

# **The SARS-CoV-2 D614G spike mutation increases transmissibility and prevalence in the United States**

**Fatimah Alghanem, JT Correy, Macie Duran, and Yaniv Maddahi**

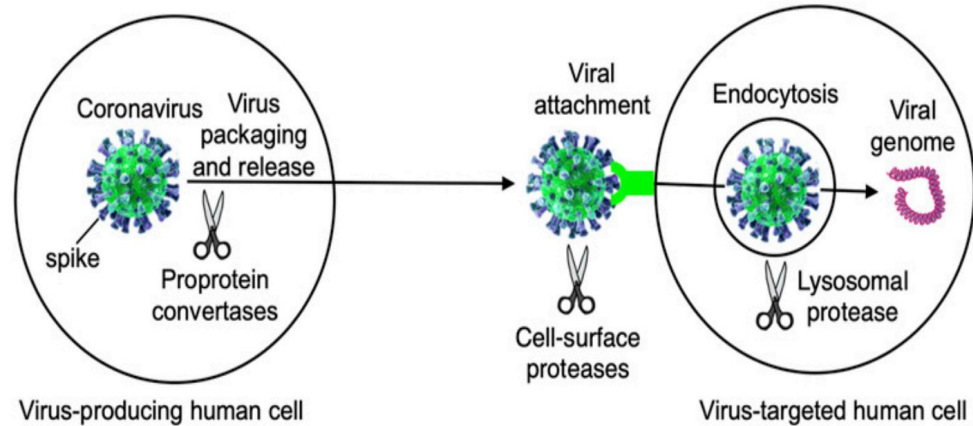
**BIOL 368: Bioinformatics Laboratory  
Loyola Marymount University  
December 10, 2020**

# Outline

- The SARS-CoV-2 spike D614G mutation increases viral transmissibility, giving it a fitness advantage.
- The G614D mutation has quickly become the dominant SARS-CoV-2 genotype in the United States.
- Tracking spike protein mutations and exploring epidemiology of SARS-CoV-2 will allow for better development of therapeutic strategies.

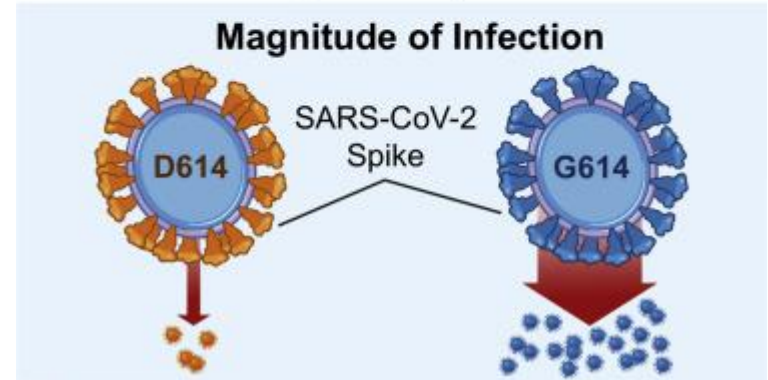
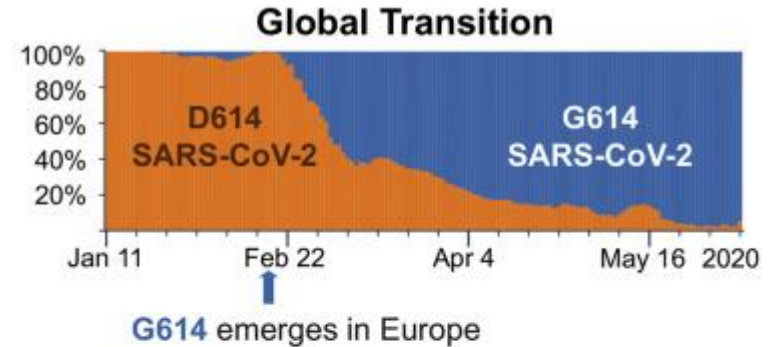
# SARS-CoV-2 is dangerous due to the way it penetrates cells

- Over 1 million global deaths and over 28 million infections
- SARS-COV-2 receptor binding domain (RBD) has a substantially higher affinity than SARS-COV.
- SARS-COV-2 is activated by furin protease which allows it to be more infectious in the lungs.
- As of July 3, 2020, 216 countries have reported COVID-19 cases, with more than 10 million confirmed cases and approximately 518,000 deaths.



