A proposed molecular mechanism for pathogenesis of severe RNA-viral pulmonary infections

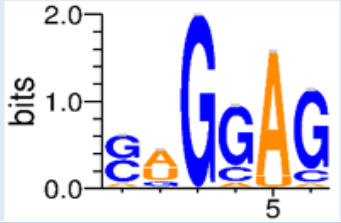
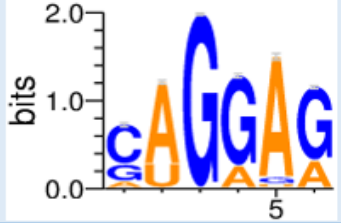
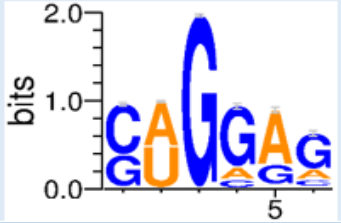
(10.6084/m9.figshare.12718799)

Peter K. Rogan (1,2), Eliseos J. Mucaki (1), Ben C. Shirley (2)

1. University of Western Ontario, 2. CytoGnomix Inc

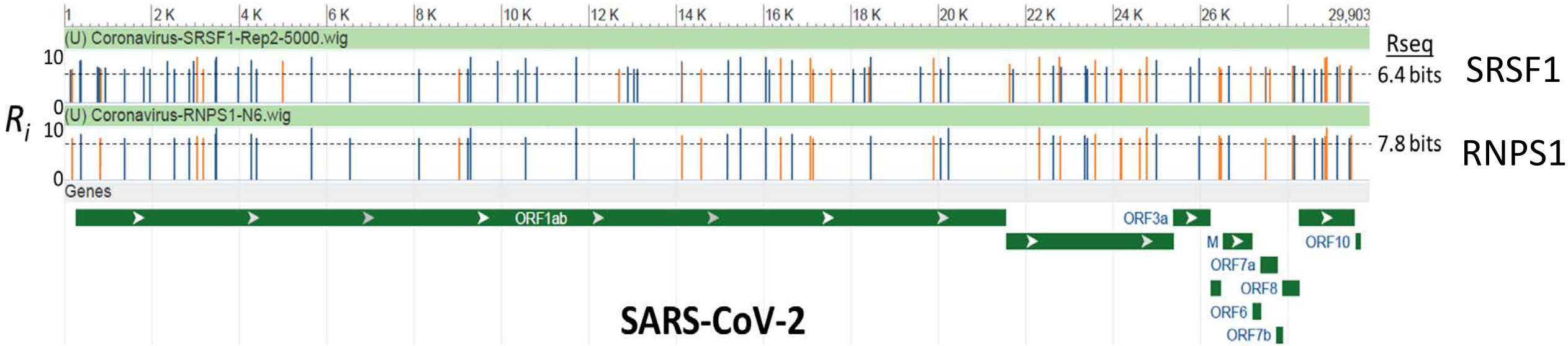
BIRS Poster Session, Tuesday, July 28th, 2020

SRSF1 and RNPS1 Information Models and Binding Sites in Host and Viral Genomes

Factor		SRSF1 [Rep1][1,2]		SRSF1 [Rep1] / RNPS1 Model Comparison		RNPS1[1]		SRSF1 [Rep2] / RNPS1 Model Comparison		SRSF1 [Rep2][1,2]	
Sequence Logo				-				-			
$R_{sequence}$ (bits)		6.7 ± 2.1		-		7.8 ± 1.9		-		6.4 ± 2.1	
Motif Similarity (E-value)[3]		-		5.0e-09		-		1.1e-09		-	
No. of Expressed Binding Sites (A549; ≥ 0 bits)[4]		1.3e08		5.4e07 (57%)[7]		9.3e07		6.4e07 (69%)		1.5e08	
No. of Expressed Binding Sites (Pneumocytes; ≥0 bits)[5]		6.8e07		2.9e07 (58%)		5.0e07		3.4e07 (69%)		7.9e07	
No. of Sites (SARS-CoV-2; +|- strand)	≥ 0 bits	860	732	435 (72%)	305 (65%)	608	466	363 (60%)	273 (59%)	810	772
	≥ $^1/_2$*Rseq*	311	232	131 (51%)	86 (44%)	256	196	155 (61%)	115 (59%)	376	358
	≥ *Rseq*	31	42	16 (46%)	10 (40%)	35	25	35 (100%)	25 (100%)	60	33
No. of Sites (Influenza A; +|- strand)[6]	≥ 0 bits	697	339	289 (61%)	118 (63%)	475	188	268 (56%)	129 (69%)	616	388
	≥ $^1/_2$*Rseq*	263	118	122 (49%)	47 (55%)	248	85	162 (65%)	65 (76%)	373	188
	≥ *Rseq*	50	23	24 (53%)	12 (75%)	45	16	45 (100%)	16 (100%)	84	35

