THE CANADIAN SCIENCE FAIR JOURNAL

YOUREKA CANADA

ARTICLE

Long-Term Effects of Influenza Vaccines on the Prevalence of Specific H3N2 Genetic Clades

Elsa George, Faraneh Moayyed, Mónica Figueroa, Will Taylor and Kieran Maheden*

Youreka Canada | yourekacanada.org

This project came in 1st place in the Youreka Canada National Competition

***Senior Author**

Influenza is a viral, highly transmittable respiratory pathogen that creates annual health and economic burdens. Flu vaccines are updated bi-annually due to drifting of the genes encoding influenza surface proteins hemagglutinin and neuraminidase. Current influenza vaccines target two fluA subtypes (H1N1 and H3N2) and two fluB lineages. The large genetic diversity of H3N2 means vaccines targeted at one genetic clade may have little effect on other H3N2 clades. Using the Francis Crick Institute's Influenza Reports, we normalized the proportion of circulating H3N2 clades and compared them to other flu seasons through a two-tailed Fisher's Exact Test. In addition, vaccine effectiveness was estimated by looking at hemagglutinin inhibition assays as performed using ferret antisera. Different patterns of persistence of H3N2 genetic clades were found in the circulating strains from 2015 to 2019. Occasionally, new strains were introduced as ancestral clades were divided into smaller subclades. Four significant changes in clade proportion were found, which occasionally matched an effective vaccine season. Sometimes, an effective vaccine did not correspond with a significant drop in proportion for that clade. Due to the multifactorial nature of seasonal flu, vaccine effectiveness can not be used solely to predict the pervasiveness of a circulating strain. Nevertheless, long term patterns are observable and can contribute to our understanding of influenza immunization over the past 5 years.

INTRODUCTION

Viral diseases represent one of the largest healthcare concerns currently facing both the developed and the undeveloped world. Influenza (flu) is a viral pathogen capable of causing both seasonal epidemics and worldwide pandemics. Flu consists of 2 main types: fluA and fluB. Two subtypes of fluA (H1N1 and H3N2) and two strains of fluB pose most significant burden to humans around the world.[4] Flu can cause mild to severe illness and lead to possibly life-threatening complications such as pneumonia. In addition to creating health issues, the flu affects global economies.

The estimated total economic burden influenza had on the United States in 2015 was $11.2 billion.[3] It is estimated for the United States that Influenza has been the cause for 9-45 million illnesses, 40-810 thousand hospitalizations and 12-61 thousand deaths annually since 2010.[4] Worldwide, estimates range up to one billion infections annually, and seasonal influenza may result in 290-650 thousand deaths each year due to respiratory complications alone. [11,14]

All influenza genomes are composed of eight single-stranded RNA segments. The fluA virion contains antigenic glycoproteins hemagglutinin (HA) and neuraminidase (NA) (**Figure 2**). Eighteen different HA subtypes and eleven different NA subtypes have been characterized for fluA, making possible a wide variety of capsid proteins. FluB forms only two antigenically distinguishable lineages, Victoria and Yamagata, based on the antigenic properties of hemagglutinin. FluB is exclusively found in humans, and mutates 2-3 times slower than fluA.[1] Notably, fluA HA subtypes H1, H2, and H3, and NA subtypes N1 and N2 have established stable lineages in humans and cause widespread human respiratory infection.[2] There have only been sporadic instances of infection from H5, H7, and H9 viruses. Although viruses such as H5N1 and H9N7 have the potential to cause severe disease and death, no sustained transmission between humans has been reported.[5] Current influenza vaccines focus on fluA subtypes H1N1 and H3N2, and the two fluB lineages, because of their high rate of transmission among humans. Seasonal epidemics of disease, commonly known as the flu season, are caused by human fluA and fluB viruses, while pandemic outbreaks have been the result of emergence of novel fluA viruses. Seasonal flu occurs bi-annually, but pandemics are relatively rare. In the last century, there were four flu pandemics: the H1N1 Spanish flu of 1918, the 1957 H2N2 virus, the 1968 H3N2 virus, and the 2009 H1N1 virus. Globally, the Spanish flu is estimated to have killed nearly 50 million people, the 1957 and 1968 viruses are each estimated to have killed a million people, and the 2009 pandemic was estimated to have killed between 151,700 and 575,400 people.[6] Flu virus antigens morph from year to year, preventing development of a single, constant vaccine (**Table 1**).

