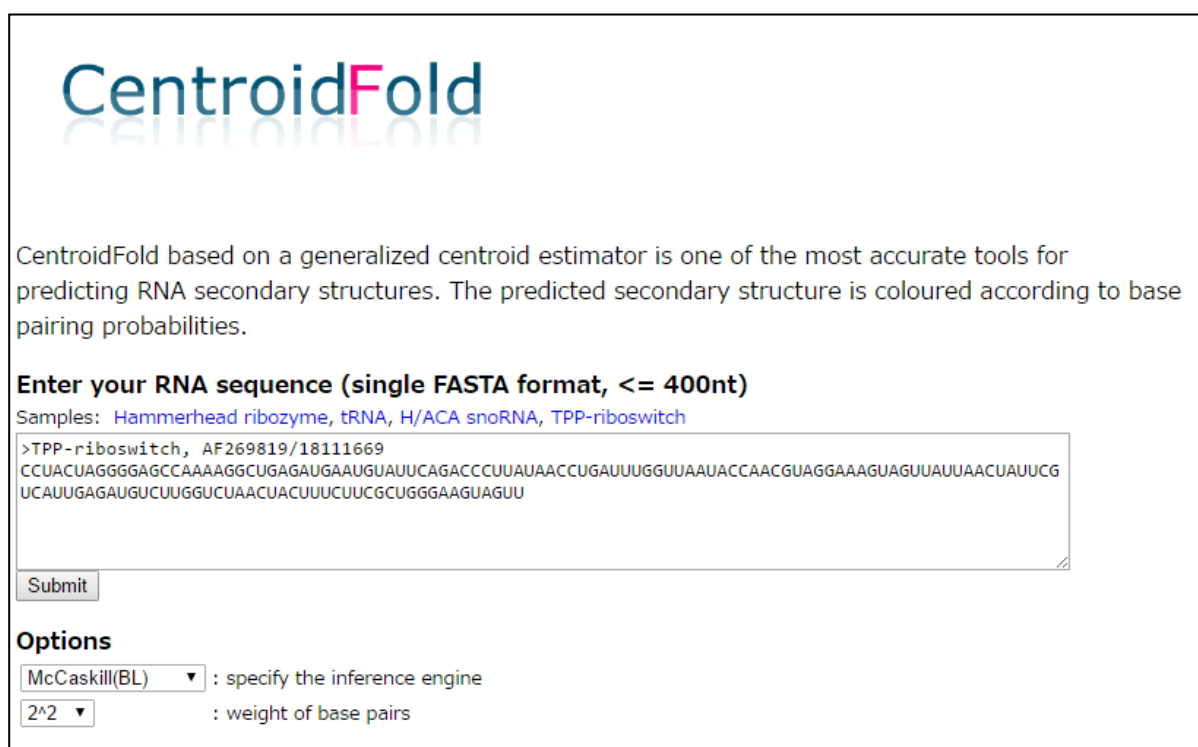
The screenshot shows the results page of the Carnac web interface. It features a navigation bar with links: home, web server, help, examples, and retrieve result with an ID. The main content area is divided into two columns. The left column lists the input data types: DNA (YASS, Magnolia, mrep, ProCARS), HTS (SortMeRNA, CRAC, Vidjil, StORM), RNA (Carnac, RNAfamily, Gardemia, Regliss, CG-seq, RNAspace), Proteins (Path, Protea, ReBLOSUM), and TFM (TFM-Explorer, TFM-Scan, TFM-Pvalue, TFM-CUDA). The right column displays the results for a specific job (job May_24_2016_21_20_00_31368). It includes a list of input sequences, a table of alignment results (showing sequence names and alignment scores), and a section for RNAfold structures. The RNAfold structures section shows the predicted secondary structure of the input sequences, represented by a diagram of a RNA molecule with stems and loops.

Ryc.3. Strona wynikowa Carnac.

