

# Evolutionary traits in Chikungunya virus untranslated regions

Michael T. Wolfinger

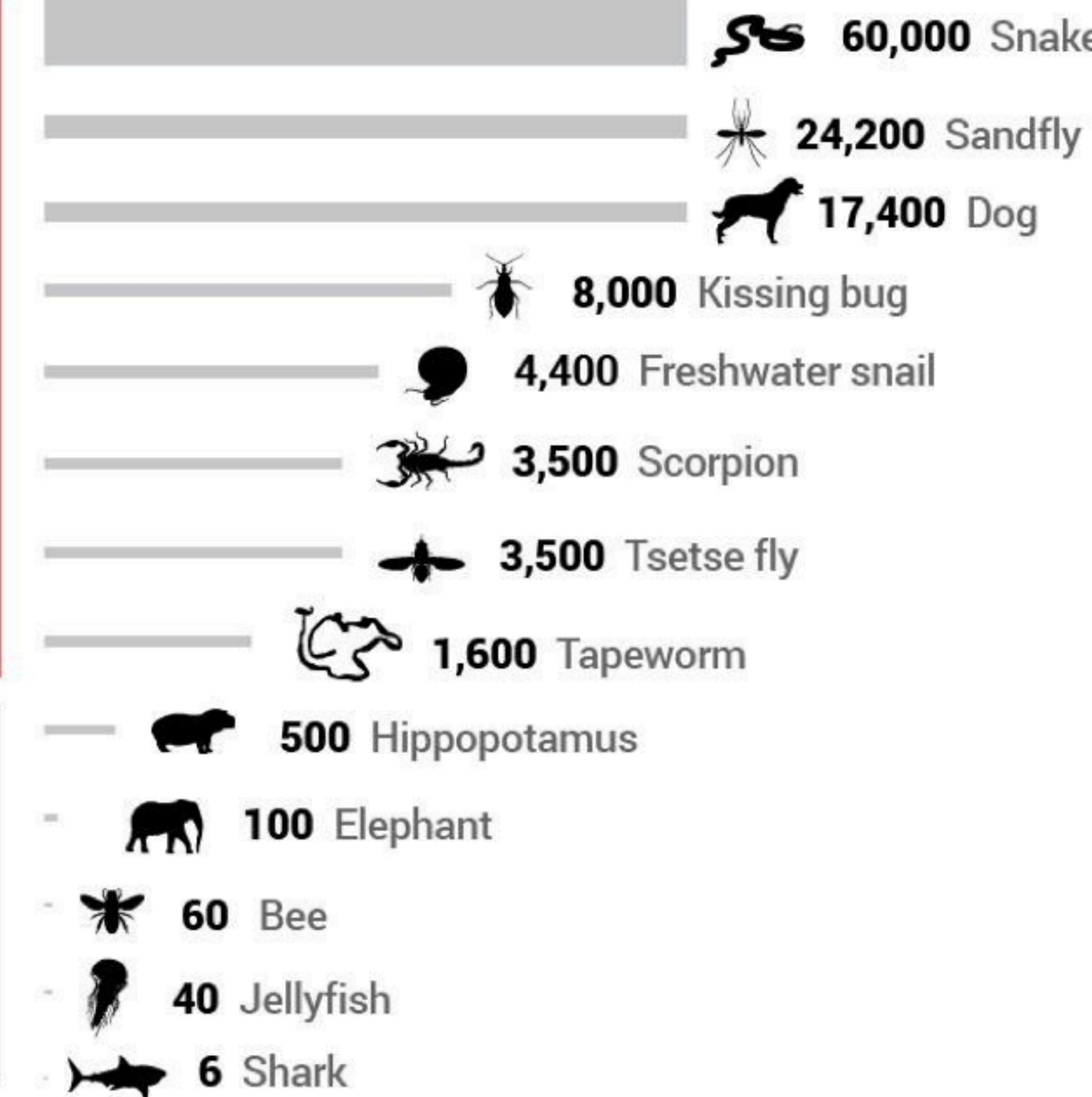
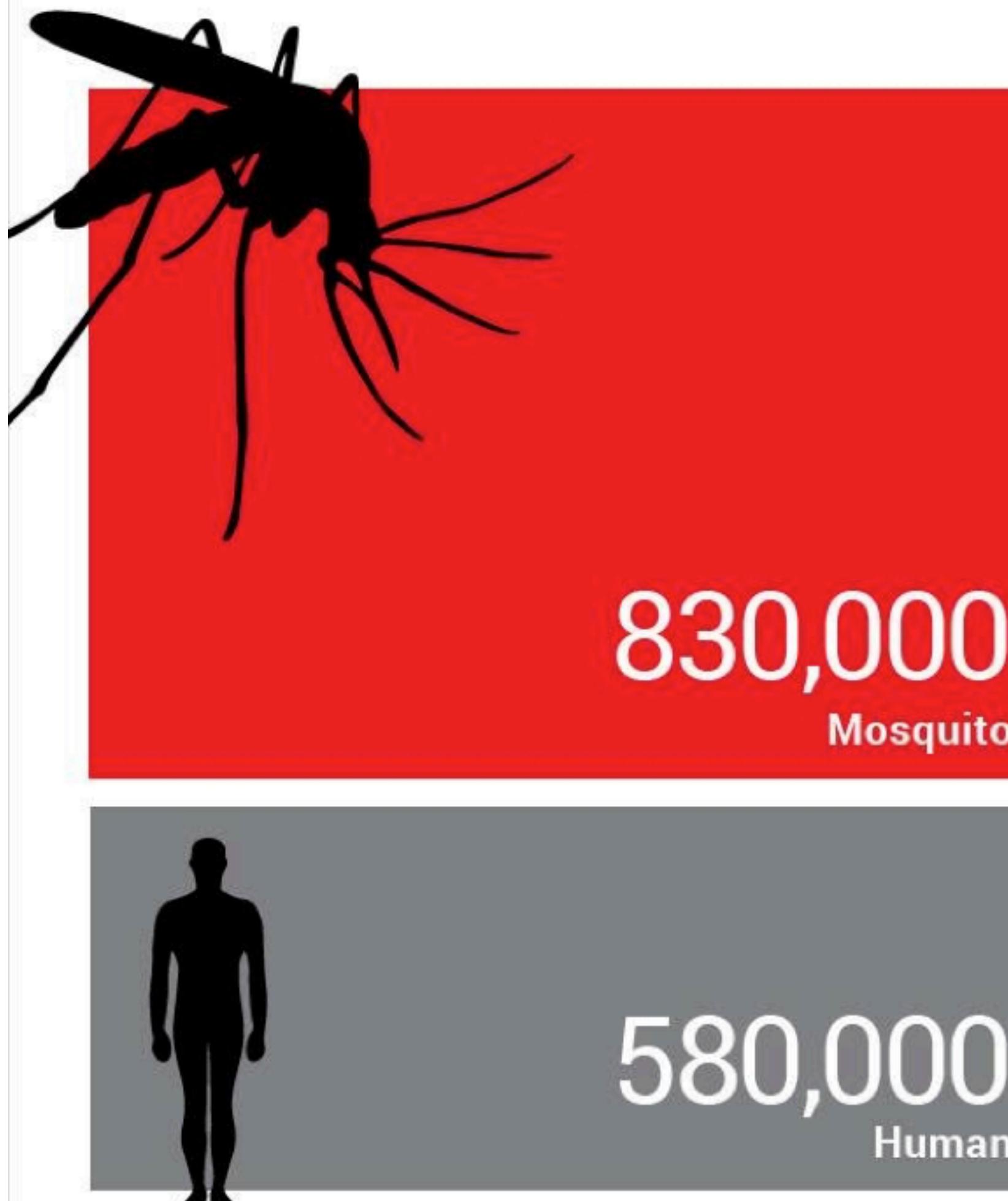
Research Group Bioinformatics and Computational Biology  
& Department of Theoretical Chemistry  
University of Vienna

4<sup>th</sup> VDS Retreat, Mondsee  
26 June 2019

# The World's Deadliest Animals

Number of people killed by animals, 2015

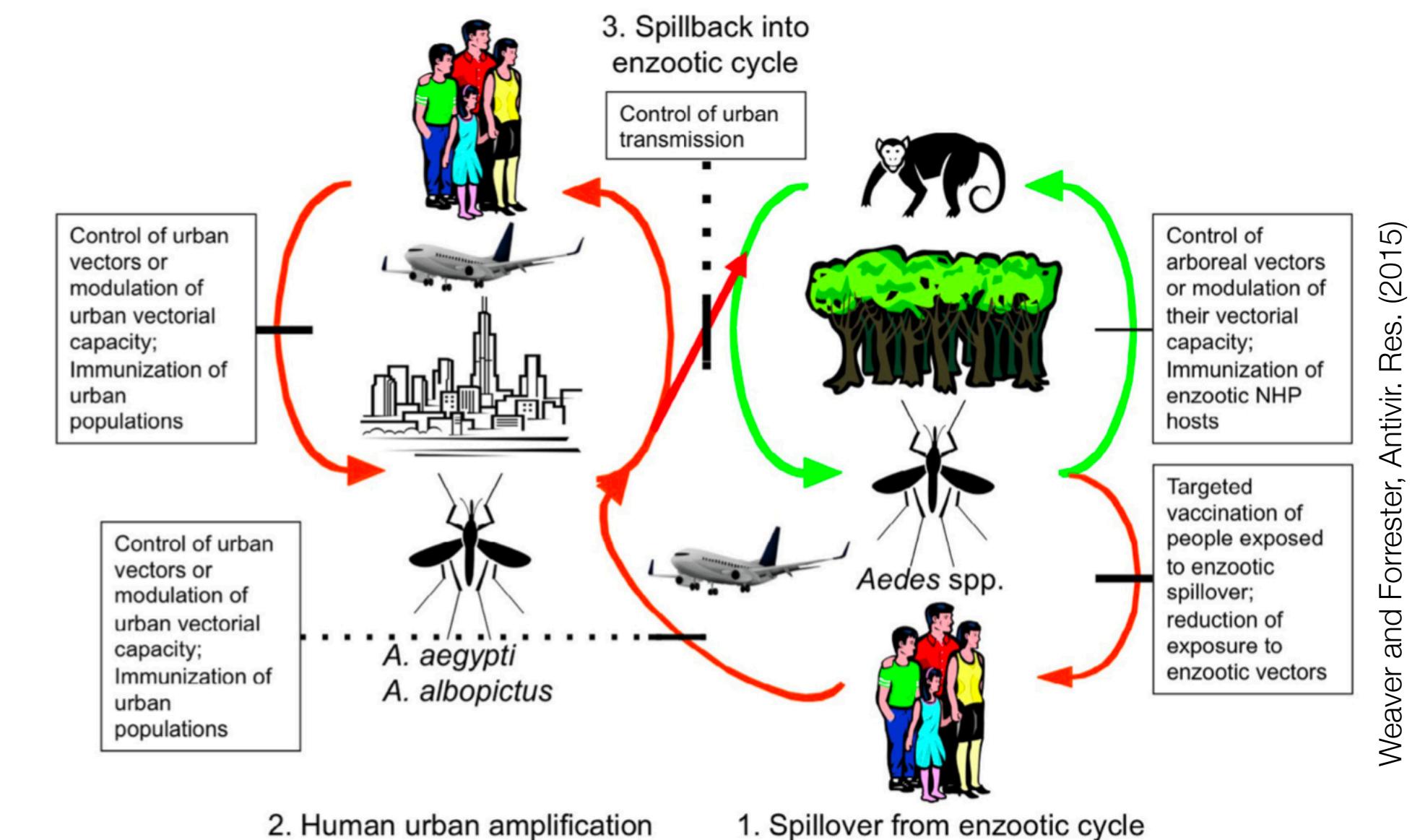
gates  
notes



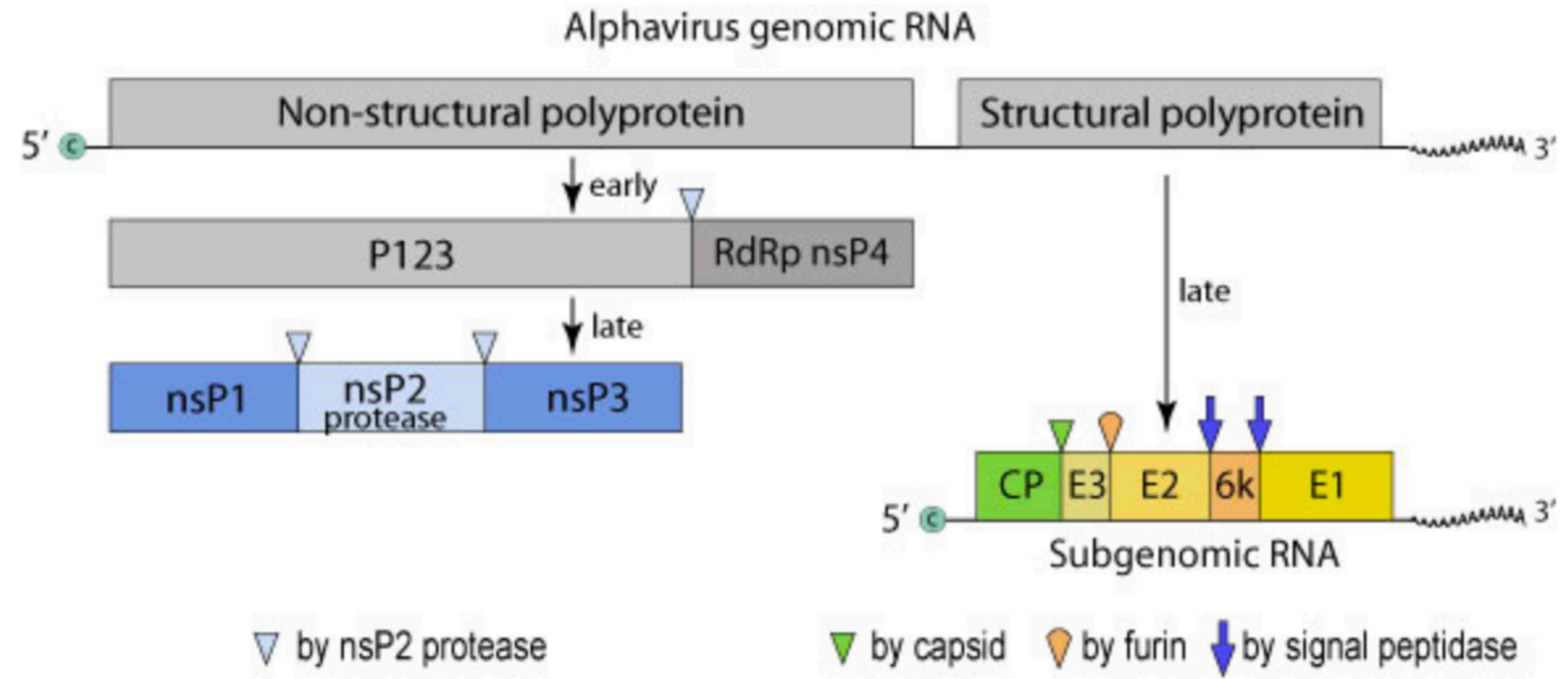
Sources: IHME, WHO, CrocBITE, FAO, Norwegian Institute for Nature Research, International Shark Attack File, National Geographic, PBS, National Science Foundation, CDC, WWF, *Wilderness & Environmental Medicine*, *Nature*, French Institute of Research for Development. All calculations have wide error margins.

# Chikungunya Virus (CHIKV)

- Family *Togaviridae* / genus *Alphavirus*; mosquito-borne (*Aedes* spp.)
- Single-stranded (+) sense RNA virus
- “Chikungunya” (Makonde language): “Disease that bend up the joints”
- Chikungunya fever: febrile illness, arthralgia, rash, rarely causes hemorrhagic complications
- Enzootic in tropical and subtropical regions of Africa
- First outbreak described 1952 in Tanzania
- No vaccine available



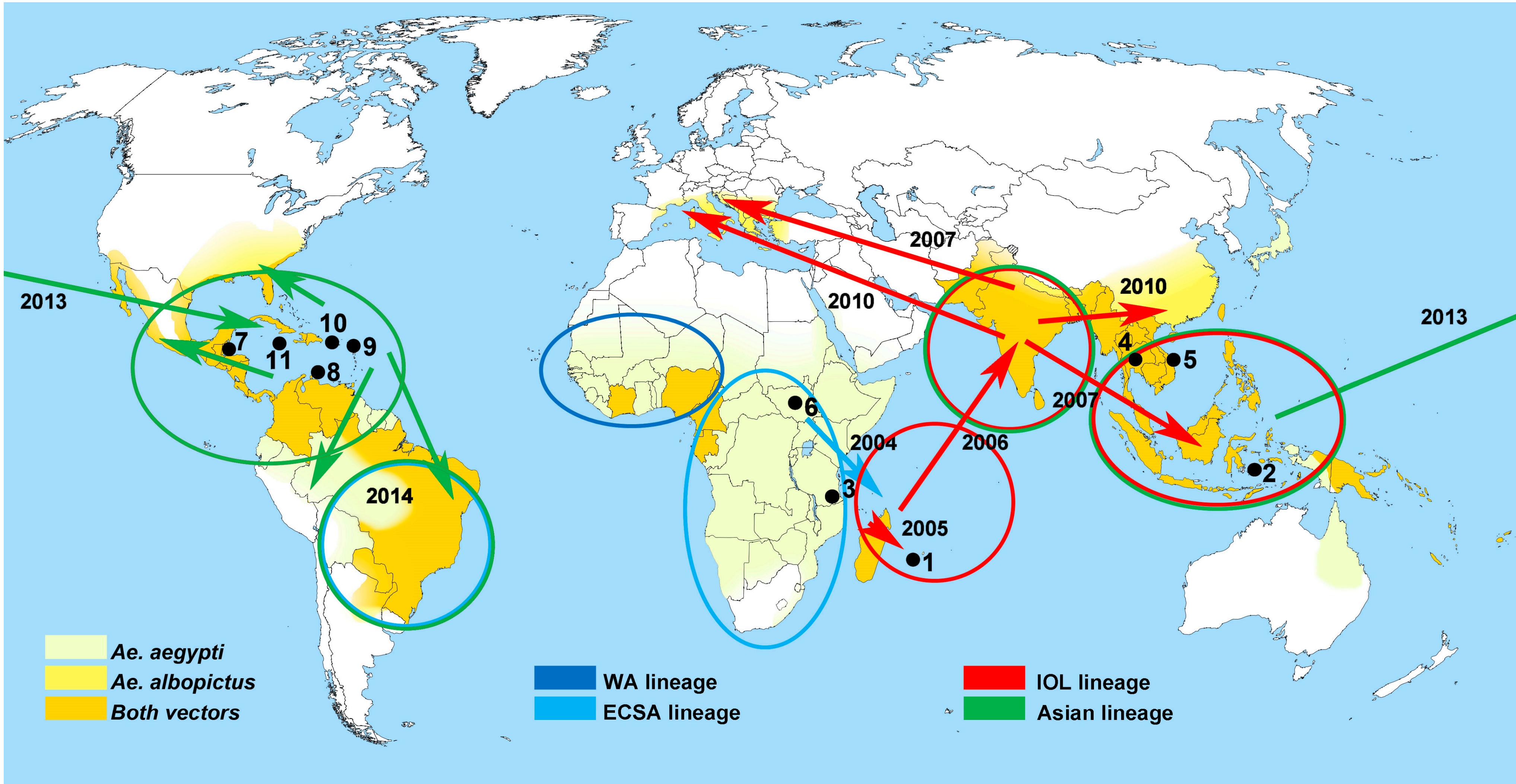
# Alphavirus Genome Organization



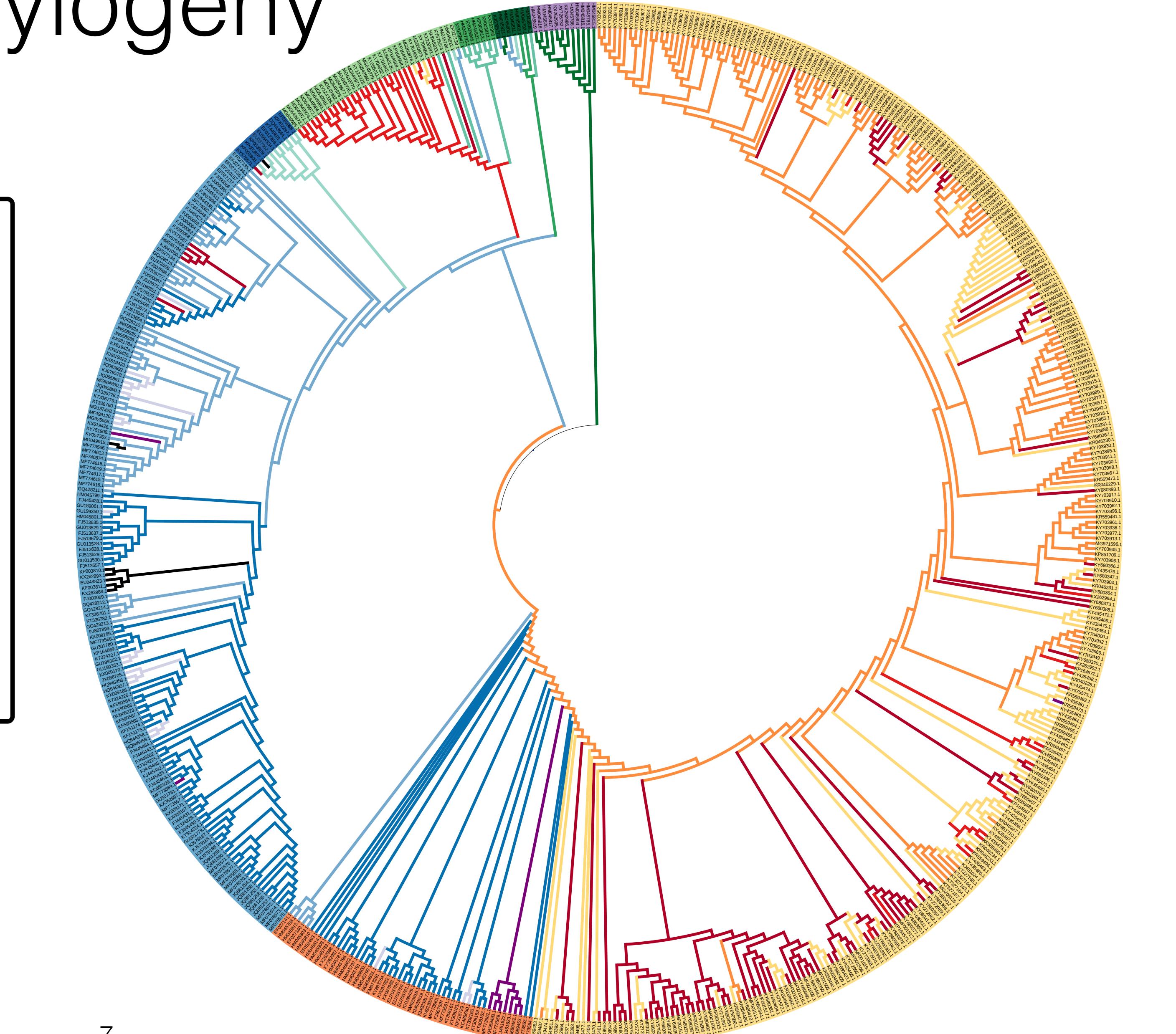
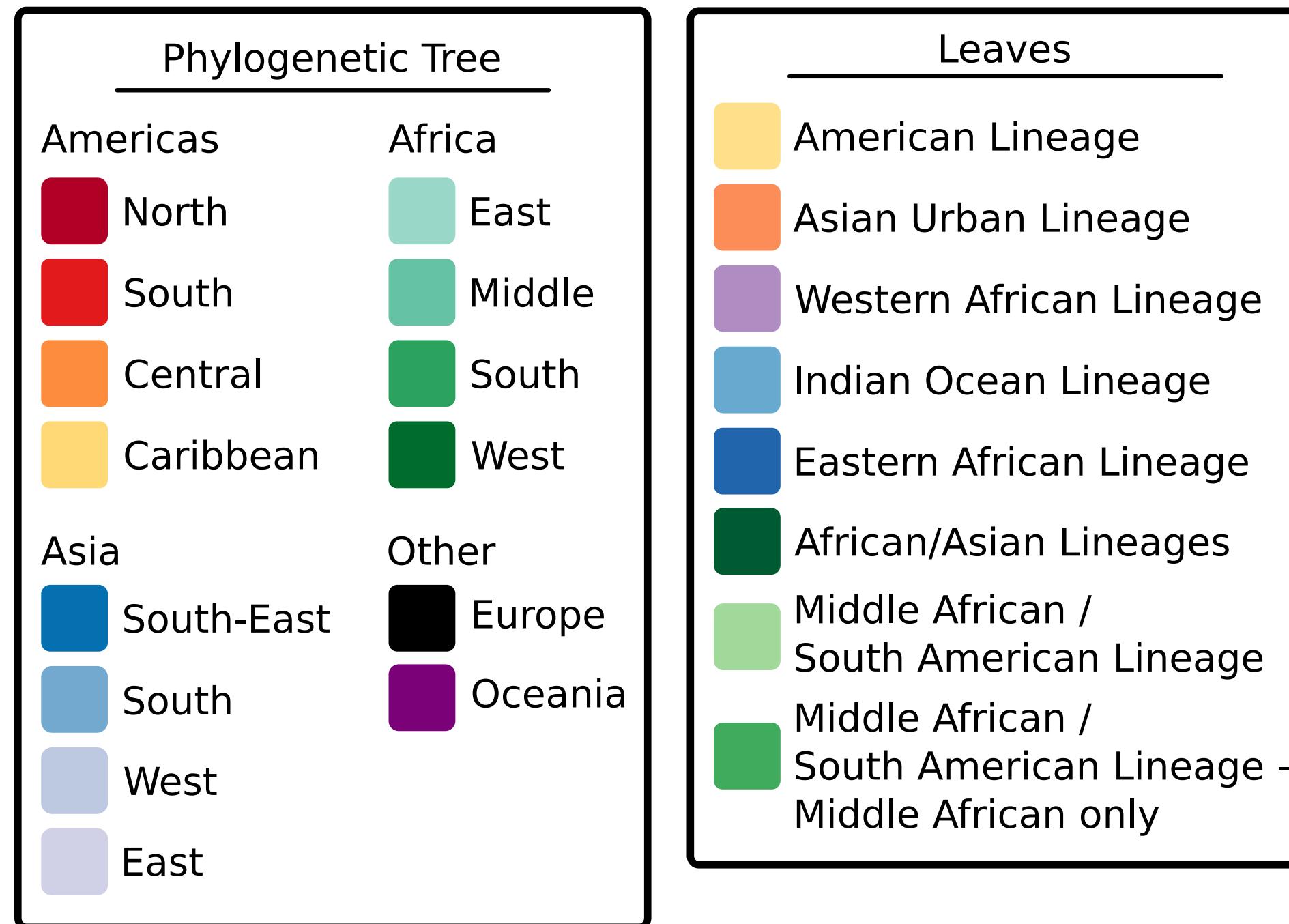
- Non-segmented, single-stranded, (+)-sense RNA genomes of 11-12kB length
- Capped and polyadenylated
- AV genomes appear to host cells as mRNA for immediate translation upon entry into the cytoplasm

