Theoretical and practical metagenomic approaches to viral discovery

Practical Session: ViennaRNA for RNA-RNA Interactions

Kevin Lamkiewicz, Manja Marz

24.10.2019

European Virus Bioinformatics Center

Why RNA-RNA Interactions?

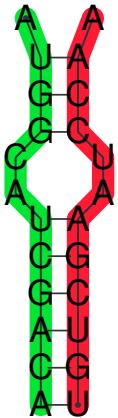
ToDo: Picture here.

Interactions for single sequences

RNAcofold

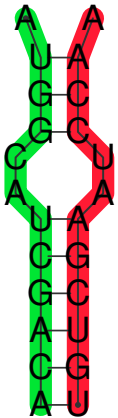
```
# RNAcofold works like RNAfold, but allows to specify two RNA sequences.
# These sequences are then allowed to form a dimer structure. In order
# to calculate the hybrid structure, it is necessary to concatenate the
# two RNA sequence, using & as a separator.

$> RNAcofold [OPTIONS] < sequences.fasta > sequences.cofold

# >seq1
# AUGGCAUCGACA
# >seq2
# UGUCGAAUCCAA

# RNAcofold Input:
# AUGGCAUCGACA&UGUCGAAUCCAA
```

RNAcofold



- First sequence is colored green
- Second sequence is colored red

RNAcofold



- First sequence is colored green
- Second sequence is colored red

- What happens, when there are more than 2 input sequences?

RNACOFOLD WITH PARTITION FUNCTION

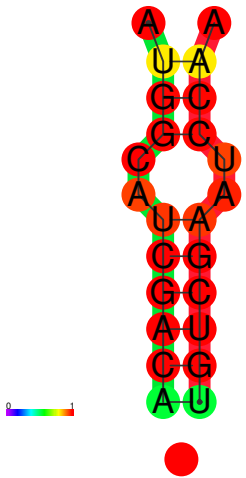
```
# We calculated the MFE structure of the interacting molecules (RNA dimer).
# RNAcofold also has the -p parameter implemented.

$> RNAcofold -p < sequences.fasta > sequences.cofold
```

RNAcofold with partition function

```
# We calculated the MFE structure of the interacting molecules (RNA dimer).
# RNAcofold also has the -p parameter implemented.

$> RNAcofold -p < sequences.fasta > sequences.cofold
```

We can use `relplot.pl` on the PostScript files as well, but it looks a bit... weird.

RNACOFOLD WITH PARTITION FUNCTION



AA, AB, BB

```
# Until now, we just looked at the heterodimer of the two sequences.
# But how do the molecules behave individually?

$> RNAcofold -a < sequences.fasta > sequences.cofold
```

AA, AB, BB

```
# Until now, we just looked at the heterodimer of the two sequences.
# But how do the molecules behave individually?

$> RNAcofold -a < sequences.fasta > sequences.cofold
```

The AA and BB dimer describe the MFE structure of two RNA molecules of sequence one and sequence two, respectively.

RNADUPLEX

```
# RNAduplex is very similar to RNAcofold. Actually,
# it is a special case of RNAcofold, where only inter-molecular
# base pairs are allowed.

$> RNAduplex [OPTIONS] < sequences.fasta > sequences.duplex

```

RNAduplex

```
# RNAduplex is very similar to RNAcofold. Actually,
# it is a special case of RNAcofold, where only inter-molecular
# base pairs are allowed.

$> RNAduplex [OPTIONS] < sequences.fasta > sequences.duplex

# Alternative:
# RNAcofold -C < sequences_constrained.fasta
# sequences_constrained.fasta
# UAGCUAGCAUGCAUCGACGAU&CGAUGCAUGCAUGCAUGCAUC
# <<<<<<<<<<<<<<<<<<<<&>>>>>>>>>>>>>>>>>>>>
```

Co-Folding with MSAs

RNAALIDUPLEX

RNAALIDUPLEX

Not much implemented...

Unfortunately, ViennaRNA does not provide many possibilities for alignment-based co-folding analyses. Indeed, only the alignment version of `RNAduplex` is implemented in `RNAaliduplex`.