CentroidFold

Narzędzie CentroidFold może być użyte do przewidywania struktury drugorzędowej RNA jednej sekwencji lub dopasowani wielu sekwencji RNA. Jest to możliwe dzięki opracowaniu algorytmu opartego na γ -centroid estymatorze. Estymator ten to pewnego rodzaju sposób dekodowania bazujący na teorii decyzji statystycznej.[7] Dzięki zastosowaniu estymatora metoda jest bardzo precyzyjna.[39]

CentroidFold dostępny jest wyłącznie online na stronie: <http://rtools.cbrc.jp/centroidfold/>. Jako dane wejściowe akceptowalne są dwa formaty: FASTA dla jednej sekwencji oraz format Clustal W dla dopasowani wielu sekwencji RNA. Nie ma potrzeby zaznaczania jaki format wybrano, ponieważ jest on automatycznie wykrywany.[7] Interfejs jest bardzo prosty i przejrzysty.



CentroidFold

CentroidFold based on a generalized centroid estimator is one of the most accurate tools for predicting RNA secondary structures. The predicted secondary structure is coloured according to base pairing probabilities.

Enter your RNA sequence (single FASTA format, <= 400nt)

Samples: [Hammerhead ribozyme](#), [tRNA](#), [H/ACA snoRNA](#), [TPP-riboswitch](#)

```
>TPP-riboswitch, AF269819/18111669
CCUACUAGGGGAGCCAAAAGGCUGAGAUGAAUGUAUUCAGACCCUUAUAACCUGAUUUGGUUAAUACCAACGUAGGAAAGUAGUUAUUAACUUAUUCG
UCAUUGAGAUGUCUUGGUCUAAACUACUUCUUCGCGGGGAAGUAGUU
```

Options

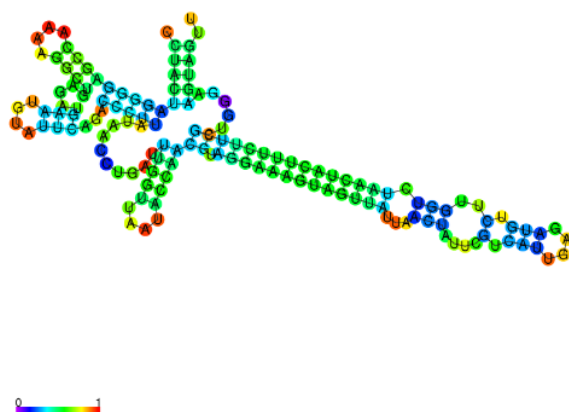
: specify the inference engine

: weight of base pairs

Ryc.4. Strona domyślna CentroidFold.

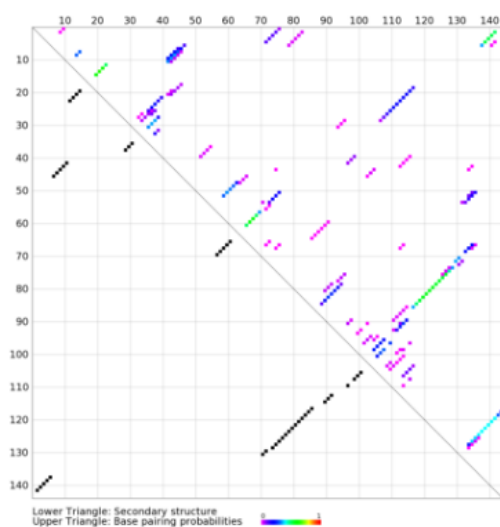
Wyniki otrzymujemy w kilku formatach, w tym także graficznych z odpowiednim kolorowaniem. Jak widać na poniższych ilustracjach mamy do wyboru formaty zapisu do pliku: TEXT, jako obrazek PNG, PDF lub z rozszerzeniem .eps.

Secondary structure



Download: [TEXT](#) [PNG](#) [PDF](#) [EPS](#)

Base pairing probability plot



Download: [TEXT](#) [PNG](#) [PDF](#) [EPS](#)

Download:TEXT (Secondary structure) is a dot-bracket format with a fasta-like header line, indicating a secondary structure.

```
>sequence_name
ACGCUGUCUGUACUUGUAUCAGUACACUGACGAGUCCCUAAAAGGACGAAACAGCGC
.(((((((.(((((((.....)))))).....((((.....)))).....)))))). (g=4,th=0.2,e=-9.14)
```

In this format, each dot represents an unpaired base, opening and closing brackets represent a base pair.

