RNA structuredness of viral genomes

Michael T. Wolfinger

Research Group Bioinformatics and Computational Biology
University of Vienna
Austria

Computational Approaches to RNA Structure and Function
Benasque
12 August 2022

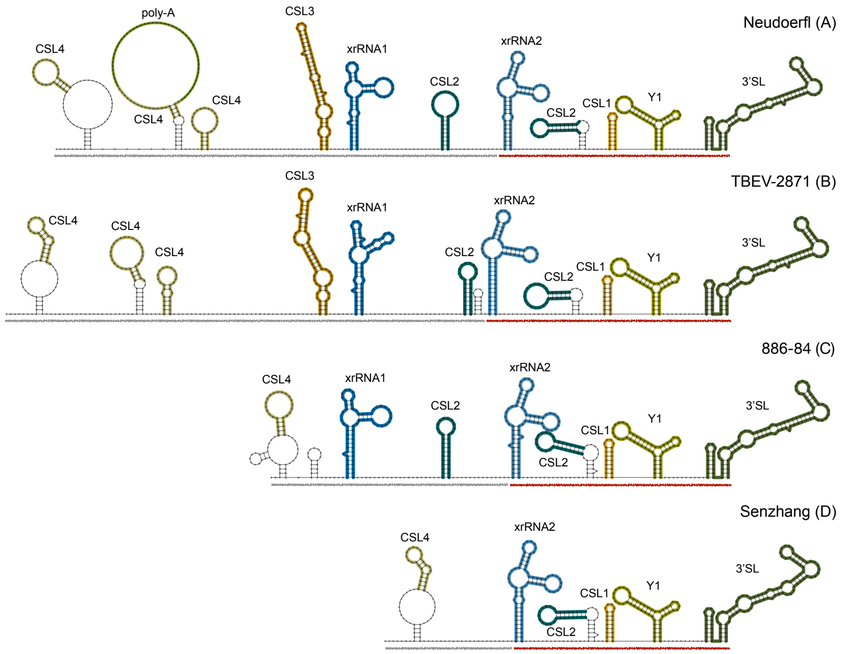
universität
wien

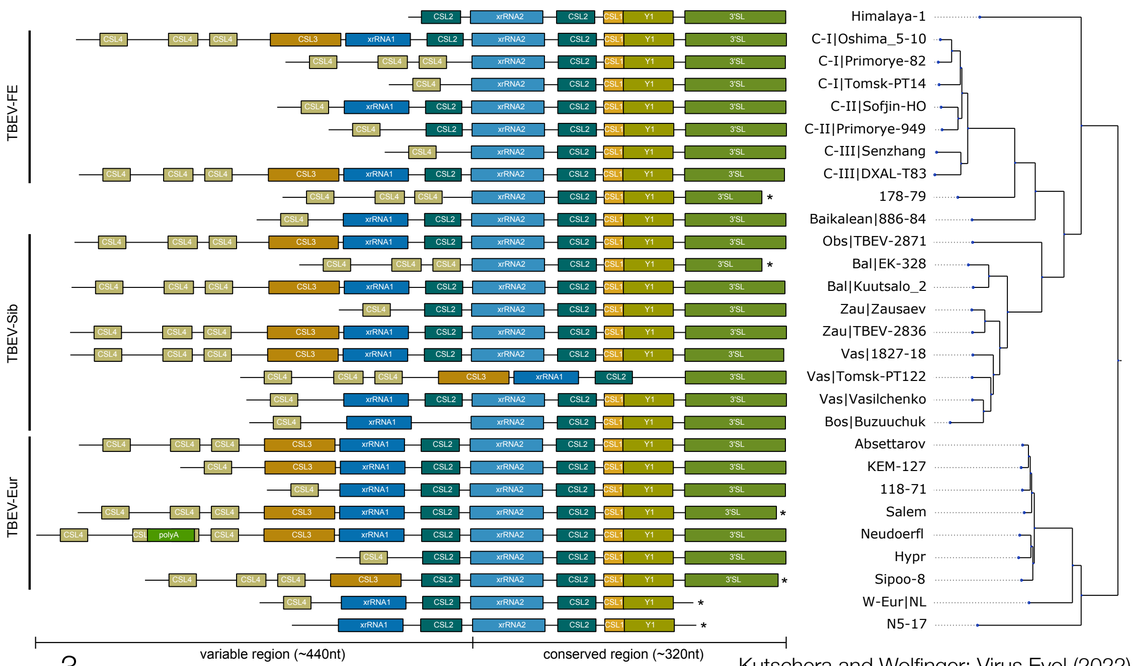
RNA structuredness of viral genomes

- Many examples of structured, functional RNAs in untranslated regions
- Some known examples of (conserved) RNA structures in coding regions
- Different evolutionary pressures on RNA structure in coding/non-coding regions

RNA structuredness of viral genomes

- Many examples of structured, functional RNAs in untranslated regions
- Some known examples of (conserved) RNA structures in coding regions
- Different evolutionary pressures on RNA structure in coding/non-coding regions

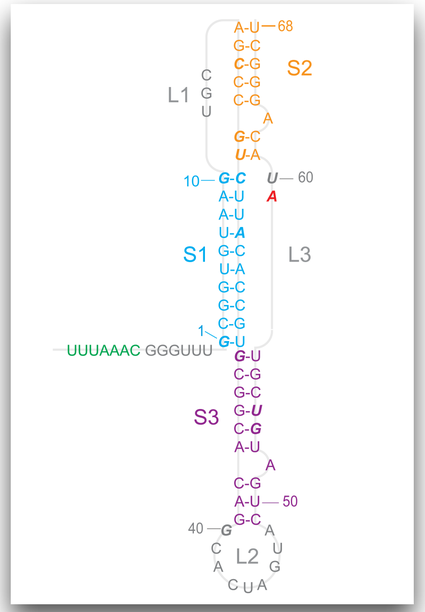




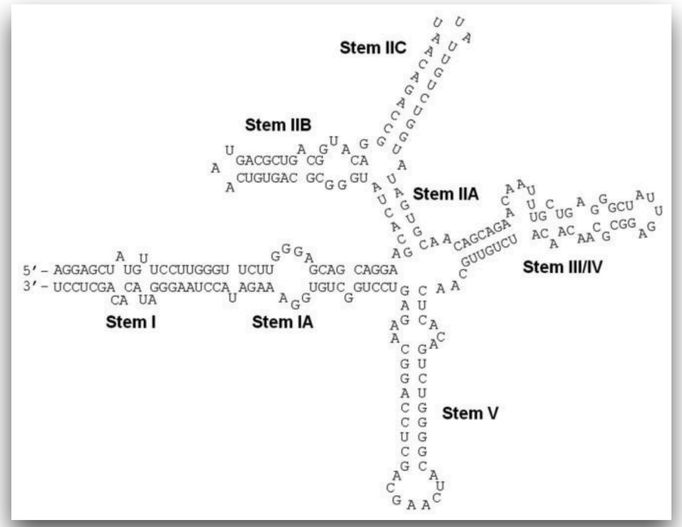
Kutschera and Wolfinger; Virus Evol (2022)

RNA structuredness of viral genomes

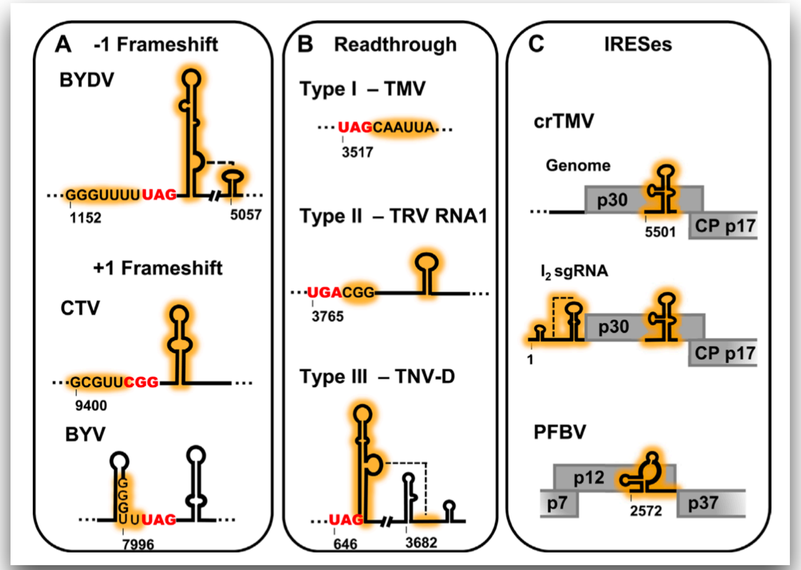
- Many examples of structured, functional RNAs in untranslated regions
- Some known examples of (conserved) RNA structures in coding regions
- Different evolutionary pressures on RNA structure in coding/non-coding regions