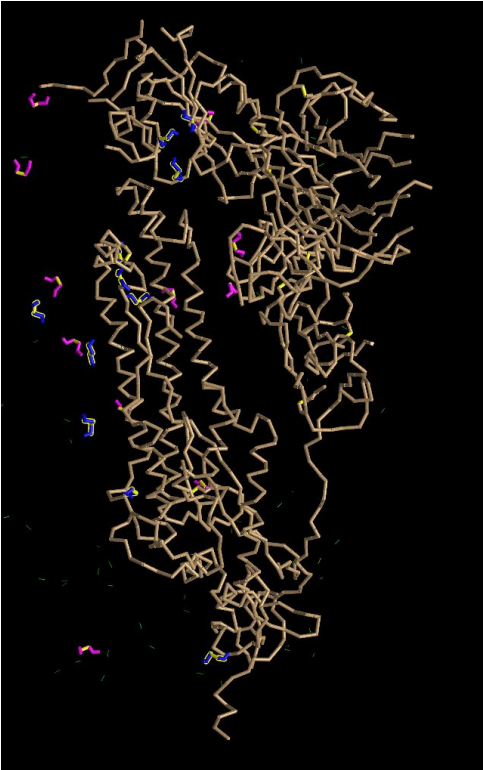
# G614 is more transmissible than D614, causing an increased prevalence of the mutation

- G614 mutant is 31% (28-34%) more transmissible than the D614 wild type.
- 17,760 D614G mutations were found in the GISAID database by June 2, 2020.
- A SARS-CoV-2 variant with Spike G614 has replaced D614 as the dominant pandemic form.
- The consistent increase of G614 at regional levels may indicate a fitness advantage.