Download:TEXT (Base pairing probability plot) contains base-pairing probabilities more than "weight of base pairs" in blank delimited format.

```
pos1 nt1 pair-pos11:prob11 pair-pos12:prob12 ...
pos2 nt2 pair-pos21:prob21 pair-pos22:prob22 ...
:
```

pos1,pos2 are sequence positions in 1-based coordinate.

nt1,nt2 are nucleotides.

pair-pos11,pair-pos12 ... are sequence positions of pairing partners in 1-based coordinate.

pron11,prob12 ... are base pairing probabilities.

CentroidHomfold

Narzędzie pochodzi z tego samego serwera co CentroidFold.

- strona internetowa: <http://rtools.cbrc.jp/centroidhomfold/>
- metoda automatycznie pobiera sekwencje homologiczne do sekwencji docelowej.
- sekwencje homologiczne pobierane są z bazy Rfam,
- pobieranie sekwencji homologicznych daje CentroidHomfold przewagę nad CentroidFold w dokładności przewidywania struktury.

CentroidHomfold

CentroidHomfold predicts RNA secondary structures by employing automatically collected homologous sequences of the target. Homologous sequences are collected from Rfam using LAST. If homologous sequences are available, CentroidHomfold can predict secondary structures for the target sequence more accurately than CentroidFold using homologous sequence information with the probabilistic consistency transformation for base-pairing probabilities.

Enter your RNA sequence (single FASTA format, <= 400nt)

Samples: [Hammerhead ribozyme](#), [tRNA](#), [H/ACA snoRNA](#), [TPP-riboswitch](#)

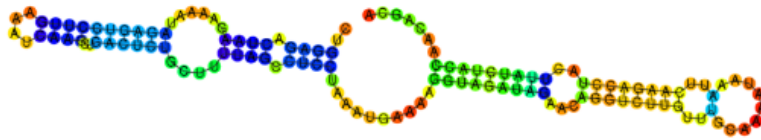
>A7609485.1/1-127
CUGGAGACUAAGAAAUAAGAGUCCUUGAAAUAAGCUGACUCUGCUUUUAGCCUCCUAAAUGAAAAGGUAGAUAGAACAGGUCUUGUUUGCAAAUA
AAUUAAGACCUACUUAUCUACCAACAGCA

Submit

Options

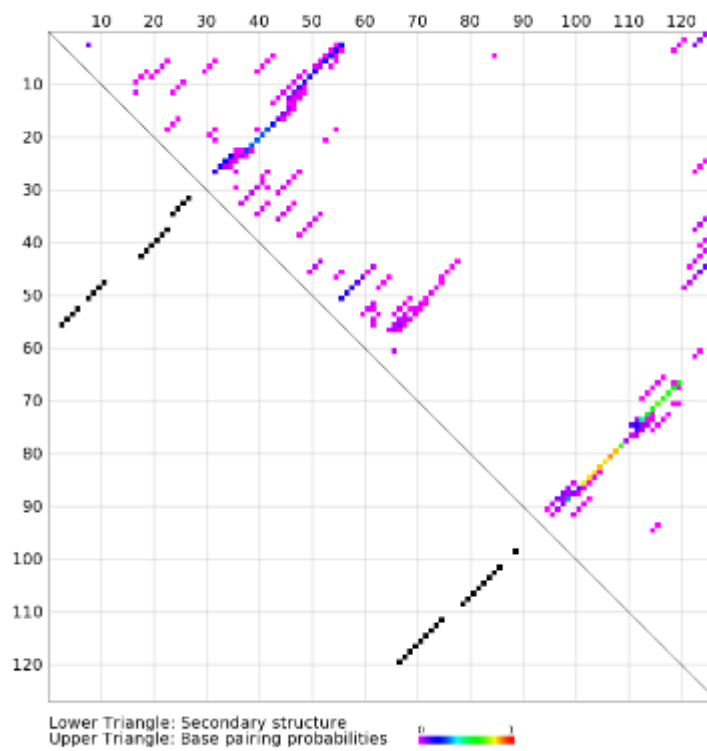
- | | |
|----------------|--|
| McCaskill(BL) | : specify the inference engine for secondary structure |
| CONTRAlign | : specify the inference engine for pairwise alignment |
| 2 ³ | : weight of base pairs |
| 0.01 | : E-value for homology search against Rfam |
| 30 | : number of homologous sequences |