Omar et al.; PLoS Comput Biol (2021)



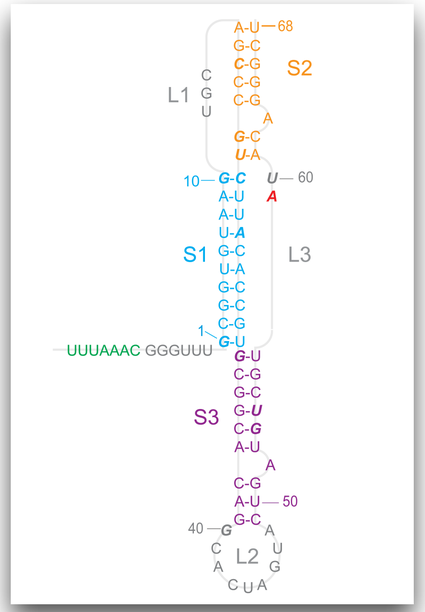
Fernandes et al.; RNA Biol (2012)



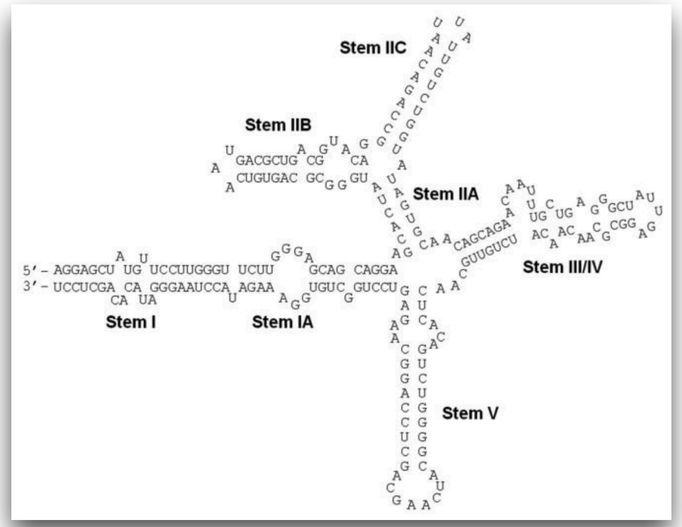
Newborn and White; Virology (2015)

RNA structuredness of viral genomes

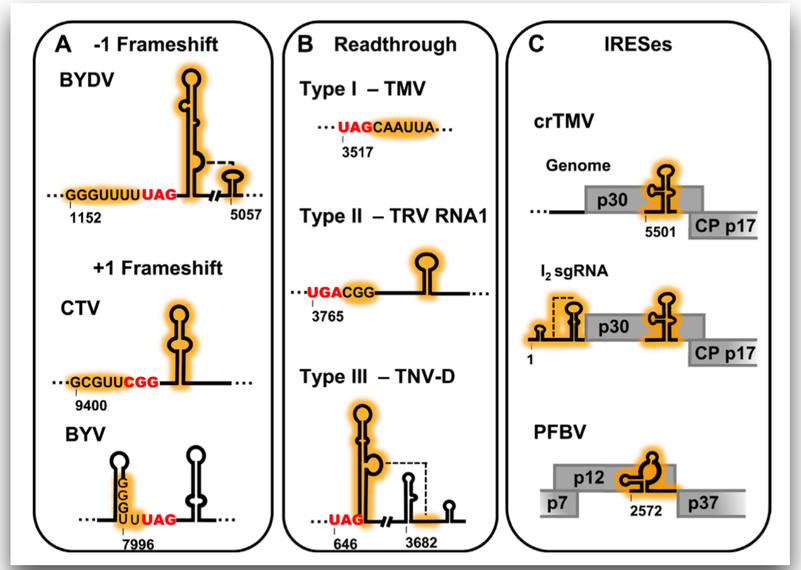
- Many examples of structured, functional RNAs in untranslated regions
- Some known examples of (conserved) RNA structures in coding regions
- Different evolutionary pressures on RNA structure in coding/non-coding regions



Omar et al.; PLoS Comput Biol (2021)



Fernandes et al.; RNA Biol (2012)

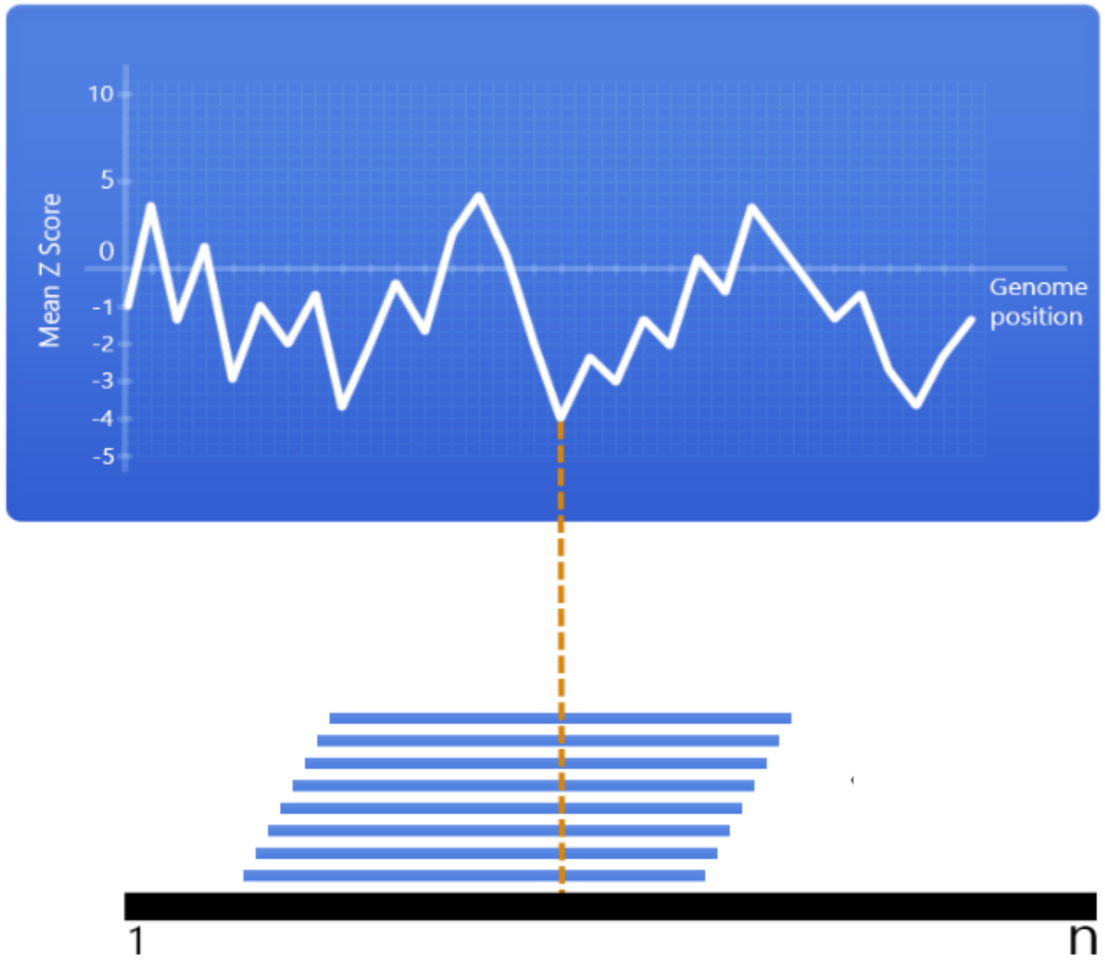


Newborn and White; Virology (2015)

How to assess global RNA structuredness?

- MFE Z scores as a proxy for RNA structuredness

$$z = \frac{m - \mu}{\sigma}$$



How to assess global RNA structuredness?