Two mechanisms for this are antigenic drift and antigenic shift.[2] Antigenic drift is caused by mutations in the genes encoding HA and NA. Albeit small, these genetic changes add up over

time, eventually reducing the familiarity of the viral antigens to human immune systems trained to recognize previous flu strains. Antigenic shift is the result of HA gene recombination creating entirely new antigenic structures, to which the human population has no pre-existing immunity. Antigenic drift is responsible for annual flu variation, while antigenic shift can lead to pandemics.[2] An ideal universal influenza vaccine would protect against antigenic drift and shift, conferring lifelong immunity. The flu transmits best in cold, dry, conditions.[10] As geographical seasons shift over time, flu outbreaks tend to follow winter (**Figure 1**). The Northern Hemisphere's flu season is December to March, while the Southern Hemisphere's is June to September. The equatorial region has no specific flu season, and outbreaks rarely occur.[7] Recently, flu vaccines contain inactivated virions of three (trivalent vaccine) or four (quadrivalent vaccine) influenza subgroups. These subgroups being H1N1, H3N2, Victoria lineage, and, if quadrivalent, Yamagata Lineage.[8] The subgroups are each represented in the vaccine by viruses prototyped off a strain recommended by the World Health Organization (WHO). WHO makes decisions based on bi-annual meetings of directors from its Collaborating Centers, which are prestigious national institutes for infectious diseases.[4] Vaccines are generally targeted towards viruses most genetically similar to those expected to circulate in the coming flu season. Egg-based vaccines are the most common and take approximately six months to manufacture on industrial scales. This leads some manufacturers to pre-empt the WHO recommendation and start production early.[4]

Due to the constant mutation of flu viruses, periodic surveillance on circulating strains is necessary to design an effective vaccine for each flu season. Currently, flu vaccines are produced by analyzing the trends in proportions of strains present in the previous flu seasons and making an educated guess as to which strains should be used in the vaccine to protect people most effectively (**Table 2**). Namely, the hemagglutinin inhibition (HI) assay is used on strain isolates from previous years to determine the creation of flu antibodies. Currently available influenza vaccines are designed to respond to a specific strain and may be ineffective against new influenza viruses. The difficulty in predicting circulating strains has frequently resulted in mismatch between the annual vaccine and circulating viruses, and despite the large amount of work on the virus, little is known on how seasonal vaccines cause changes over long timeframes. Analyzing the behaviour of the virus over multiple seasons may provide valuable data to make more informed decisions, as well as understand how the virus shifts. While the ultimate goal of influenza vaccination strategies is to generate a vaccine targeting a universally present epitope, the work outlined below serves towards an interim strategy. By analysing previously reported genotypic data and corresponding annual antigenic data in long timeframes retroactively, we gain a better understanding of influenza complexity as well as an opportunity to make intelligent vaccine decisions.

MATERIALS AND METHODS

We examined circulating flu strains over a span of five years (February 2015 - September 2019) and organized the data in a proportion versus time graph. The data on circulating clades of H3N2 was collected from the publicly available Francis Crick Institute's Interim Influenza Reports. These clade proportions were then analysed using Fisher's Exact Test. Significant changes between years were indicated with * for significance at α=0.001. Data was normalized prior to analysis by standardizing the total counts per year and setting the lowest value per strain per year to 1. On the other hand, the antigenic analyses of H3N2 viruses reported by the Crick Institute was reviewed to examine the vaccine efficacy of any given year, inferred from antigenic data generated by HI assays. Lastly, a phylogenetic tree of A(H3N2) clades and subclades was generated using the Francis Crick Institute's information on clade-defining mutations.

Antigenic data reported by the Crick Institute originates from in-house HI assays. The viruses were propagated using egg-based systems, then used to test for red blood cell agglutination. HA titers were determined by titrating sera obtained from ferrets vaccinated with the given year's vaccine strain, and the fold of the inhibition delivered by the sera was reported. Samples were sequenced and organized into clades according to defined mutations in the HA gene. Samples for both antigenic data and genetic sequencing were collected randomly, and therefore amenable to statistical analysis.