Secondary structure



Download: [TEXT](#) [PNG](#) [PDF](#) [EPS](#)

Base pairing probability plot



Download: [TEXT](#) [PNG](#) [PDF](#) [EPS](#)

ContraFold

- strona internetowa: <http://contra.stanford.edu/contrafold/>
- zaimplementowany w C++,
- metoda oparta warunkowych modelach logarytmiczno-liniowych, probabilistycznych modelach uogólniających stochastyczną gramatykę bezkontekstową (SCFGs) poprzez funkcje wyższych punktacji,
- metoda osiąga wysoką dokładność predykcji pojedynczego ciągu,
- metoda łączy techniki probabilistyczne z fizycznymi. Jest to odpowiedź na lukę pomiędzy metodami probabilistycznymi a termodynamicznymi udowadniając, że procedury uczenia statystycznego mogą stanowić skuteczną alternatywę dla empirycznego pomiaru parametrów termodynamicznych do przewidywania struktury drugorzędowej RNA,
- wyniki można przesłać sobie bezpośrednio na wcześniej podanego maila lub otworzyć je w przeglądarce od razu po wykonaniu,
- mamy dwie opcje do wyboru pod względem parowania się nukleotydów. Możemy wybrać opcję domyślną typu Watsona-Cricka lub dopuścić parowanie się wszystkich możliwych.

DotKnot

- strona internetowa: <http://dotknot.csse.uwa.edu.au/>
- metoda heurystyczna,

HotKnots

- strona internetowa: <http://www.cs.ubc.ca/labs/beta/Software/HotKnots/>
- metoda heurystyczna,
- zaimplementowana w 2005 przez Jihong Ren,
-

TurboKnot

- strona internetowa: <http://rna.urmc.rochester.edu/>
- dostępny jako część RNAstructure

CMfinder

- strona internetowa:
<http://bio.cs.washington.edu/CMfinderWeb/CMfinderInput.pl>

CRWrnafold

CyloFold

Mastr

MaxExpect

Multialign

PETfold

- strona internetowa: <http://rth.dk/resources/petfold/submit.php>

PPfold

- strona internetowa: <http://www.daimi.au.dk/~compbio/pfold/downloads.html>

ProbKnot

RNAalifold

RNA Sampler

- strona internetowa: <http://stormo.wustl.edu/RNASampler/index.html>

RNAshapes

RNASLOpt

RNAwolf

- strona internetowa: <http://www.tbi.univie.ac.at/software/rnawolf/rnawolf.html>
- strona internetowa: <http://hackage.haskell.org/package/RNAwolf>

Sfold

- strona internetowa: <http://sfold.wadsworth.org/cgi-bin/index.pl>

taveRNA

TurboFold

Opcje narzędzi

Nazwa metody	Dostępność	Dostęp	Systemy operacyjne, na których działa	Input	Output	Struktury	Język	Ograniczenia
Afold	darmowe oprogramowanie dostępne dla wszystkich	przez serwer ftp	Linux, Windows	FASTA lub format Genbank	plik „Afold.res”	pętle wewnętrzne	C/C++	do 28 000 nt, min. 2GB RAM
Carnac	darmowe oprogramowanie dostępne dla wszystkich	online (web server) lub lokalnie na własnym komputerze	Linux, Windows	FASTA	CT, PostScript, JPEG, dot-bracket, list of stems, archiwum .zip	elementy struktury drugorzędowej rodzin homologicznych niekodującego RNA		do 300 nt (kilka s) do 2000 nt (kilka min)
CentroidFold	darmowe oprogramowanie dostępne dla wszystkich	online	-	FASTA	TEXT, PNG, PDF, EPS	wszystkie elementy struktury drugorzędowej		do 400 nt
CentroidHomfold	darmowe oprogramowanie dostępne dla wszystkich	online	-	FASTA	TEXT, PNG, PDF, EPS	wszystkie elementy struktury drugorzędowej		do 400 nt
ContraFold	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Linux	FASTA	BPSEQ		C++	do 1000 nt