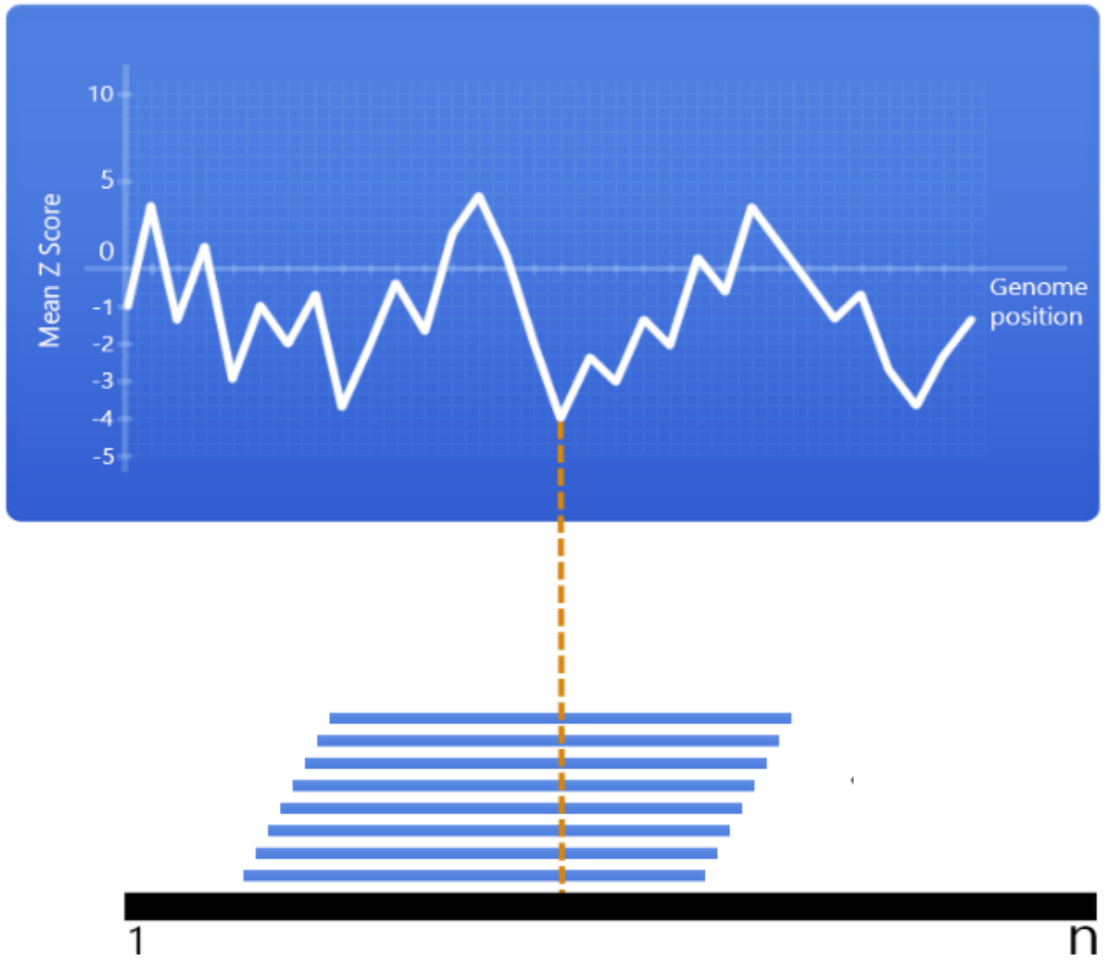
- MFE Z scores as a proxy for RNA structuredness

$$z = \frac{m - \mu}{\sigma}$$

- Opening energy

$$\Delta G_{\text{open}} = -RT \ln P(\text{unpaired})$$

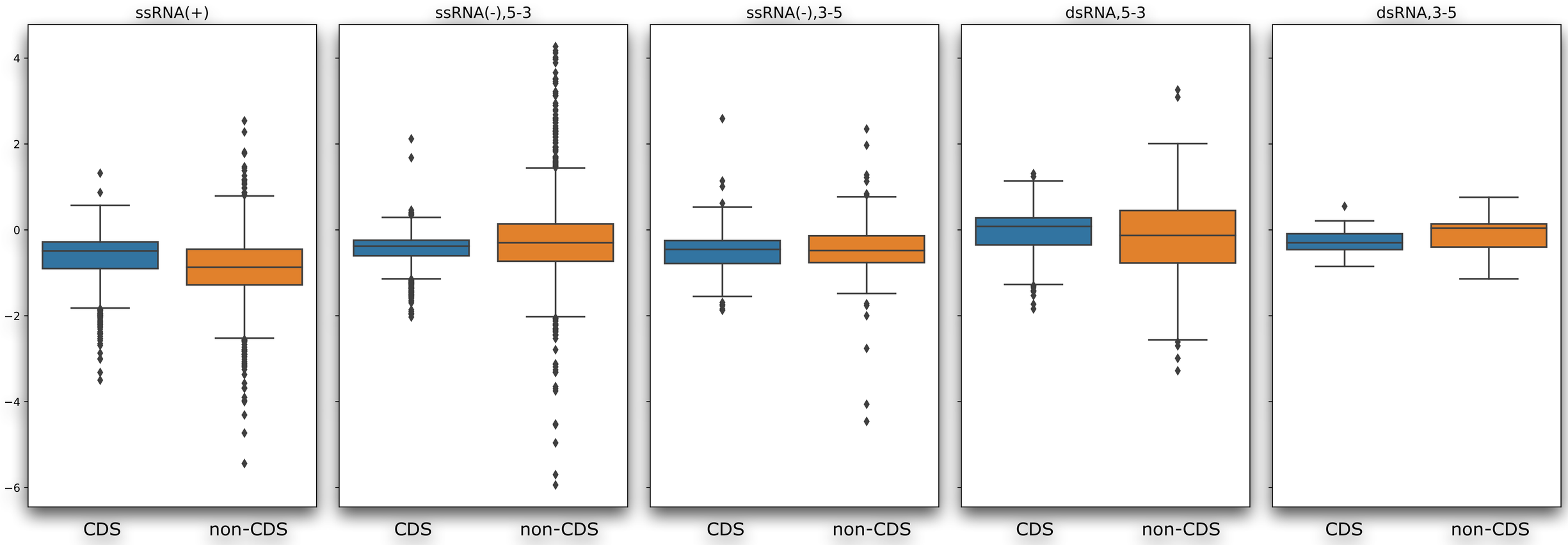
- GC content



Data Set

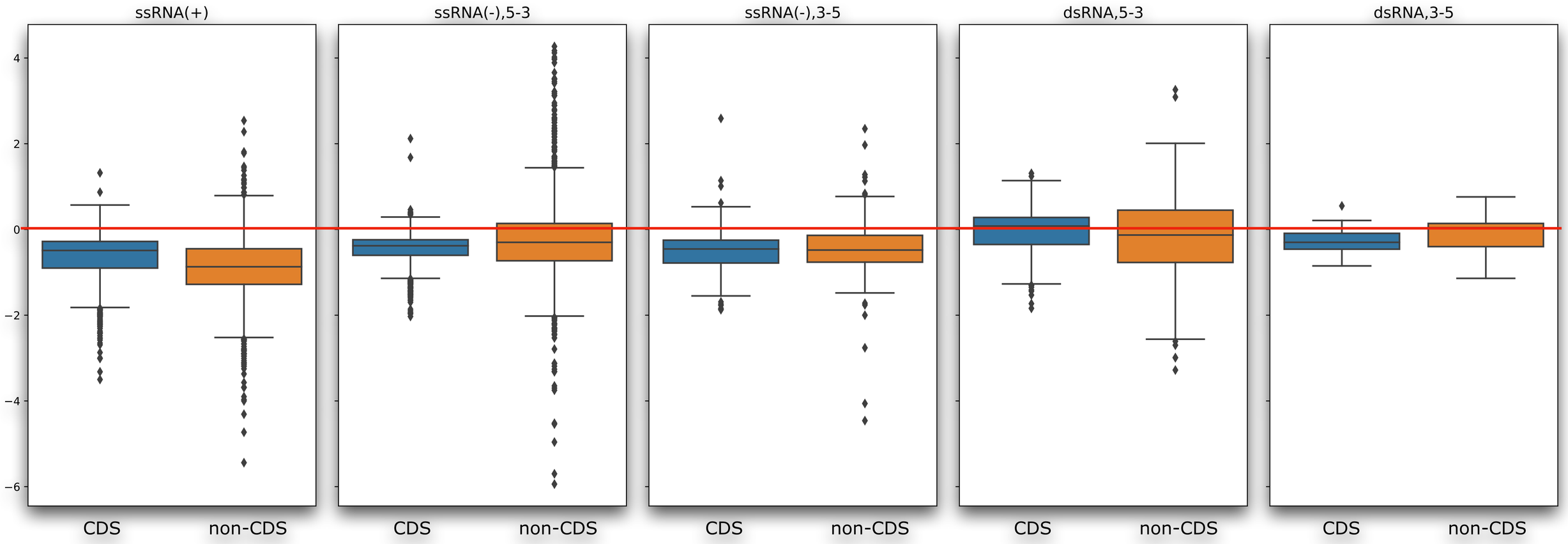
Baltimore classification	Unsegmented	Segmented	Total
ssRNA(+)	1333	373	1706
ssRNA(-)	355	1118	1473
dsRNA	73	890	963
dsDNA	714	0	714
	2475	2381	4856