RESULTS

Different patterns of persistence are seen in the proportions of a circulating strain (**Figure 3**). Some remain at a low level over the entire period while some appear for only 3 seasons at a very high proportion. Occasionally, new strains are introduced as ancestral clades are divided into smaller subclades. 3C.3a is the only clade remaining present throughout all flu seasons analysed, with higher proportions on 2015.5-2017. 3C.3a increased in 2015.5-2016 and decreased in 2016-2017. It remained on low proportions until 2019, when it increased again. Overall, 3C.3a exhibits low proportions in comparison to other clades. Clade 3C.2a was the predominant strain from 2015 to 2016.5, and it disappeared after 2018. It increased from 2015 to 2016.5 and from 2017 to 2018. There is an observable decrease in proportion of 3C.2a between 2016.5 and 2017. Clade 3C.2a1 appeared for the first time in 2017, decreased in proportion for one year and disappeared afterwards. 3C.2a1 is the ancestral clade of 3C.2a1b (**Figure 5**), which appeared for the first time in 2018.5 and persisted onwards as a predominant clade.

There is an observable increase in proportion for 3C.2a1b from 2018.5 to 2019 and a decline from 2019 to 2019.5. 3C.2a is the ancestral clade of 3C.2a2 (**Figure 5**), which appeared for the first time in 2018.5 with high proportion, and decreased in the following two seasons. There is an observable decrease in proportion from 2018.5 to 2019, and the clade remained in low proportion in

the following season. Clade 3C.2a3 appeared for the first time in 2018.5, and it has persisted onwards with low proportions. Clade 3C.2a4 appeared for the first time in 2019, and it has had low proportions for two seasons. Clade 3C.3b appeared for the first time in 2015, it declined consistently until it disappeared in 2016. Clade 3.3 appeared for the first time in 2015.5, it decreased for two seasons and disappeared in 2016.5.

Proportions were analysed by comparing each year to the preceding year via two-tailed Fisher's Exact Test. Four significant changes between influenza seasons emerged. Between 2015.5 to 2016, the proportions of 3C.3a and 3C.2a increased and the other two strains declined. In this season 3C.3a was vaccinated against. In the period from 2016.5 to 2017, the proportions of both 3C.2a and 3C.3a decreased, with 3C.2a being vaccinated against. Between 2018 to 2018.5, the proportions changed significantly because two clades (3C.2a and 3C.2a1) ceased to appear on the graph, and two clades emerged in the following season. 3C.2a disappeared into its subclade 3C.2a2, and 3C.2a1 disappeared into the subclade 3C.2a1b. Between 2018.5 to 2019, the proportion of 3C.2a2 declined observably and the proportion of 3C.2a1b grew.

There is variability in vaccine efficacy as predicted by HI assay (**Figure 4**). Some seasons (2015, 2015.5, 2018.5, and 2019) exhibit high vaccine effectiveness (percentage of >8 fold) while other seasons varied between moderate (4 fold) to low (<2 fold) vaccine effectiveness. In 2015, 94.1% of isolates were >8 fold decrease in titre as compared to unvaccinated sera. This high degree of inhibition indicates a more effective vaccine in 2015. There is an even greater degree of inhibition in 2015.5, with 98% at titres within >8 fold which corresponds to a greater effectiveness of the vaccine for 2015.5. In 2016, vaccine effectiveness decreased as only 45.7% were at titres within >8-fold. In 2016.5 and 2017, the isolates that were at titres within 4-fold had the highest percentage, 47.3% and 48.2% respectively. Additionally, in those same seasons, the isolates at titres within <2 fold had a larger percentage than the isolates at titres within >8 fold. This increase of isolates at titres within <2 and 4-fold corresponds to a less effective vaccine for 2016.5 and 2017. The least effective vaccine can be seen from the results of 2017.5 where 48% of isolates were at titres within <2-fold. The percentage of isolates at titres within >8-fold was 26.6% which is similar to the 27.4% of isolates that were at titres within 4-fold. In the following season, 2018, the percentage of isolates at titres within 4-fold and >8-fold increased to 46.7% and 36.6%, respectively. The greater percentages indicate that there was a higher vaccine effectiveness for that season. In 2018.5, the percentage of isolates at titres within >8-fold spikes to 80.6% with only 1.5% at titres within <2-fold. Compared to the previous season, the results of 2018.5 indicate a higher vaccine effectiveness. The highest vaccine effectiveness occurs during the 2019 season as 98.1% of isolates were at titres within >8-fold and there were no isolates recorded at titres within <2-fold. The 2019.5 season sees a decrease in the effectiveness of the vaccine as the percentage of isolates at titres within >8-fold decreased to 53.6%.

