SilentMutations (SIM): a tool for analyzing long-range RNA-RNA interactions in viral genomes and structured RNAs

Daniel Desirò

RNA Bioinformatics and High-Throughput Analysis, Jena
Friedrich Schiller University, Jena
European Virus Bioinformatics Center, Jena

daniel.desiro@uni-jena.de

February 11, 2019
Long-range RNA-RNA Interactions

Translation Initiation

- 3’cap-independent translational enhancers
- internal ribosome entry sites

Long-range RNA-RNA Interactions

Translation Initiation
- 3’cap-independent translational enhancers
- internal ribosome entry sites

Translational Recoding
- ribosomal frameshifting
- stop codon readthrough

Long-range RNA-RNA Interactions

Translation Initiation
- 3’cap-independent translational enhancers
- internal ribosome entry sites

Translational Recoding
- ribosomal frameshifting
- stop codon readthrough

Genome Replication
- autonomous
- protein factors

Long-range RNA-RNA Interactions

Translation Initiation
- 3’cap-independent translational enhancers
- internal ribosome entry sites

Translational Recoding
- ribosomal frameshifting
- stop codon readthrough

Genome Replication
- autonomous
- protein factors

sgmRNA Transcription
- discontinuous template synthesis
- premature RdRp termination

Long-range RNA-RNA Interactions

Translation Initiation
- 3’cap-independent translational enhancers
- internal ribosome entry sites

Translational Recoding
- ribosomal frameshifting
- stop codon readthrough

Genome Replication
- autonomous
- protein factors

sgmRNA Transcription
- discontinuous template synthesis
- premature RdRp termination

Conserved LRIs in Hepatitis C Viruses

HCV Genome
- ssRNA+ ~10kb
- single polyprotein
- UTRs highly structured

[Fricke et al., 2015, Conserved RNA secondary structures and long-range interactions in hepatitis C viruses.]
Conserved LRIs in Hepatitis C Viruses

HCV Genome
- ssRNA+ ~10kb
- single polyprotein
- UTRs highly structured

Replication
- 3’ X-tail structure
- LRI between 3’SLII and 5BSL3.2

[Fricke et al., 2015, Conserved RNA secondary structures and long-range interactions in hepatitis C viruses.]
[Friebe et al., 2004, Kissing-loop interaction in the 3’ end of the hepatitis C virus genome essential for RNA replication]
Conserved LRIs in Hepatitis C Viruses

HCV Genome
- ssRNA+ ~10kb
- single polyprotein
- UTRs highly structured

Replication
- 3’ X-tail structure
- LRI between 3’SLII and 5BSL3.2

LRIscan
- detected multiple LRIs
- verify several known
- identify several novel possible

[Fricke et al., 2015, Conserved RNA secondary structures and long-range interactions in hepatitis C viruses.]
[Friebe et al., 2004, Kissing-loop interaction in the 3’ end of the hepatitis C virus genome essential for RNA replication]
Influenza A Virus (IAV)

- polymerase basic proteins (PB2,PB1)
- polymerase acidic protein (PA)
- nucleoproteins (NP) + vRNA
- nuclear export proteins (NS)
- matrix protein (M) + lipid bilayer
- neuraminidase (NA) + hemaglutin (HA)

[Eisfeld et al., 2015, At the centre: influenza A virus ribonucleoproteins.]
Influenza A Virus (IAV)

- polymerase basic proteins (PB2, PB1)
- polymerase acidic protein (PA)
- nucleoproteins (NP) + vRNA
- nuclear export proteins (NS)
- matrix protein (M) + lipid bilayer
- neuraminidase (NA) + hemaglutin (HA)

[Eisfeld et al., 2015, At the centre: influenza A virus ribonucleoproteins.]
Verification Experiment

Interaction

- avian H5N2
- 4 silent, trans-complementary mutations
- PB1 and NS
- reassort wild-type and mutant

[Gavazzi et al., 2013, A functional sequence-specific interaction between influenza a virus genomic RNA segments.]
Verification Experiment

Interaction

- avian H5N2
- 4 silent, trans-complementary mutations
- PB1 and NS
- reassort wild-type and mutant

LRIscan

-26.70 kcal/mol

[Gavazzi et al., 2013, A functional sequence-specific interaction between influenza a virus genomic RNA segments.]
Verification Experiment