[1] RNPS1 model derived from publicly available iCLIP data (E-MTAB-4215; ArrayExpress), while SRSF1 models were derived from eCLIP data (ENCSR456FVU; ENCODE Data Coordination Center); [2] SRSF1 [Rep1] and [Rep2] were derived from eCLIP dataset replicate 1 [50,000 peaks] and replicate 2 [5,000 peaks], respectively; [3] RNA binding motifs were compared using STAMP (34) using the Pearson Correlation Coefficient distance metric (74); [4] A549 cell line expression from GSE141171 dataset; [5] Primary type II pneumocyte expression from GSE86618 dataset; [6] Influenza A virus H3N2 strain (Ontario/104-25/2012). [7] RNPS1 sites used as denominator for all percentages.

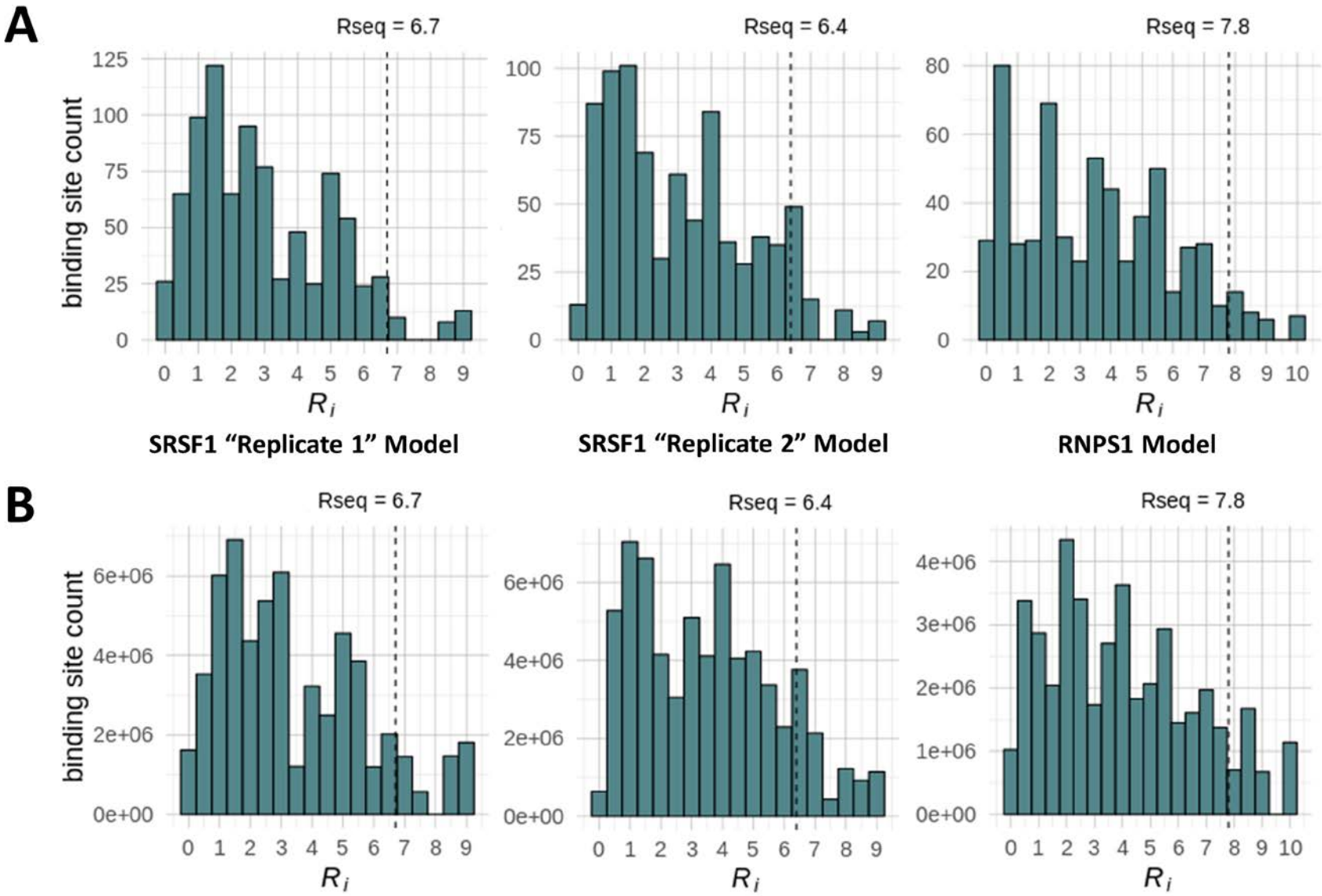
Distribution of SRSF1 and RNPS1 Binding Sites Across SARS-CoV-2