The significant changes in clade proportions occasionally match an effective vaccine season. Sometimes, an effective vaccine does not correspond with a significant drop in proportion for that clade. The first statistically significant season 2016 where there was a decrease in proportion for clades 3.3 and 3C.3b but an increase in proportion for clades 3C.2a and 3C.3a (**Figure 3**). The strain targeted by the vaccine that year was 3C.3a and it had a vaccine effectiveness as 45.7% of isolates were at titres within >8-fold. This is a season where the vaccine effectiveness does not correspond to a decrease in the proportion of 3C.3a that is circulating. There was statistical significance in 2017 where the only two circulating strains, 3C.3a and 3C.2a, decreased in proportion. However, this does not correspond with the vaccine efficacy as predicted by HI assay (**Figure 4**) in that only 14.3% of isolates were at titres within >8-fold. There is a drastic drop in the proportion of 3C.2a in 2017 (**Figure 3**) although there is not a high vaccine effectiveness for that season. 2017 may also be statistically significant because while all circulating strains are decreasing, there is an addition of a new clade, 3C.2a1, the following year. The two-tailed Fisher's Exact Test indicated that there was statistical significance in the 2018.5 season. Based on the data from the Crick's Influenza Report, there is only one circulating strain, 3C.3a, in that season. Significance may be marked by the disappearance of the strains 3C.2a and 3C.2a1. During that same season, the targeted vaccine was 3C.2a1 and 80.6% of isolates were at titres within >8-fold. This is significant because there was no proportion of the strain targeted by the recommended vaccine.

Later in 2019, two new strains are named, 3C.2a1b and 3C.2a2 which replace the strains 3C.2a1 and 3C.2a, respectively, because the parental clades disappear into their subclades (**Figure 5**). 2019 was a season with statistically significant change in strain proportion. In this season, the targeted strain for the recommended vaccine was 3C.2a2 and the graph shows a drastic drop in the proportion of 3C.2a2 viruses circulating. In addition, the decrease in the proportion of 3C.2a2 viruses corresponds to the high vaccine effectiveness indicated by the 98.1% of isolates that were at titres within >8-fold.

DISCUSSION

Due to the multifactorial nature of seasonal flu, vaccine effectiveness can not be used solely to predict the pervasiveness of a circulating strain. Long term patterns are observable, with strains persisting throughout the 5 year period and rising to abundance in later years (**Figure 3**, 3C.3a). Throughout the time period, there is observably a high degree of variability, with strains increasing in prevalence despite sometimes being the vaccination target. As

well, some strains disappear without a target vaccine. Throughout the 5 year period, new strains emerge. Some continue to circulate after being introduced while some either disappear into their subclades or stop circulating entirely. This high degree of variation, at times seemingly unaffected by vaccination strategies, may originate from the global scope of our data, which represents many geographically different locals which vary in terms of density and climate. It is also difficult to rule out the effectiveness of implemented vaccines at curbing specific strains, as there is no available data examining cohorts of individuals who are unvaccinated for which specific strains establish infections. Indeed, due to the nature of our data source (isolates taken from infected individuals worldwide), we may expect individuals who have not been vaccinated successfully to be overrepresented in our data. That being the case, the population of individuals who have gained protection against specific strains would still be a factor in influencing the overall proportions of virus present at any given time.

Recently, there has been a considerable amount of research done on the factors that affect vaccine effectiveness. It was observed that in the previous 10 seasons, patients who had received no vaccinations had the highest vaccine effectiveness for the current season.[15] This result suggests that reduced vaccine effectiveness can be a result of repeated vaccination. Although current seasonal vaccinations are likely to provide some protection against influenza regardless of the number of vaccines, it is possible that repeated vaccinations were a contributing factor to the varying proportions seen in our results. However, it is difficult to relate the proportion of viruses in each H3N2 clade with the impact of repeated vaccinations because the vaccine and medical history of individuals remains unclear. On the other hand, the concept of original antigenic sin, first proposed by Thomas Francis, Jr. in 1960, can be used to explain why some circulating strains prevailed despite being targeted by a recommended vaccine with the majority of their isolates at titres within >8-fold. The concept suggests that when a pathogen mutates slightly, the immune system uses the memory of the earlier infection to respond to the new pathogen. However, mounting another response to the new pathogen would allow for stronger responses in the future.[13] This idea can be used to describe what occurs in the immune system of people who receive a vaccine for a new strain of the H3N2 virus. It may be possible that instead of adapting and altering its immunological response, the vaccinated individual relies on their memory of the previous, antigenically similar strain to mount a response. Lastly, H3N2 viruses mutate rapidly and vaccine strain selection has proven to be a difficult process as antigenic characterization of circulating viruses has become more challenging. As a result, vaccine effectiveness in some studies was as low as 9% for the flu season from November 2018 to May 2019. 16 Additionally, it was found that in the 2018-2019 season in the United States, the recommended virus was 3C.2a1 but the predominant clade was 3C.3a.[16] This indicates how a mismatch in the vaccine strain can lead to a low vaccine effectiveness during the season. Although vaccine efficacy as predicted by HI assay can correspond to the effectiveness of a vaccine (**Figure 4**), proportion of the clades such as 3C.3a can start increasing. Therefore, mismatches in vaccine strain selection can result in the varying proportions of circulating H3N2 clades.