# Part I: Phylogeography

# CHIKV Epidemic Spread



# Updated CHIKV Phylogeny

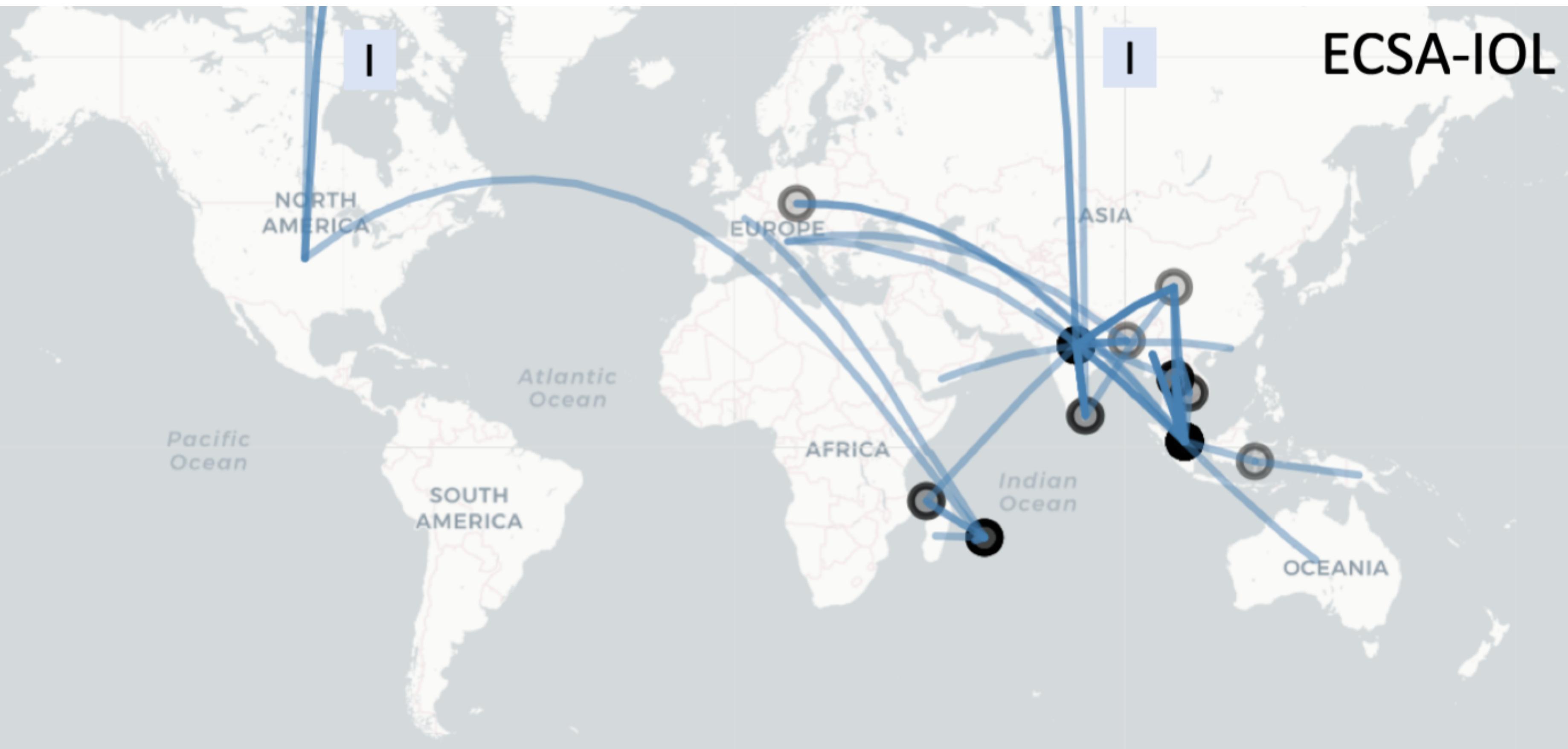


- 590 CHIKV genomes
- iq-tree / SH-aLRT

# CHIKV Phylogeography AUL



# CHIKV Phylogeography ECSA-IOL

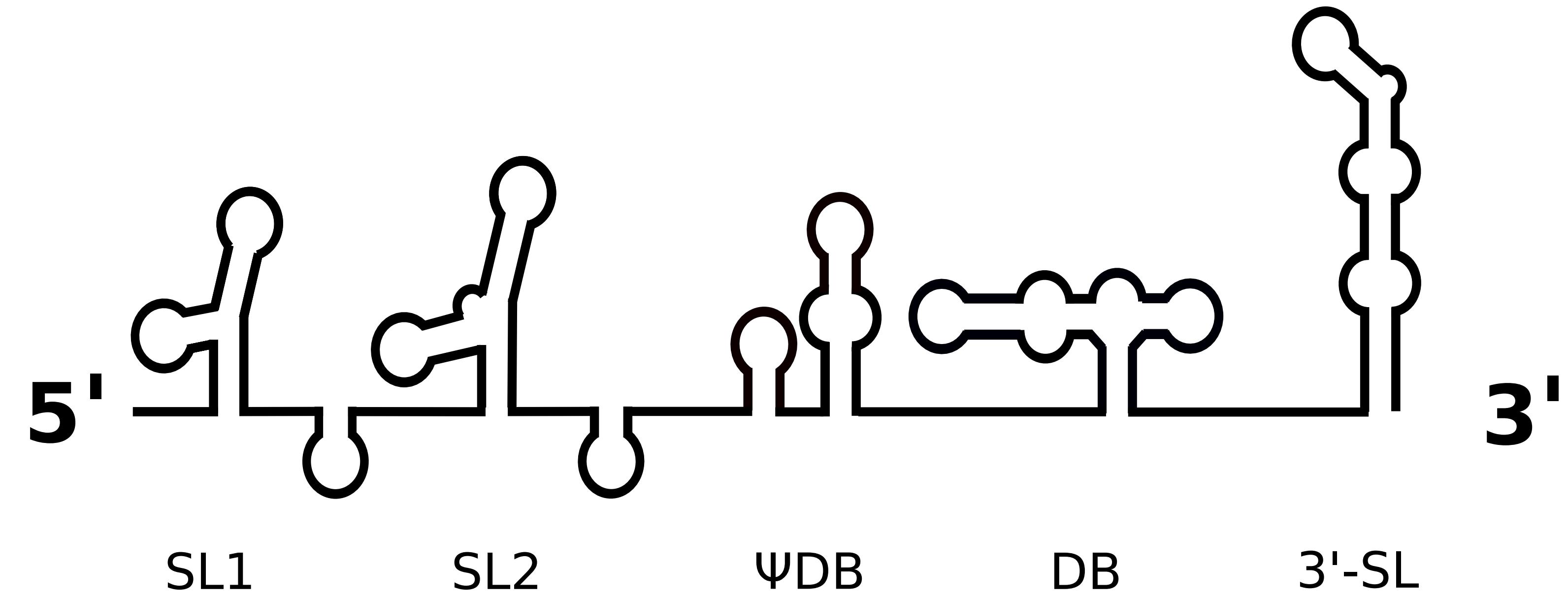
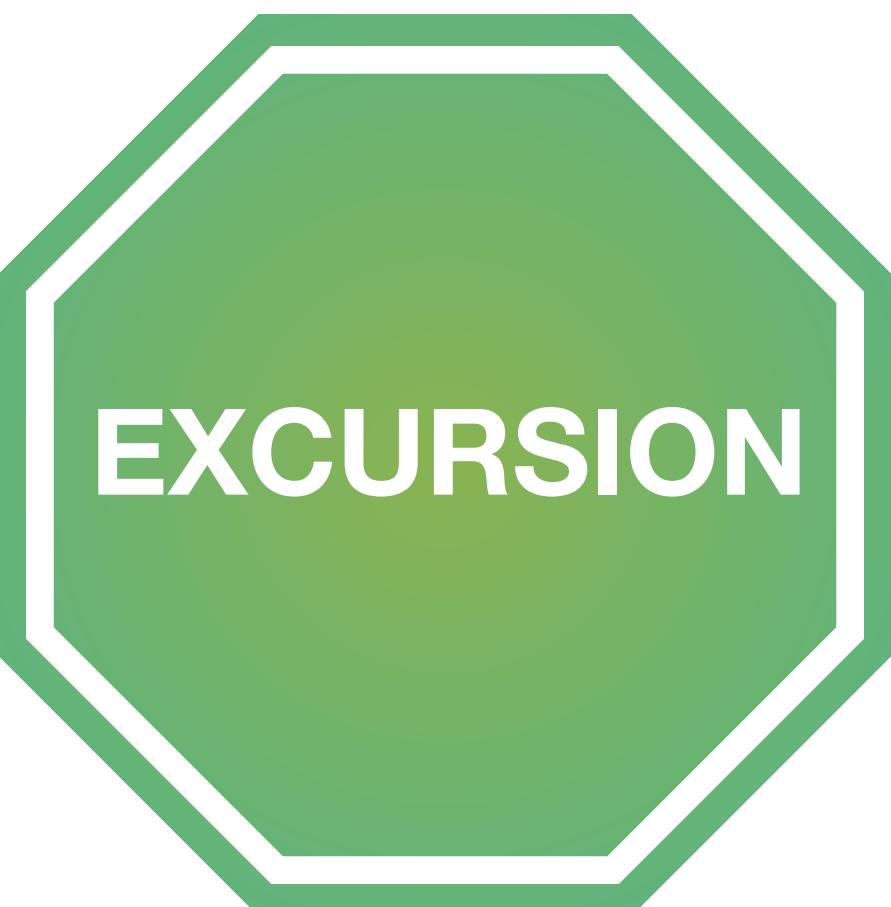


# CHIKV Phylogeography ECSA-MASA

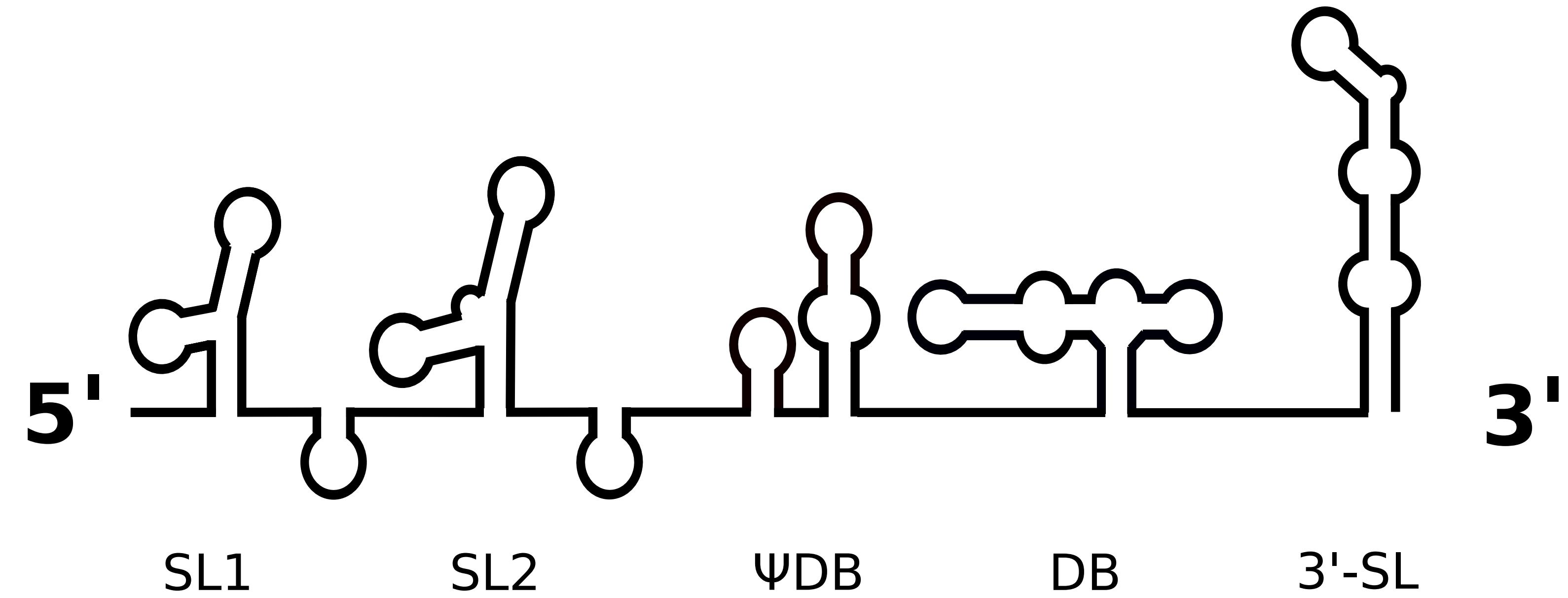


# Part II: Conserved ncRNAs

# Conserved RNAs in Viral 3'UTR: ZIKV

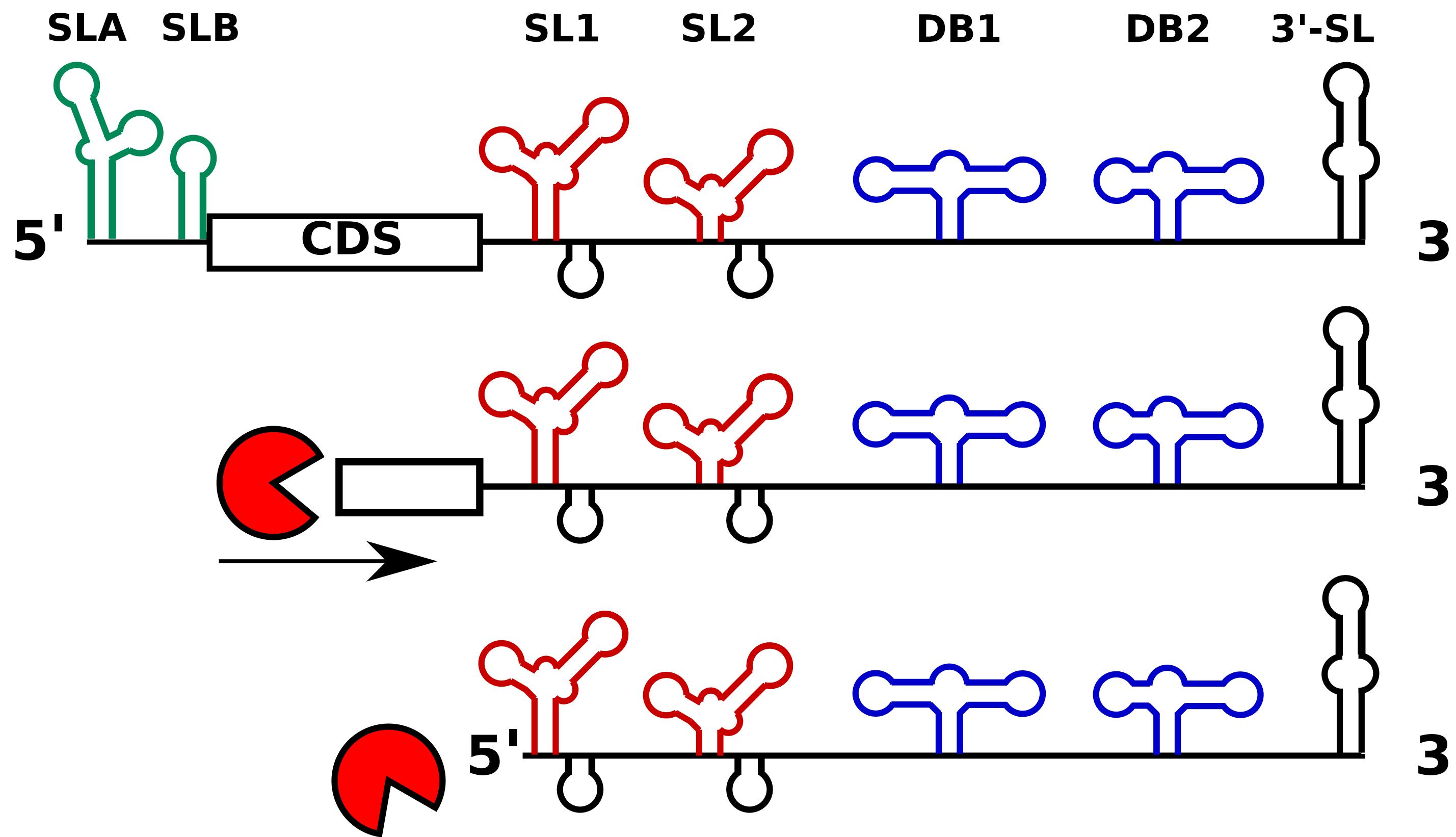


# Conserved RNAs in Viral 3'UTR: ZIKV



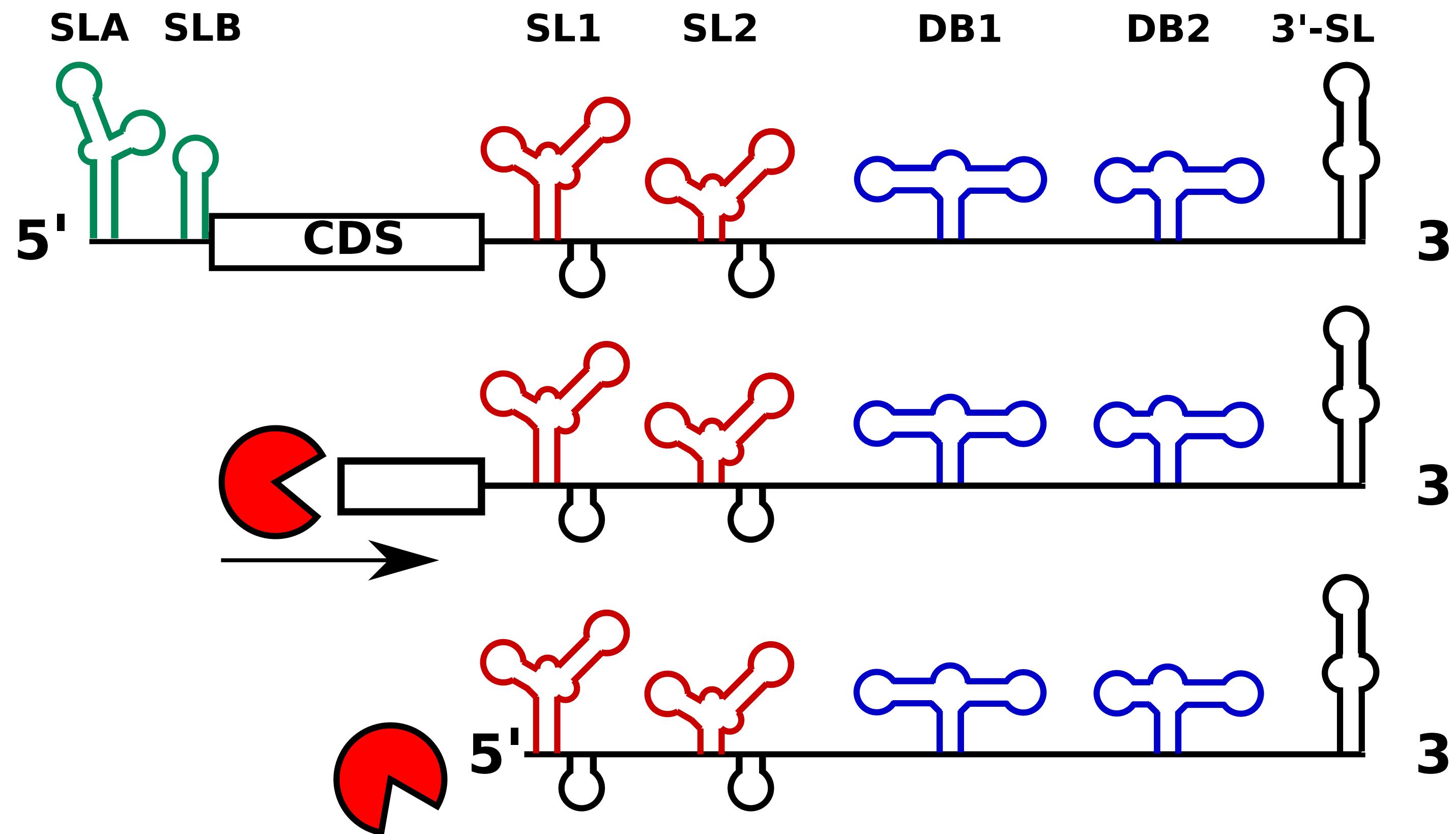
- Accumulation of short flavivirus RNA (sfRNA) upon infection
- Stable decay intermediates produced by partial exoribonuclease degradation
- Xrn1 is efficiently stalled at conserved xrRNA structures

# Conserved RNAs in Viral 3'UTR



- Accumulation of short flavivirus RNA (sfRNA) upon infection
- Stable decay intermediates produced by partial exoribonuclease degradation
- Xrn1 is efficiently stalled at conserved xrRNA structures

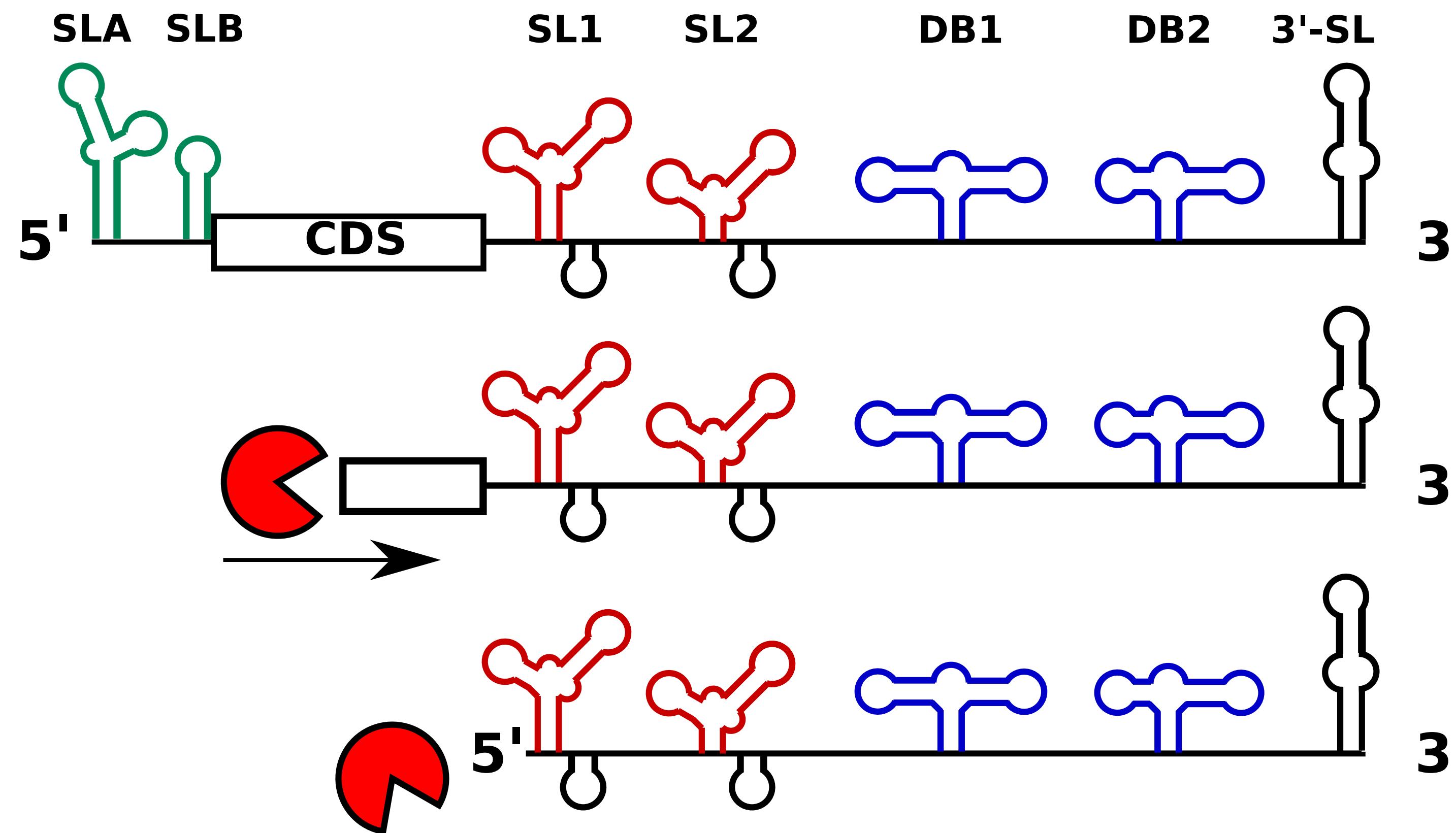
# Conserved RNAs in Viral 3'UTR



**sfRNA mediates  
pathogenicity !**

- Accumulation of short flavivirus RNA (sfRNA) upon infection
- Stable decay intermediates produced by partial exoribonuclease degradation
- Xrn1 is efficiently stalled at conserved xrRNA structures

# Conserved RNAs in Viral 3'UTR



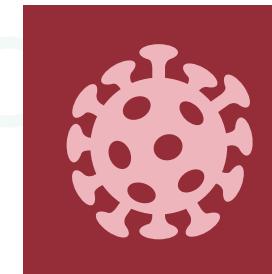
**sfRNA mediates  
pathogenicity !**

**How to find them?**

- Accumulation of short flavivirus RNA (sfRNA) upon infection
- Stable decay intermediates produced by partial exoribonuclease degradation
- Xrn1 is efficiently stalled at conserved xrRNA structures

# Conserved RNAs in Viral 3'UTR

EXCURSION



*viruses*

<https://doi.org/10.3390/v11030298>



Article

## Functional RNA Structures in the 3'UTR of Tick-Borne, Insect-Specific and No-Known-Vector Flaviviruses

Roman Ochsenreiter <sup>1</sup>, Ivo L. Hofacker <sup>1,2</sup> and Michael T. Wolfinger <sup>1,2,\*</sup>

<sup>1</sup> Department of Theoretical Chemistry, University of Vienna, Währingerstraße 17, 1090 Vienna, Austria; romanoch@tbi.univie.ac.at (R.O.); ivo@tbi.univie.ac.at (I.L.H.)