Nazwa metody	Dostępność	Dostęp	Systemy operacyjne, na których działa	Input	Output	Struktury	Język	Ograniczenia
DotKnot	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Windows, Linux, Mac OS-X	sekwencja RNA w formacie FASTA	dot-bracket	pseudowęzły	Python	do 1000 nt dla pseudowęzłów typu H do 400 nt dla kissing harpins
HotKnots	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Windows, Linux, Mac OS-X	sekwencja RNA w formacie FASTA lub tekstowym	dot-bracket	pseudowęzły	C/C++	do 500 nt lokalnie do 100 nt online
TurboKnot	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Windows, Linux, Mac OS-X	FASTA		pseudowęzły	C/C++, Java	
CMFinder	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Windows, Linux, Mac OS-X	FASTA	archiwum .zip na podany adres mailowy	motywy		do 60 sekwencji po 500 nt
CRWrnaFold	darmowe oprogramowanie dostępne dla wszystkich	online	-	wybór konkretnych informacji	archiwum .zip	spinki, pętle wewnętrzne	C#	brak
CyloFold	darmowe oprogramowanie dostępne dla wszystkich	online	-	FASTA lub Raw Sequence	JPG, dot-bracket, CT	pseudowęzły	C++, Groovy	do 300 nt
Mastr	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Linux	FASTA	PostScript	odcinki non-coding RNAs	C++	od 70 do 250 nt

Nazwa metody	Dostępność	Dostęp	Systemy operacyjne, na których działa	Input	Output	Struktury	Język	Ograniczenia
MaxExpect	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Linux, Windows	FASTA	FASTA, SVG, JPEG, PDF, PostScript, CT	wszystkie elementy struktury drugorzędowej		do 3000 nt
Multialign	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Linux, Windows	FASTA	SVG, JPEG, PDF, PostScript, CT	wspólna struktura drugorzędowa o najniższej energii dla kilku sekwencji	C++	powyżej 1500 nt
PETfold	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	dowolny	FASTA, Newick, notacja dot-bracket	PostScript, PDF, Newick, dotplot file	wszystkie elementy struktury drugorzędowej	Perl	
PPfold	darmowe oprogramowanie dostępne dla wszystkich	lokalnie na własnym komputerze	Linux	FASTA, Newick	CT, pik .seq	wszystkie elementy struktury drugorzędowej	Java 6.0	do ok. 9000 nt
ProbKnot	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Windows, Linux, Mac OS-X	FASTA	SVG, JPEG, PDF, PostScript, CT	pseudowęzły	C++, Java	do ok 3000 nt
RNAalifold	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Linux	FASTA lub Clustal W	Vienna Format, CT, PDF, EPS, JPG	wszystkie elementy struktury drugorzędowej		
RNA Sampler	darmowe oprogramowanie dostępne dla wszystkich	lokalnie na własnym komputerze	Linux, Windows	FASTA	własny format	wszystkie elementy struktury drugorzędowej	C	