Working through the data presented, a number of imperfections were undertaken to obtain a tangible result. Specifically, the H3N2 virus is prone to mutating while replicating inside the fertilized chicken egg the virus is cultured in.[14] This means that both HI assays and final flu vaccines are not strictly using the genetics of their prototypical strain. One possible way to overcome this is to culture the H3N2 virus in mammalian cells. These vaccines exist, however the majority of global flu vaccines are egg-based. Since our project is dealing with samples of the human population, we were required to restrict ourselves to analysing the vaccine type most relevant to trends in global data. In recent years, many circulating H3N2 strains have become less effective at agglutinating red blood cells, greatly reducing the accuracy of HI assays.[14] Ferrets are used as test subjects for fluA vaccines, due to a respiratory tract resembling humans'. However, the ferrets used are antigenically naïve, as they have not been exposed to the flu or flu vaccine previously. As discussed previously, this is a sharp distinction from humans receiving the vaccine, as people have likely been exposed numerous times. Unfortunately, we were only able to find one source that samples worldwide for genetic and antigenic data, the Crick Institute. The Crick's sample size from the seasons analyzed ranged from 70 to 350 isolates. A larger sample size would be more representative of the human population, reducing sampling bias and enabling more accurate statistical analyses.

As outlined above, numerous hurdles present themselves before the field of flu vaccinology. Due to both technical and temporal limitations, an idealized and completely effective vaccination strategy hinging on global surveillance and industrial scale vaccine production may never come to be. Using our work as a foundation, however, many new avenues of research are available that allow us to work towards that ideal. When considering the needs of research that aims towards this goal, finding more reliable data sources would be at the top of the list. While our main data source was useful in providing both antigenic and genetic data, the sample size of the sequenced and assayed groups varied widely from year to year. One way to address this would be to find a second source to corroborate similar data and to compare their reports - unfortunately at the time of writing, no other global and publicly available source was found. Future work building off of this project would aim to validate many of the conclusions outlined above. As our work considered global data as a whole, one method of validation would be to zoom in to smaller populations, such as a single country, and track data in a similar manner as above. Further work characterizing the effect of repeated vaccination on vaccine effectiveness also exists as an unanswered question in the field. *In vivo* work aiming at uncovering a mechanism leading to lower vaccine success in subsequent years may help us better plan

mass-scale vaccination, for flu and as well as other diseases. A final question to ask would probe the causes of the appearance of novel strains of H3N2. As represented in our data, the appearance of a novel strain often coincided with a targeted vaccine against its parental strain, possibly indicating a causal link. Research into the selective pressures generated by currently used vaccine strategies may allow the development of vaccines that limit the escape of parental strains into their ancestors via mutation.

Overall, our work builds a foundation of data that allows for a retrospective look at the success of influenza immunization over the past 5 years, as well as highlighting the complexity behind globally coordinating H3N2 vaccination. This work is a step in the direction of more tailored global influenza strategies, further limiting the effect of this viral pathogen.