<sup>2</sup> Research Group BCB, Faculty of Computer Science, University of Vienna, Währingerstraße 29, 1090 Vienna, Austria

\* Correspondence: michael.wolfinger@univie.ac.at

How to find them?

- Accurately predict flavivirus RNA (sfRNA) upon infection

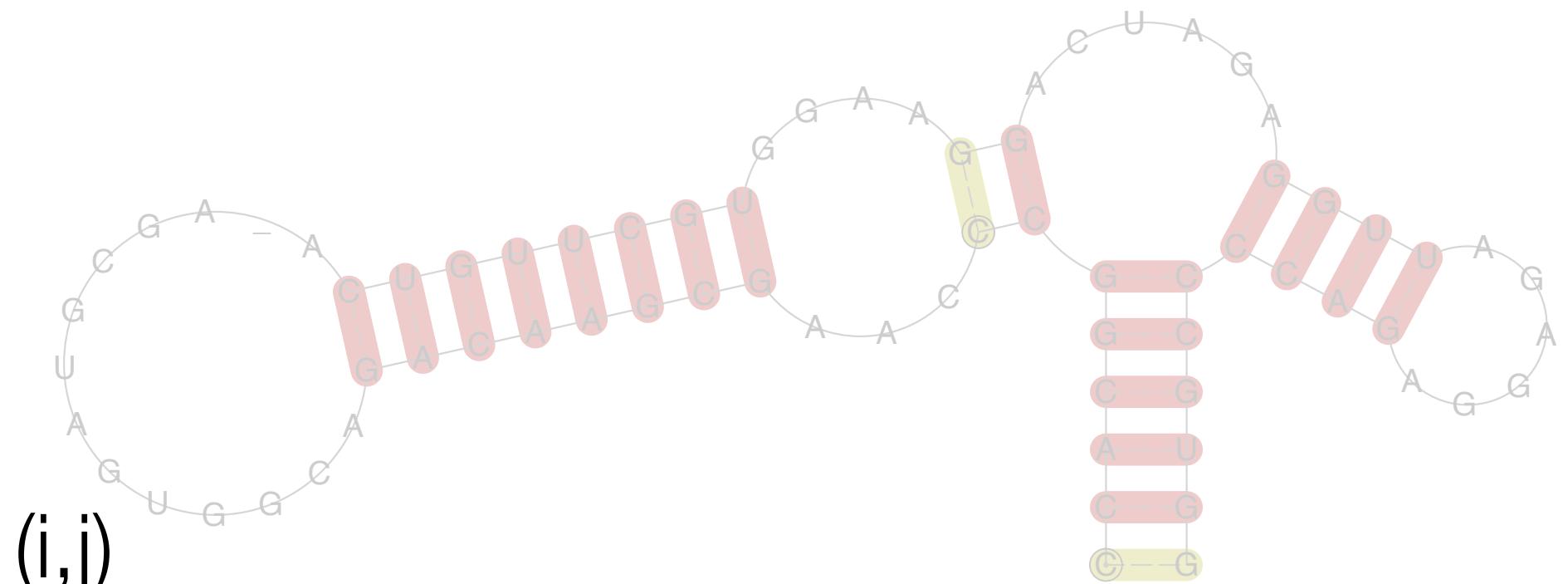
- Stable decay intermediates produced by partial exoribonuclease degradation

Received: 1 March 2019; Accepted: 20 March 2019; Published: 24 March 2019

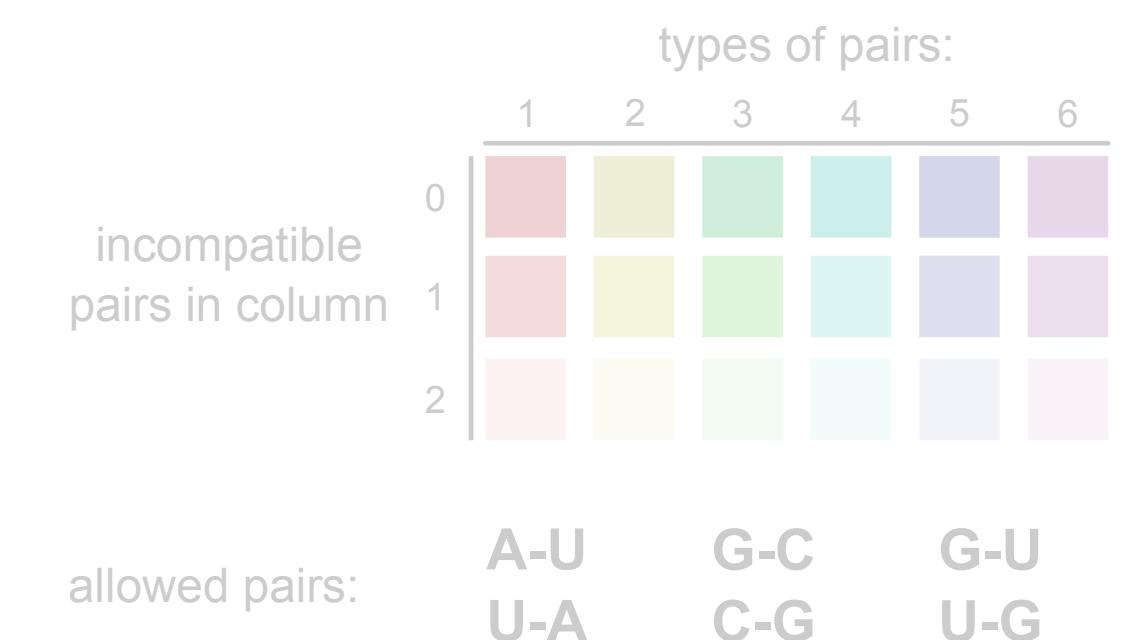


# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)

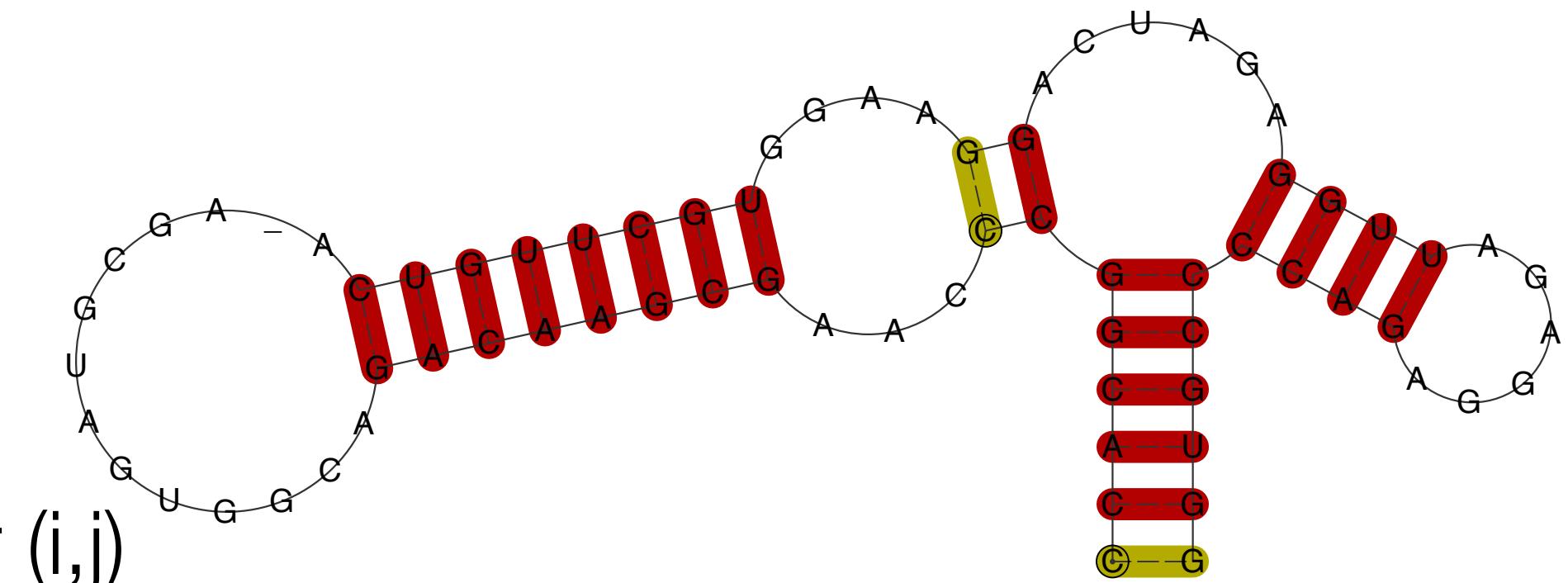


USUV.10 ((((((((...(((((((.....))))))))....)).).(((((.....))))))))  
USUV.11 CCACGGCUCAA GCGAACAGACGGUGAUGCAG-A CUGUUCGU GGAAAG GACUAGA GGUUAGAGGA GACCCCGUGG 72  
UCACGGCCCAAGCGAACAGACGGUGAUGCAG-A CUGUUCGU GGAAAG GACUAGA GGUUAGAGGA GACCCCGUGG 72  
.....10.....20.....30.....40.....50.....60.....70..

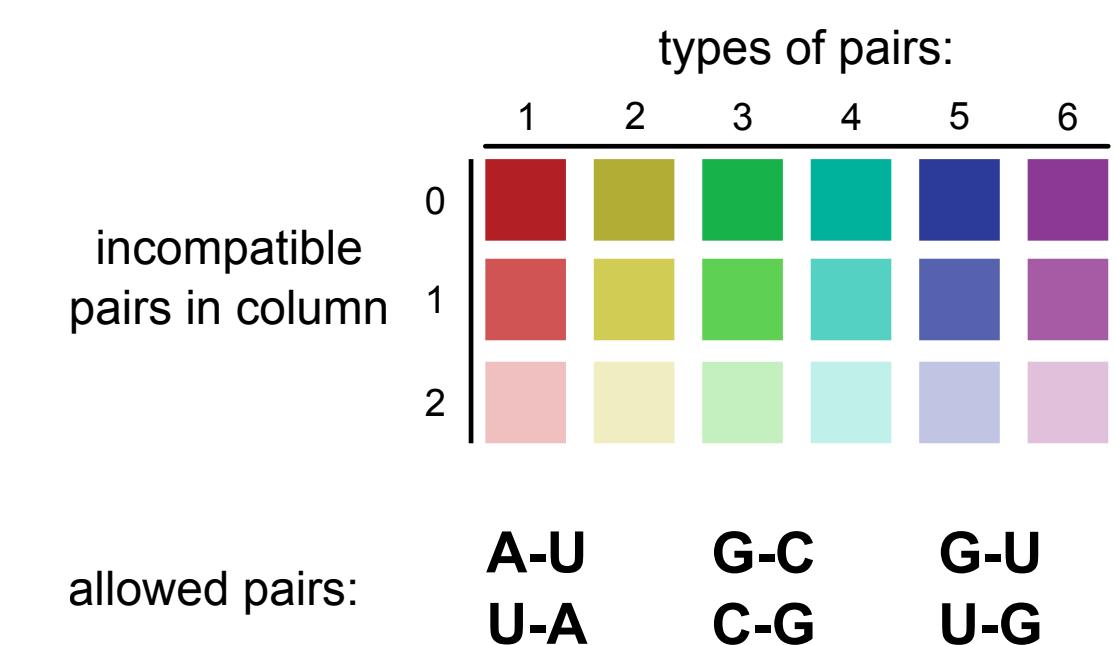


# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)

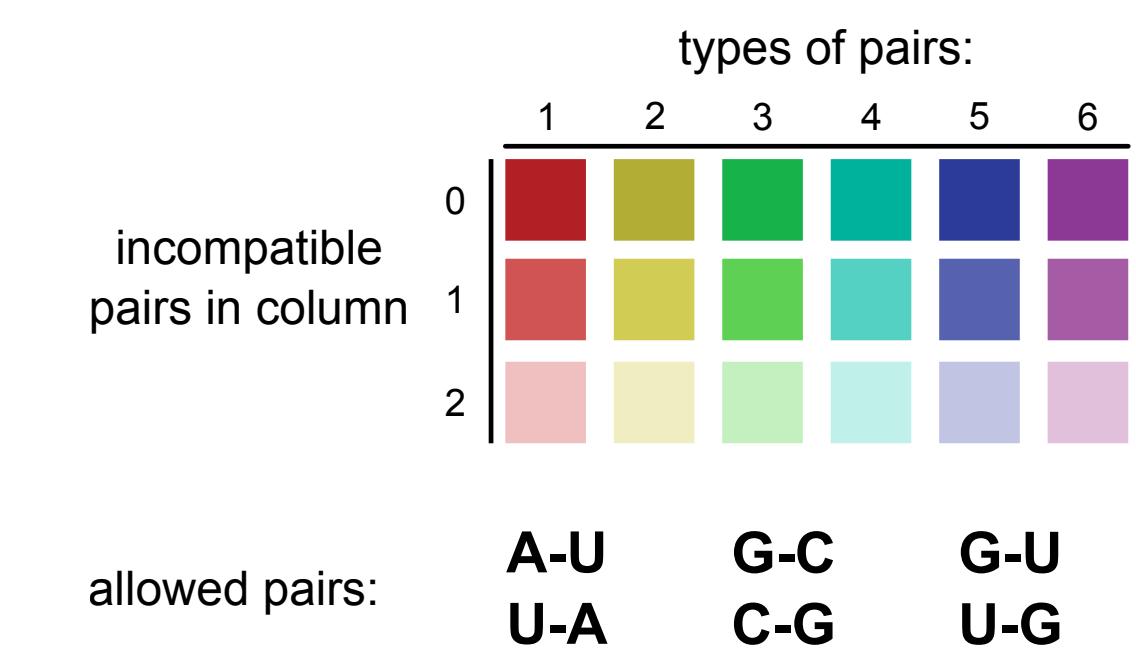
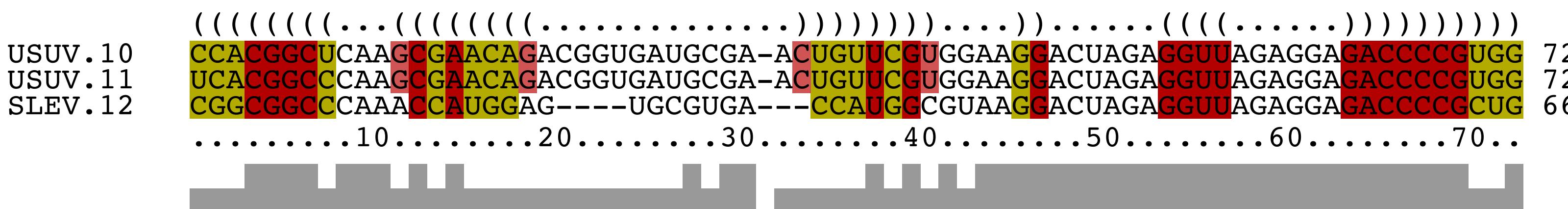
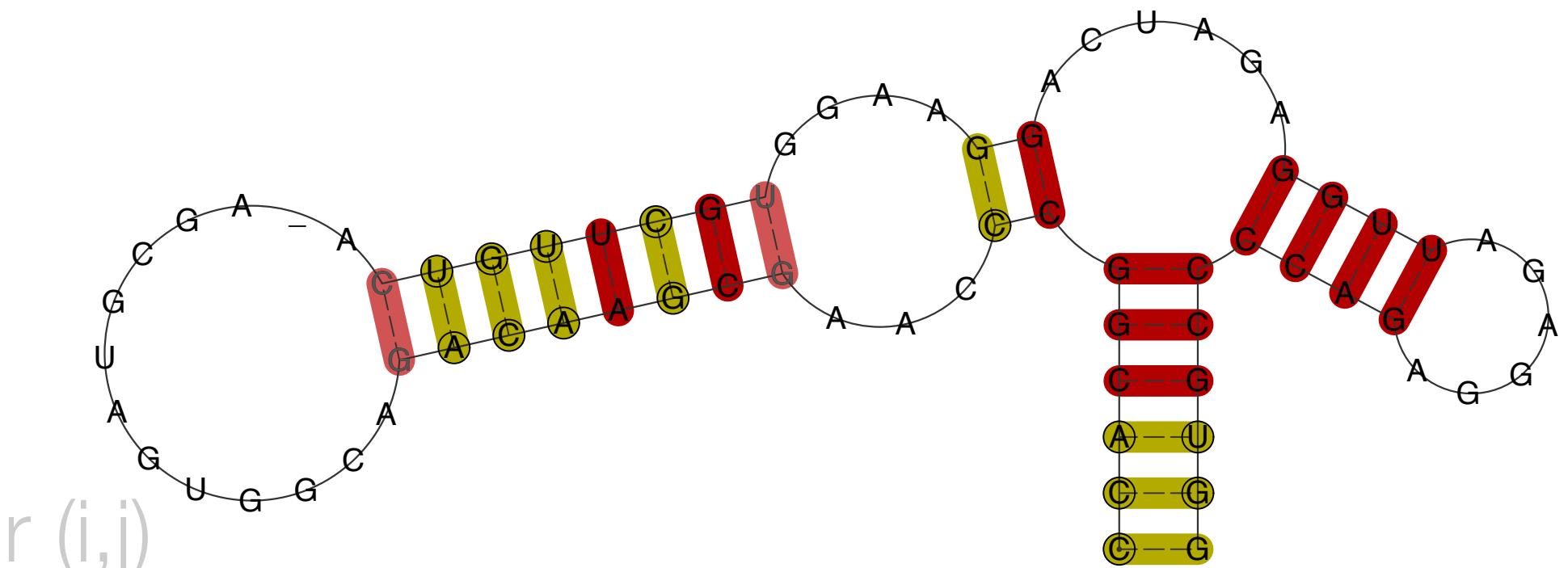


USUV.10	<code>(((((.....((((.....))))))).....((.....))))))</code>	
USUV.11	<code>CCACGGCUCAA</code> <b>GCGAACAGACGGUGAUGC</b> GA-A <b>CUGUUCGU</b> GGAA <b>GCACUAGA</b> GGUAGAGGA <b>GACCCC</b> GU <code>UCACGGGCCAAG</code> CGAACAGACGGUGAUGC-GA-A <b>CUGUUCGU</b> GGAA <b>GCACUAGA</b> GGUAGAGGA <b>GACCCC</b> GU	72 72
	.....10.....20.....30.....40.....50.....60.....70..	



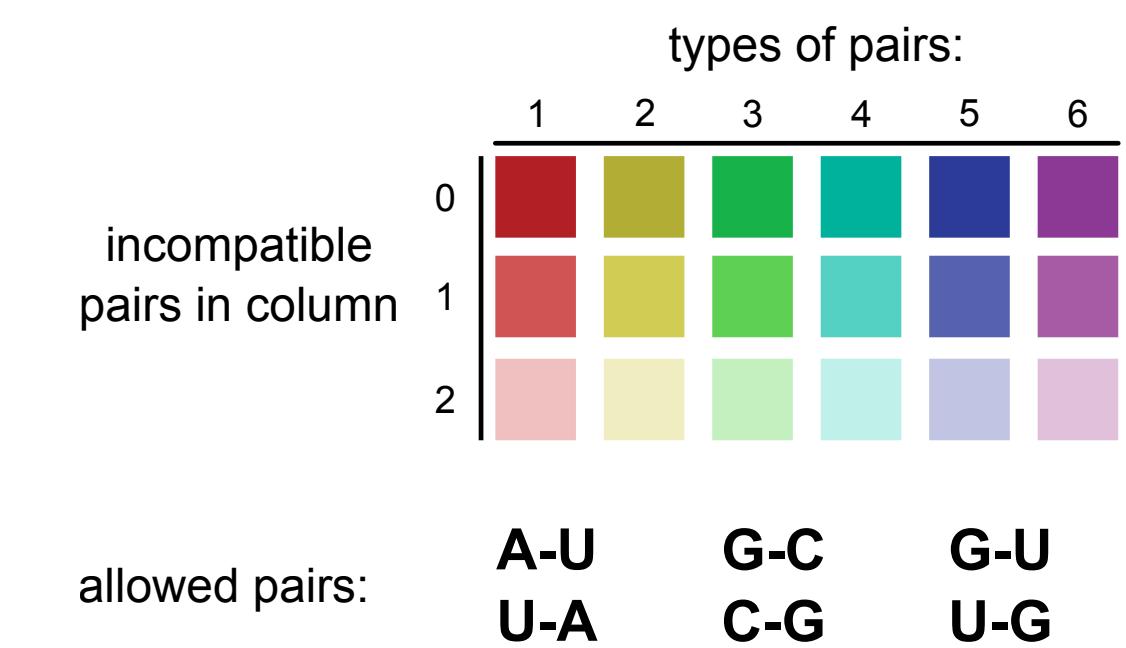
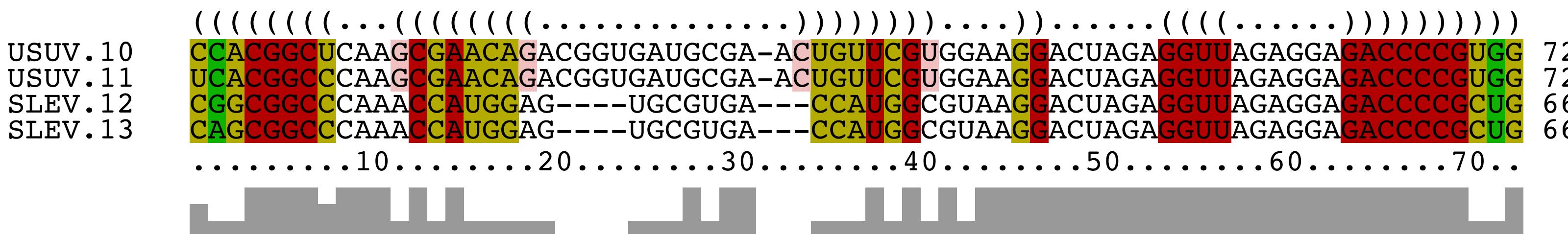
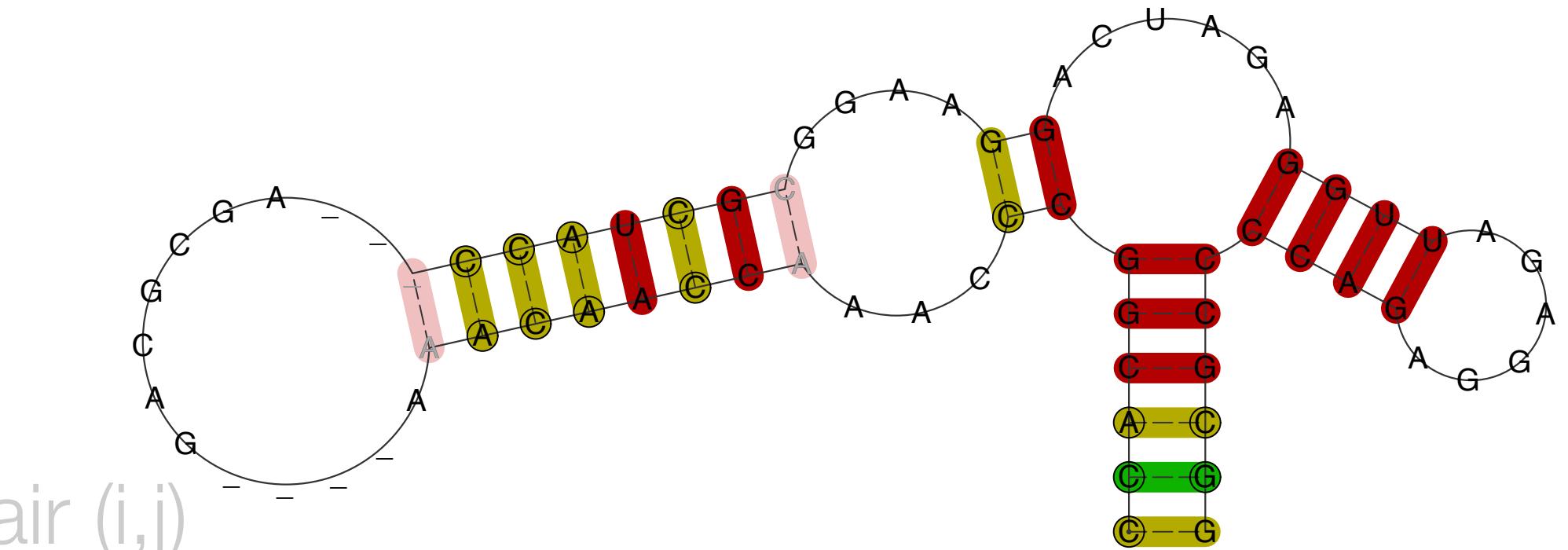
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)