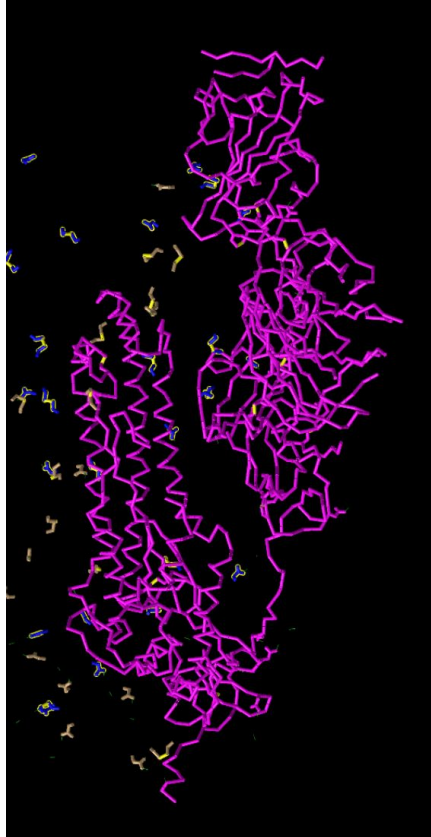


# Structure of the SARS-CoV-2 Spike Glycoprotein

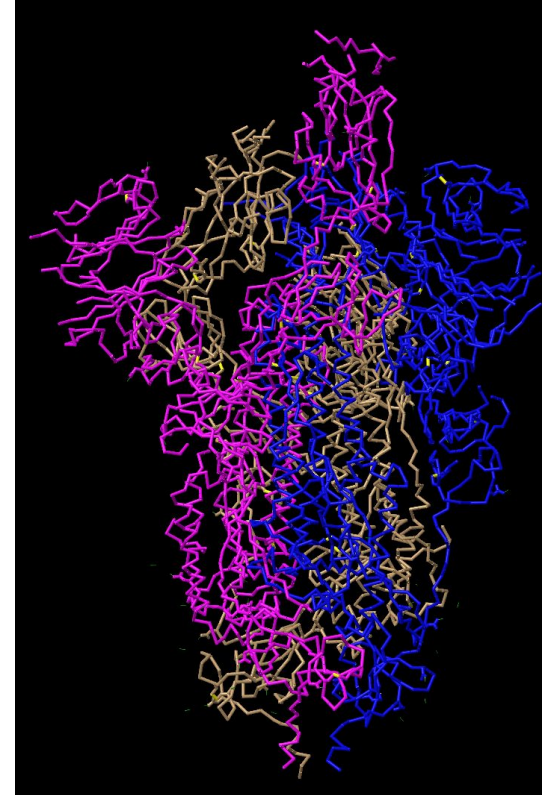
Closed protein



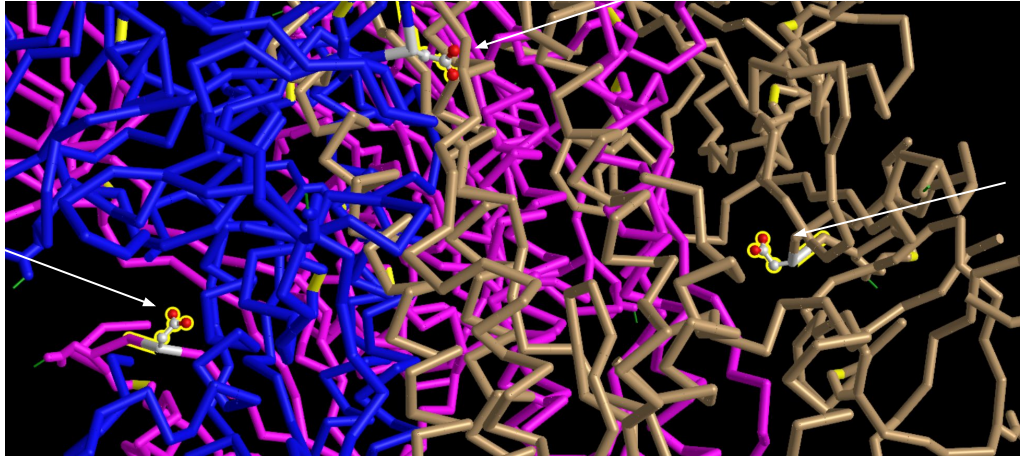
Open protein



Full spike



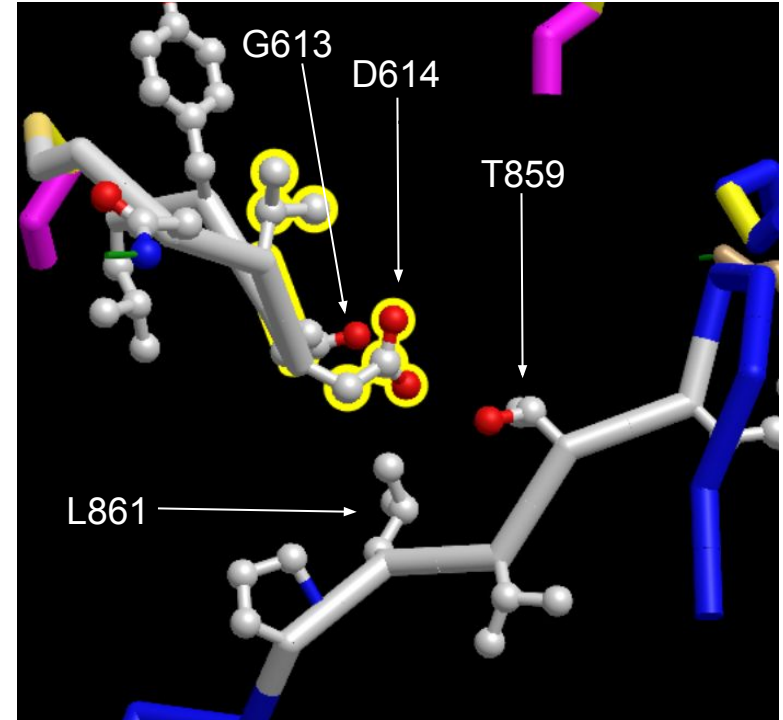
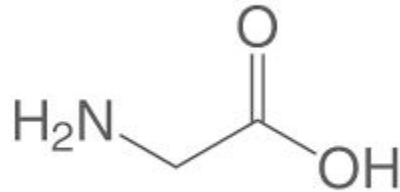
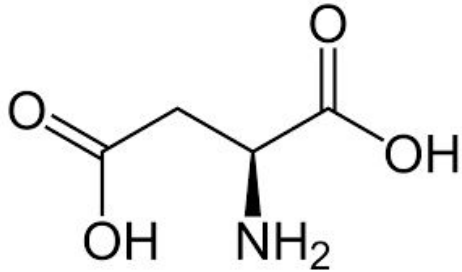
# Location of the D614G mutation in the SARS-CoV-2 Spike Glycoprotein