Positive Strand Binding Sites

Negative Strand Binding Sites

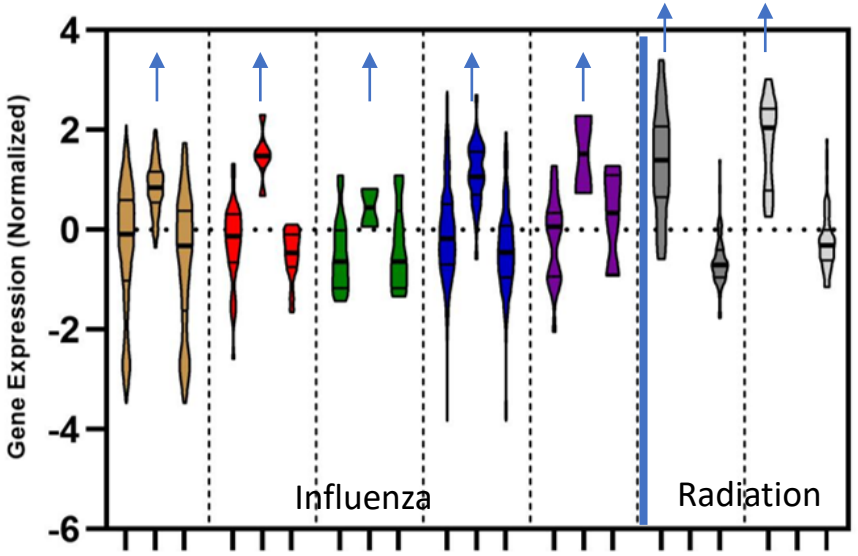
R_i of SRSF1 and RNPS1 Binding Sites in the SARS-CoV-2 and Human Genome