Structuredness coding/non-coding regions



MFE Z scores

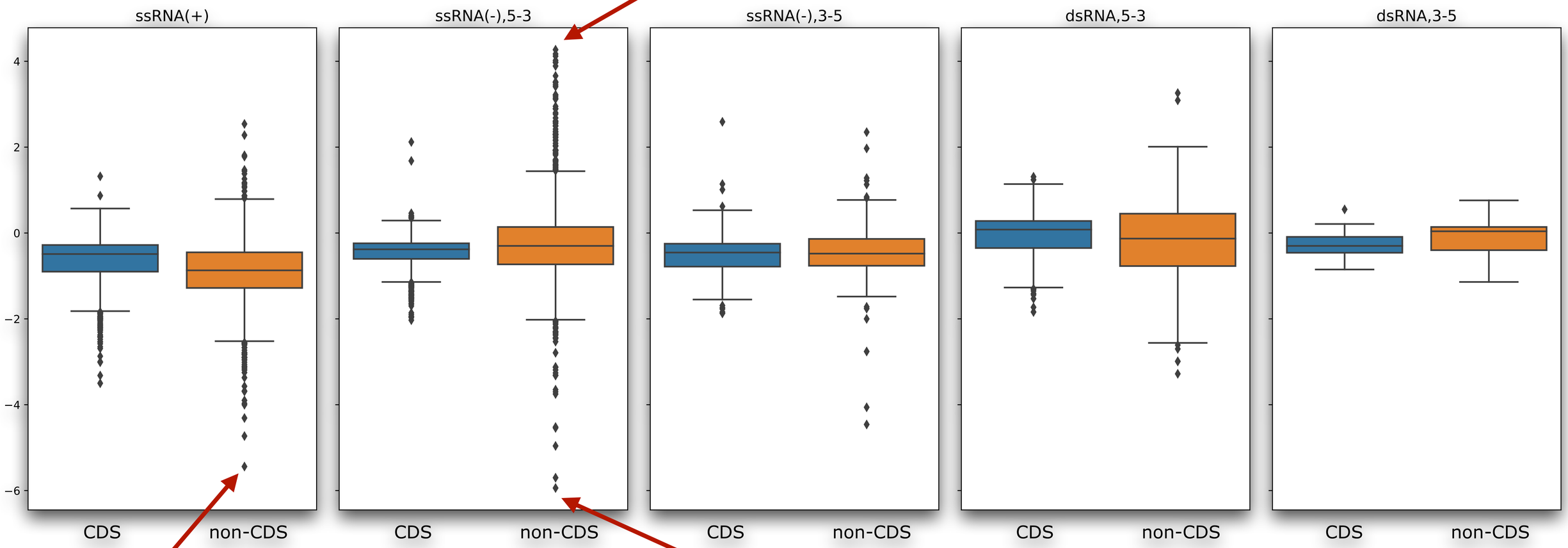
Structuredness coding/non-coding regions



MFE Z scores

Structuredness coding/non-coding regions

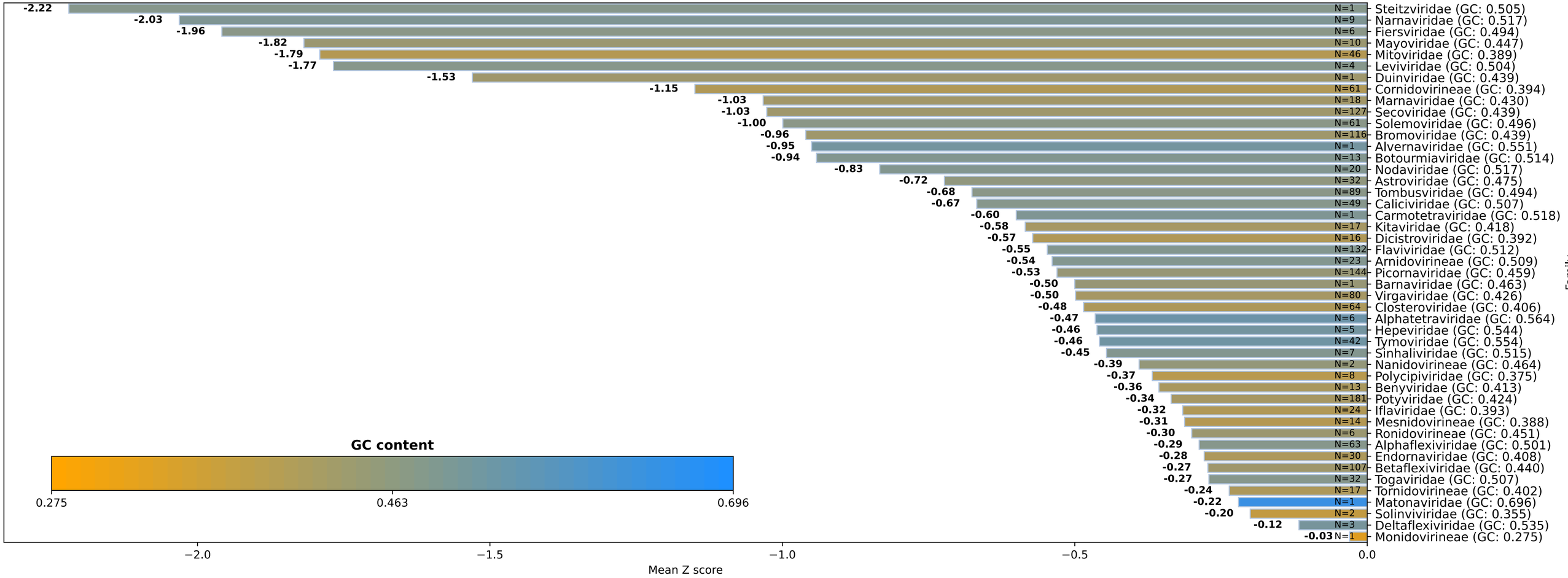
Anhembi virus, segment L (Bunyavirus)



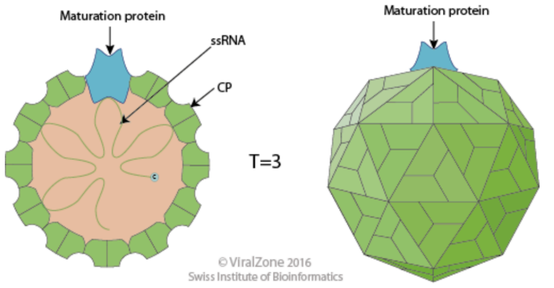
Blechmonas luni narnavirus 1

Uukuvirus (Bunyavirus)

Mean Z score of ssRNA(+) families



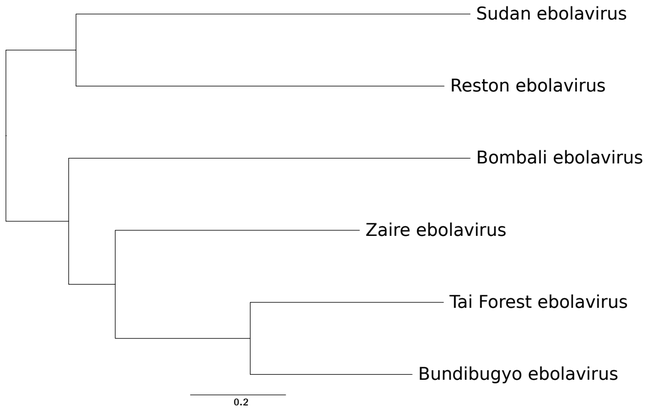
Phage example [ssRNA(+)]



https://viralzone.expasy.org/163

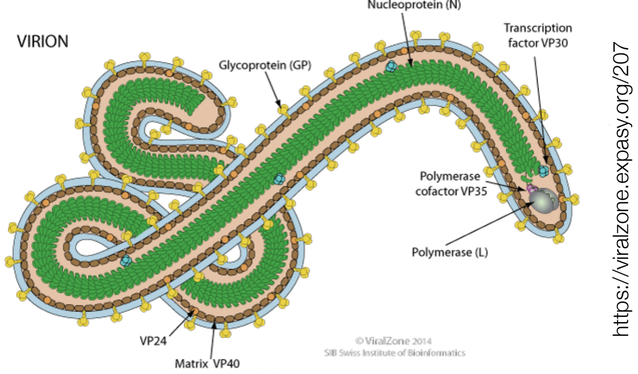
Viruses|Orthornavirae|Lenarviricota|Leviviricetes|Norzivirales|Fiersviridae|Enterobacteria phage M

