# Influenza Virus and PA-X: Looking At The Sequence Specificity of RNA Cleavage

**Amrita Iyer | Tufts Summer Scholars Program 2020 | Gaglia Lab**

## Introduction

- Influenza A Virus (flu) affects millions of people
  - According to the CDC, there are between 9-45 million cases of Influenza in the United States alone every year
  - Only limited treatment options for the virus, so by understanding how the virus works → can develop better treatments for people
- The Gaglia lab studies PA-X, a virus-encoded ribonuclease (RNAse) that degrades host RNAs and blocks the expression of genes involved in cellular immune response (Jagger *et al.,* 2012; Gaucherand *et al*., 2019)
  - PA-X activity is tied to RNA splicing, a processing step of RNA in which parts of the RNA, known as the introns, are cut out (Gaucherand et al., 2019) but this is not the only determinant for specificity
  - Mapping PA-X cut sites in 3 RNAs revealed that a "GCTG" sequence may also define cut sites specificity (Gaucherand, unpublished)
    - Since "GCTG" sequence is ubiquitous sequence → may only be part of the criteria for PA-X activity

## Objectives

### Overall: Is PA-X Activity Sequence Specific?

To try and answer this question, I broke my research down into the following questions:

Q1: Is there a correlation between the frequency of a sequence in RNA and its degradation level?

Q2: Is there evidence in the data that splicing influences PA-X activity?

Q3: Is there evidence of a specific sequence?

## Methods

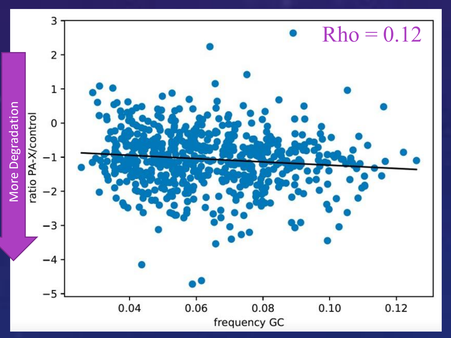
### Q1: Is there a correlation between the frequency of a sequence in RNA and its degradation level?

Using Python, I analyzed RNA sequences to determine frequency of "k-mers":

- Dimers → Groups of 2 bases (AT)
- Trimers → Groups of 3 bases (CTG)
- Tetramers → Groups of 4 bases (GCTG)

The Gaglia lab already measured total RNA levels in cells that express PA-X versus control cells and found that cells that express PA-X have lower levels of RNA overall (Gaucherand et al., 2019)

- Lower RNA ratio in PA-X / control for a specific RNA→ indicates more PA-X degradation activity on that RNA
- Negative correlation (using Spearman statistic) between the frequency of a sequence and RNA ratios→ may indicate that PA-X acts on that sequence (as it is broken down more)




### Q2: Is there evidence in the data that splicing influences PA-X activity?

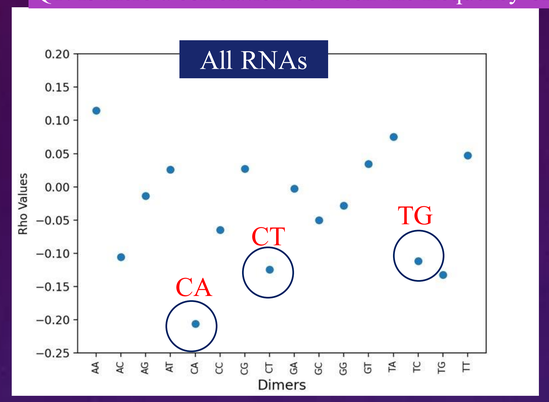
- Plotted Spearman rho values for correlations between "k-mers" and RNA degradation levels in a set of RNAs with a specific number of exons versus in an equally sized random subset of RNAs (w/ random number of exons) → to see if splicing was a confounding variable

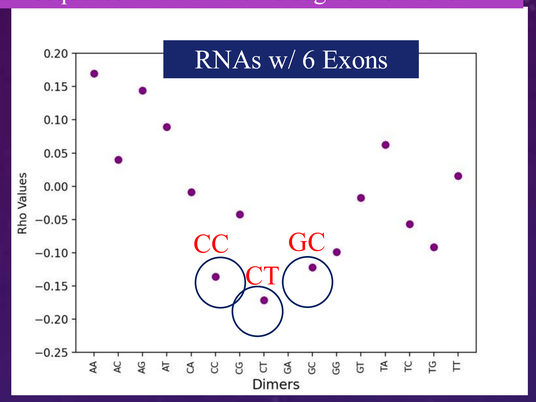
### Q3: Is there evidence of a specific sequence?

- Took "k-mers" with most negative Spearman rho values and generated a proposed PA-X recognition sequence of CTGCTGGGCA
  - Found the RNAs with perfect or near perfect matches to sequence (Hamming Distance of 0, 1 or 2 → 0, 1, 2 bases off from sequence)
    - Plotted the ratio values of RNAs w/ this sequence or close to this sequence to see whether these RNAs are preferentially degraded by PA-X

## Results

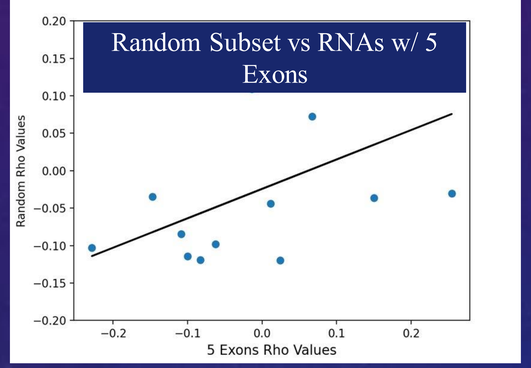
### Q1: Is there a correlation between the frequency of a sequence in RNA and its degradation level?