Interaction
- avian H5N2
- 4 silent, trans-complementary mutations
- PB1 and NS
- reassort wild-type and mutant

LRIscan

-26.70 kcal/mol

-14.20 kcal/mol

[Gavazzi et al., 2013, A functional sequence-specific interaction between influenza a virus genomic RNA segments.]
Verification Experiment

Interaction
- avian H5N2
- 4 silent, trans-complementary mutations
- PB1 and NS
- reassort wild-type and mutant

LRIscan

\[ \begin{array}{c}
\text{257} \\
\text{277} \\
\text{289} \\
\text{309} \\
\end{array} \]

-26.70 kcal/mol
-14.20 kcal/mol
-16.10 kcal/mol

[Gavazzi et al., 2013, A functional sequence-specific interaction between influenza a virus genomic RNA segments.]
Verification Experiment

Interaction

- avian H5N2
- 4 silent, trans-complementary mutations
- PB1 and NS
- reassort wild-type and mutant

LRIscan

\[ \begin{align*}
-26.70 \text{ kcal/mol} \\
-28.50 \text{ kcal/mol} \\
-14.20 \text{ kcal/mol} \\
-16.10 \text{ kcal/mol}
\end{align*} \]

[Gavazzi et al., 2013, A functional sequence-specific interaction between influenza a virus genomic RNA segments.]
Verification Experiment

Interaction

- avian H5N2
- 4 silent, trans-complementary mutations
- PB1 and NS
- reassort wild-type and mutant

LRIscan

[Gavazzi et al., 2013, A functional sequence-specific interaction between influenza a virus genomic RNA segments.]
SIM Workflow

(a) preprocessing

seq1:2:29-45
5' .......................... CUA UUC CGG AGG ACC AGC ACU CAG AGC GAC GUG UUU GGU UUU GGU UUU GGU UUU 3'
seq2:1:21-42
5' .......................... CUA UUC UUC AAG GAG ACC UGC UUC CAA AAG ACA UAG UGU UAG 3'

pre-snip1
5' .......................... CCA UUC ACU ACU CAA UUCUGC CAU AGCGAA AAGUUC GGA UUC 3'
pre-snip2
3' .......................... UUG UUCG GAA AAG

snip1
5' .......................... ACU GED AGC GAG GGC GAG UUU 3'
3' .......................... CAA UUC UGCCAU AGC GAA AAG

[ Lorenz et al., 2011, ViennaRNA package 2.0. ]
(b) permutation

codons1 (reverse complement)

pre-perms1 (P*)

pre-perms2 (Q*)

perms1 (P): 2684

perms2 (Q): 1446

pre-perms1 (P*): 10368

pre-perms2 (Q*): 2304

codons2 (reverse complement)

[Lorenz et al., 2011, ViennaRNA package 2.0.]
(c) attenuation

\[ l = \left( \text{mfe}(\text{snip}1, \text{mut}2) + \text{mfe}(\text{mut}1, \text{snip}2) \right) \times 0.5 \times (1 - \text{mrg}) \]

\[ u = \left( \text{mfe}(\text{snip}1, \text{mut}2) + \text{mfe}(\text{mut}1, \text{snip}2) \right) \times 0.5 \times (1 + \text{mrg}) \]

[Lorenz et al., 2011, ViennaRNA package 2.0.]
(d) recovery

\begin{align*}
\text{snip1} & \quad \cdot \text{ldv} \quad \geq \\
\text{mut1}_i (\in U) & \\
\text{snip2} & \\
\text{mut2}_j (\in V) & \\
\text{snip1} & \quad \cdot \text{udv} \\
\end{align*}

\begin{align*}
\text{snip1} & \\
\text{mut1}_i (\in U) & \\
\text{snip2} & \\
\text{mut2}_j (\in V) & \\
\text{snip1} & \quad \cdot \text{udv} \\
\end{align*}