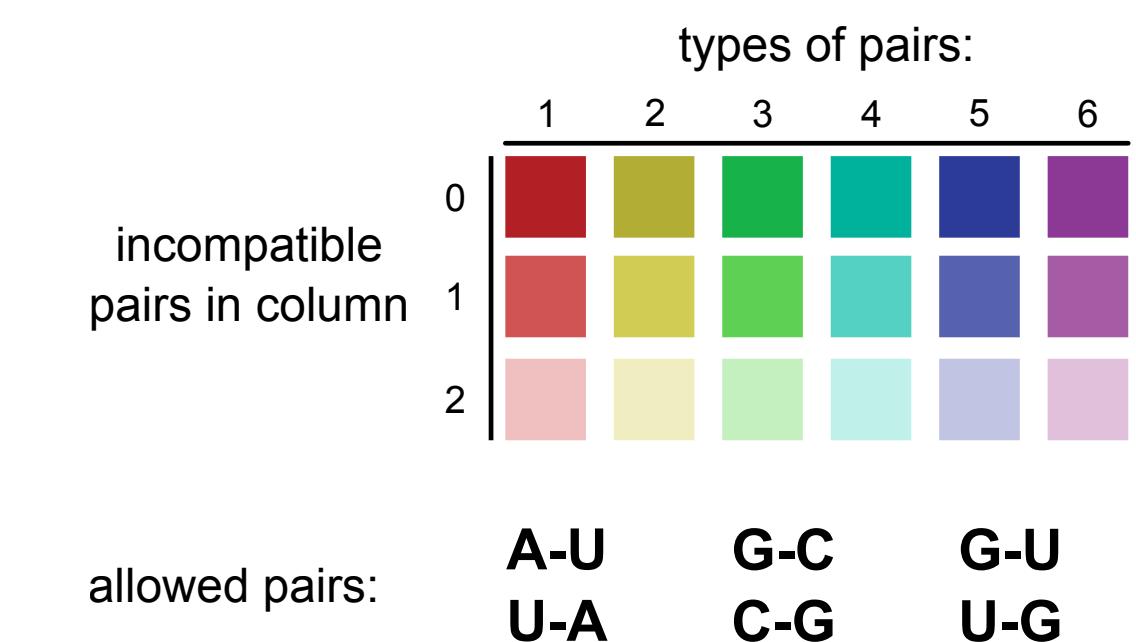
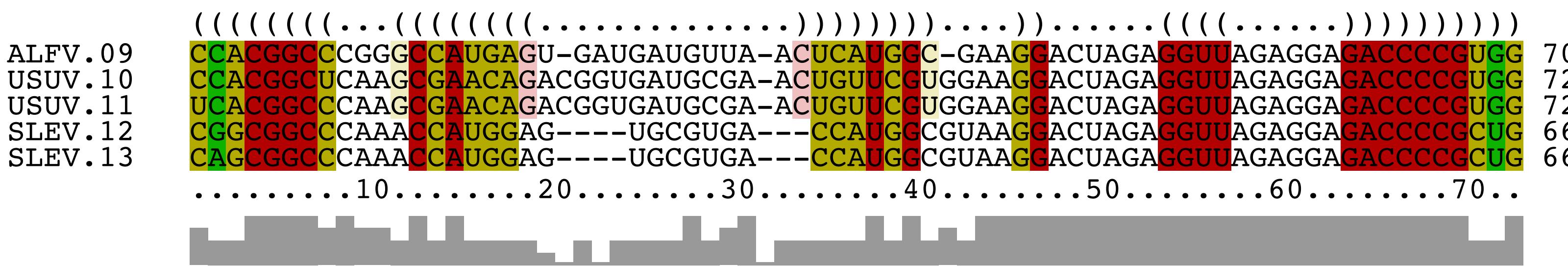
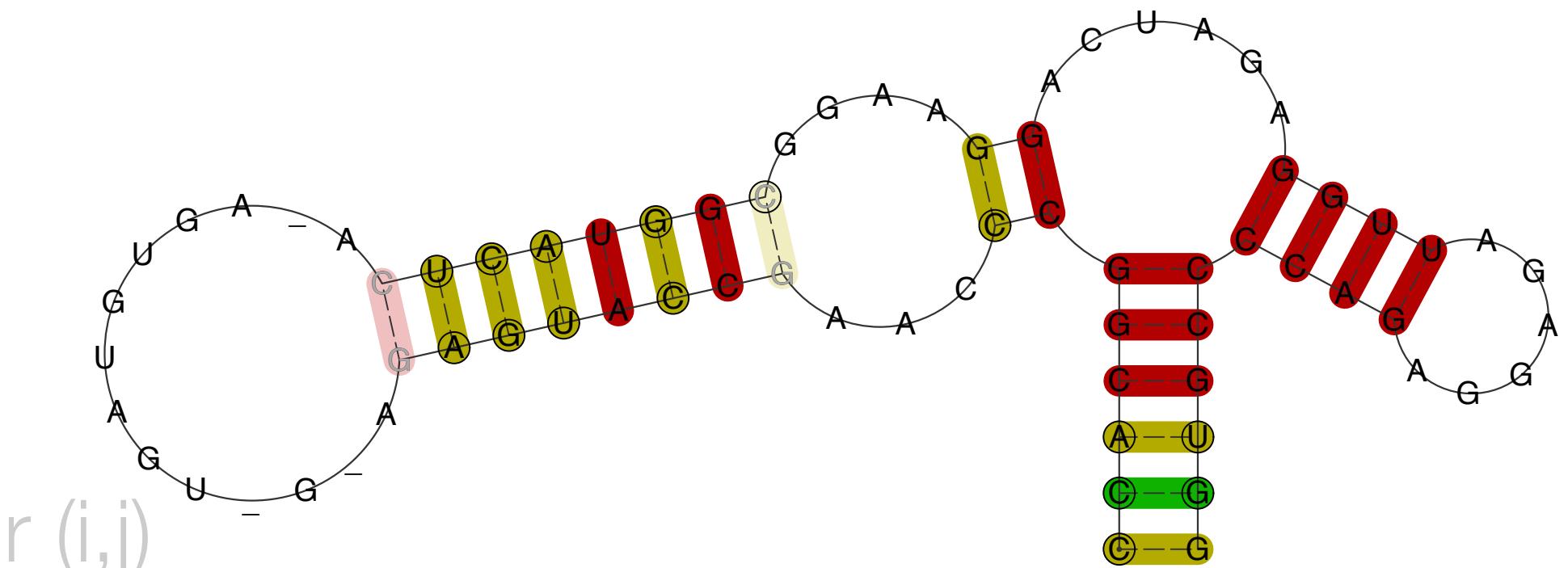
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)



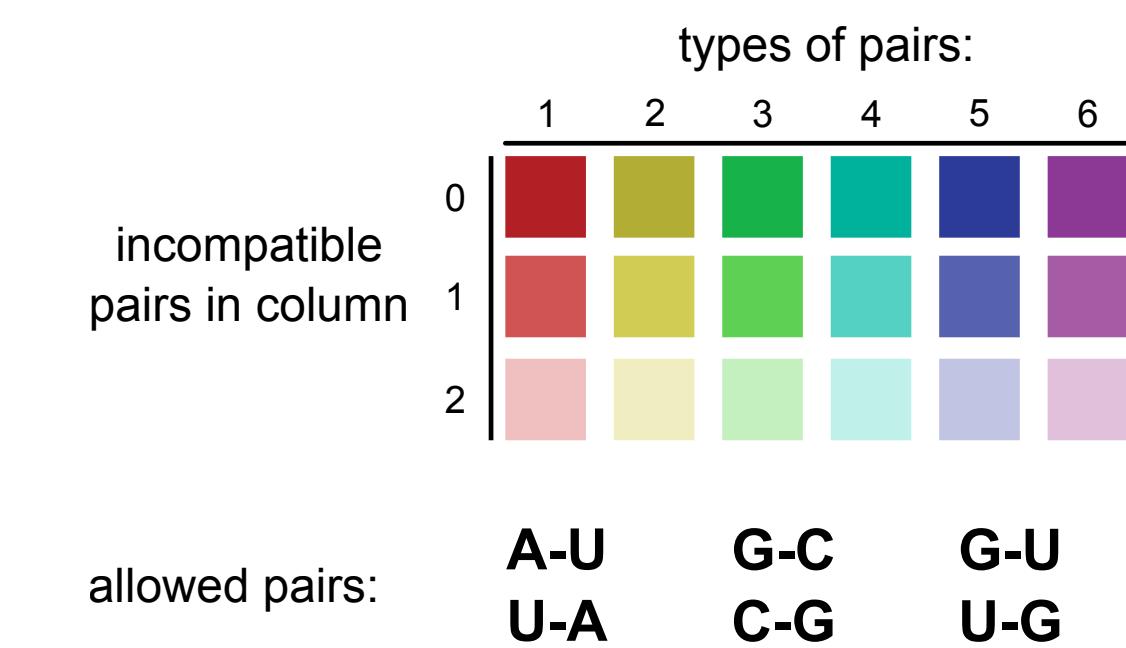
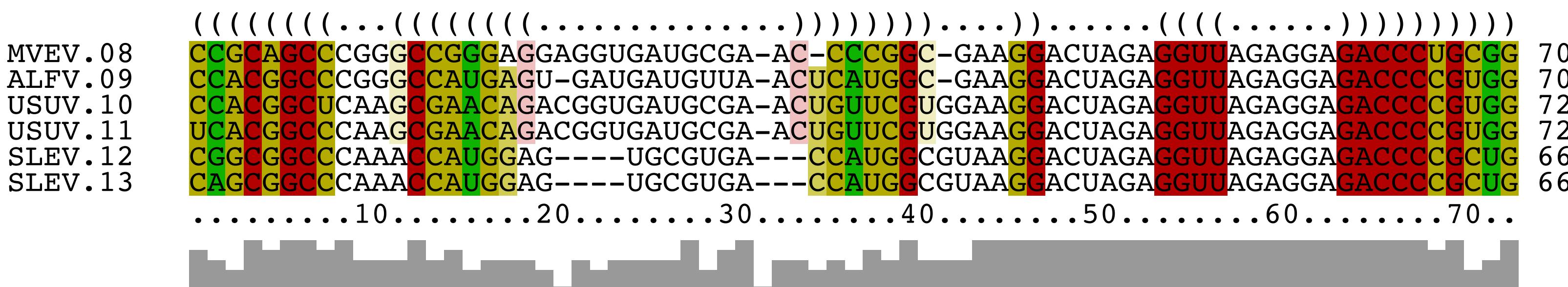
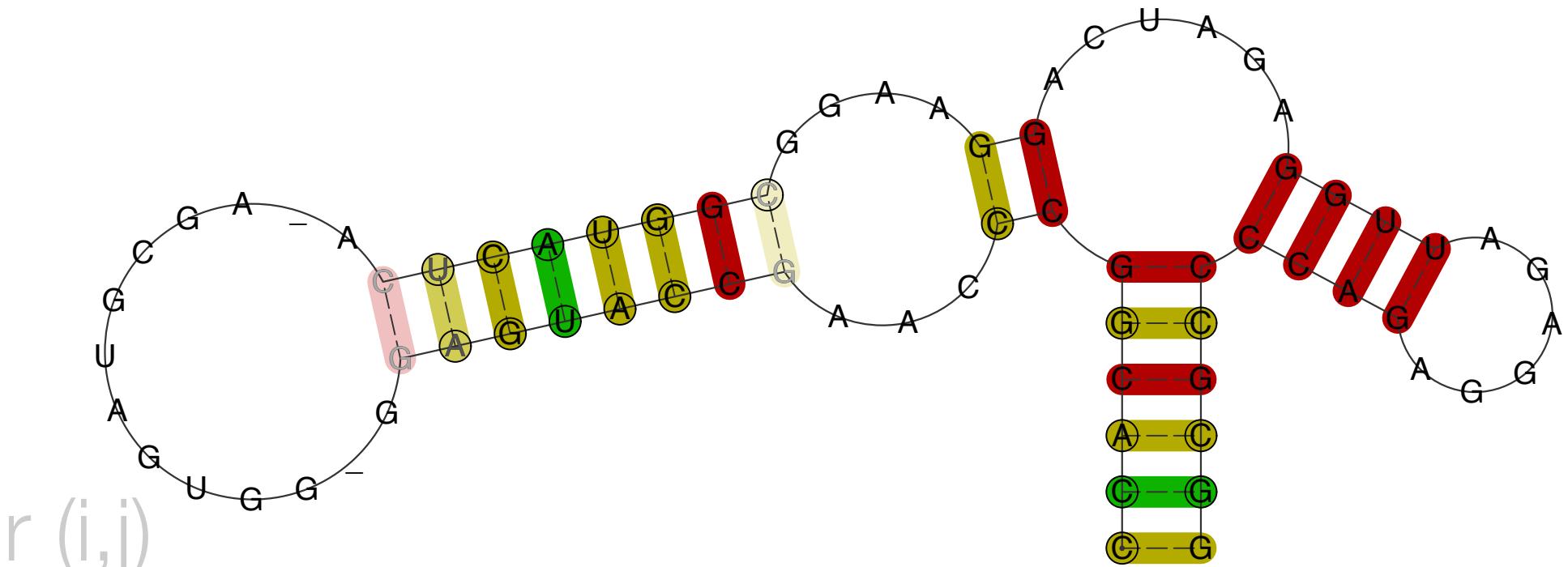
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)



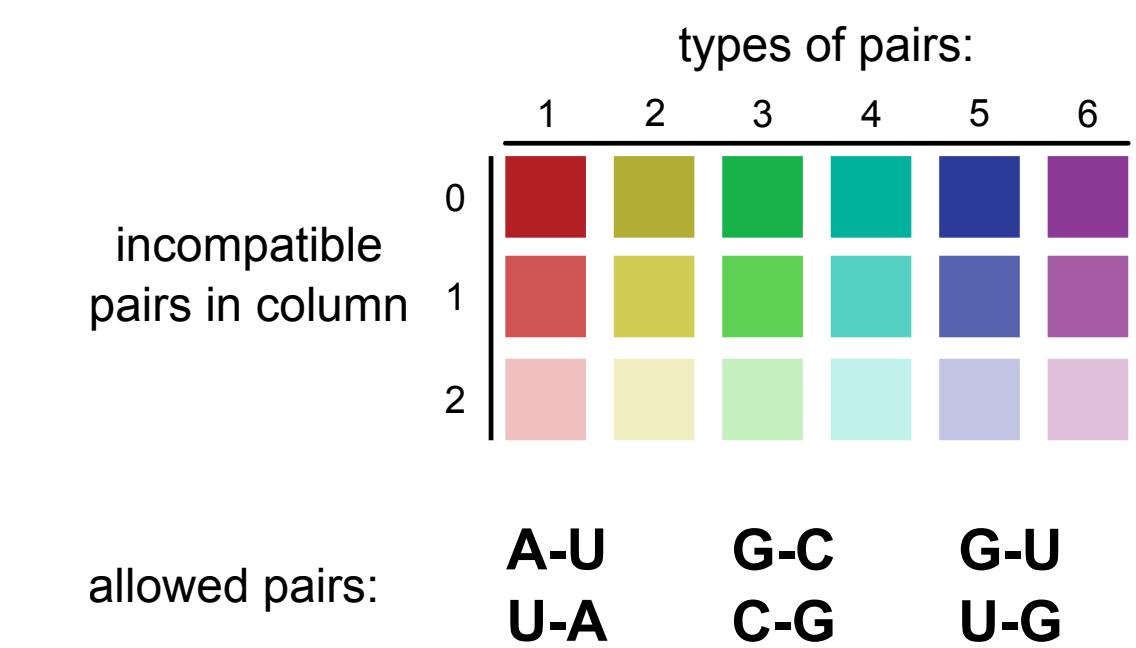
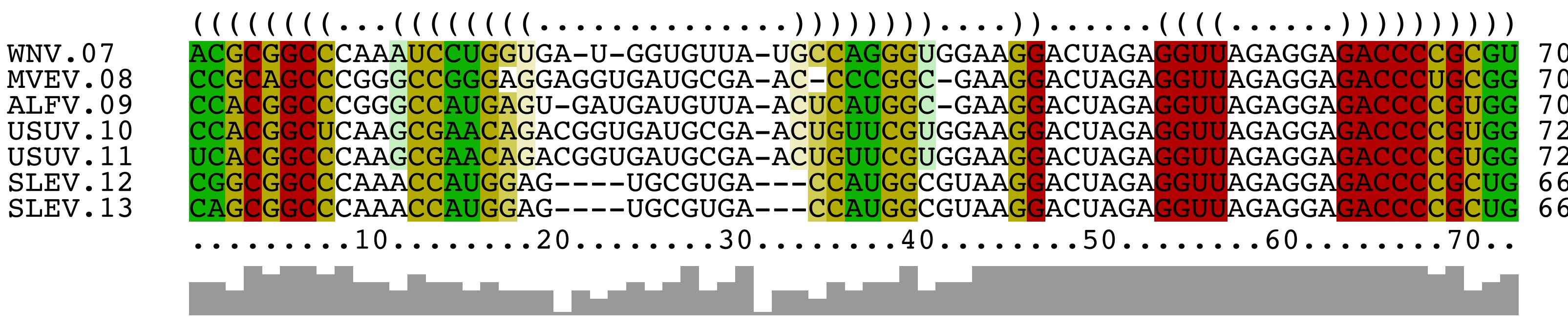
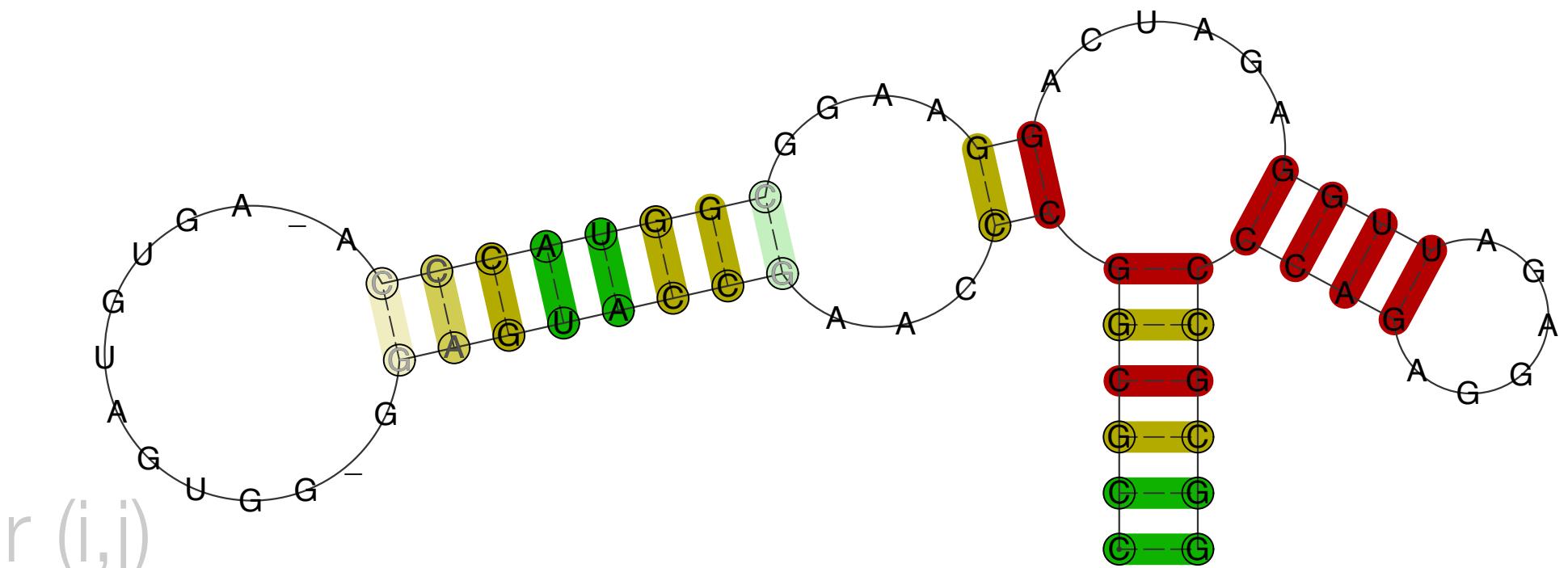
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)