# Interactions around the D614G mutation

- Both sequences are closely surrounded by polar sidechains
- Switching to the uncharged glycine reduces the interaction with other side chains



# Outline

- The SARS-CoV-2 spike D614G mutation increases viral transmissibility, giving it a fitness advantage.
- The G614D mutation has quickly become the dominant SARS-CoV-2 genotype in the United States.
- Tracking spike protein mutations and exploring epidemiology of SARS-CoV-2 will allow for better development of therapeutic strategies.



# Mutation G614 became more prevalent than D614 in the United States

- 2 amino acid positions showed high conservation (5 and 243).
- Amino acid 614 is the only position that shows weak conservation.
- Alignment indicates that the D614G mutation became more prevalent in the United States around March or April.

```
QMS48295.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY May
YP_0097243      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY Wuhan-1
QH060594.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY January
QH071973.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY January
QH082464.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY January
QID21068.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY February
QJI54050.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY February
QKU53745.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY February
QOF12173.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY March
QIZ15909.1      GTNTSNQVAVLYODYNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY March
QOE81952.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY April
QIZ16077.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY March
QLC92636.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY April
QLY88731.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY April
QOC77751.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY May
QPG83379.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY May
QLI50233.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY June
QNR99264.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY June
QMT48961.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY June
QNL23817.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY July
QNL11630.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY July
QNO40800.1      GTNTSNQVAVLYGGNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSY July
*****
```

# Mutation G614 became more prevalent than D614 in the United States

- Sequences with the D614G mutation were being reported to GISAID as early as February.
- The prevalence of G614 surpassed that of D614 by March.

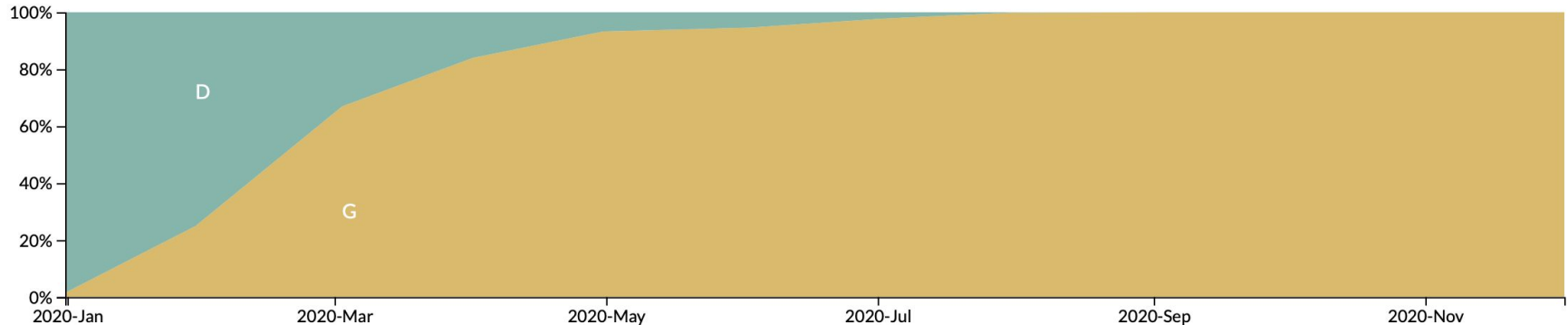
**Table 1: Prevalence of D614G mutations in sequences collected and reported to GISAID by month**