(Lorenz et al., 2011, ViennaRNA package 2.0.)
SIM Workflow

(e) sampling

\[
\begin{align*}
\text{l-mrg} &= (mfe(\text{snip1}, \text{mut2}_i) + mfe(\text{mut1}_i, \text{snip2})) \cdot 0.5 \cdot (1 - \text{mrg}) \\
\text{u-mrg} &= (mfe(\text{snip1}, \text{mut2}_i) + mfe(\text{mut1}_i, \text{snip2})) \cdot 0.5 \cdot (1 + \text{mrg})
\end{align*}
\]

\[\text{snip1} = 5' \text{AAC CUU CGU UAC CGA} 3' \]
\[\text{snip2} = 5' \text{AAC CUU CGU UAC CGA} 3' \]

\[\text{mut1}_i, (\in \text{M}) \]

\[\text{mut2}_i, (\in \text{M}) \]

\[\text{Lorenz et al., 2011, ViennaRNA package 2.0.} \]
HCV Result

mfe(5BSL3.2_{WT}, SLII_{WT}) = -9.70 kcal/mol
mfe(5BSL3.2_{mut}, SLIimut) = -9.90 kcal/mol
mfe(5BSL3.2_{WT}, SLIImut) = -2.40 kcal/mol
mfe(5BSL3.2_{mut}, SLII_{WT}) = -2.10 kcal/mol

[Darty et al., 2009, VARNA: Interactive drawing and editing of the RNA secondary structure..]
mfe(NS\textsubscript{wt}, PB1\textsubscript{wt}) = -24.60 kcal/mol

mfe(NS\textsubscript{mut}, PB1\textsubscript{mut}) = -26.40 kcal/mol

mfe(NS\textsubscript{mut}, PB1\textsubscript{mut}) = -23.40 kcal/mol

[Darty et al., 2009, VARNA: Interactive drawing and editing of the RNA secondary structure.]

Daniel Desirò

SIM: analyzing LRIs
mfe(NS_{mut}, PB1_{WT}) = -11.80 \text{ kcal/mol}

mfe(NS_{WT}, PB1_{mut}) = -14.00 \text{ kcal/mol}

mfe(NS_{WT}, PB1_{mut}) = -10.10 \text{ kcal/mol}

mfe(NS_{mut}, PB1_{WT}) = -11.30 \text{ kcal/mol}

[Darty et al., 2009, VARNA: Interactive drawing and editing of the RNA secondary structure..]
Discussion

Dependencies
- Python (v3.6.5)
- ViennaRNA Package (v2.4)
- VARNA (v3.93) (optional)

Run Time
- worst case: $O(c^3 + l^3)$
- highly dependent on filter steps
  - Example: $10368 \cdot 2304$ to $1029 \cdot 36$

Summary
- 4 way verification experiment
  - ssRNA+ and ssRNA–
    - coding and non-coding

Application
- predict LRIs (LRIscan)
- simulate mutations (SIM)
- create mutants (Lab)

[Fricke and Marz, 2016, Prediction of conserved long-range RNA-RNA interactions in full viral genomes.]
Discussion

Dependencies
- Python (v3.6.5)
- ViennaRNA Package (v2.4)
- VARNA (v3.93) (optional)

Run Time
- worst case: $O\left(c^{\frac{k}{3}} + \frac{l}{3}\right)$
- highly dependent on filter steps
- Example: 10368 \cdot 2304 to 1029 \cdot 36

[Fricke and Marz, 2016, Prediction of conserved long-range RNA-RNA interactions in full viral genomes.]
Discussion

Dependencies

- Python (v3.6.5)
- ViennaRNA Package (v2.4)
- VARNA (v3.93) (optional)

Summary

- 4 way verification experiment
- ssRNA+ and ssRNA–
- coding and non-coding

Run Time

- worst case: $O(c^{\frac{k}{3}} + \frac{l}{3})$
- highly dependent on filter steps
- Example: 10368 · 2304 to 1029 · 36

[Fricke and Marz, 2016, Prediction of conserved long-range RNA-RNA interactions in full viral genomes.]
Discussion

Dependencies
- Python (v3.6.5)
- ViennaRNA Package (v2.4)
- VARNA (v3.93) (optional)

Run Time
- worst case: \( O(c^3 + l^3) \)
- highly dependent on filter steps
- Example: 10368 \cdot 2304 to 1029 \cdot 36

Summary
- 4 way verification experiment
- ssRNA+ and ssRNA–
- coding and non-coding

Application
- predict LRIs (LRIsan)
- simulate mutations (SIM)
- create mutants (Lab)

[Fricke and Marz, 2016, Prediction of conserved long-range RNA-RNA interactions in full viral genomes.]
Acknowledgements

- Martin Hölzer
- Bashar Ibrahim
- Manja Marz