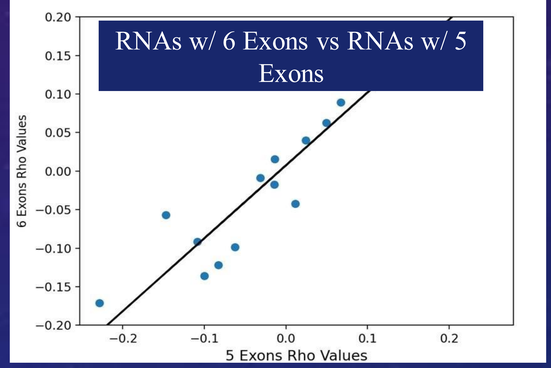



- There are correlations between frequency of specific dimers and their degradation levels
- Correlations for all RNAs and RNAs w/ 6 Exons are different → indicating that splicing may be confounding variable + influences PA-X activity
- GC dimer shows negative correlation → similar to observed "GCTG" sequence
- Obtained similar results for trimers and tetramers

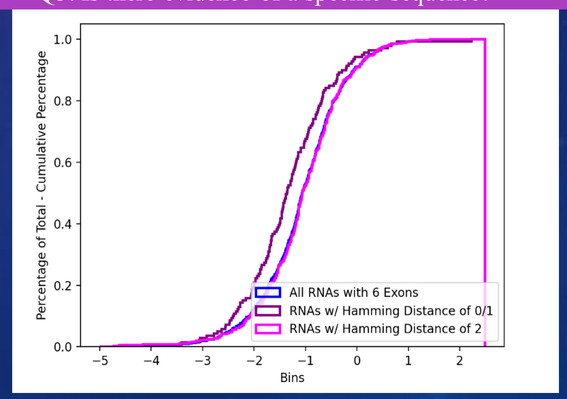
### Q2: Is there evidence in the data that splicing influences PA-X activity?




- Rho values for RNAs w/ 6 Exons versus a random subset of similar size (where # of exons are not controlled for) are different
- Rho values for RNAs w/ 6 Exons versus RNAs w/ 5 Exons are more similar → these rho values point to sequence specificity and are not random
- Obtained similar results for trimers and tetramers

### Q3: Is there evidence of a specific sequence?




- Based on the graph → shows RNAs w/ this sequence or sequences w/ 1 difference have lower ratio values, indicating they are degraded more readily than RNAs w/ two or more differences
  - PA-X may act on this sequence

## Conclusion and Future Directions

- Rho values of different groups of RNAs show that the number of exons influences PA-X activity
- Correlations in the exon analysis consistent with idea that these correlations have some significance
- Preliminary cumulative probability graph supports idea that the CTGCTGGGCA sequence may be required for PA-X activity
  - Still need to further analyze these data by also plotting random subset of RNAs to see how graphs compare → to determine if this difference in ratio values is significant + refine cut sequence according to results

## References


- "Disease Burden of Influenza." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 17 Apr. 2020, www.cdc.gov/flu/about/burden/index.html.
- Freeman, S., Quillan, K., Allison, L., Black, M., Podgorsky, G., Taylor, E. and Carmichael, J. (2017). "**Biological Science"**, 6th Edition, Pearson Education, London, UK.
- Gaucherand, L., Porter, B. K., Levene, R. E., Price, E. P., Schmaling, S. K., Rycroft, C. H., Kevorkian, Y., Mccormick, C., Khaperskyy, D. A. and Gaglia, M. M. (2019). **"The Influenza A Virus Endoribonuclease PA-X Usurps Host MRNA Processing Machinery to Limit Host Gene Expression."** *Cell Reports*, vol. 27, no. 3, 2019, pp. 776–792.e7**.**
- Jagger, B.W., Wise, H.M, Kash, J. C., Walters, K-A., Wills, N. M.,Xiao, Y.- L., Dunfee, R. L., Schwartzman, L. M., Ozinsky, A., Bell, G. L., Dalton, R. M., Lo, A., Efstathiou, S., Atkins, J F., Firth, A. E., Taubenberger, J. K., and Digard, P. **"An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response.** " *Science*, vol. 337, no. 6091, 2012, pp. 199–204.
- Levene, R. E. and Gaglia, M. M. (2018) **"Host Shutoff in Influenza A Virus: Many Means to an End."** *Viruses*, vol. 10, no. 9, 2018, p. 2018.
- Nogales, A., Martinez-Sobrido, L., Topham, D. J., and DeDiego, M. L**.. (2018) "Modulation of Innate Immune Responses by the Influenza A NS1 and PA-X Proteins."** *Viruses*, vol. 10, no. 12, 2018, p. 708.
- Rivas, H.G., Schmaling, S. K. and Gaglia, M. M. (2016) **. "Shutoff of Host Gene Expression in Influenza A Virus and Herpesviruses: Similar Mechanisms and Common Themes."** *Viruses-Basel*, vol. 8, no. 4, 2016, p. 102.


## Acknowledgments


- Prof. Gaglia and Lea Gaucherand along with the rest of the members of the Gaglia Lab
- Tufts Summer Scholars Program