<b>Month collected</b>	<b>Total sequences collected and reported</b>	<b>Sequences with D614G</b>	<b>Proportion of sequences with D614G</b>
January	15	0	0.000
February	116	8	0.069
March	11,327	8,336	0.736
April	8,895	7,719	0.868
May	6,062	5,756	0.950
June	8,178	8,040	0.983
July	4,017	3,998	0.995

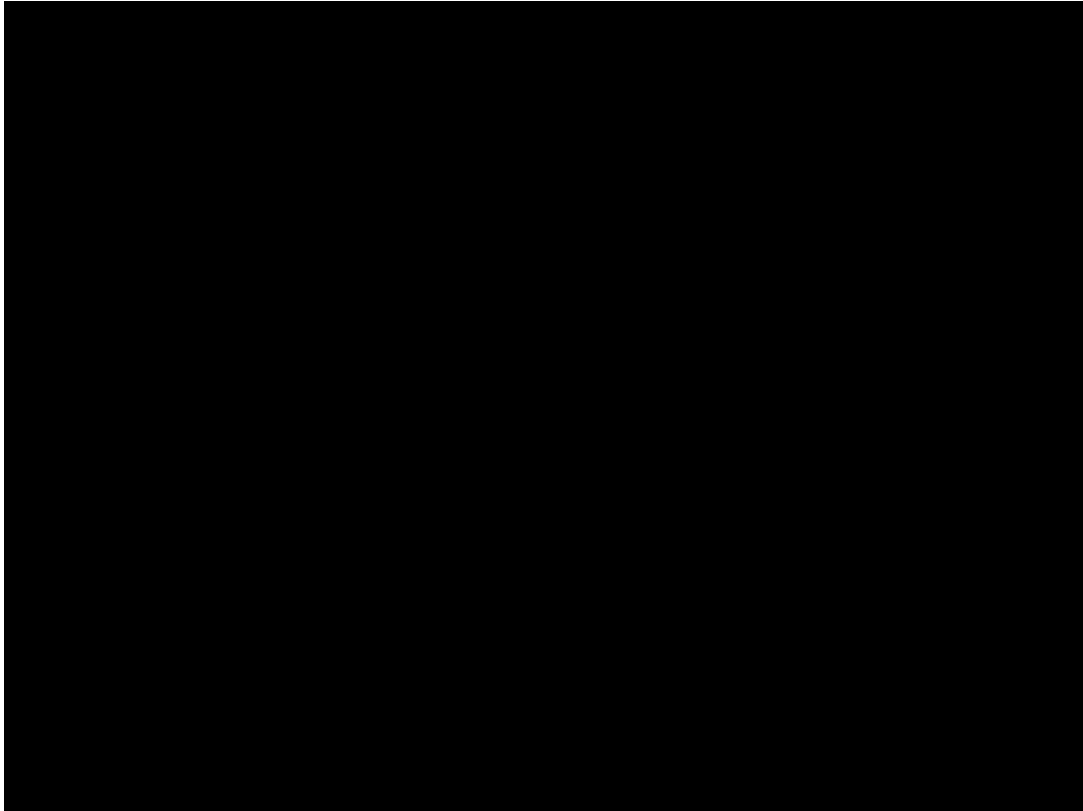
# G614 has become the dominant genotype in the United States

- The database Nextstrain, which utilizes sequences from GISAID, indicates that prevalence of G614 surpassed D614 by early March.

Frequencies (colored by Genotype at S pos 614 and normalized to 100% at each time point for 1893 out of a total of 3685 tips)