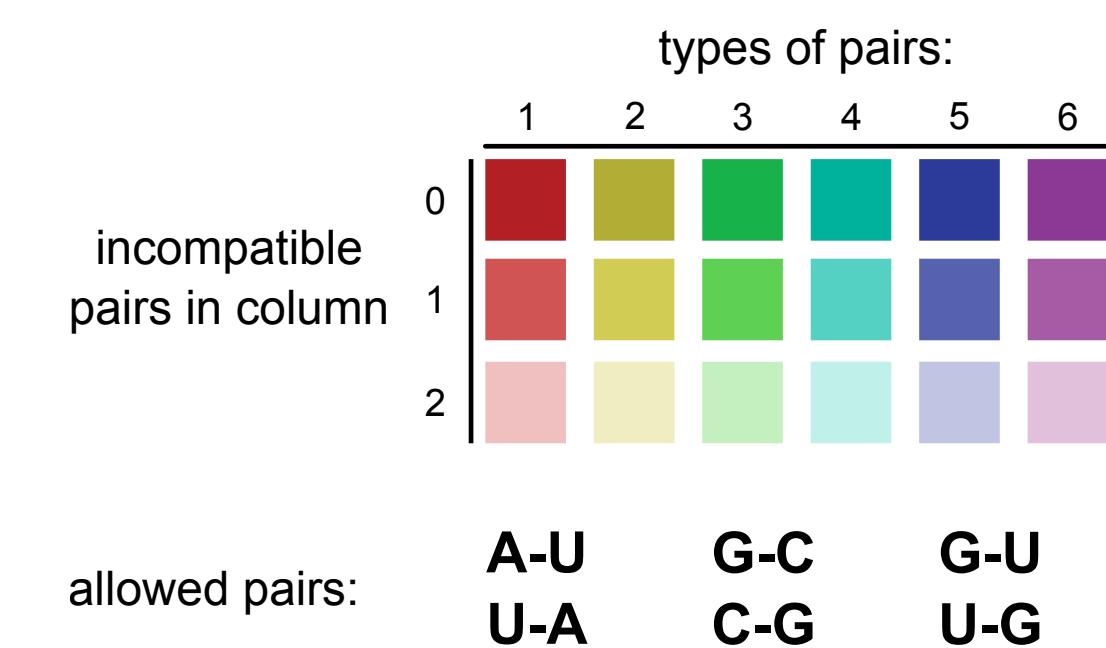
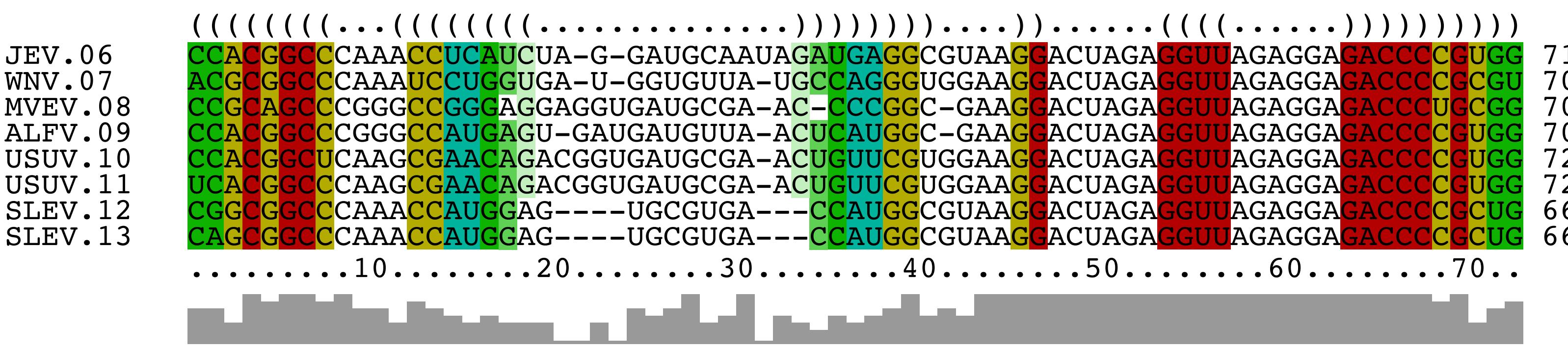
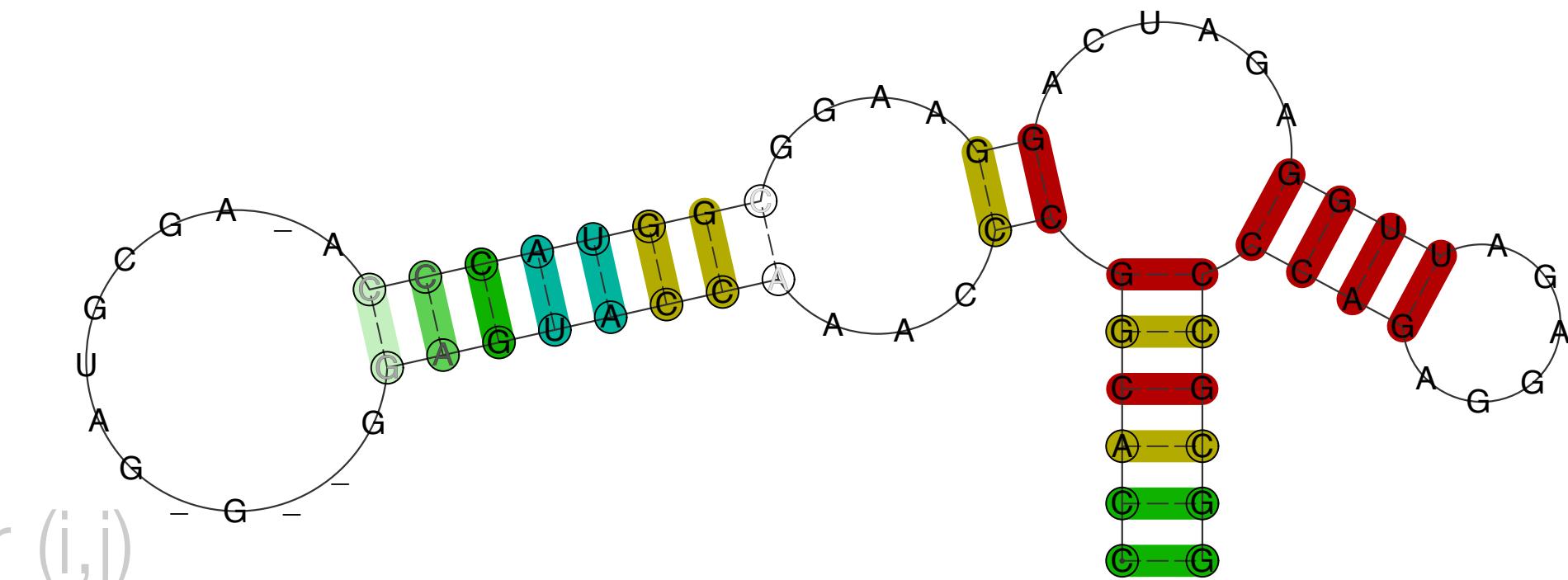


# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)

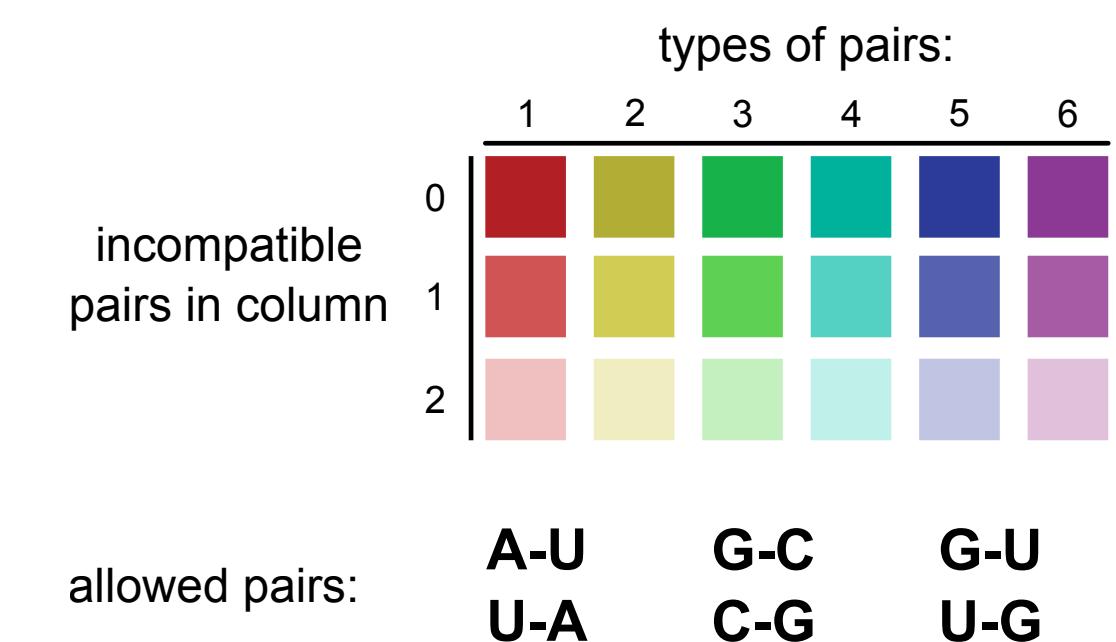
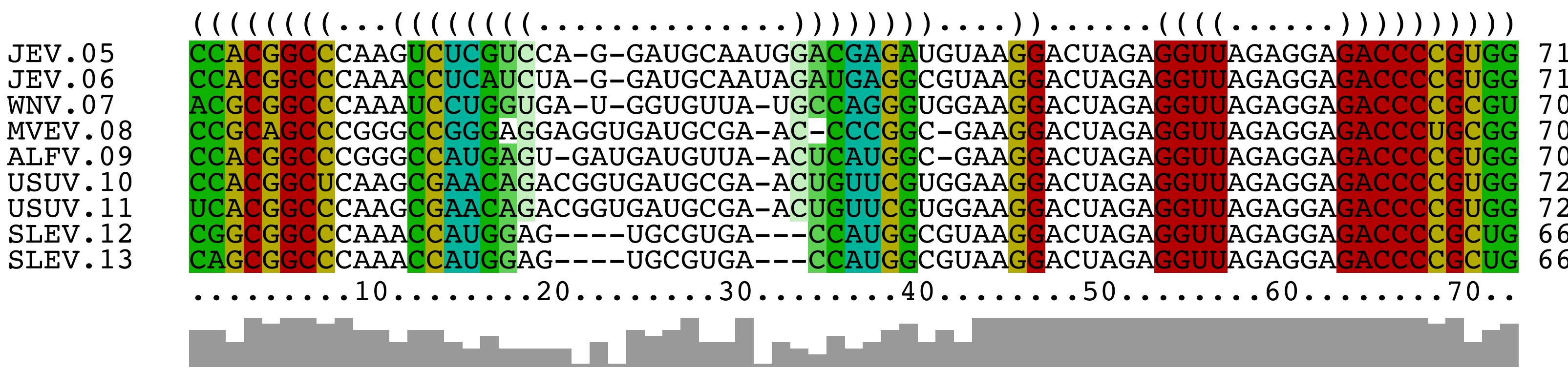
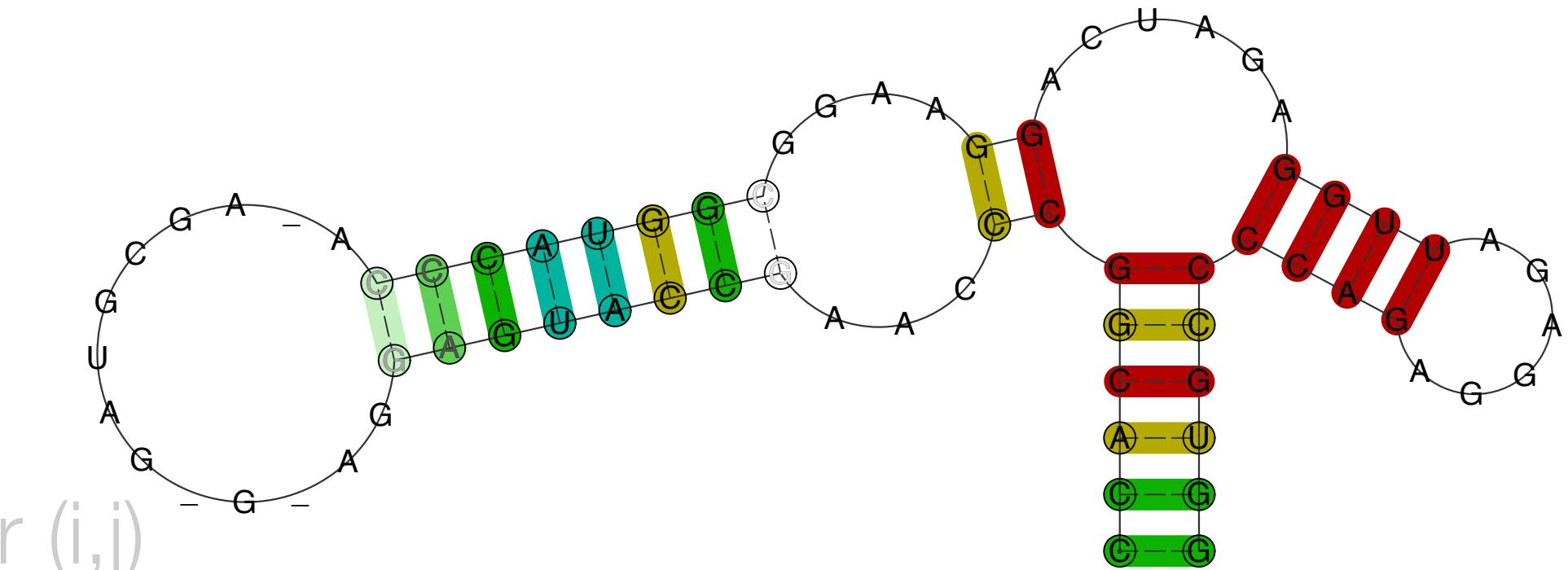


# RNA Covariation as Evolutionary Trait



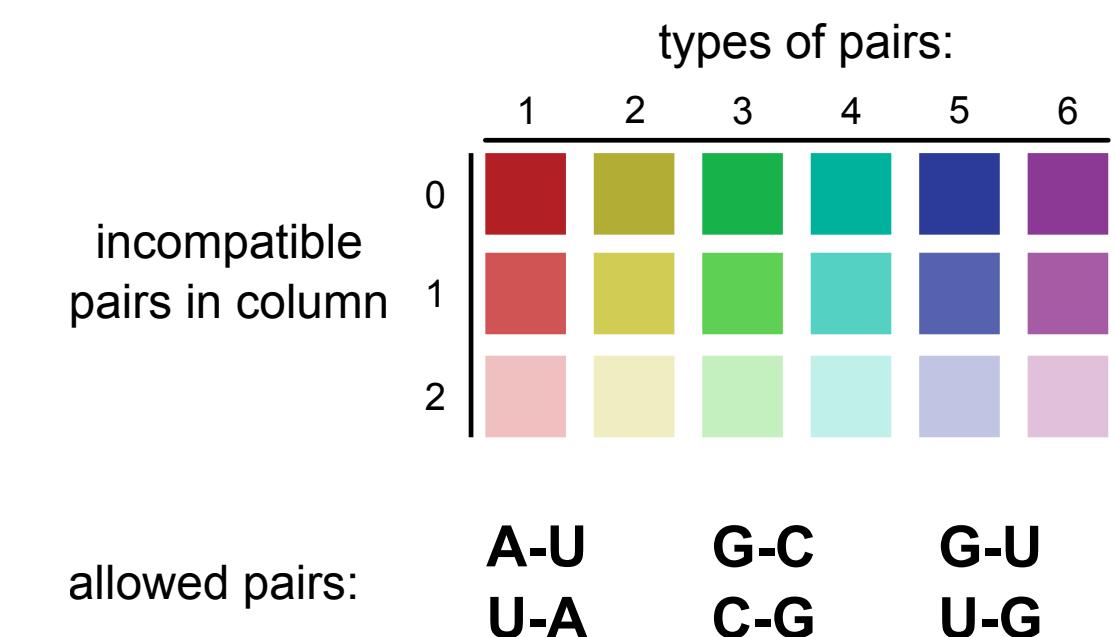
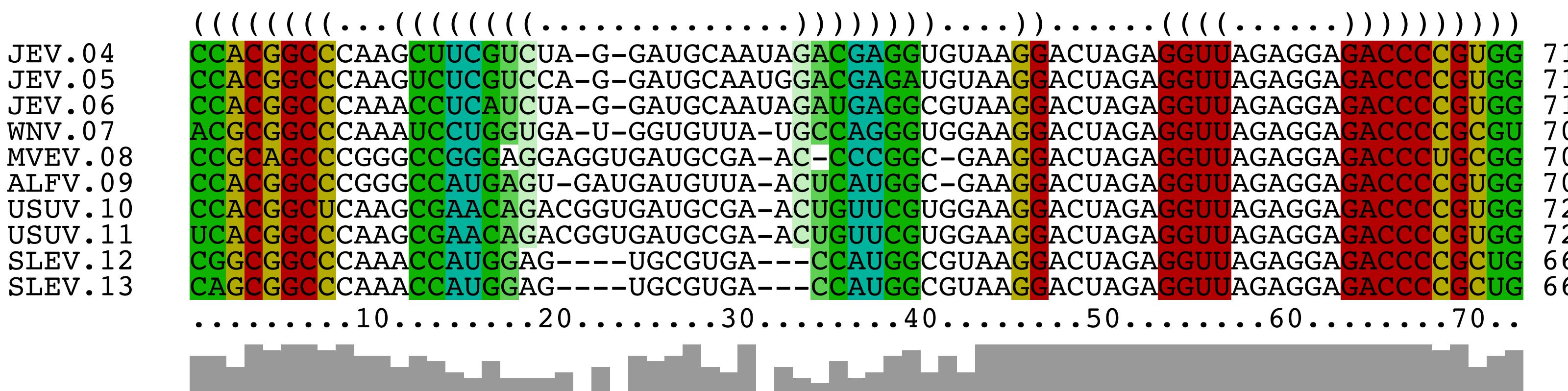
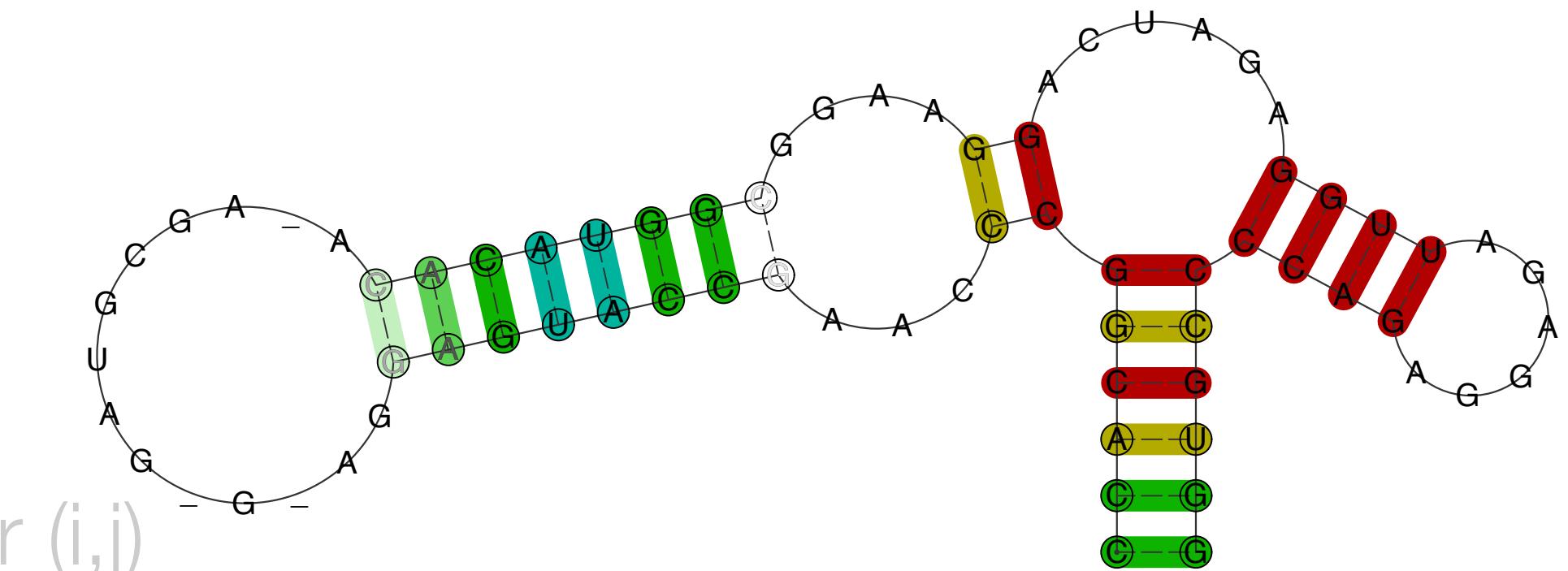
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)



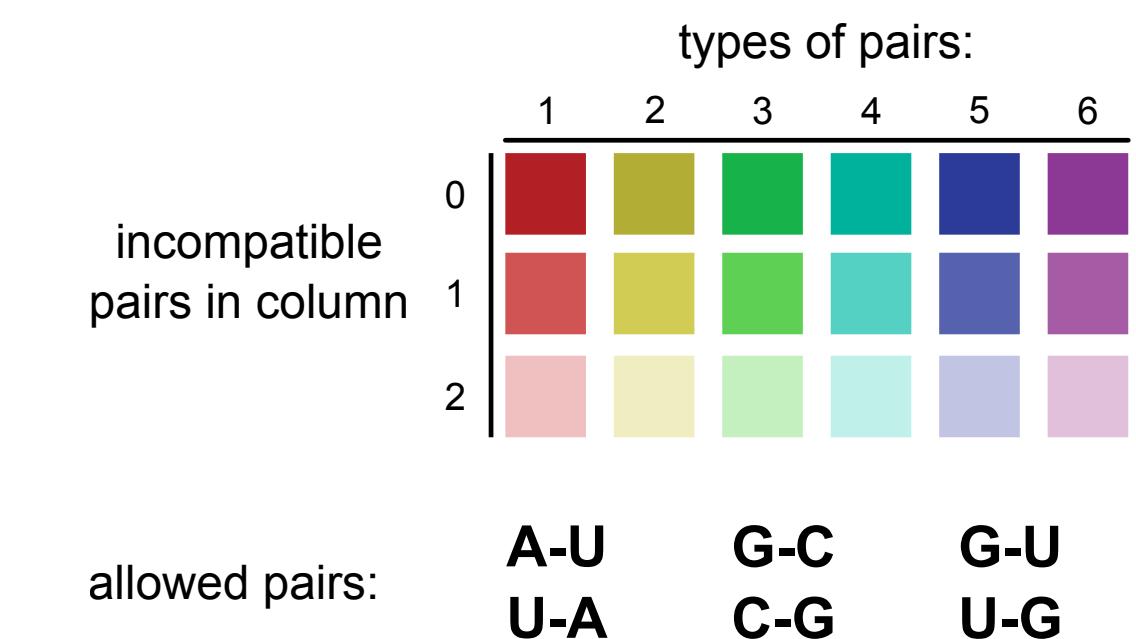
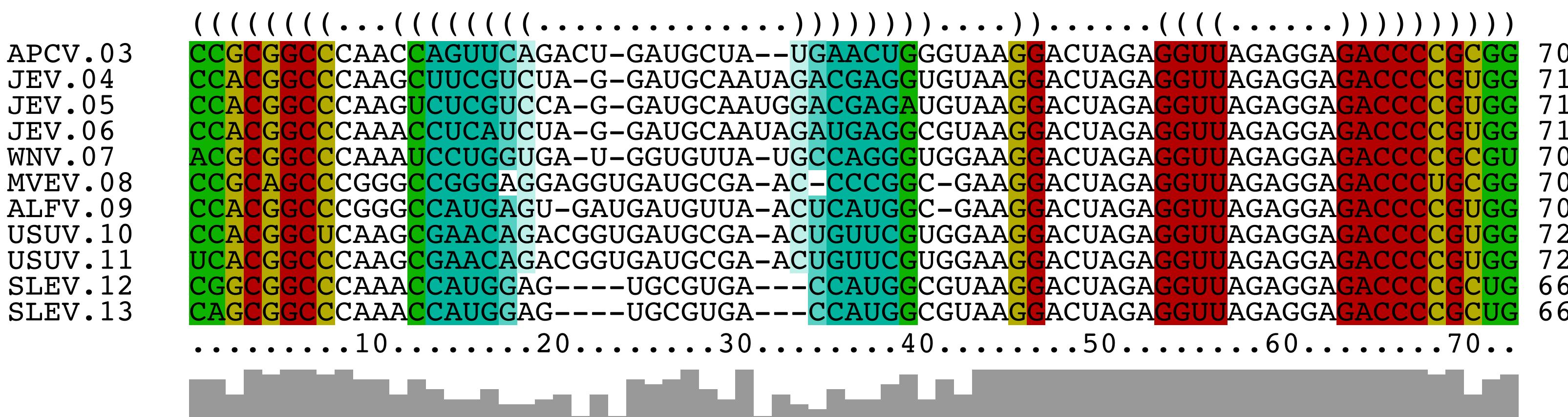
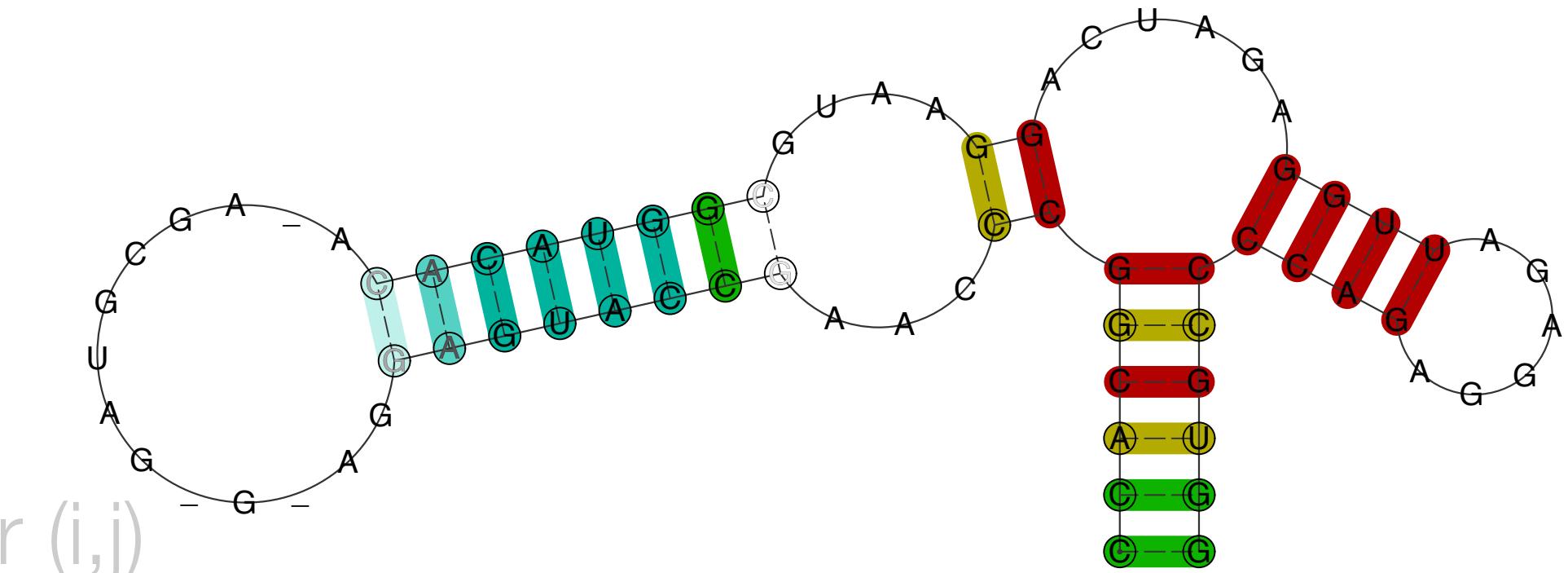
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)