Ebolavirus example [ssRNA(-)]

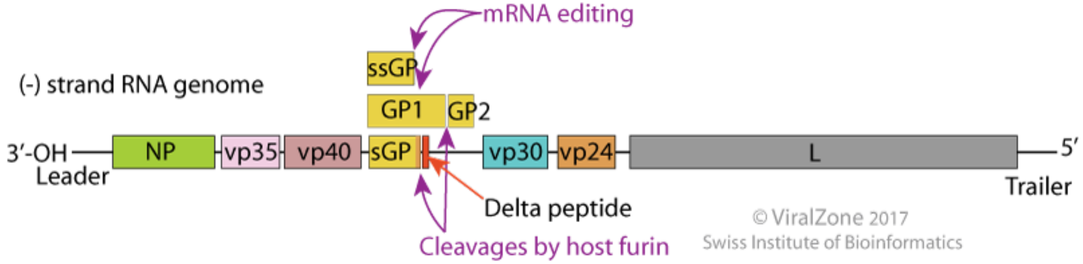


	BDBV	BOMV	EBOV	RESTV	SUDV	TAFV
TAFV	0.678	0.603	0.634	0.598	0.596	1
SUDV	0.597	0.578	0.6	0.594	1	
RESTV	0.597	0.582	0.602	1		
EBOV	0.633	0.605	1			
BOMV	0.602	1				
BDBV	1					

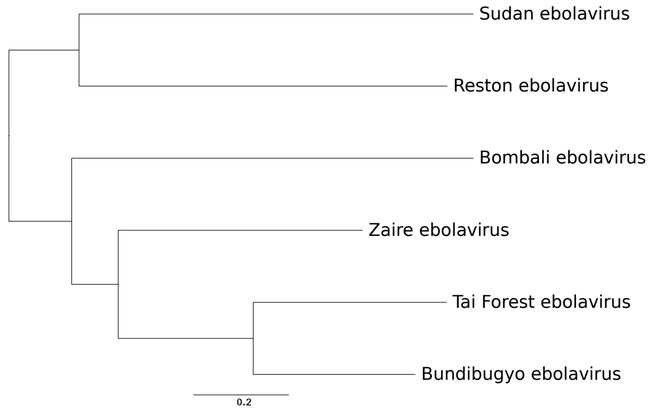
nucleotide identity
0.6 0.7 0.8 0.9 1.0

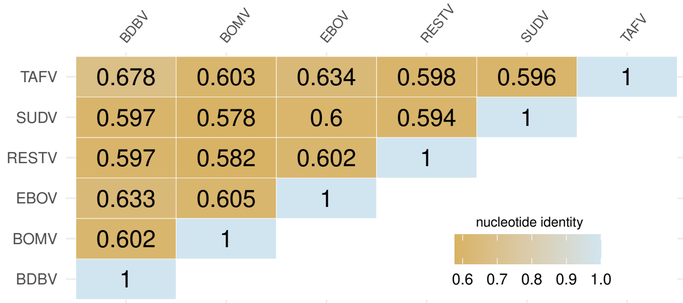


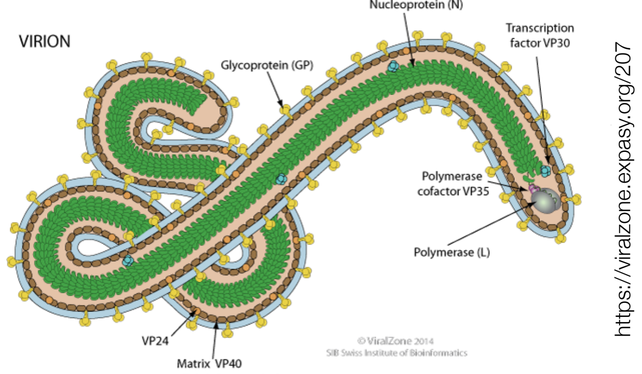
https://viralzone.expasy.org/207



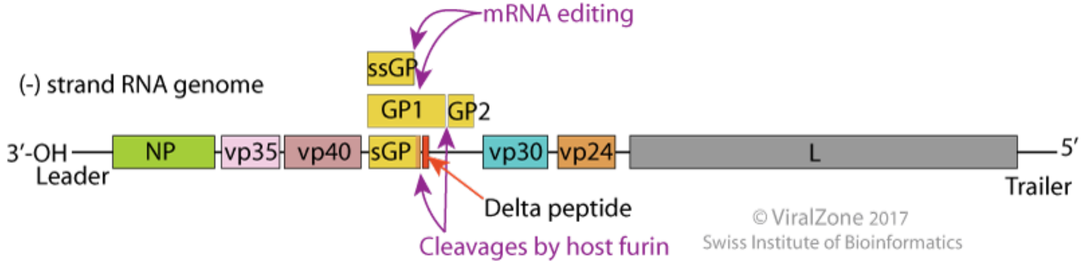
Ebolavirus example [ssRNA(-)]