RNAaliduplex

Not much implemented...

Unfortunately, ViennaRNA does not provide many possibilities for alignment-based co-folding analyses. Indeed, only the alignment version of `RNAduplex` is implemented in `RNAaliduplex`.

```
RNAaliduplex [OPTIONS] <file1.aln> <file2.aln>

# RNAaliduplex expects two input files (both CLUSTAL alignments)
# and predicts optimal and suboptimal binding sites.
# However, only inter-molecular base pairs are taken into account.
```

ALIGNMENT-BASED INTRA-MOLECULAR BASE PAIRS?

ALIGNMENT-BASED INTRA-MOLECULAR BASE PAIRS?

What would you do?

Discuss, play around, try to make some examples - I will go around and answer questions, discuss your ideas and help you as good as I can.

ALIGNMENT-BASED INTRA-MOLECULAR BASE PAIRS!

ALIGNMENT-BASED INTRA-MOLECULAR BASE PAIRS!

Different ways to do it

Most commonly, you'd want to do the following:

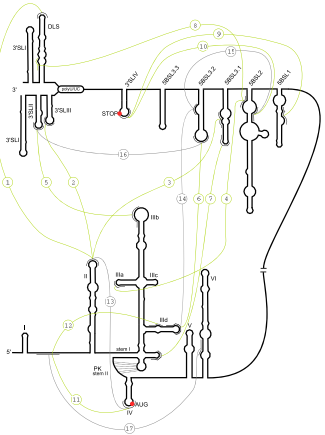
1. Extract your sequences and align them individually
2. Merge the alignments, use `'NNNNN'` as a separator
3. Apply `RNAalifold` on the alignment

Long-Range Interactions

Why RNA-RNA Interactions?

ToDo: Picture here.

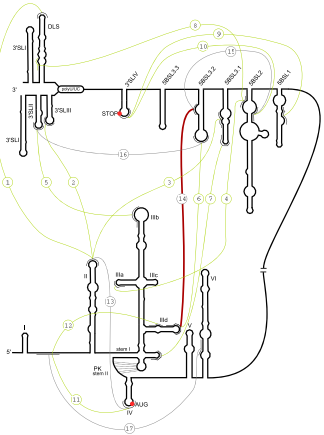
RNA-RNA INTERACTIONS ARE CRUCIAL FOR RNA VIRUSES



Fricke, M. *et al.* (2015). Conserved RNA secondary structures and long-range interactions in hepatitis C viruses. RNA, http://doi.org/10.1261/rna.049338.114

- Hepatitis C virus: (+)ssRNA
- around 9 kb in size

RNA-RNA INTERACTIONS ARE CRUCIAL FOR RNA VIRUSES

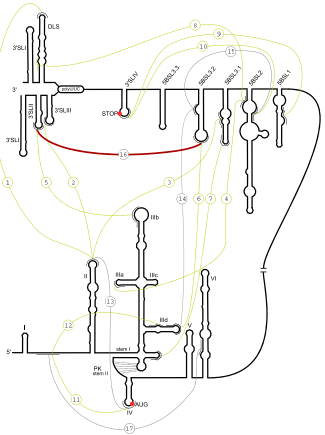


Fricke, M. *et al.* (2015). Conserved RNA secondary structures and long-range interactions in hepatitis C viruses. RNA, http://doi.org/10.1261/rna.049338.114

- Hepatitis C virus: (+)ssRNA
- around 9 kb in size

- Initiation of Translation

RNA-RNA INTERACTIONS ARE CRUCIAL FOR RNA VIRUSES



Fricke, M. *et al.* (2015). Conserved RNA secondary structures and long-range interactions in hepatitis C viruses. RNA, http://doi.org/10.1261/rna.049338.114

- Hepatitis C virus: (+)ssRNA
- around 9 kb in size

- Initiation of Translation
- Initiation of Replication

Exercise:

Take any LRI from HCV, described in the following paper, and try to reconstruct / predict it with the ViennaRNA package.

Fricke, M. *et al.* (2015). Conserved RNA secondary structures and long-range interactions in hepatitis C viruses. RNA, http://doi.org/10.1261/rna.049338.114