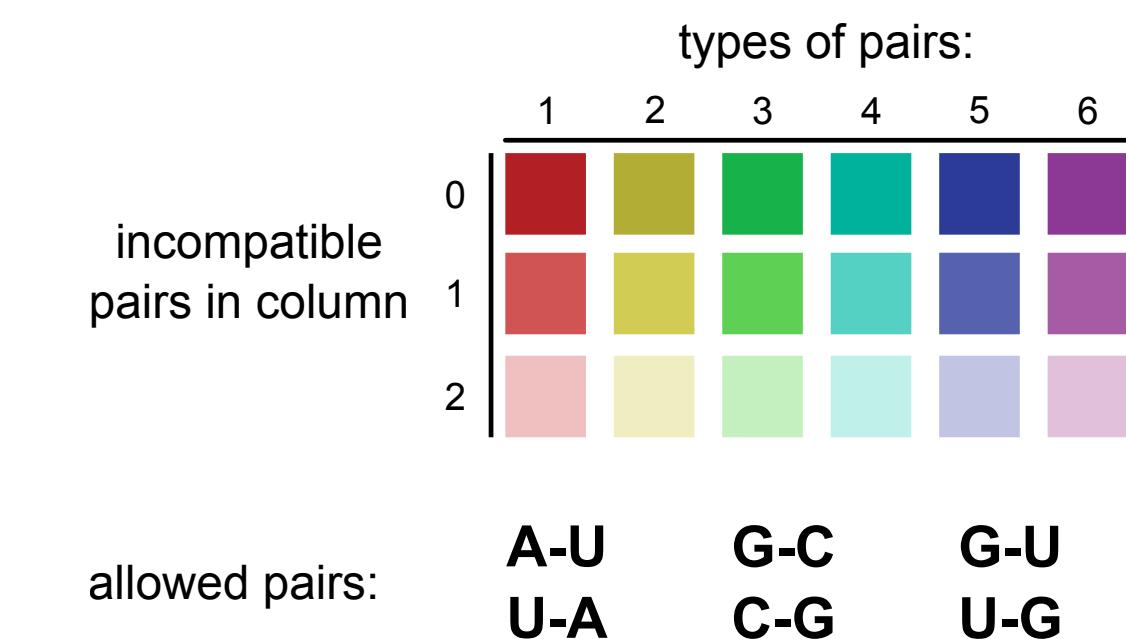
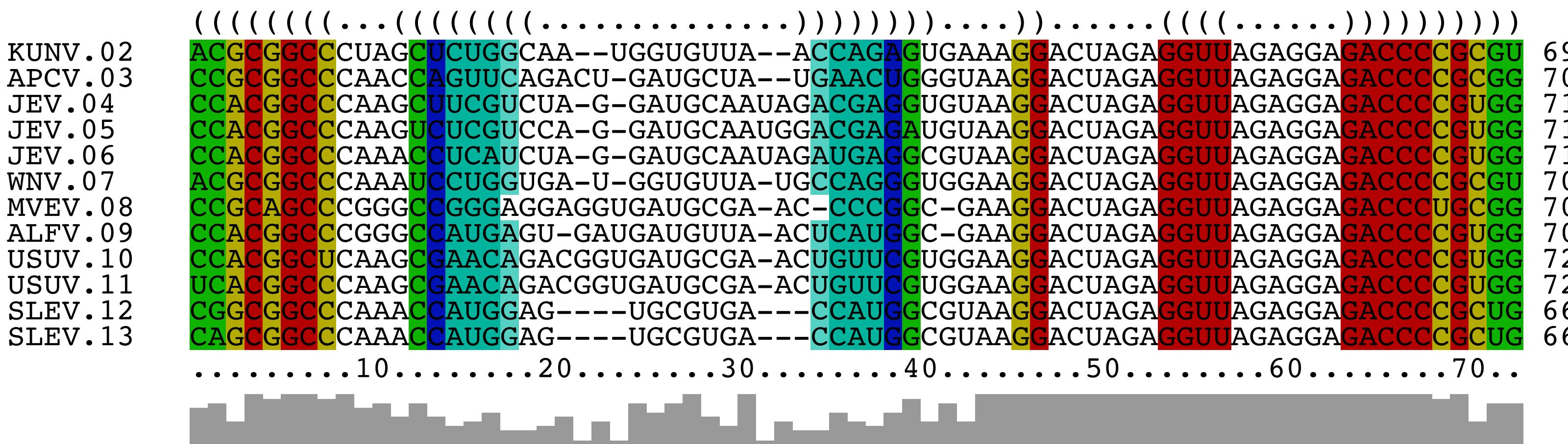
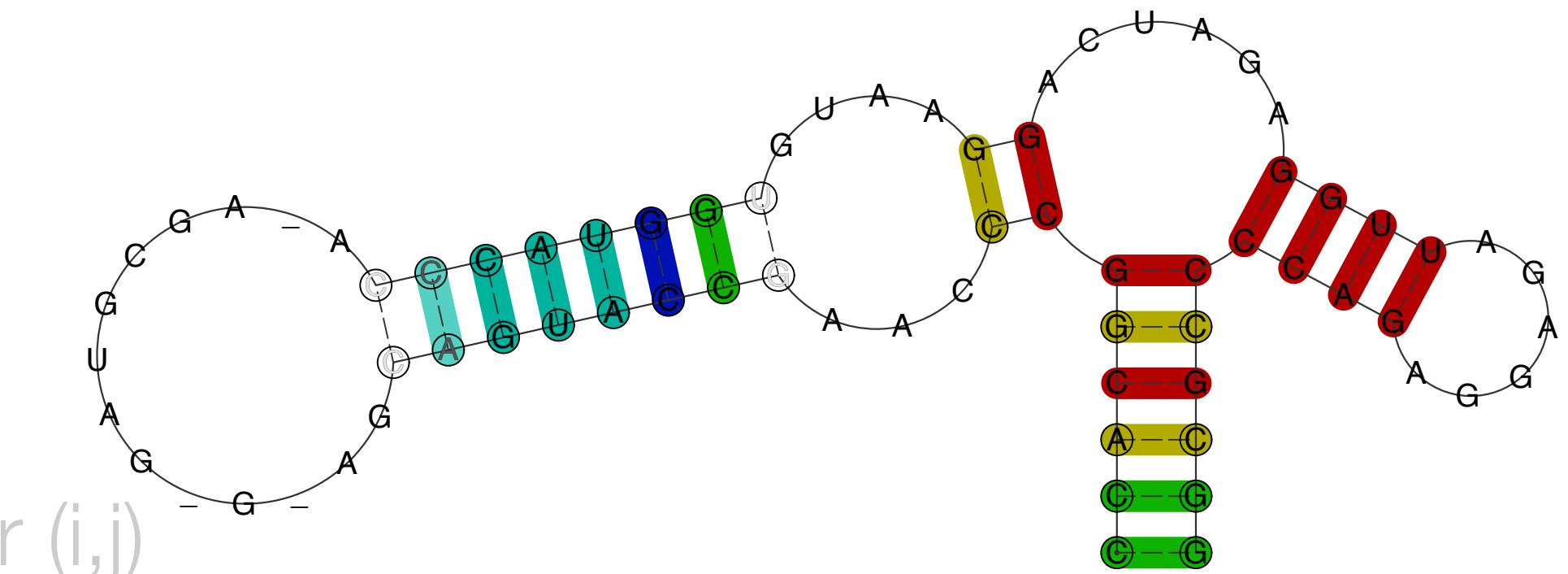
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)



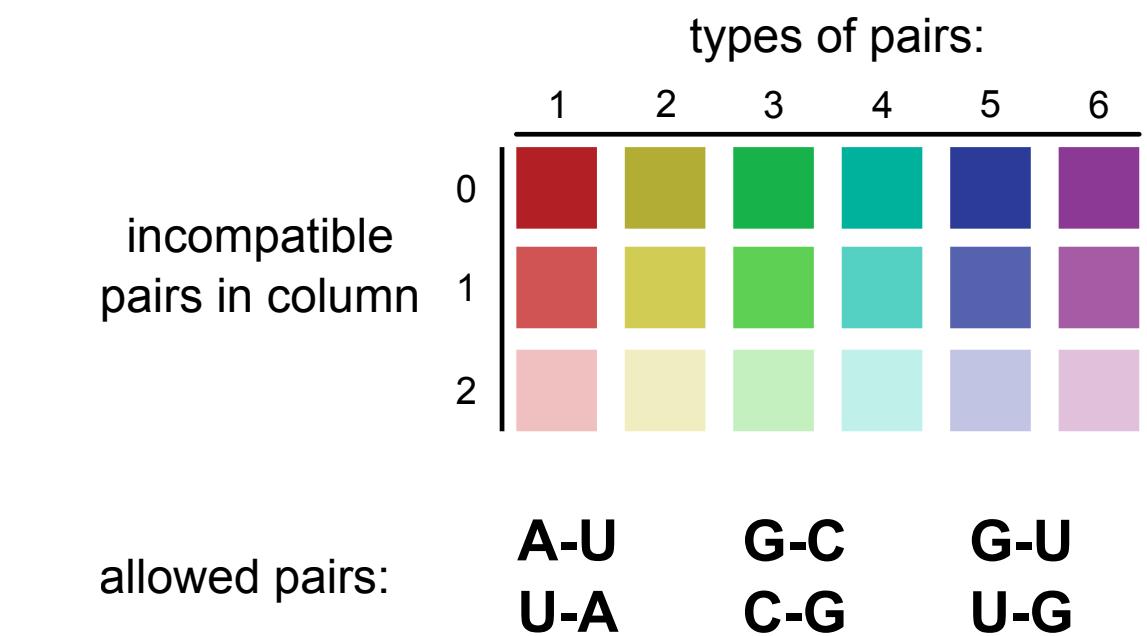
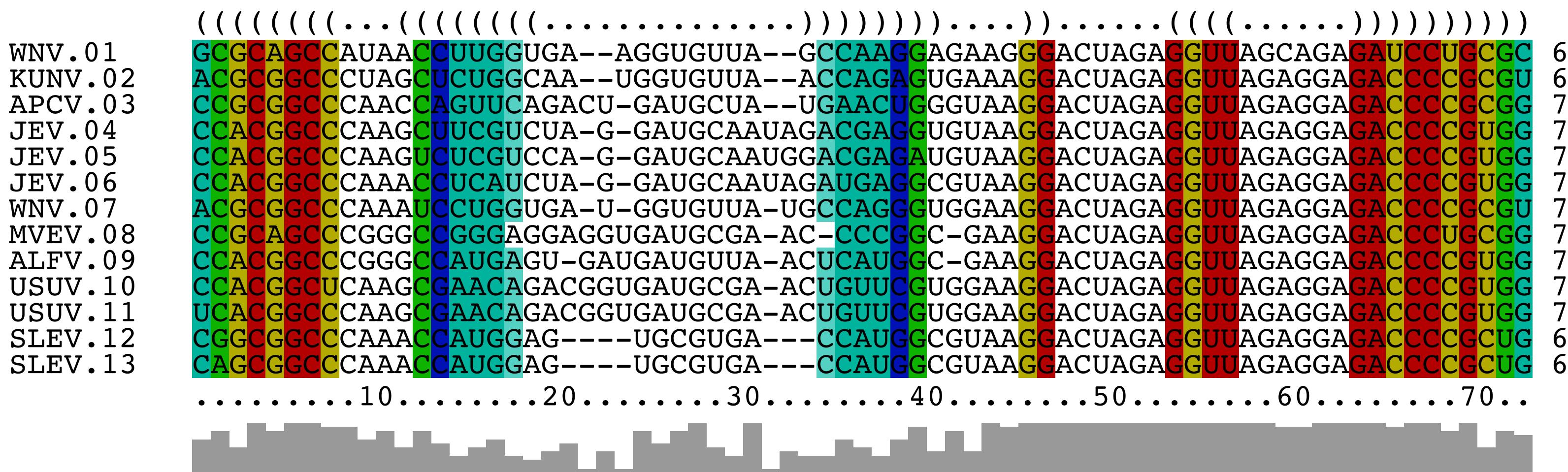
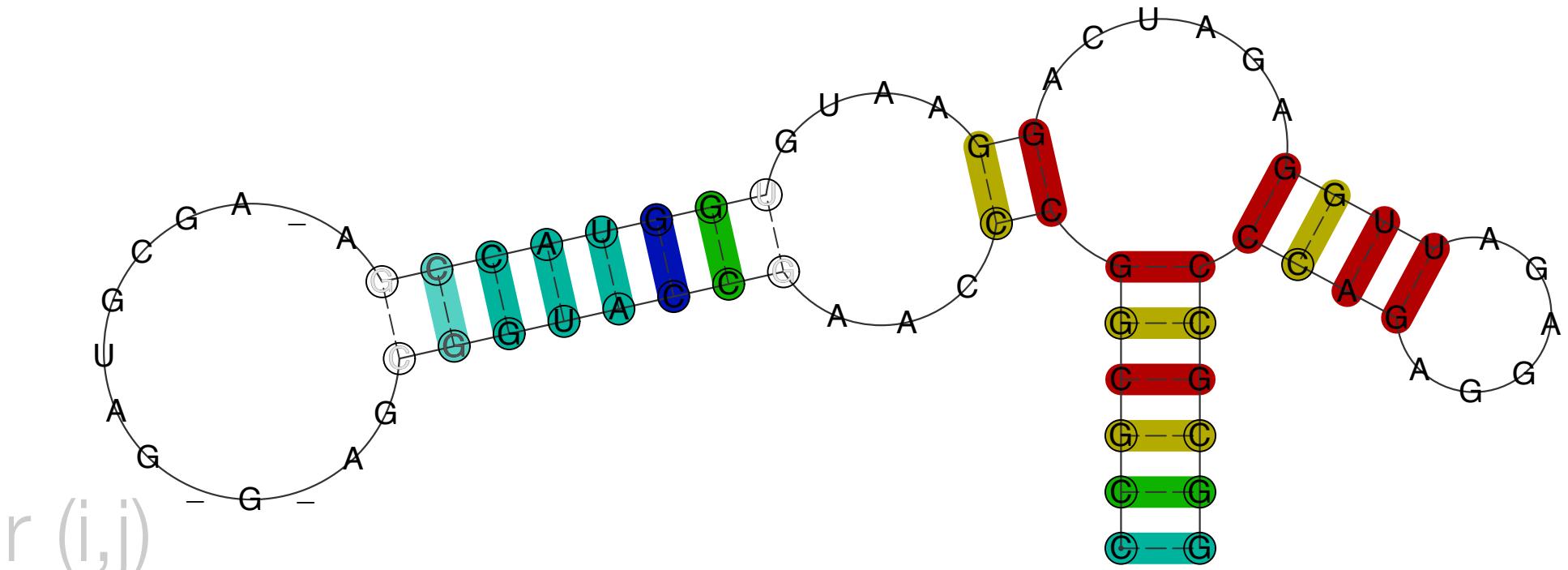
# RNA Covariation as Evolutionary Trait

- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)

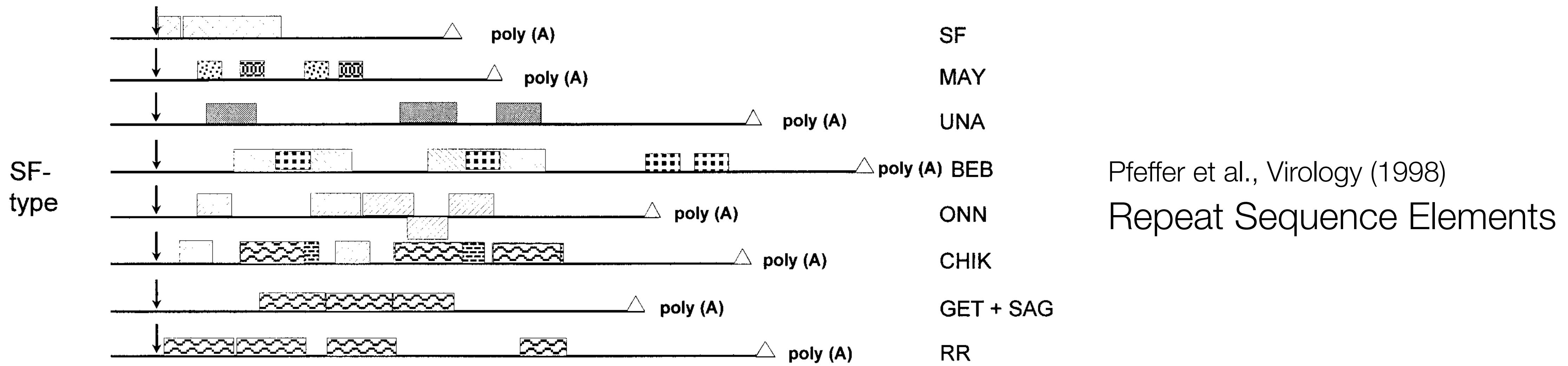


# RNA Covariation as Evolutionary Trait

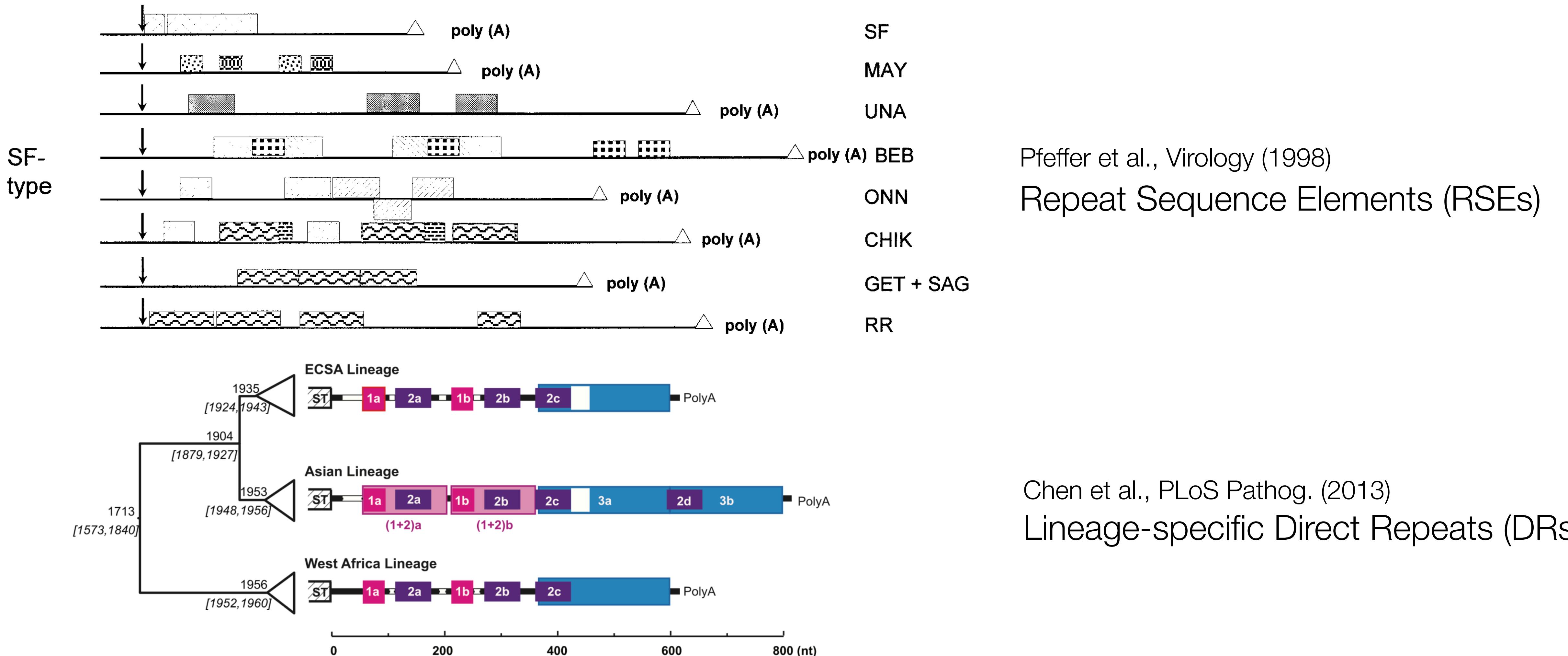
- High mutation rate in RNA viruses due to error-prone RdRP
- For base pair (i,j): GC/CG/AU/UA/GU/UG
- Consistent mutation: different standard combinations
- Compensatory mutation: both positions are mutated
- Presence of both strongly supports predicted base pair (i,j)



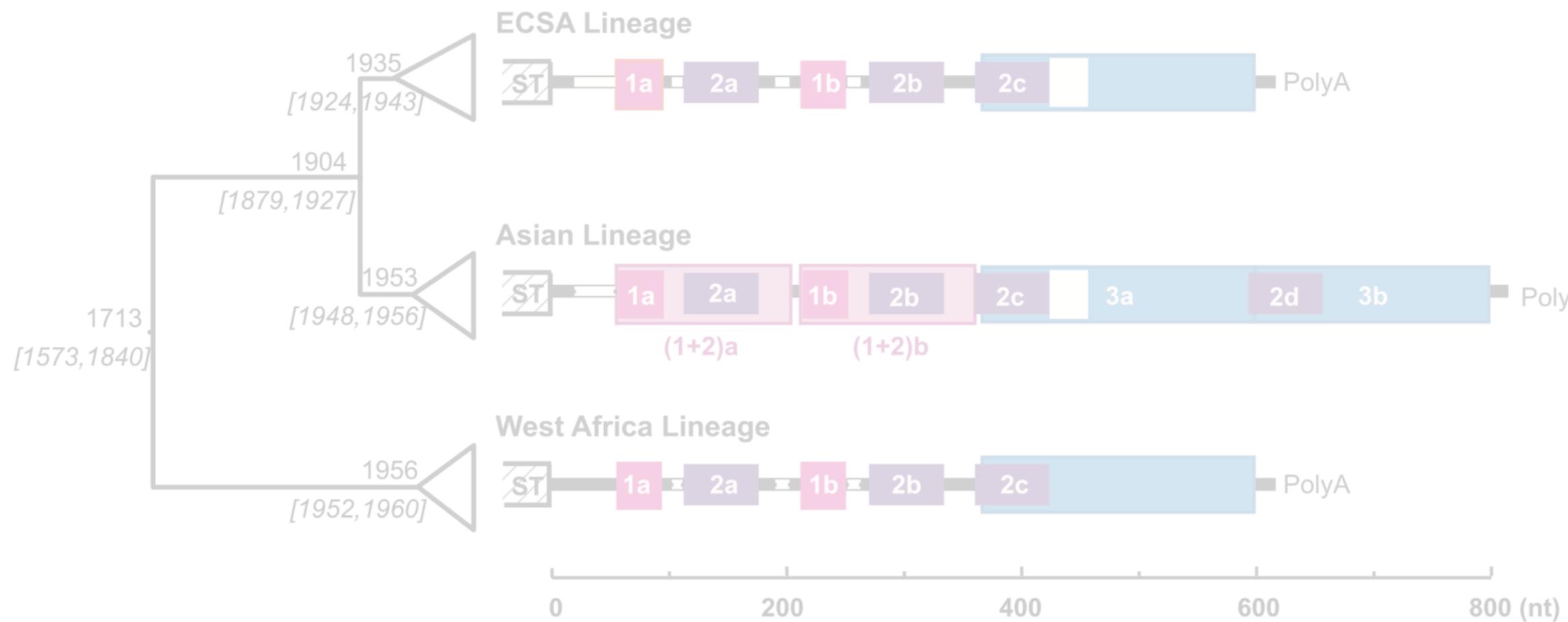
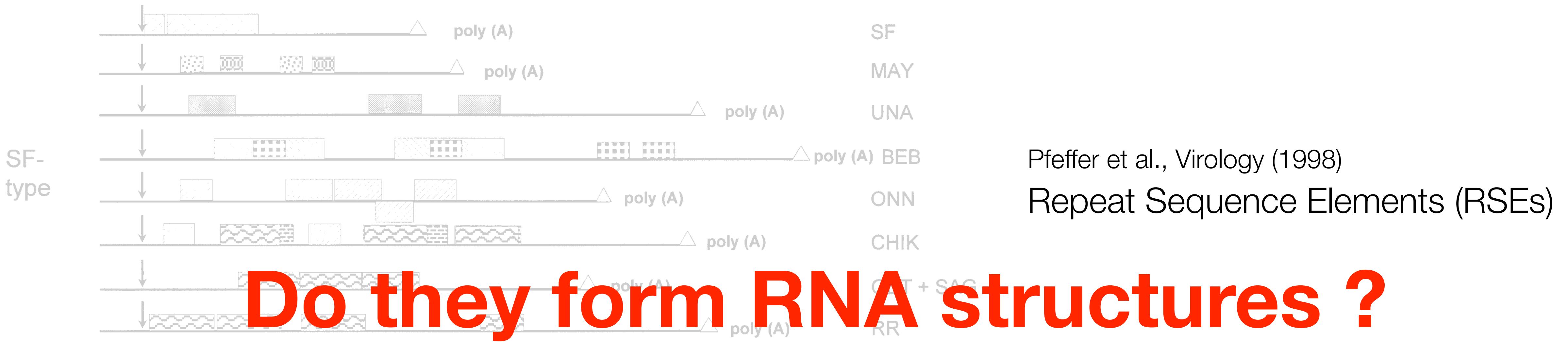
# CHIKV 3'UTR: Sequence Repeats