Nazwa metody	Dostępność	Dostęp	Systemy operacyjne, na których działa	Input	Output	Struktury	Język	Ograniczenia
RNAshapes	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Linux, Windows	FASTA	dot-bracket	wszystkie elementy struktury drugorzędowej	C	do 1500 nt
RNASLOpt	darmowe oprogramowanie dostępne dla wszystkich	lokalnie na własnym komputerze	Linux	sekwencja RNA	optymalne reprezentacje w notacji dot-bracket wraz z ich energiami	wszystkie elementy struktury drugorzędowej		do 250 nt
RNAwolf	darmowe oprogramowanie dostępne dla wszystkich	lokalnie na własnym komputerze	Linux	sekwencja RNA		triplety i niekanoniczne pary zasad	Haskell	
Sfold	darmowe oprogramowanie dostępne dla wszystkich	online	-	FASTA, GenBank	PostScript, PDF, TXT, archiwum .zip	wszystkie elementy struktury drugorzędowej		do 200 nt online, do 5000 nt lokalnie
taveRNA	darmowe oprogramowanie dostępne dla wszystkich	online lub niektóre z narzędzi możliwe do pobrania	Linux	FASTA	CT, PostScript, GIF	wszystkie elementy struktury drugorzędowej	C/C++	do 150 nt dla Stacked Pair Model, do 500 nt dla Loop Models
TurboFold	darmowe oprogramowanie dostępne dla wszystkich	online lub lokalnie na własnym komputerze	Linux, Windows	FASTA	SVG, JPEG, PDF, PostScript, CT	wszystkie elementy struktury drugorzędowej		

Podsumowanie

Bibliografia

- [1] Lubert Stryer *Biochemia* Wydawnictwo Naukowe PWN (1999)
- [2] Suzanne Clancy, Ph.D. *Chemical Structure of RNA* Nature Education (2008)
- [3] Ronny Lorenz, Michael T. Wolfinger, Andrea Tanzer, Ivo L. Hofacker
Predicting RNA secondary structures from sequence and probing data Elsevier (2016)
- [4] Aleksey Y. Ogurtsov, Svetlana A. Shabalina, Alexey S. Kondrashov, Mikhail A. Roytberg
Analysis of internal loops within the RNA secondary structure in almost quadratic time Bioinformatics (2006)
- [5] Woods, D.A., Batzoglou, S.
CONTRAFold: RNA Secondary Structure Prediction without Energy-Based Models. Bioinformatics (2006)
- [6] Helene Touzet, Olivier Perriquet
CARNAC: folding families of related RNAs (2004)
- [7] Kengo Sato, Michiaki Hamada, Kiyoshi Asai, Toutai Mituyama
CENTROIDFOLD: a web server for RNA secondary structure prediction (2009)
- [8] Michiaki Hamada, Koichiro Yamada, Kengo Sato, Martin C. Frith, Kiyoshi Asai
CentroidHomfold-LAST: accurate prediction of RNA secondary structure using automatically collected homologous sequences (2011)
- [9] Zizhen Yao, Zasha Weinberg, Walter L. Ruzzo
CMfinder - a covariance model based RNA motif finding algorithm Bioinformatics (2005)
- [10] David P. Gardner, Pengyu Ren, Stuart Ozer, Robin R. Gutell
Statistical Potentials for Hairpin and Internal Loops Improve the Accuracy of the Predicted RNA Structure (2011)
- [11] Eckart Bindewald¹, Tanner Kluth, Bruce A. Shapiro
CyloFold: secondary structure prediction including pseudoknots (2010)
- [12] Kengo Sato, Yuki Kato, Tatsuya Akutsu, Kiyoshi Asai, Yasubumi Sakakibara
DAFS: simultaneous aligning and folding of RNA sequences via dual decomposition Bioinformatics (2012)
- [13] Jana Sperschneider, Amitava Datta
DotKnot: pseudoknot prediction using the probability dot plot under a refined energy model (2010)
- [14] Jihong Ren, Baharak Rastegardi, Anne Condon, Holger H. Hoos

HotKnots: Heuristic prediction of RNA secondary structures including pseudoknots
Bioinformatics (2005)

[15] Kengo Sato, Yuki Kato, Michiaki Hamada, Tatsuya Akutsu, Kiyoshi Asai
IPknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming Bioinformatics

[16] Stinus Lindgreen, Paul P. Gardner, Anders Krogh
MASTR: multiple alignment and structure prediction of non-coding RNAs using simulated annealing Bioinformatics (2007)