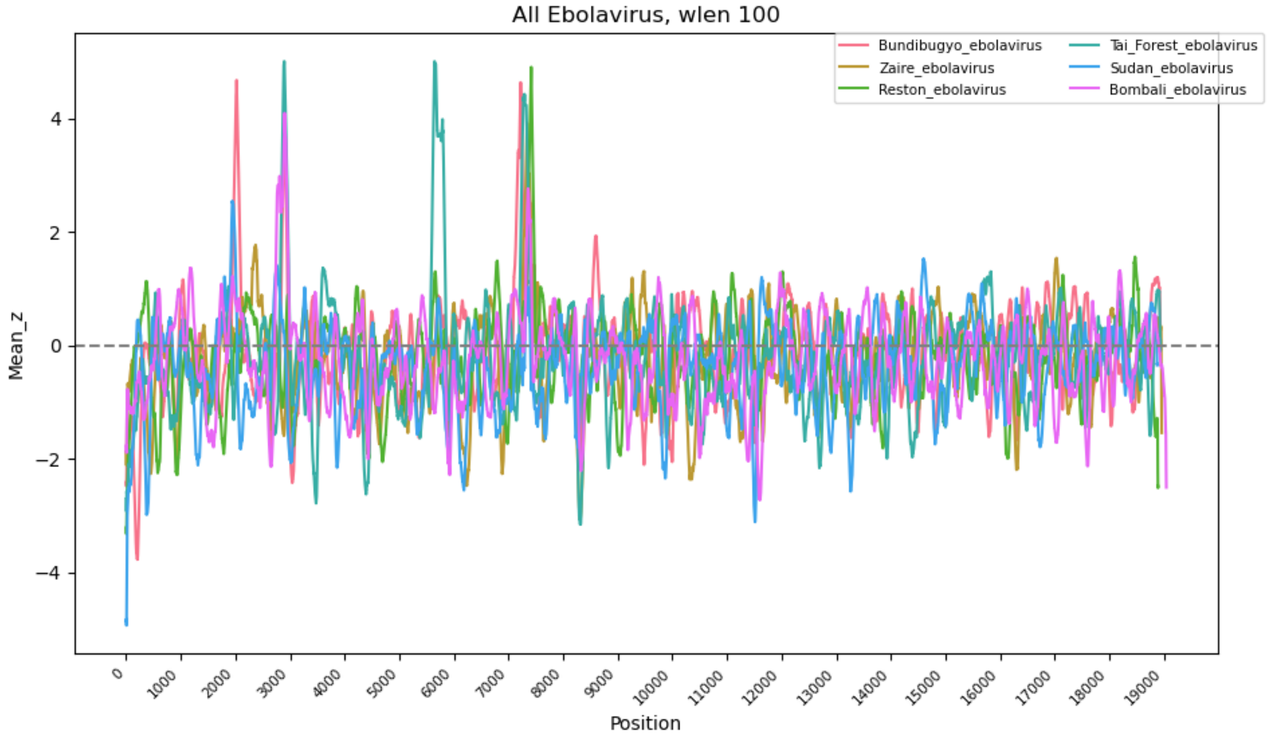


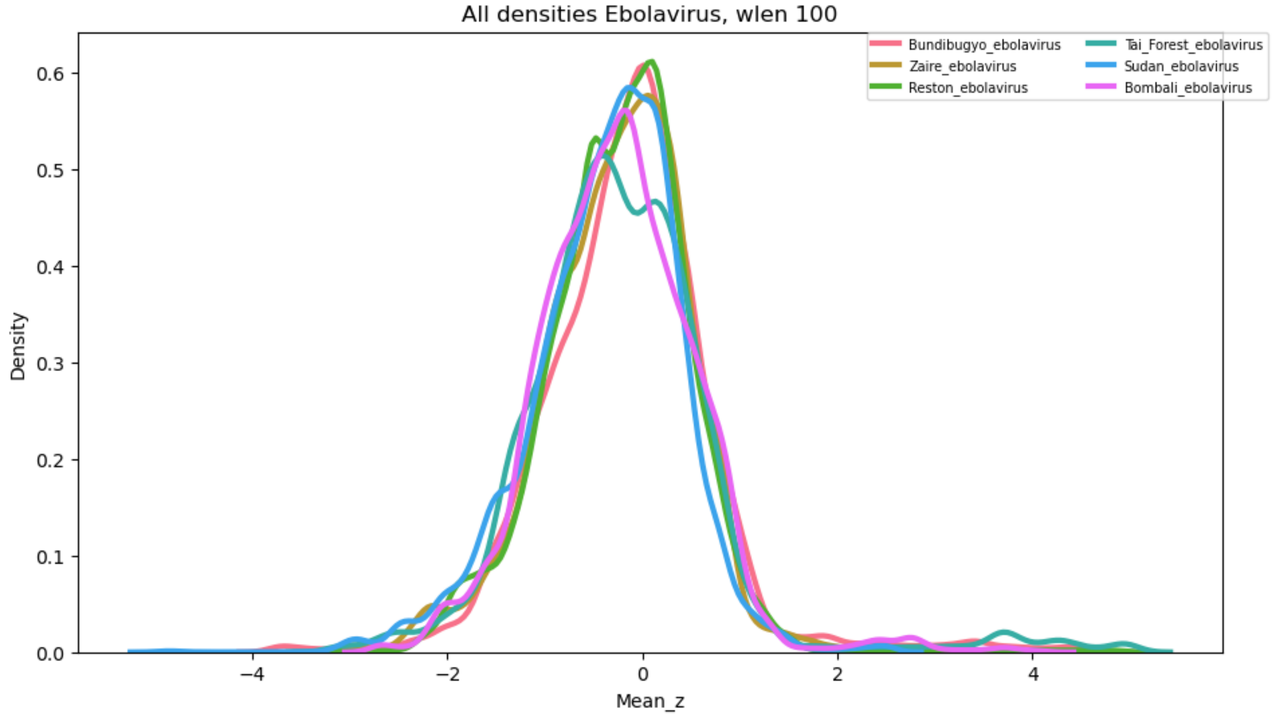




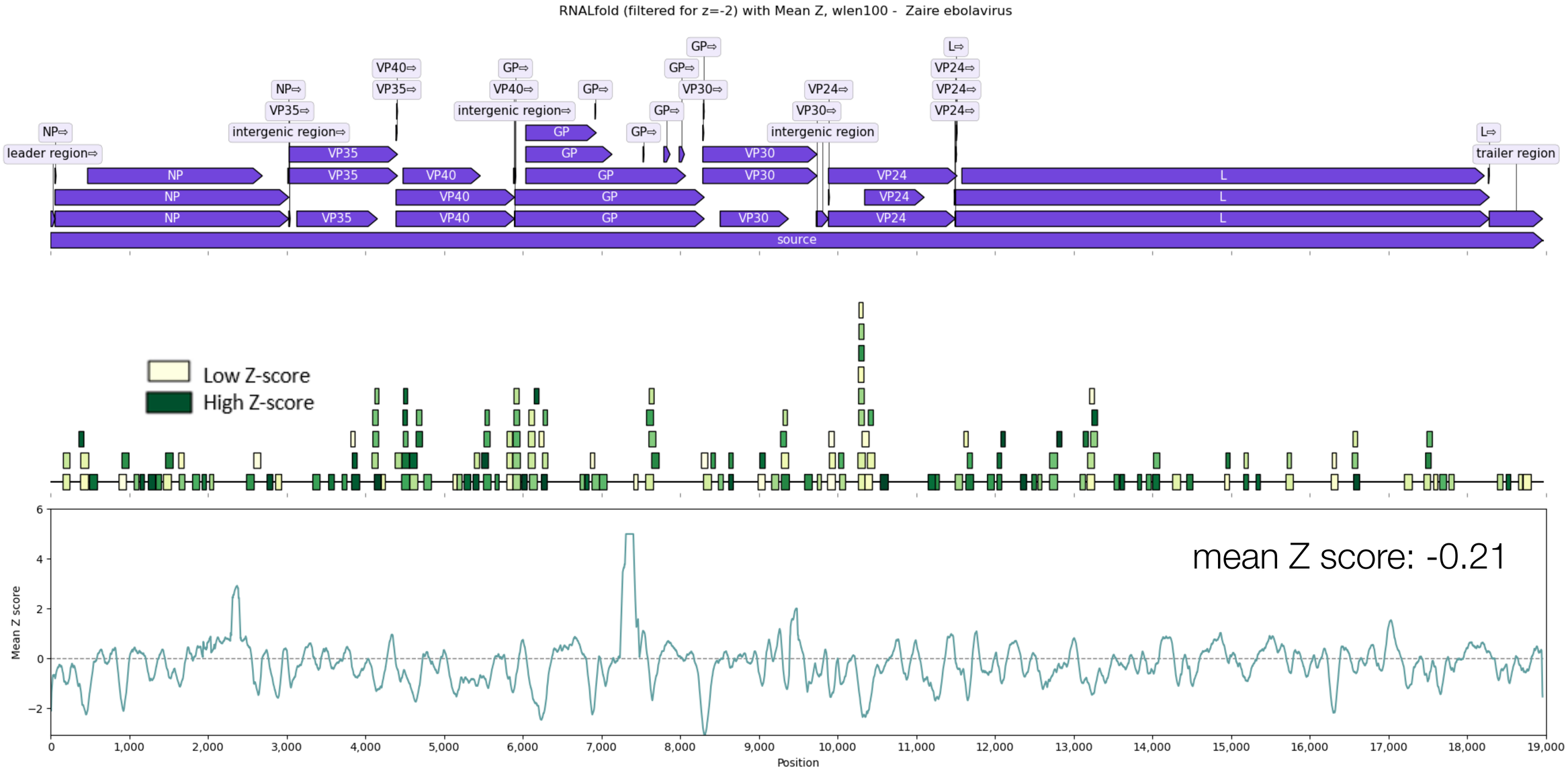
https://viralzone.expasy.org/207





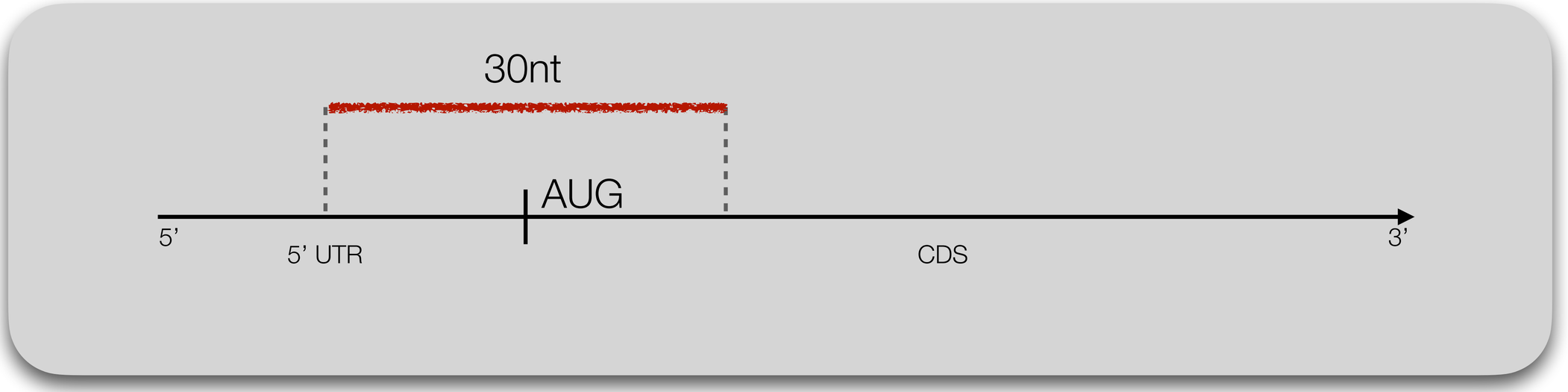


Ebolavirus example [ssRNA(-)]