# CHIKV 3'UTR: Sequence Repeats



# CHIKV 3'UTR: Sequence Repeats



Chen et al., PLoS Pathog.(2013)  
Lineage-specific Direct Repeats (DRs)

# CHIKV 3'UTR



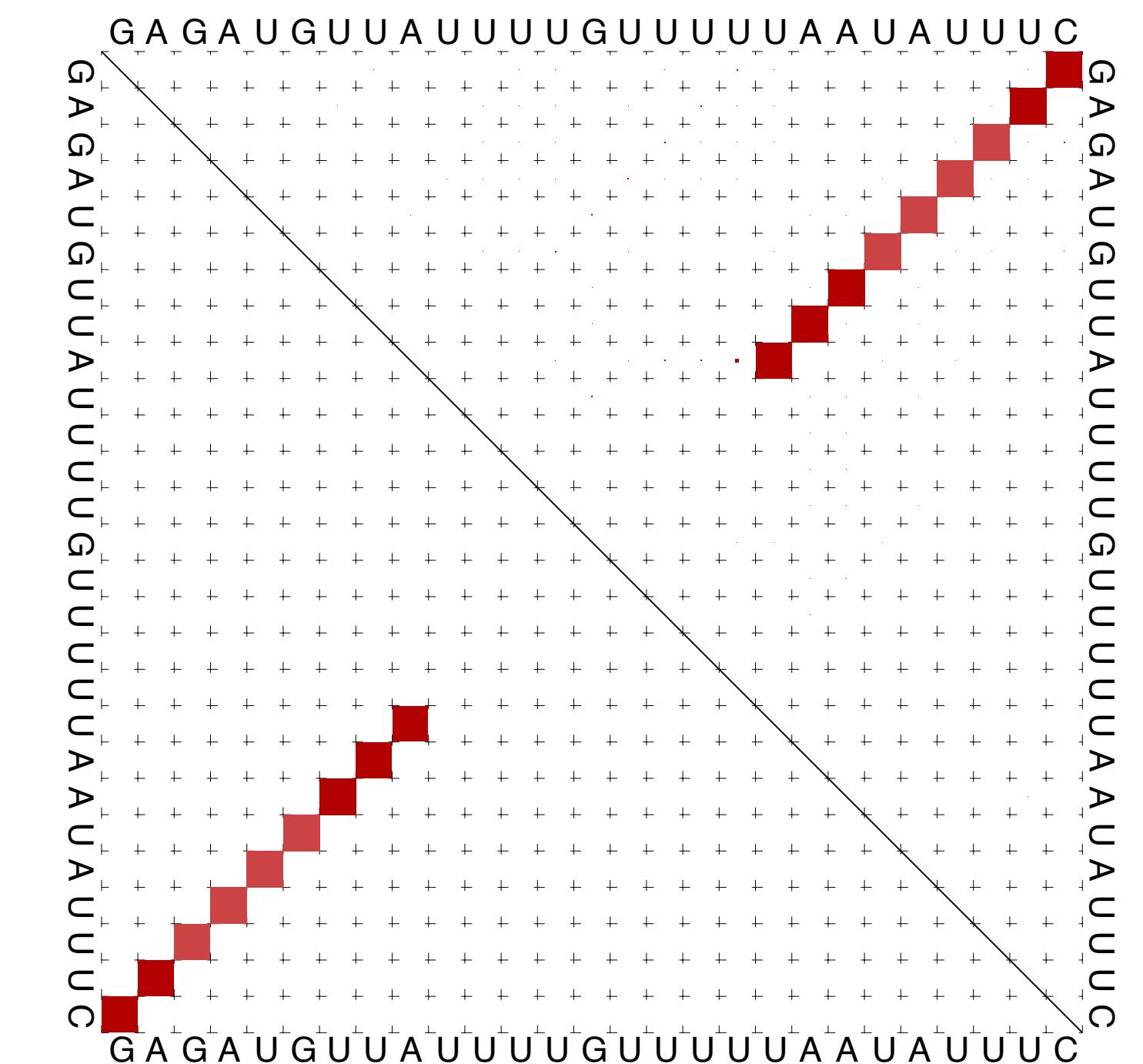
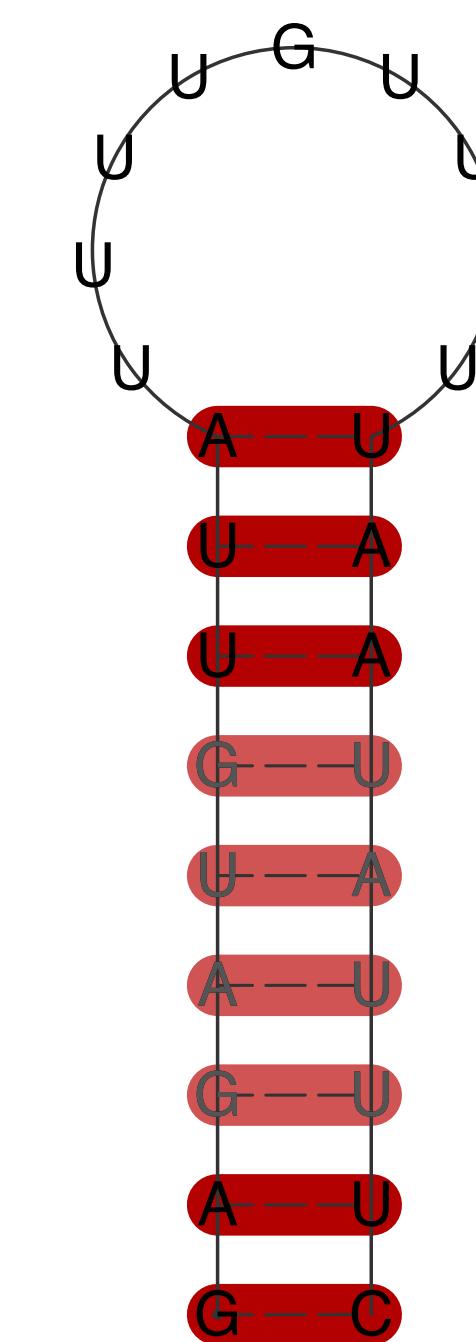
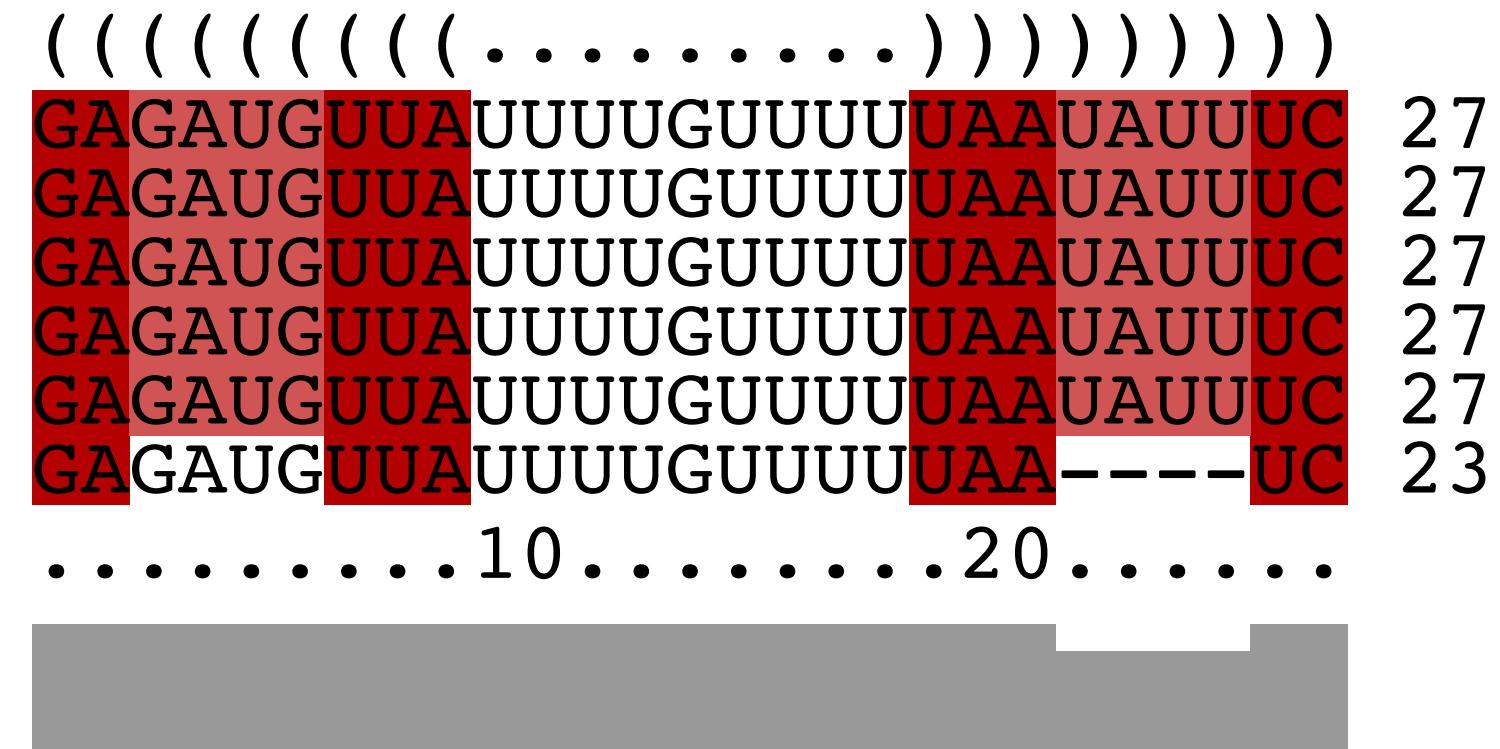
- 110 full length CHIKV 3'UTRs
- Variable length (510nt – 930nt)
- Structural alignments + Covariance models
- Thermodynamic modelling based on ViennaRNA Package

# CHIKV 3'UTR: Conserved RNA Structures

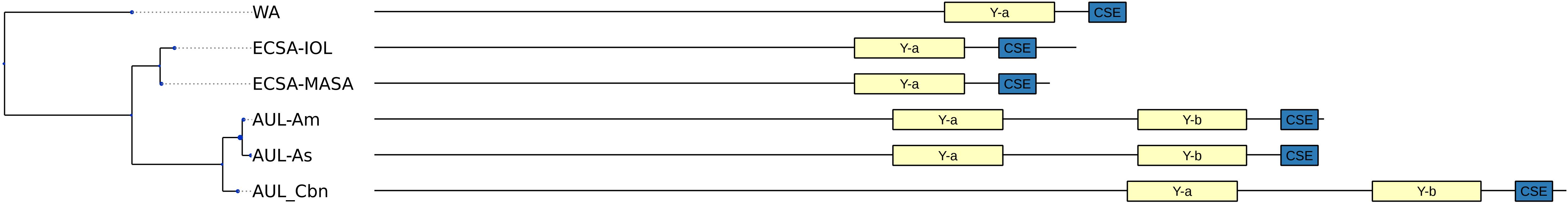


Conserved Sequence Element (CSE)

KF318729.1/688-714  
KT327163.2/862-888  
JF274082.1/472-498  
KY038946.1/472-498  
AY726732.1/540-566  
MF001517.1/685-707

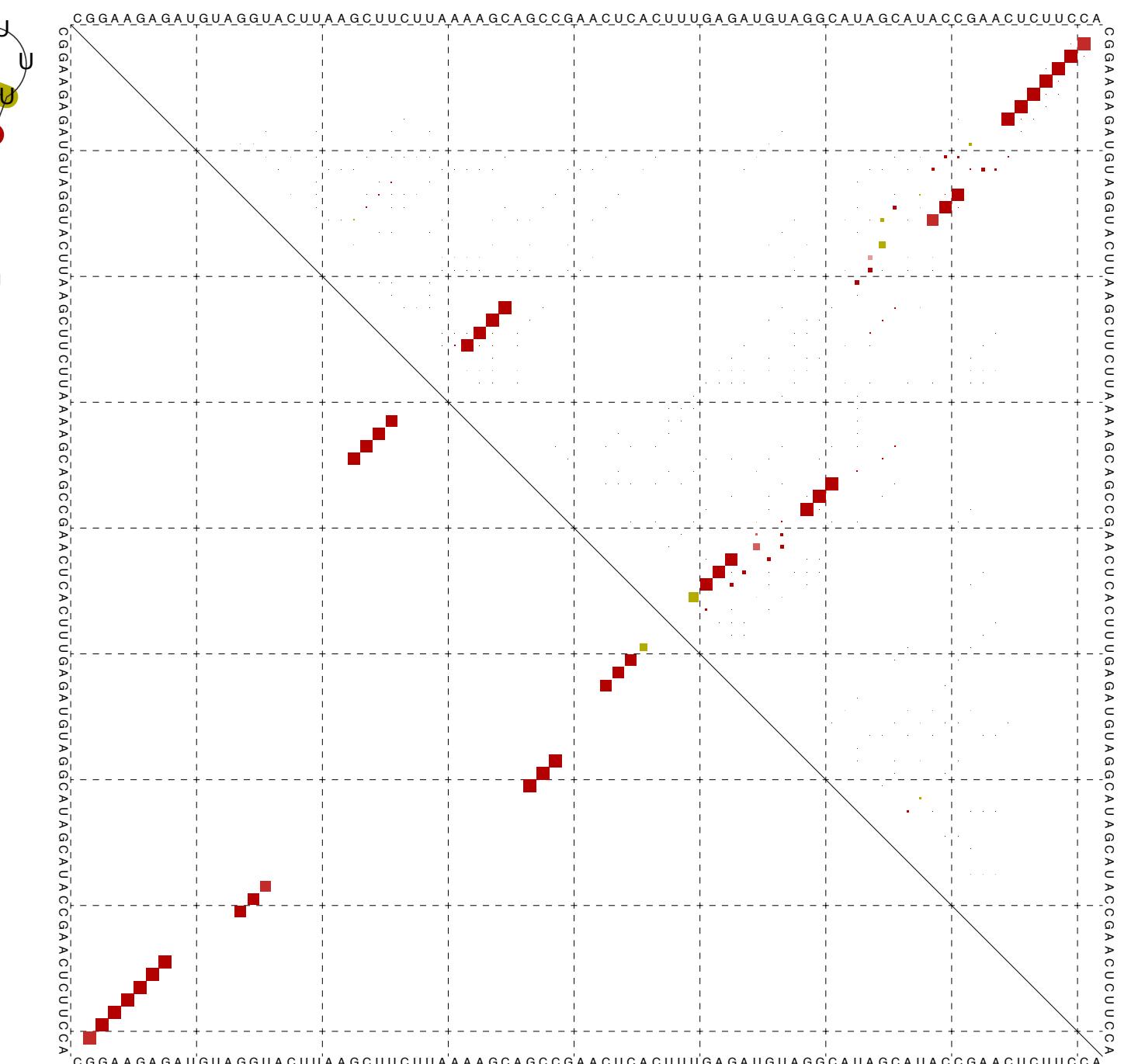
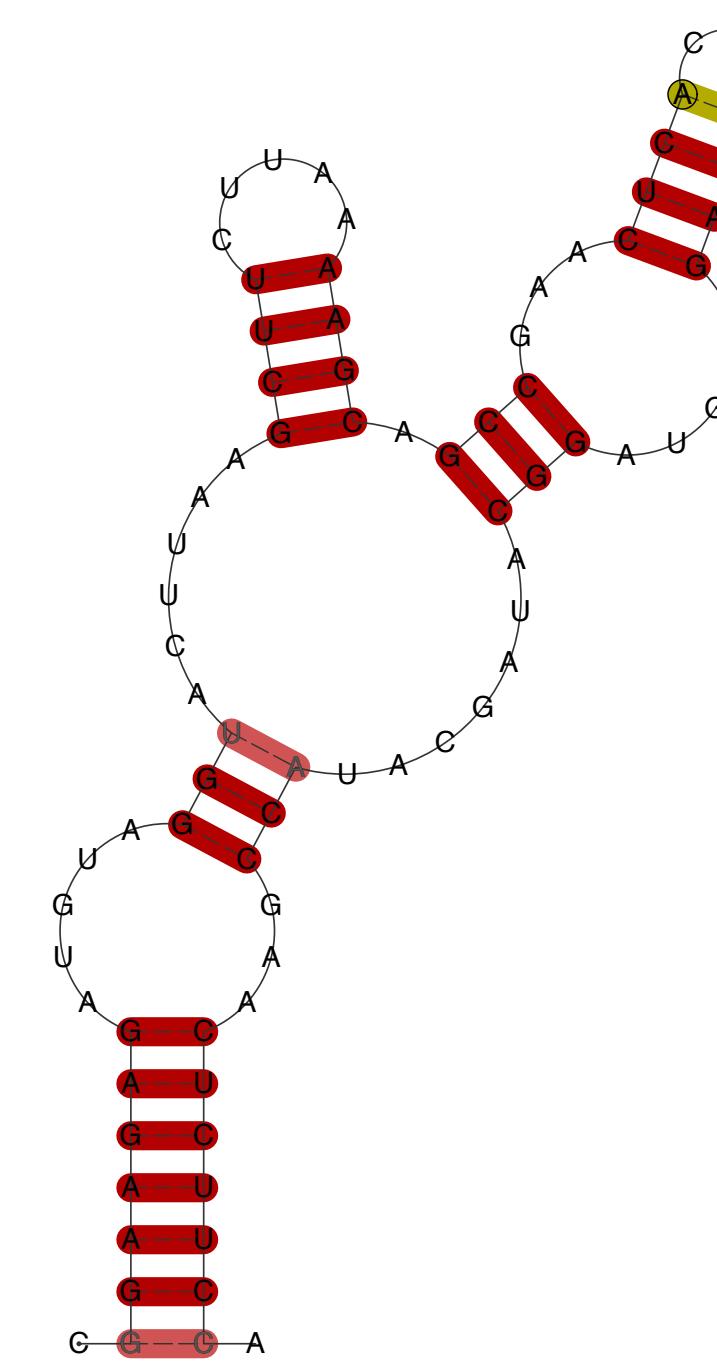
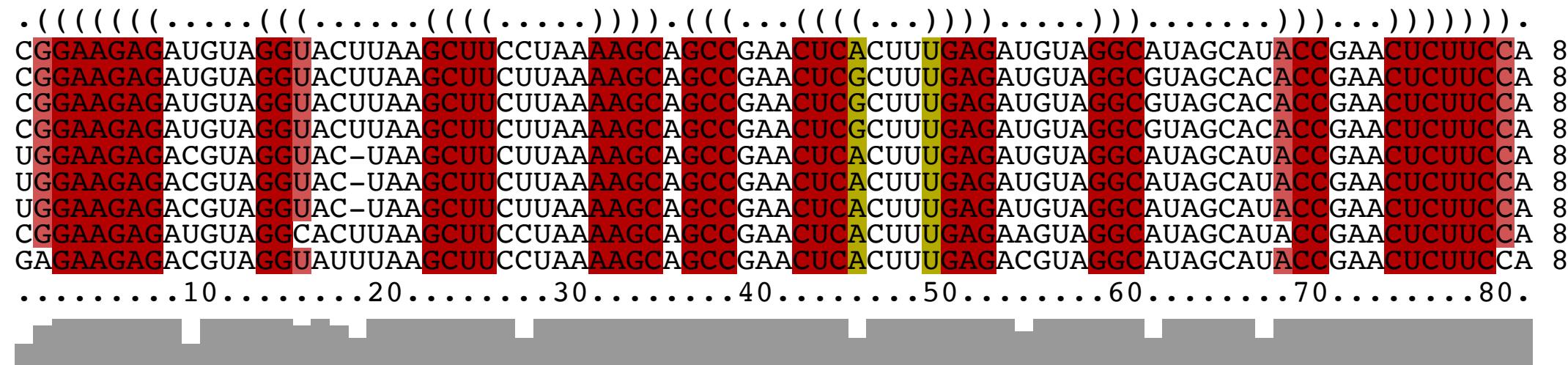


# CHIKV 3'UTR: Conserved RNA Structures

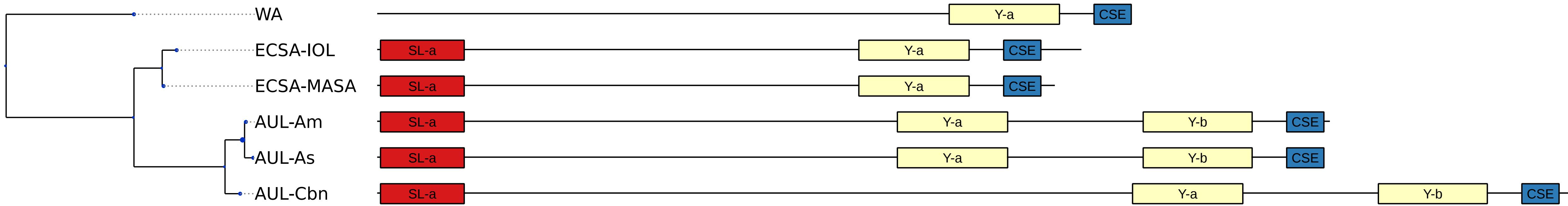


# Y-shaped Element (SL-Y)

KY038946.1/363-444  
KF318729.1/395-476  
MF001517.1/392-473  
KT327163.2/569-650  
KF318729.1/580-660  
MF001517.1/577-657  
KT327163.2/754-834  
JF274082.1/363-444  
AY726732.1/431-512

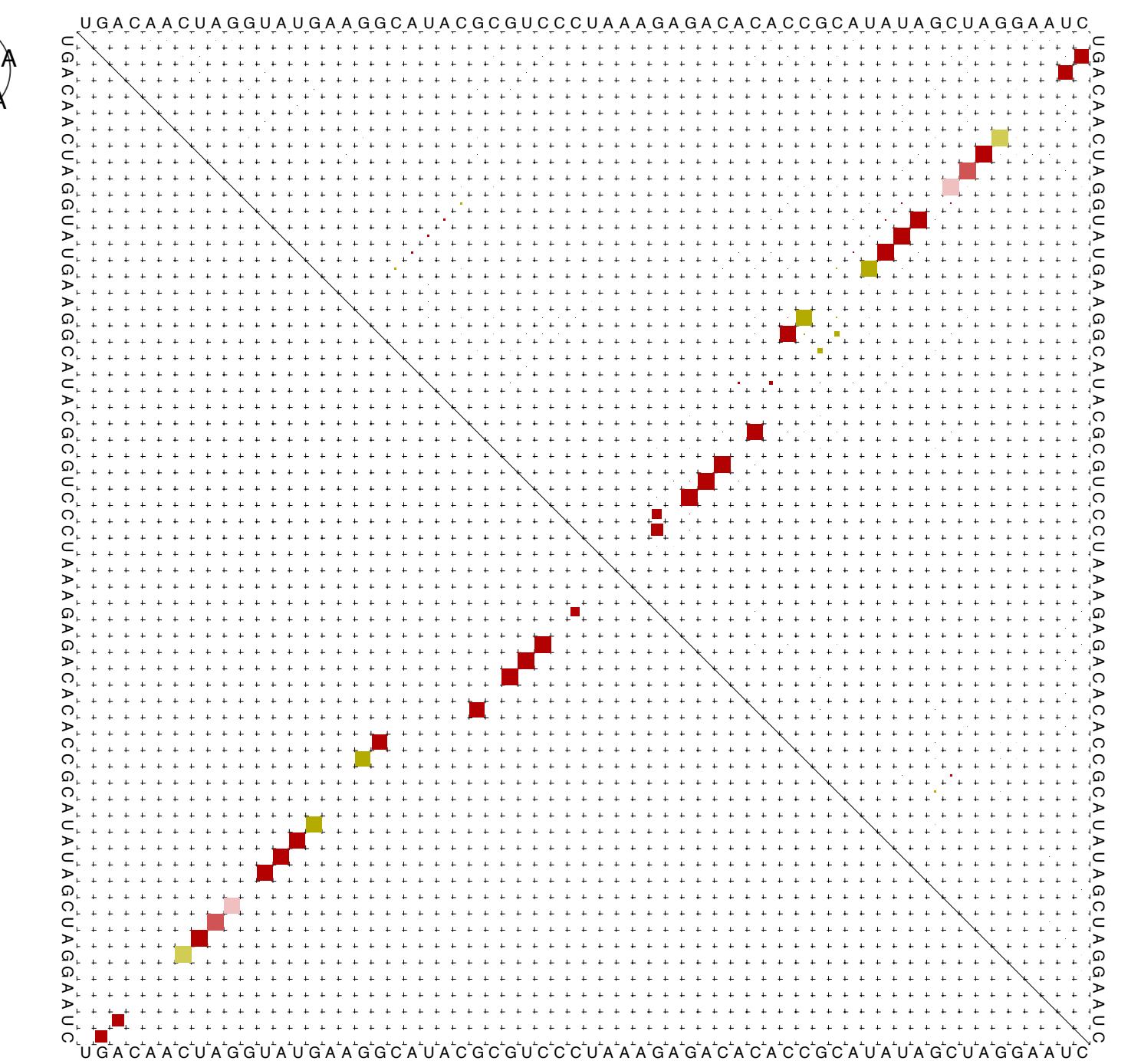
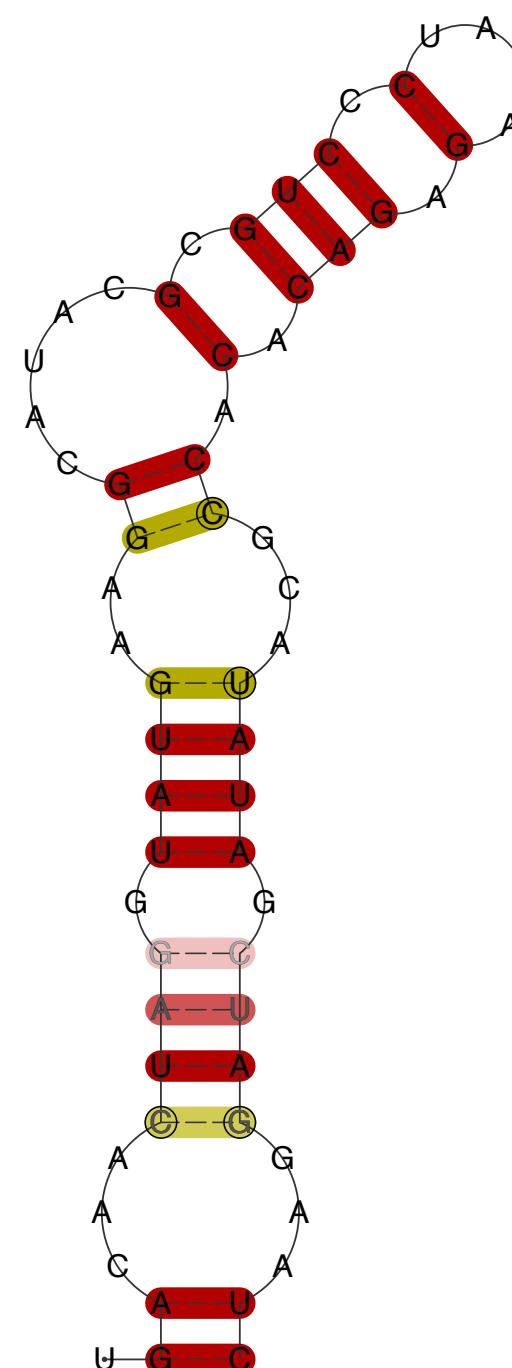