[17] Zhi John Lu, Jason W. Gloor, David H. Mathews
Improved RNA secondary structure prediction by maximizing expected pair accuracy
Bioinformatics (2009)

[18] Zhenjiang Xu, David H. Mathews
Multalign: an algorithm to predict secondary structures conserved in multiple RNA sequences Bioinformatics (2010)

[19] Hisanori Kiryu, Yasuo Tabei, Taishin Kin, Kiyoshi Asai
Murlet: a practical multiple alignment tool for structural RNA sequences
Bioinformatics (2007)

[20] Yasuo Tabei, Kiyoshi Asai
A local multiple alignment method for detection of non-coding RNA sequences
Bioinformatics (2009)

[21] Stefan E. Seemann, Jan Gorodkin, Rolf Backofen
Unifying evolutionary and thermodynamic information for RNA folding of multiple alignments (2008)

[22] Jens Reeder, Peter Steffen, Robert Giegerich
pknotsRG: RNA pseudoknot folding including near-optimal structures and sliding windows (2007)

[23] Zsuzsanna Sükösd, Bjarne Knudsen, Morten Værum, Jørgen Kjems, Ebbe S Andersen
Multithreaded comparative RNA secondary structure prediction using stochastic context-free grammars BMC Bioinformatics (2011)

[24] Stanislav Bellaousov, David H. Mathews
ProbKnot: Fast prediction of RNA secondary structure including pseudoknots
Bioinformatics (2010)

[25] Xiaomin Ying, Hong Luo, Jingchu Luo, Wuju Li
RDfolder: a web server for prediction of RNA secondary structure (2004)

[26] Stephan H Bernhart, Ivo L Hofacker, Sebastian Will, Andreas R Gruber, Peter F Stadler
RNAalifold: improved consensus structure prediction for RNA alignments

BMC Bioinformatics (2008)

[27] Xing Xu, Yongmei Ji, Gary D. Stormo
RNA Sampler: a new sampling based algorithm for common RNA secondary structure prediction and structural alignment Bioinformatics (2007)

[28] Peter Steffen, Björn Voß, Marc Rehmsmeier, Jens Reeder, Robert Giegerich
RNAshapes: an integrated RNA analysis package based on abstract shapes Bioinformatics (2005)

[29] Yuan Li, Shaojie Zhang
Finding stable local optimal RNA secondary structures Bioinformatics (2011)

[30] Jessica S Reuter, David H Mathews
RNAstructure: software for RNA secondary structure prediction and analysis BMC Bioinformatics (2010)

[31] Christian Höner zu Siederdissen, Stephan H. Bernhart, Peter F. Stadler Ivo L. Hofacker
A folding algorithm for extended RNA secondary structures Bioinformatics

[32] Junilda Spirollari, Jason T.L. Wang, Kaizhong Zhang, Vivian Bellofatto2, Yongkyu Park, Bruce A. Shapiro
Predicting Consensus Structures for RNA Alignments Via Pseudo-Energy Minimization (2009)

[33] Ye Ding, Chi Yu Chan, Charles E. Lawrence
Sfold web server for statistical folding and rational design of nucleic acids (2004)

[34] Cagri Aksay, Raheleh Salari, Emre Karakoc, Can Alkan, S. Cenk Sahinalp
taveRNA: a web suite for RNA algorithms and applications (2007)

[35] Arif O Harmanci, Gaurav Sharma, David H Mathews
TurboFold: Iterative probabilistic estimation of secondary structures for multiple RNA sequences BMC Bioinformatics (2011)

[36] Matthew G. Seetin, David H. Mathews
TurboKnot: rapid prediction of conserved RNA secondary structures including pseudoknots Bioinformatics (2012)

[37] <http://iimcb.genesilico.pl/comparna/methods/>

[38] <http://bioinfo.lifl.fr/RNA/carnac/>

[39] <http://rtools.cbrc.jp/centroidfold/>

[40] <http://rtools.cbrc.jp/centroidhomfold/>