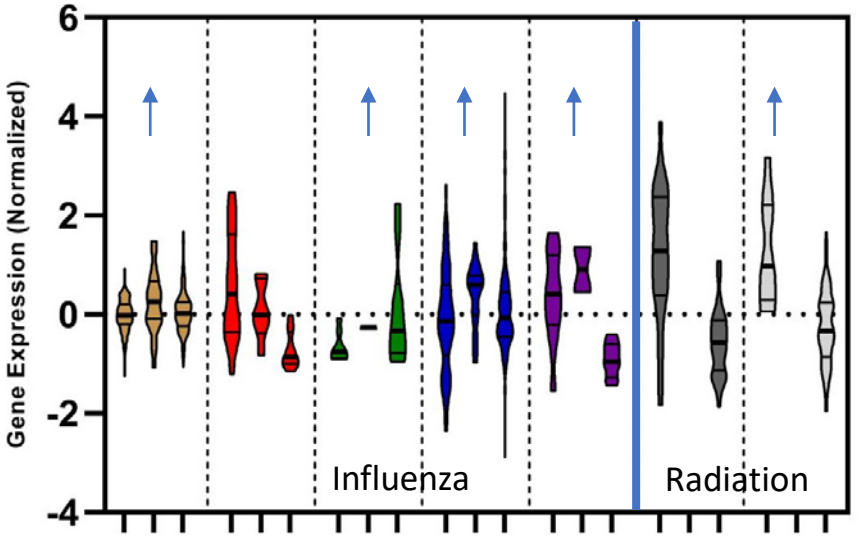
The SARS-CoV-2 Viral Genome

Transcribed Regions of the Human Genome

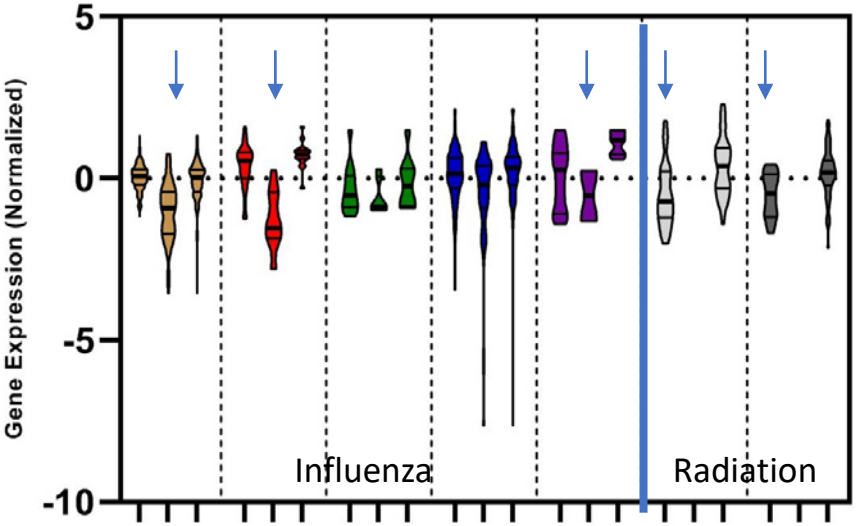
Expression of Rad-signature Genes in Influenza-infected and Radiation-Exposed Leukocytes