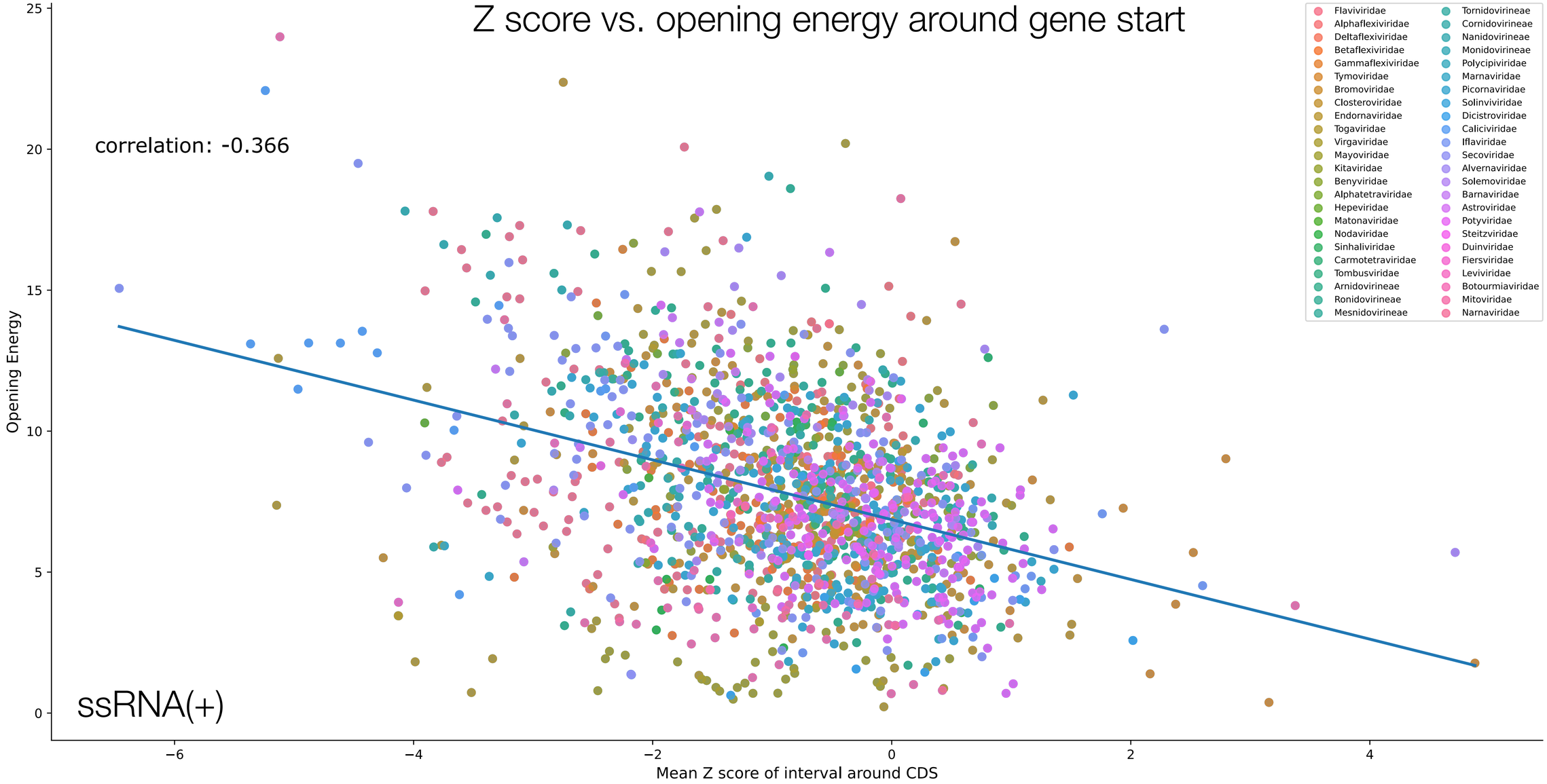


Structuredness of gene start regions

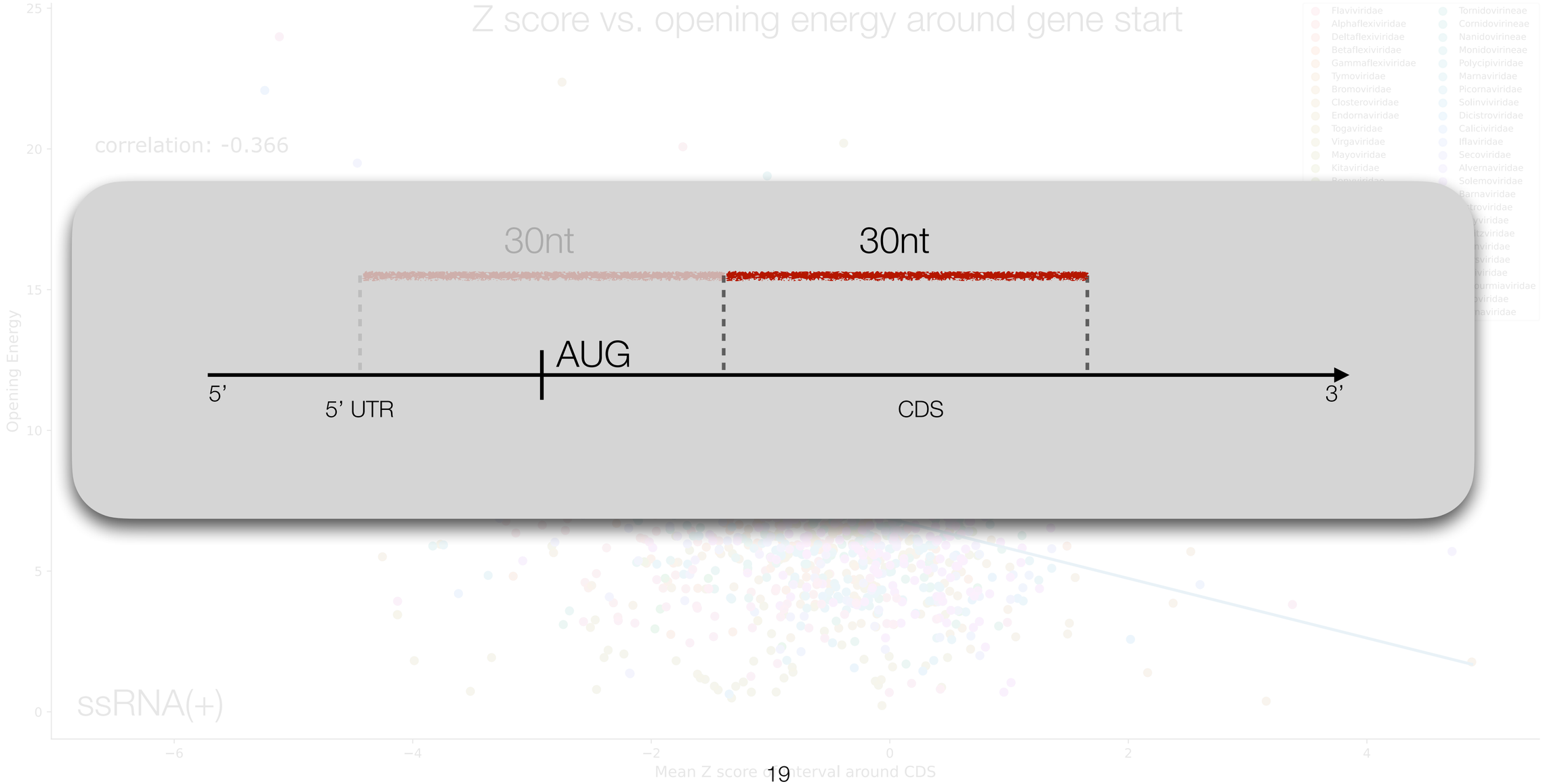
Are they accessible?