# CHIKV 3'UTR: Conserved RNA Structures

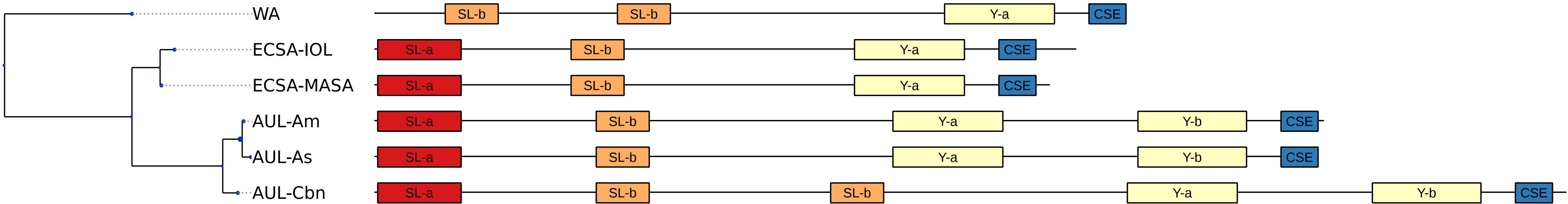


# Stem-loop a (SL-a)

KF318729.1/3-64  
MF001517.1/3-64  
KT327163.2/3-64  
JF274082.1/3-64  
KY038946.1/3-64



# CHIKV 3'UTR: Conserved RNA Structures

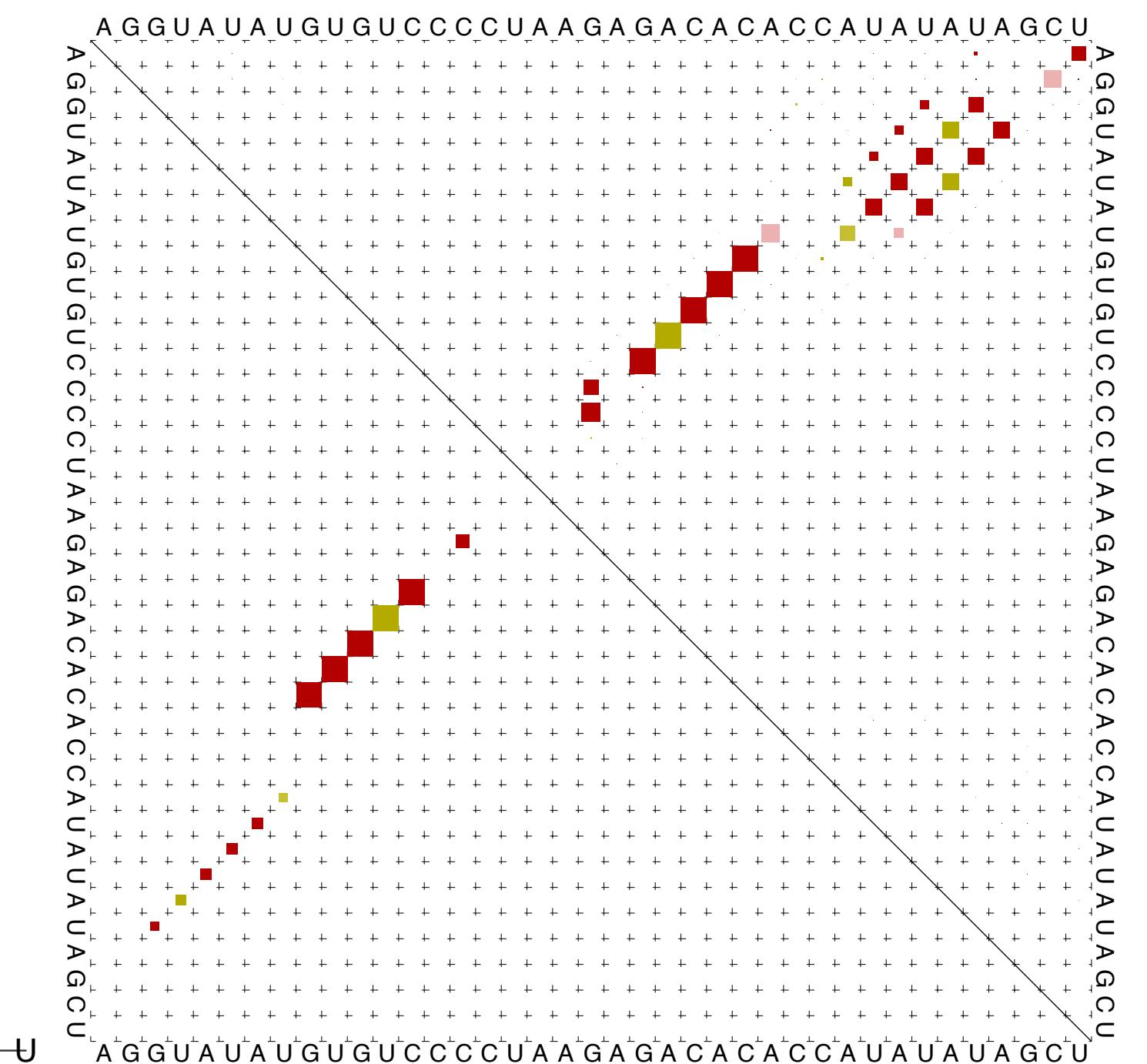
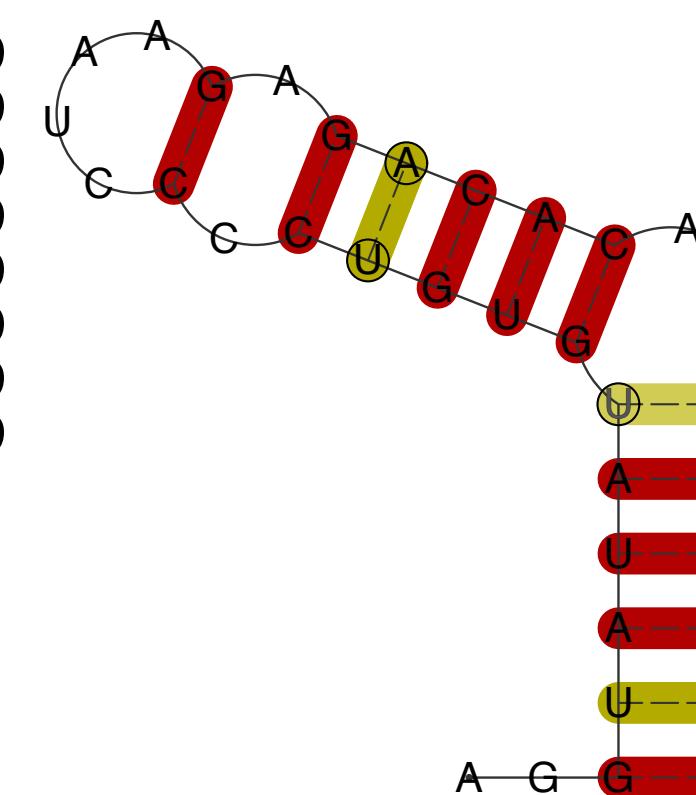


Stem-loop b (SL-b)

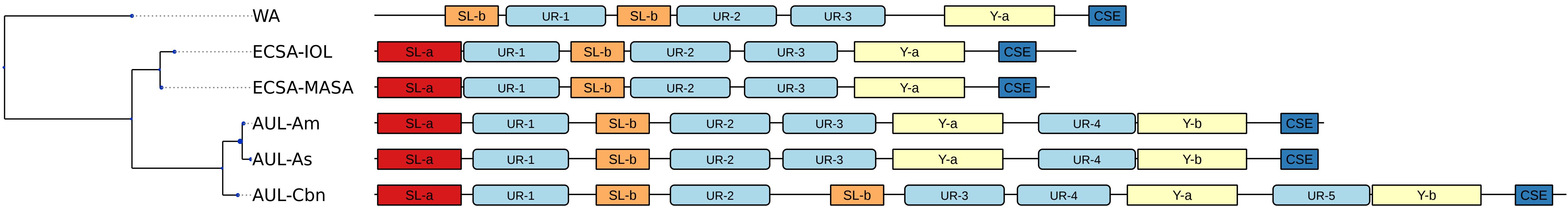
KF318729.1/171-209  
MF001517.1/168-206  
KT327163.2/168-206  
KT327163.2/345-383  
JF274082.1/149-187  
KY038946.1/149-187  
AY726732.1/184-222  
AY726732.1/54-92

..(((((((((.((....))))))))....))....)  
AGGUUAUAUGUGUCCCCUAAGGACACACUAAAUAGCU 39  
AGGUUAUAUGUGUCCCCUAAGGAACACUAAAUAGCU 39  
AGGUUAUAUGUGUCCCCUAAGGAACACUAAAUAGCU 39  
AGGUUAUAUGUGUCCCCUAAGGAACACUAAAUAGCU 39  
AGGUUAACGUGUCCCCUAAGGAACACUAAUAGGU 39  
AGGUUAACGUGUCCCCUAAGGAACACUAAUAGGU 39  
AGGUUAAAGUGUCCCCUAAGGAACACUAAUAGGU 39  
AGGUUAAAGUGUCCCCUAAGGAACACUAAUAGGU 39

.....10.....20.....30.....



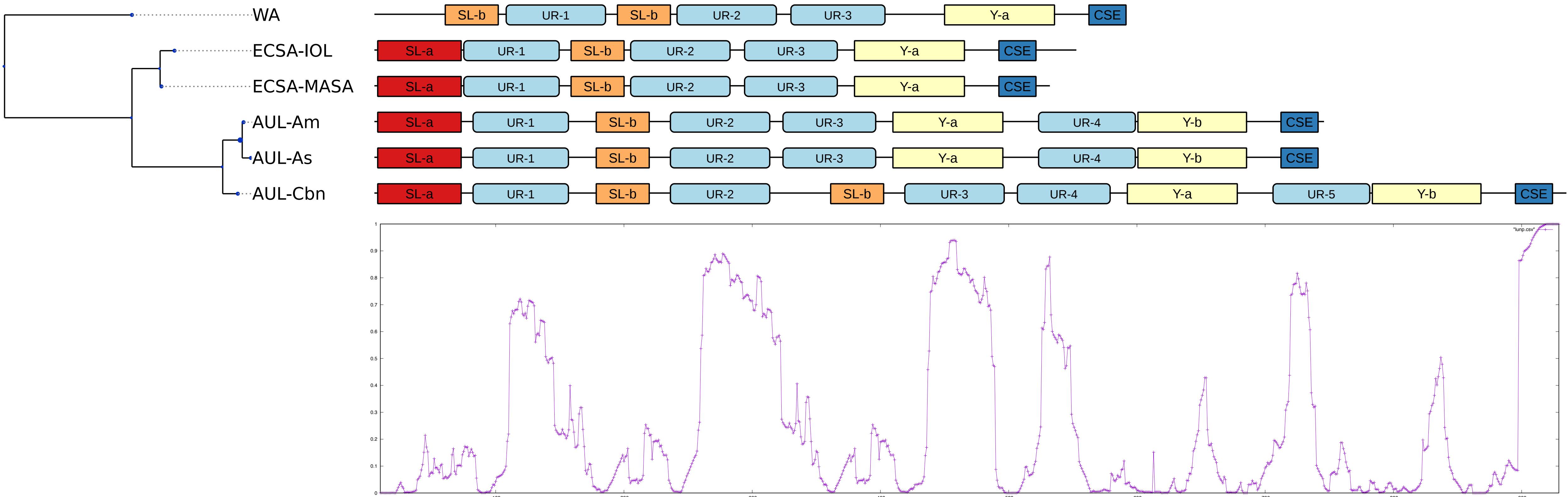
# CHIKV 3'UTR: Conserved RNA Elements



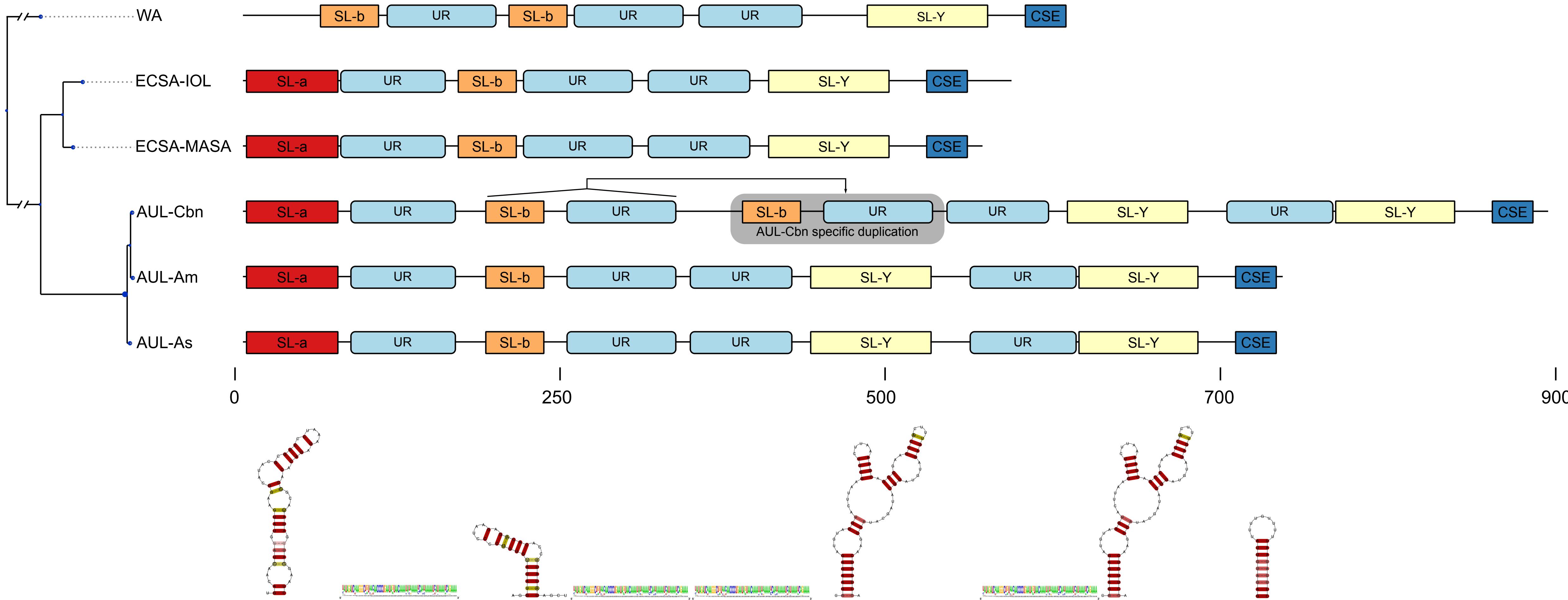
Unstructured repeat (UR)



# CHIKV-Cbn Accessibility

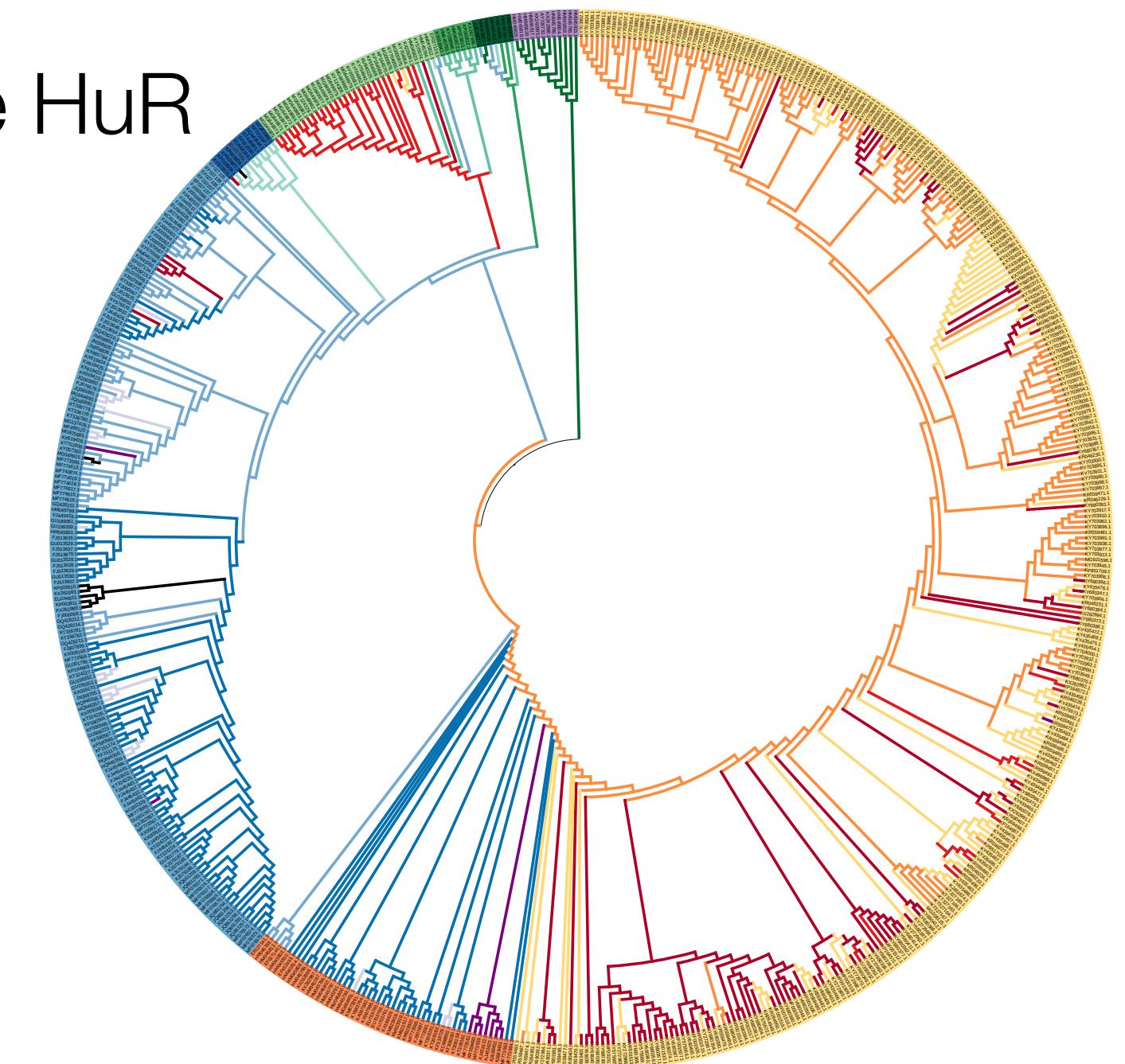
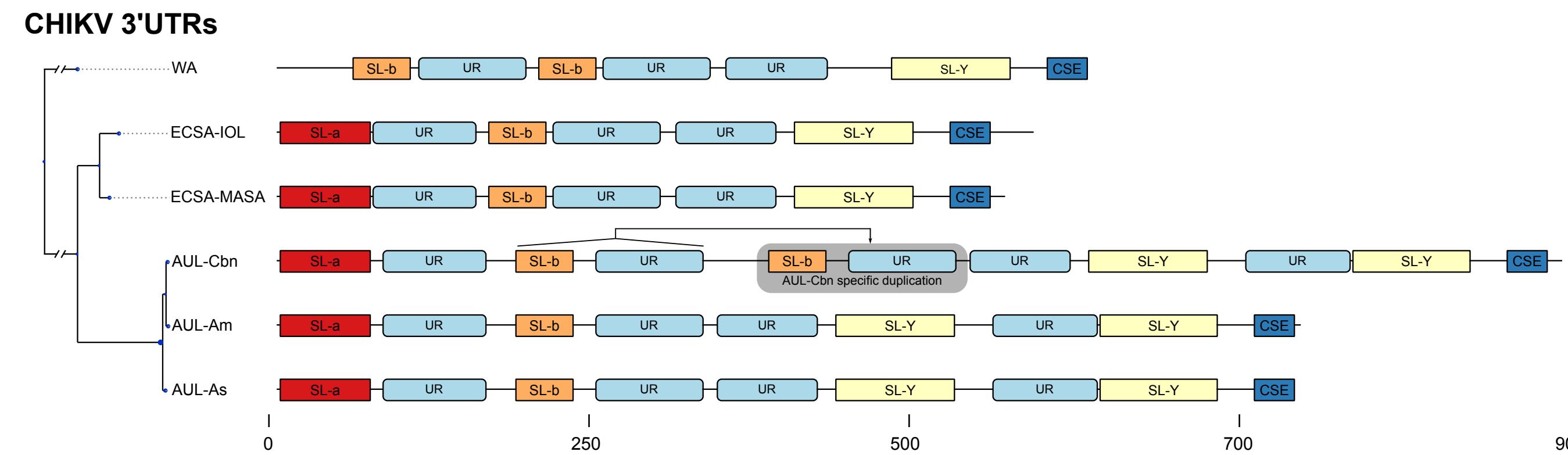


# CHIKV 3'UTR Overview



# Summary

- Lineage-specific non-coding RNA architecture
- Conclusive mapping of RSEs / DRs to structured and unstructured regions
- Individual lineages show varied replication potential in mosquito / vertebrate cells
- CHIKV 3'UTR stem-loops do not stall exoribonucleases
- A-rich unstructured repeats are likely sponge RNA-binding proteins like HuR
- Whether or not miRNAs are encoded needs to be established



# Acknowledgements

- Roman Ochsenreiter (TBI)
- Ivo L. Hofacker (TBI)
- Adriano de Bernardi Schneider (UCSD)

## Updated phylogeny of Chikungunya virus reveals lineage-specific non-coding RNAs

**Adriano de Bernardi Schneider<sup>1,†</sup>, Roman Ochsenreiter<sup>2,†</sup>, Reilly Hostager<sup>1</sup>,  
Ivo L. Hofacker<sup>2,3</sup>, Daniel Janies<sup>4</sup>, and Michael T. Wolfinger<sup>2,3,✉</sup>**

<sup>1</sup>AntiViral Research Center, Department of Medicine, University of California San Diego, San Diego, 92103, United States of America

<sup>2</sup>Department of Theoretical Chemistry, University of Vienna, Währingerstraße 17, 1090 Vienna, Austria

<sup>3</sup>Research Group BCB, Faculty of Computer Science, University of Vienna, Währingerstraße 29, 1090 Vienna, Austria

<sup>4</sup>Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, 28223, United States of America

<sup>†</sup>These authors have contributed equally to this work.



**bioRxiv**  
THE PREPRINT SERVER FOR BIOLOGY