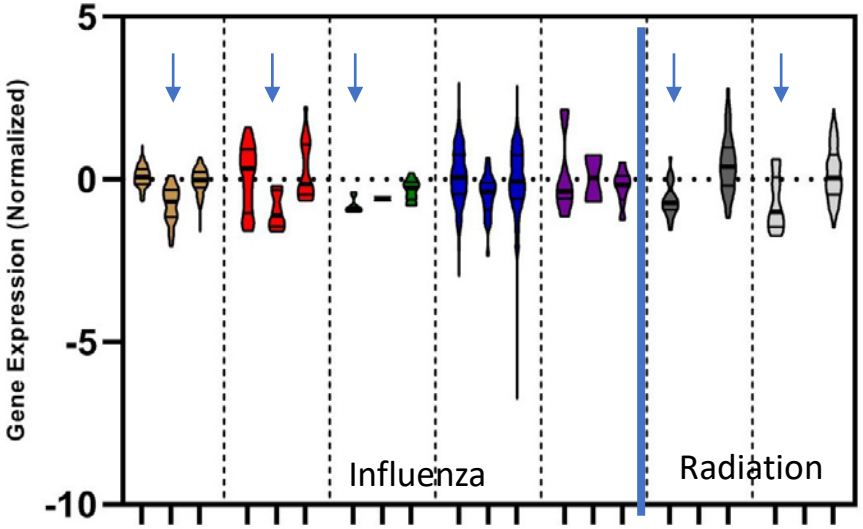
Expression of *DDB2*



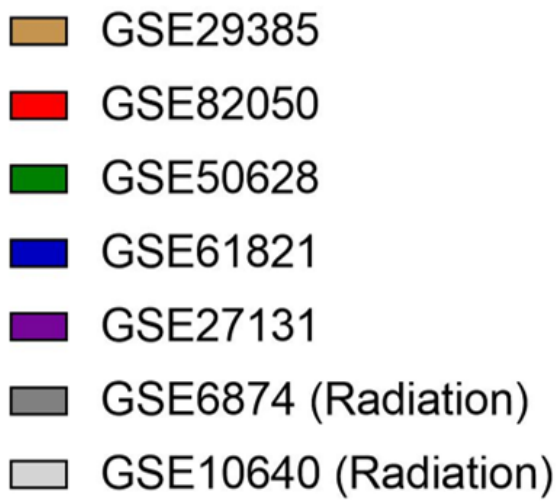
Expression of *PCNA*



Expression of *PRKCH*



Expression of *GTF3A*

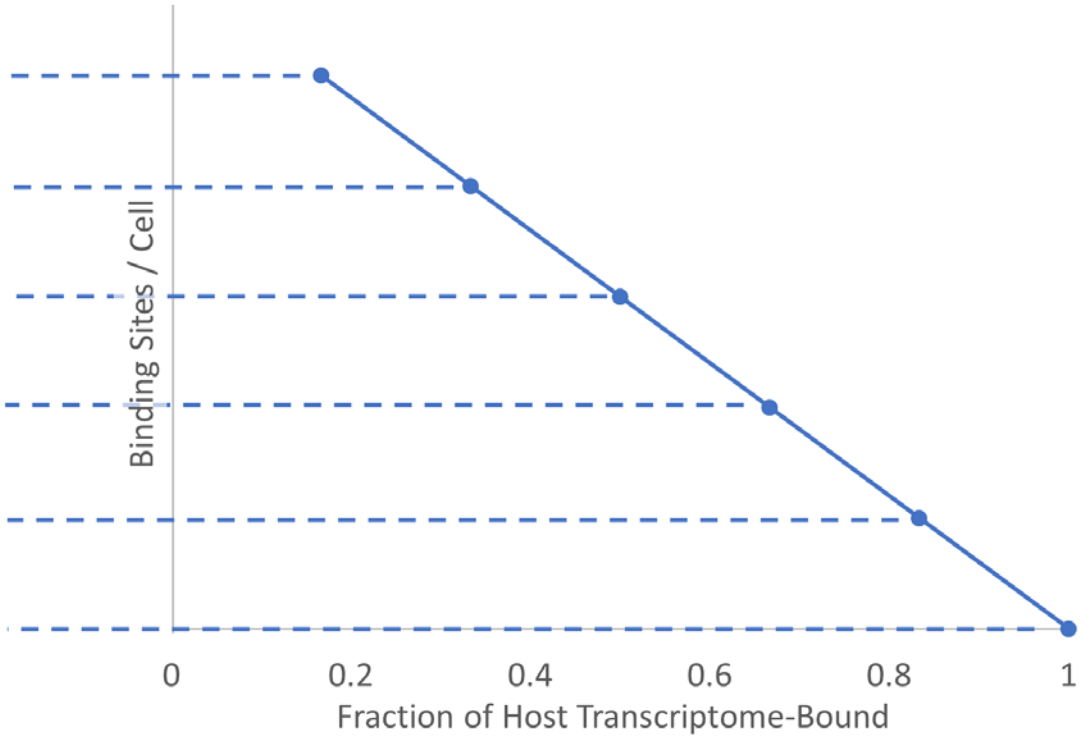


<u>Infection Datasets</u>:
Left = True Negatives
Middle = False Positives
Right = Controls

<u>Radiation Datasets</u>:
Left = Irradiated
Right = Non-Irradiated Controls

Inhibition of Host SRSF1 Binding by Viral Genome Replication

Virus : Host Binding Site Ratio	Theoretical (K_D)		Observed No. of Expressed Binding Sites in Host			
			Pneumocytes		A549 Cell Line	
	# Viral Genomes	Doubling Time (hours)	# Viral Genomes	Doubling Time (h)	# Viral Genomes	Doubling Time (h)
5 : 1	26,042	34.8	781,884	46.4	1,520,676	48.7
2 : 1	20,833	34.8	312,753	44.1	608,271	46.4
1 : 1	15,625	32.5	156,377	41.8	304,135	44.1
1 : 2	10,417	32.5	78,188	39.4	152,068	41.8
1 : 5	5,208	30.2	31,275	34.8	60,827	37.1
0 : 1	0	-	0	-	0	-