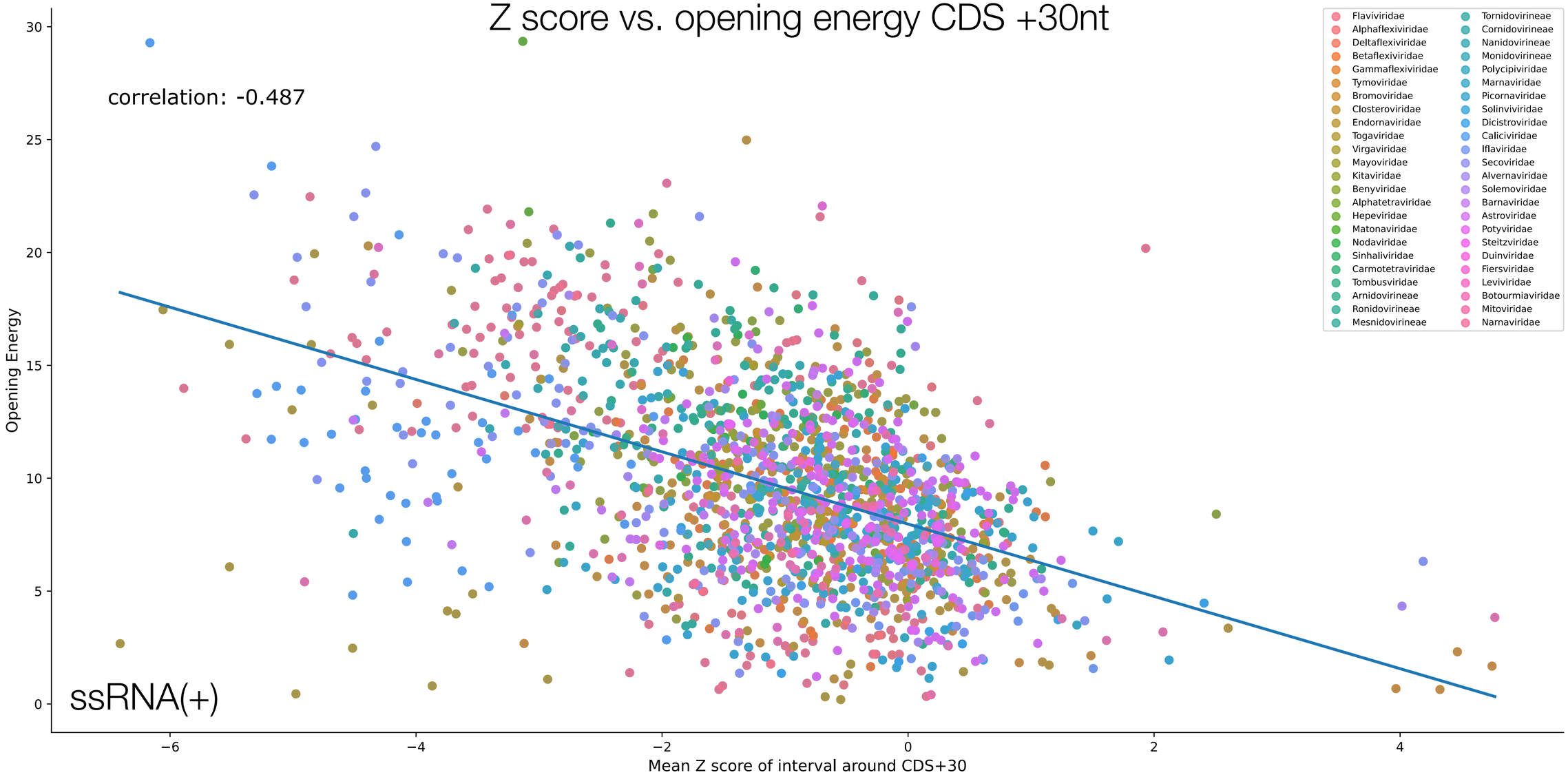
Structuredness of gene start regions



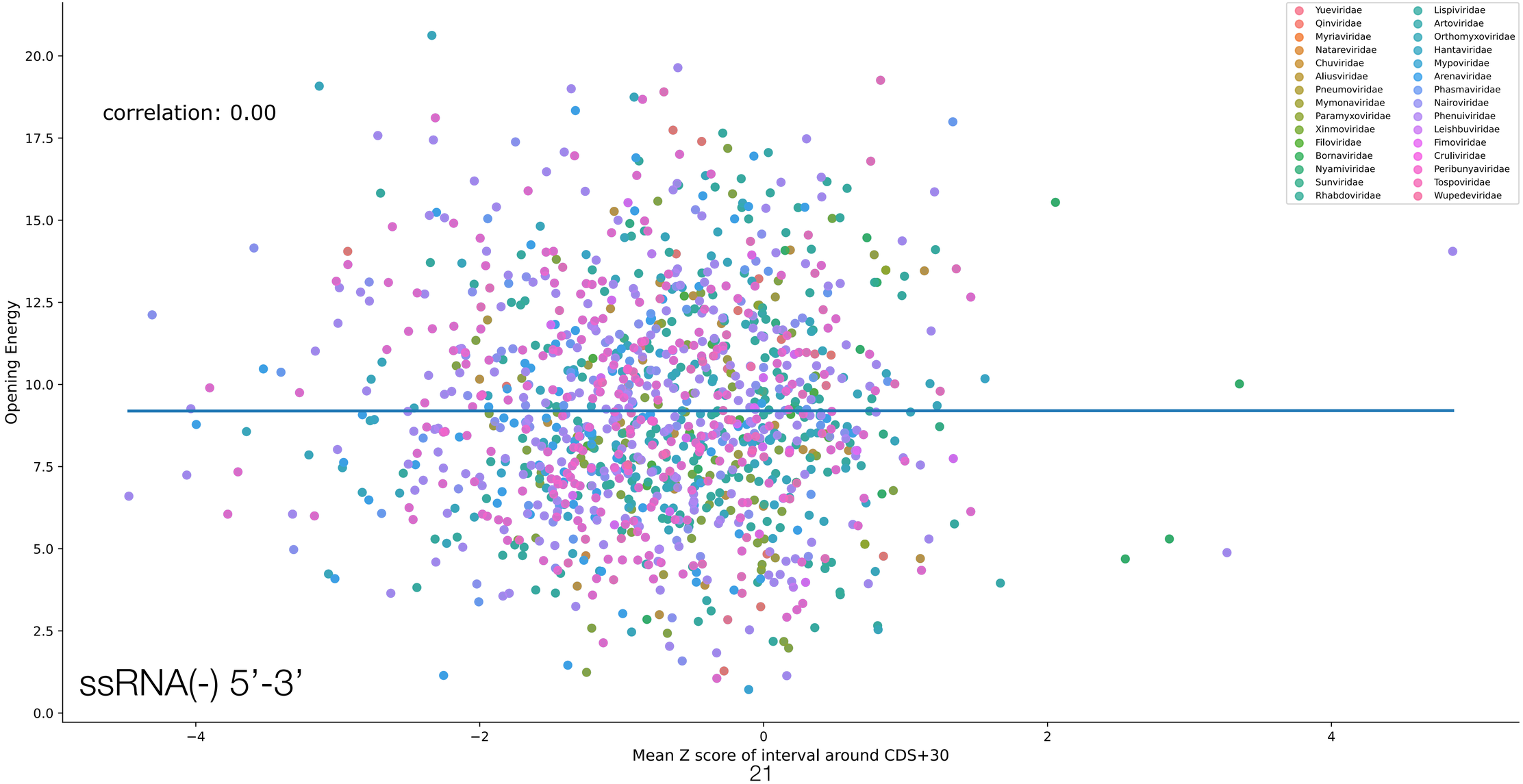
Structuredness of gene start regions



Structuredness of gene start regions



Structuredness of gene start regions



Where we are

- Viruses differ in their RNA structuredness; many viruses are more structured than expected
- GC content is not always a proxy for RNA structuredness
- Some viruses achieve high structuredness despite low GC content

The next steps

- Analyse structuredness of human mRNAs
- Assess the impact of codon usage bias on RNA structuredness
- Synbio: Study the impact of alternative genetic codes on RNA structuredness

Acknowledgements

TBI Vienna

Teodora Bucaciuc Mracica
Ivo Hofacker
Ronny Lorenz

Boston University

Elke Mühlberger

Auburn University

Joanna Sztuba-Solinska

University of Lethbridge

Trushar Patel

University of Leipzig

Mario Mörl

michaelwolfinger.com

@mtwolfinger