REFERENCES

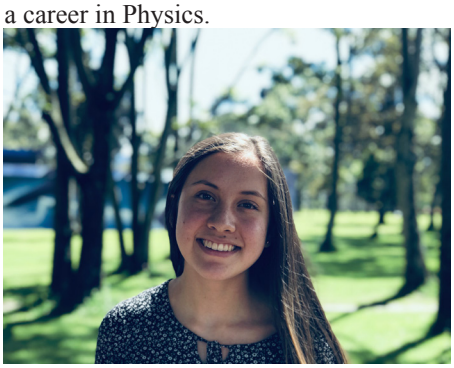
1. Nobusawa E, Sato K. Comparison of the mutation rates of human influenza A and B viruses. *J Virol.* 2006;*80*(7):3675–3678. doi:10.1128/JVI.80.7.3675-3678.2006
2. Berlanda Scorza F, Tsvetnitsky V, Donnelly JJ. Universal influenza vaccines: Shifting to better vaccines. *Vaccine* .2016;*34*(26):2926–2933. doi:10.1016/j.vaccine.2016.03.085
3. Putri WC, Muscatello DJ, Stockwell MS, Newall AT. Economic burden of seasonal influenza in the United States. *Vaccine* .2018;*36*:3960–3966. doi:10.1016/j.vaccine.2018.05.057 .
4. Disease Burden of Influenza. Centers for Disease Control and Prevention. https://www.cdc.gov/flu/about/burden/index.html
5. de Vries RD, Herfst S, Richard M. Avian Influenza A Virus Pandemic Preparedness and Vaccine Development. *Vaccines* (Basel) . 2018;*6*(3):46. Published 2018 Jul 25. doi:10.3390/vaccines6030046
6. Past Pandemics. Centers for Disease Control and Prevention. www.cdc.gov/flu/about/viruses/change.htm.
7. Krammer, F., Smith, G.J.D., Fouchier, R.A.M. et al. Influenza. Nature Reviews Disease Primers 4, 3, 2018. https://doi.org/10.1038/s41572-018-0002-y
8. World Health Organization Bi-Annual Reports for Recommended composition of influenza virus vaccines 9. How Influenza (Flu) Vaccines Are Made. Centers for Disease Control and Prevention. (CDC, 2019) https://www.cdc.gov/flu/prevent/how-fluvaccine-made.htm
10. Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog* . 2007;*3*(10):1470–1476. doi:10.1371/journal.ppat.0030151
11. Iuliano AD, Roguski KM, Chang HH, et al. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study [published correction appears in Lancet. 2018 Jan 19;:]. *Lancet* . 2018;*391*(10127):1285–1300. doi:10.1016/S0140-6736(17)33293-2
12. Thommes EW, Kruse M, Kohli M, Sharma R, Noorduyn SG. Review of seasonal influenza in Canada: Burden of disease and the cost-effectiveness of quadrivalent inactivated influenza vaccines. *Hum Vaccin Immunother* . 2017;*13*(4):867–876. doi:10.1080/21645515.2016.125153
13. Vatti, A., Monsalve, D. M., Pacheco, Y., Chang, C., Anaya, J.-M., & Gershwin, M. E. Original antigenic sin: A comprehensive review. *Journal of Autoimmunity*. 2017 Sep;*83*:12-21. doi: 10.1016/j.jaut.2017.04.008. Epub 2017 May 5.
14. Jorquera PA, Mishin VP, Chesnokov A, et al. Insights into the antigenic advancement of influenza A(H3N2) viruses, 2011-2018. *Sci Rep* . 2019;*9*(1):2676. Published 2019 Feb 25. doi:10.1038/s41598-019-39276-1
15. Kwong JC, Chung H, Jung JK, et al. The impact of repeated vaccination using 10-year vaccination history on protection against influenza in older adults: a test-negative design study across the 2010/11 to 2015/16 influenza seasons in Ontario, Canada. *Euro Surveill* . 2020;*25*(1):1900245. doi:10.2807/1560-7917.ES.2020.25.1.1900245
16. Flannery B., Kondor R.J.G., Chung J.R., Gaglani M., Reis M., Zimmerman R.K., Nowalk M.P., Jackson M.L., Jackson L.A., Monto A.S., Martin E.T., Belongia E.A., McLean H.Q., Kim S.S., Blanton L., Kniss K., Budd A.P., Brammer L., Stark T.J., Barnes J.R., Wentworth D.E., Fry A.M., Patel M. Spread of antigenically drifted influenza A(H3N2) viruses and vaccine effectiveness in the United States during the 2018-2019 season. *J Infect Dis.* 2020 Jan 1;*221*(1):8-15. doi: 10.1093/infdis/jiz543 .

ABOUT THE AUTHORS

MONICA FIGUEROA

Age: 17 | Edmonton, Alberta

Monica Figueroa grew up in Colombia and moved to Canada in early 2019. She will begin a Bachelor of Science in the Fall (2020) at the University of Alberta. Monica loves to cook, read and play the violin. She is interested in science and research, and will pursue a career in Physics.