Summary

- The proposed mechanism of RNA viral infection-induced apoptosis is supported by bioinformatic analysis of RNA binding events
- Host RNA binding protein (RBP) binding sites are frequent and conserved among different strains of RNA viral genomes
- DNA repair and apoptotic proteins (such as *DDB2*, *PCNA*, *PRKCH* and *GTF3A)* were differentially expressed in influenza A and dengue-infected cells
- Information-dense binding site clustering of strong RBP binding sites coincides with the distribution of RNA-DNA hybridization sites across the genome
- Most binding sites and information-dense clusters do not overlap DRIP- and DRIPc-seq intervals, which are regions in the human genome with evidence of R-loop formation
- SRSF1 and RNPS1 have similar binding motifs which explains why DNA damage from SRSF1 depletion is complimented by RNPS1 overexpression

RESEARCH ARTICLE

A proposed molecular mechanism for pathogenesis of severe RNA-viral pulmonary infections

Peter K. Rogan [1,2], Eliseos J. Mucaki[1], Ben C. Shirley[2]

[1]Biochemistry, University of Western Ontario, London, Ontario, N6A 2C8, Canada
[2]CytoGnomix Inc, London, Ontario, N5X 3X5, Canada

v1 **First published:** N/A, **N/A**: N/A N/A
Latest published: N/A, **N/A**: N/A N/A

Abstract

Background: Certain riboviruses can cause severe pulmonary complications leading to death in some infected patients. We propose that DNA damage induced-apoptosis accelerates viral release, triggered by depletion of host RNA binding proteins (RBPs) from nuclear RNA bound to replicating viral sequences.

Methods: Information theory-based analysis of interactions between RBPs and individual sequences in the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), influenza A (H3N1), HIV-1, and dengue viral genomes identifies strong RBP binding sites in viral genomes. Replication and expression of viral sequences are expected to increasingly sequester RBPs - SRSF1 and RNPS1. Ordinarily, RBPs bound to nascent host transcripts prevents their annealing to complementary DNA. Their depletion induces destabilizing R-loops. Chromosomal breakage occurs when an excess of unresolved R-loops collide with incoming replication forks, overwhelming the DNA repair machinery. We estimated stoichiometry of inhibition of RBPs in host nuclear RNA by counting competing binding sites in replicating viral genomes and host RNA.

Results: Host RBP binding sites are frequent and conserved among different strains of RNA viral genomes. Similar binding motifs of SRSF1 and RNPS1 explain why DNA damage resulting from SRSF1 depletion is complimented by expression of RNPS1. Clustering of strong RBP binding sites coincides with the distribution of RNA-DNA hybridization sites across the genome. SARS-CoV-2 replication is estimated to require 32.5-41.8 hours to effectively compete for binding of an equal proportion of SRSF1 binding sites in host encoded nuclear RNAs. Significant changes in expression of transcripts encoding DNA repair and apoptotic proteins were found in an analysis of influenza A and Dengue-infected cells in some individuals.

Conclusions: R-loop-induced apoptosis indirectly resulting from viral replication could release significant quantities of membrane-associated virions into neighboring alveoli. These could infect adjacent pneumocytes and other tissues, rapidly compromising lung function, multiorgan system failure and the other described symptoms.

Keywords

SARS-CoV-2, Influenza A, HIV-1, Dengue Virus, Apoptosis, R-loop, DNA damage, RNA binding protein