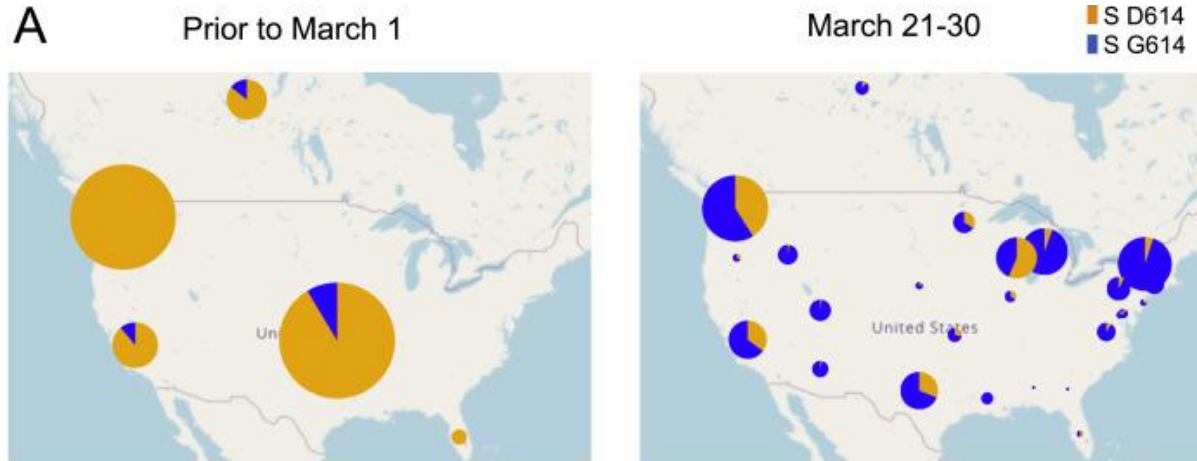
(Nextstrain, 2020)



Nextstrain, 2020

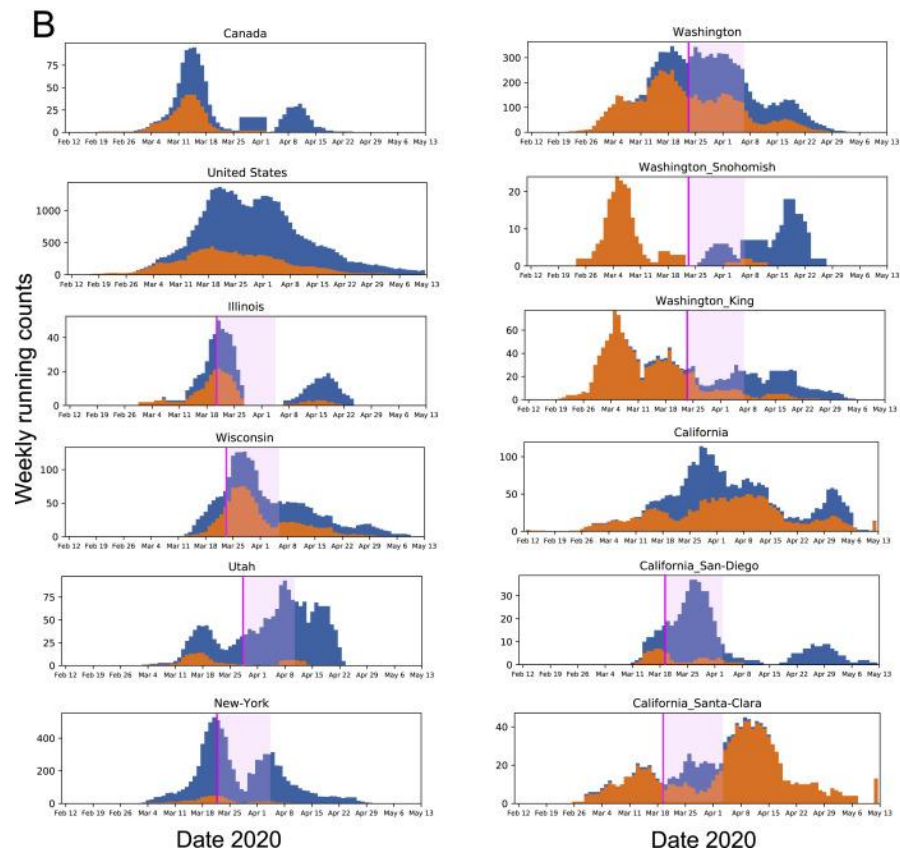
# G614 has become the dominant genotype in the United States

- Maps of the relative frequencies of D614 and G614 in North America in two different time windows