KIERAN MAHEDEN

Age: 21 | Edmonton, Alberta

A recent graduate of the University of Alberta, Kieran completed his B.Sc. in Immunology and Infection while completing his honours thesis in the lab of Dr. Matthias Götte. Following long-harboured interests in the realm of systems and stem cell biology, Kieran will be starting graduate school (University of British Columbia) this coming Fall. When not in the lab (or quarantined at home) you can probably find him bouldering on campus.

ABOUT THE AUTHORS

FARANEH Y-MOAYYED

Age: 17 | Edmonton, Alberta

I recently graduated from Strathcona High School in Edmonton, AB. In the fall I will attending the University of British Columbia, pursuing a bachelor of science majoring in biochemistry. I am looking forward to gaining more research experience and exploring new topics such as oncology and psychology. In my spare time, you can find me on the field playing ultimate frisbee or hiking in the mountains.

WILLIAM TAYLOR

Age: 16 | Edmonton, Alberta

William Taylor is a student at Strathcona High School interested in many mathematical, scientific, and political areas of learning.

ABOUT THE AUTHORS

ELSA GEORGE

Age: 16 | Edmonton, Alberta

Elsa George is a Grade 11 student attending Archbishop MacDonald High School. She is interested in pursuing a science degree and becoming a specialized physician while contributing to her community. Elsa has received Honours with Distinction Awards throughout junior high and the Highest Academic Honours Award in Grade 9. She placed second school-wide in the Pascal Mathematics Contest.