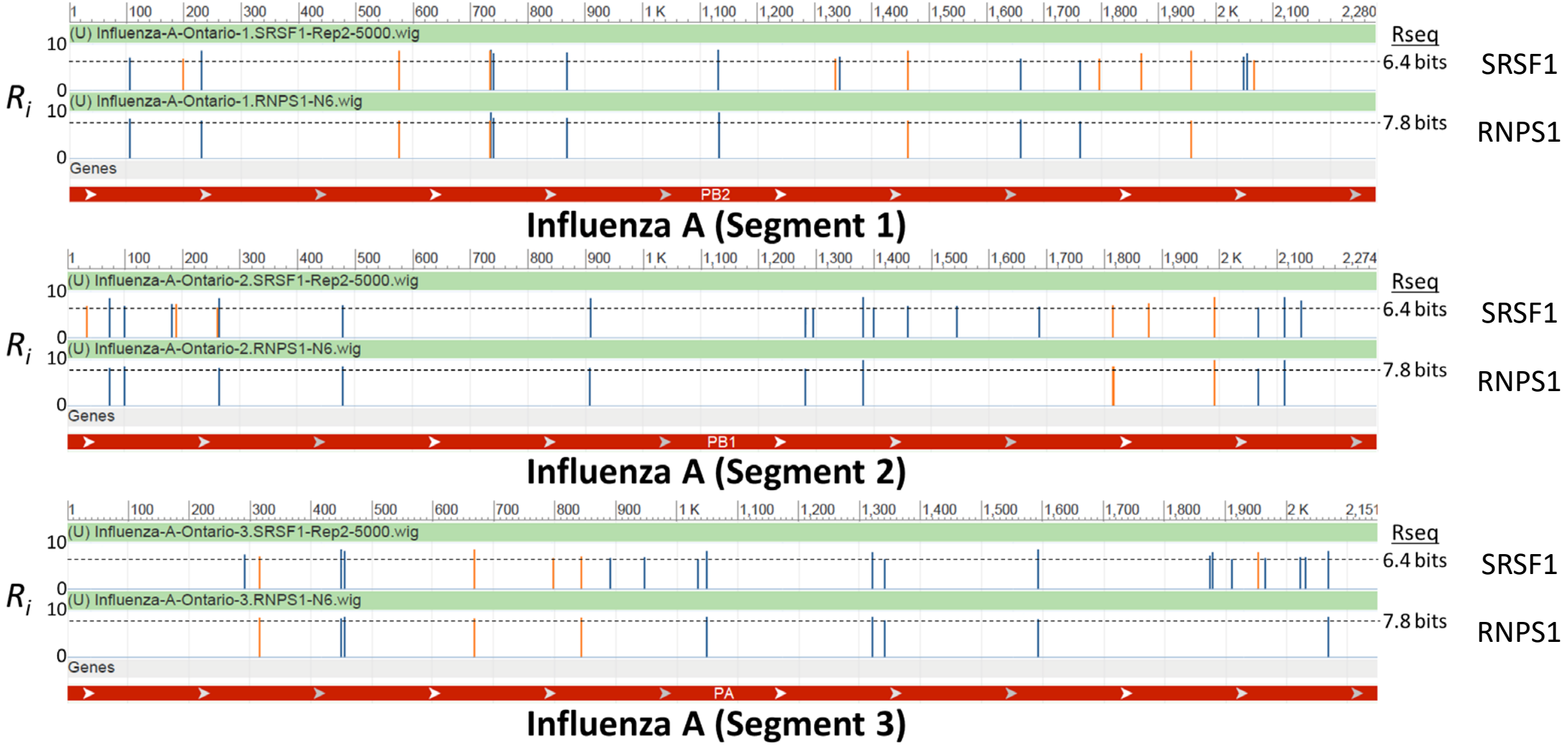
Open Peer Review

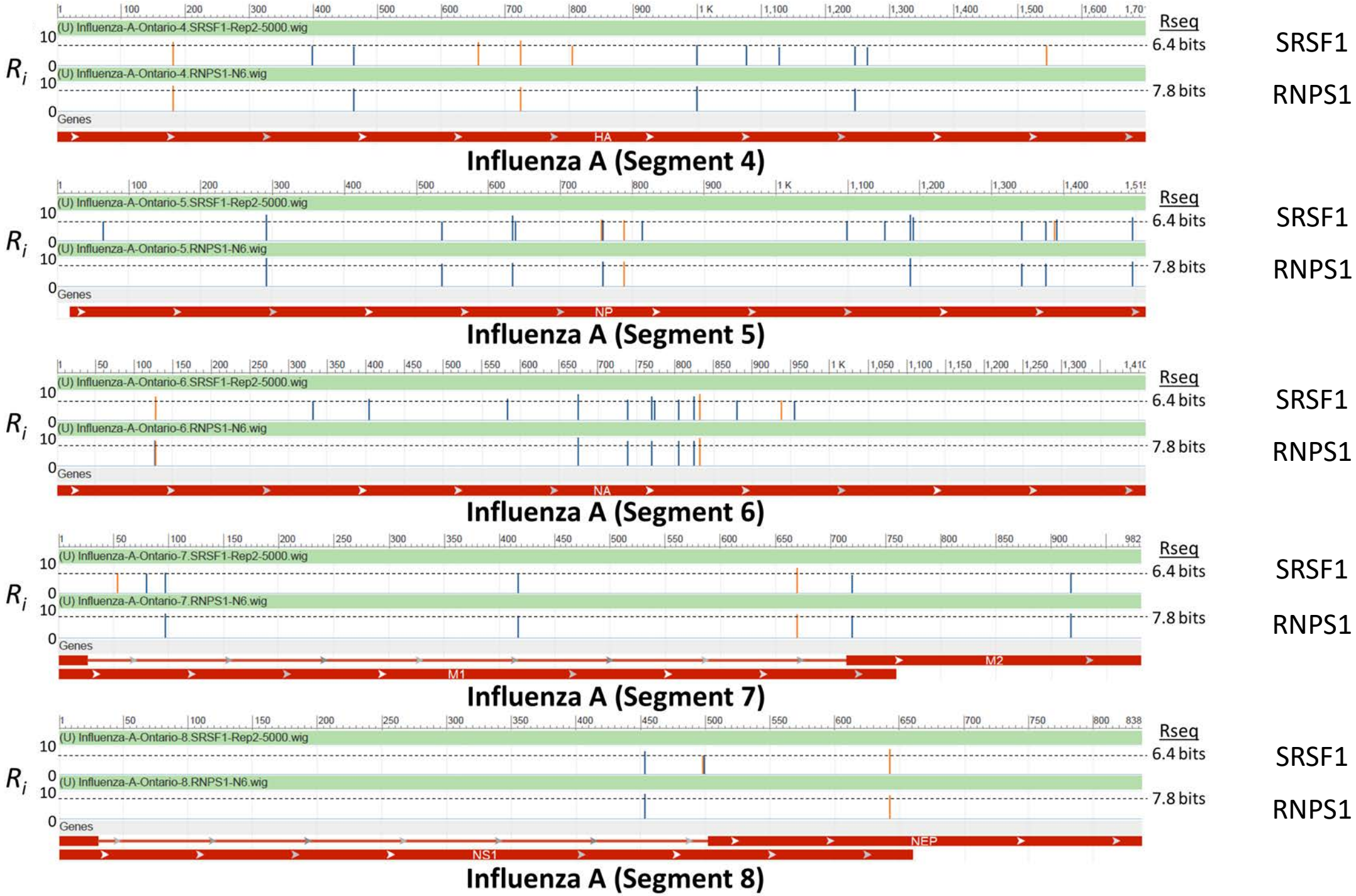
Reviewer Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

Alternative Versions of Slides

Distribution of SRSF1 and RNPS1 Binding Sites Across Influenza A




Rseq
6.4 bits
7.8 bits
SRSF1
RNPS1
(U) Influenza-A-Ontario-4.SRSF1-Rep2-5000.wig
(U) Influenza-A-Ontario-4.RNPS1-N6.wig
Genes
HA
Influenza A (Segment 4)
(U) Influenza-A-Ontario-5.SRSF1-Rep2-5000.wig
(U) Influenza-A-Ontario-5.RNPS1-N6.wig
NP
Influenza A (Segment 5)
(U) Influenza-A-Ontario-6.SRSF1-Rep2-5000.wig
(U) Influenza-A-Ontario-6.RNPS1-N6.wig
NA
Influenza A (Segment 6)
(U) Influenza-A-Ontario-7.SRSF1-Rep2-5000.wig
(U) Influenza-A-Ontario-7.RNPS1-N6.wig
M1
M2
Influenza A (Segment 7)
(U) Influenza-A-Ontario-8.SRSF1-Rep2-5000.wig
(U) Influenza-A-Ontario-8.RNPS1-N6.wig
NS1
NEP
Influenza A (Segment 8)