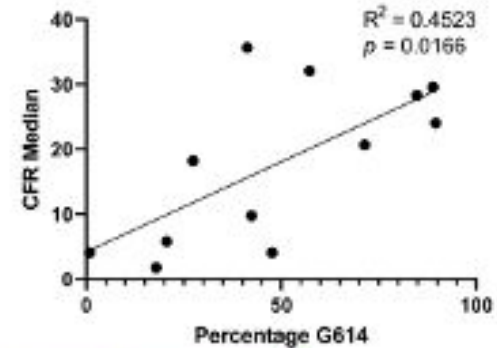
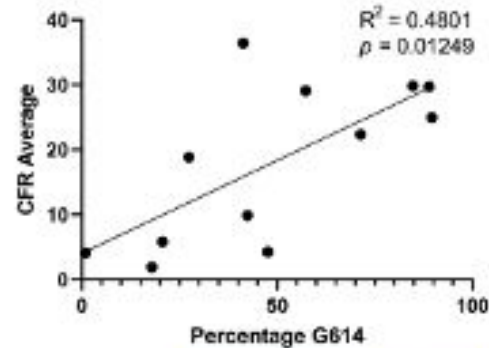
# G614 has become the dominant genotype in North America

- Blue represents G614.
- Orange represents D614.
- Weekly running counts of G614 illustrating the timing of its spread in North America.
- Stay-at-home order date is shown as a pink line.
- The introduction of G614 led to the decline in D614 cases.



# Case Fatality Rate (CFR) and percentage of virus exhibiting G614

- Linear regression of average case fatality rate with percentage of viruses exhibiting a glycine (G) at amino acid position 614.
- China exhibiting G614 was less than 1% and this country exhibited the second lowest average or median CFR.



Country	Percentage G614 Mutation	CFR Average	CFR Median
Belgium	88.89	29.73	29.57
France	84.82	29.83	28.32
China	0.89	4.07	4.07
Germany	47.62	4.25	4.08
Netherlands	57.30	29.07	32.07
Brazil	71.43	22.36	20.67
Canada	42.40	9.89	9.76
Italy	89.66	24.93	24.04
Australia	17.91	1.92	1.77
Japan	20.62	5.82	5.82
Spain	41.30	36.43	35.66
US	27.42	18.85	18.20

(Becerra-Flores, 2020)



# Outline

- The SARS-CoV-2 spike D614G mutation increases viral transmissibility, giving it a fitness advantage.
- The G614D mutation has quickly become the dominant SARS-CoV-2 genotype in the United States.
- Tracking spike protein mutations and exploring epidemiology of SARS-CoV-2 will allow for better development of therapeutic strategies.

## **An understanding of the epidemiology of SARS-CoV-2 may further explain the dominance of the G614 mutation**

- Because viral transmission is dependent on whom it infects, it is important to further explore the epidemiology of SARS-CoV-2.
- Increased infectiousness may be solely responsible for the increased prevalence, but it is uncertain if other factors play a role.
- Tracking mutations in the spike protein, especially in the RBD region, is important during the development of therapeutic strategies.

# Summary

- The SARS-CoV-2 spike D614G mutation increases viral transmissibility, giving it a fitness advantage.
- The G614D mutation has quickly become the dominant SARS-CoV-2 genotype in the United States.
- Tracking spike protein mutations and exploring epidemiology of SARS-CoV-2 will allow for better development of therapeutic strategies.

# Acknowledgments

We would like to thank Dr. Dahlquist, our TA Annika, and the rest of our Bioinformatics Laboratory team for helping out with this project.

We would also like to thank Loyola Marymount University for making this project possible.



# References

- Becerra-Flores, M., & Cardozo, T. (2020). SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. *International Journal of Clinical Practice*.
- Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., ... & Hastie, K. M. (2020). Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*, 182(4), 812-827.
- Leung, K., Pei, Y., Leung, G. M., Lam, T. T., & Wu, J. T. (2020). Empirical transmission advantage of the D614G mutant strain of SARS-CoV-2. *medRxiv*.
- Nextstrain. (2020). Genomic epidemiology of novel coronavirus - North America-focused subsampling. <https://nextstrain.org>
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C., Abiona, O., . . . McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367(6483), 1260-1263. doi:10.1126/science.abb2507