APPENDIX

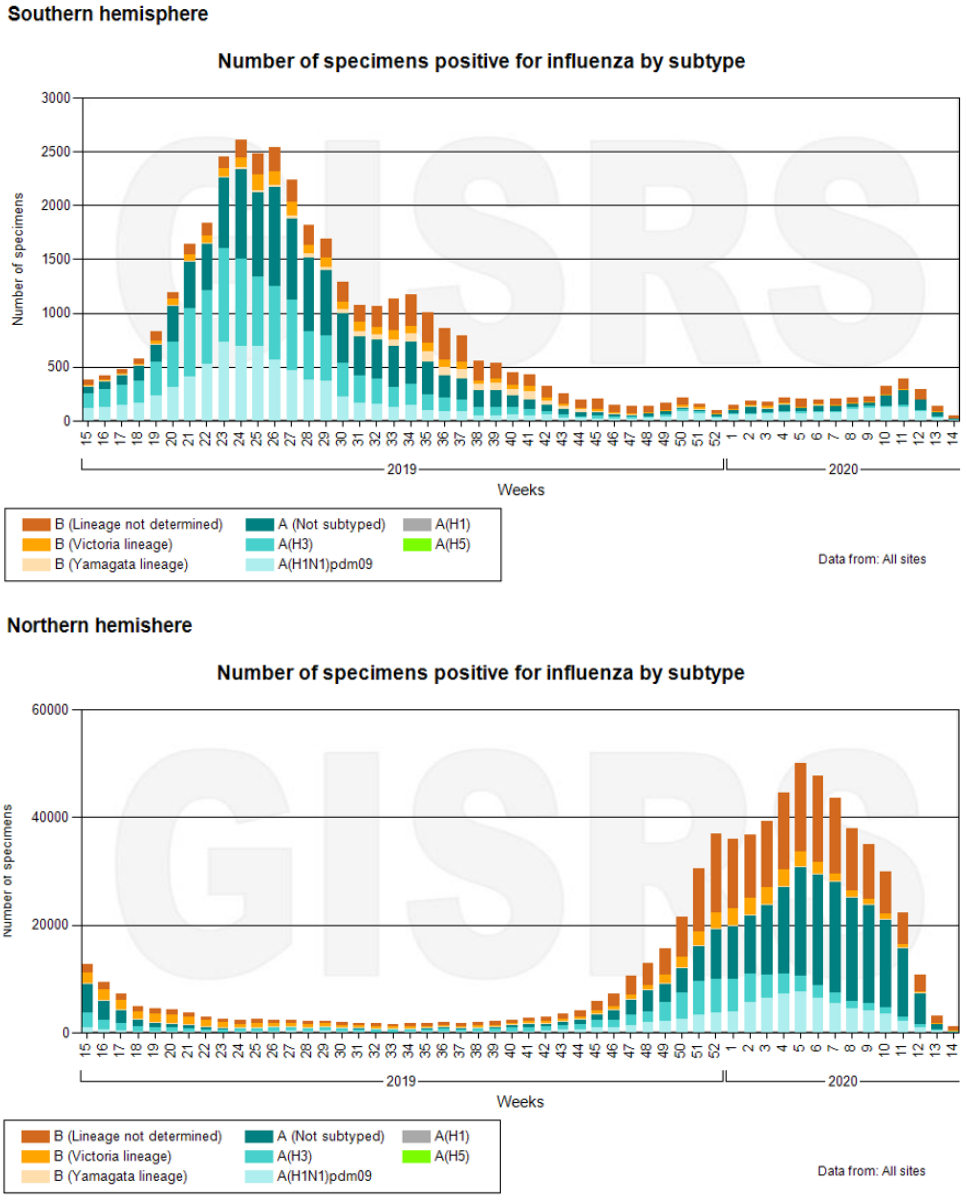


Figure 1: Influenza seasonality in the Northern and Southern Hemispheres. Flu season varies throughout the globe. The number of positive specimens peaks during flu season, and gradually declines as flu season ends. Adapted from WHO - FluNet, Global Influenza Surveillance and Response System (GISRS).

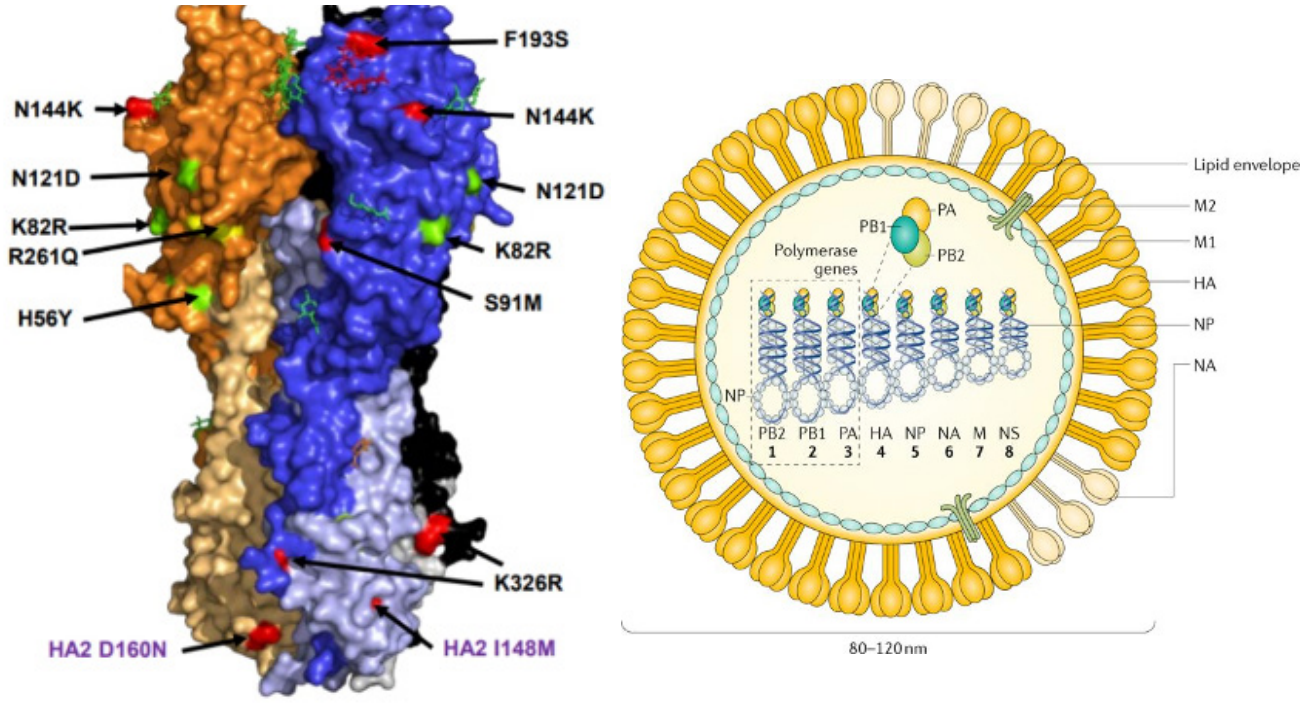


Figure 2: Left - H3N2 clade 3C.2a hemagglutinin, indicating amino acid substitutions which define the genetic clade. Right - Typical influenza virion, markers HA and NA indicate hemagglutinin and neuraminidase glycoproteins respectively. Adapted from the Francis Crick Institute Seasonal Influenza Report.

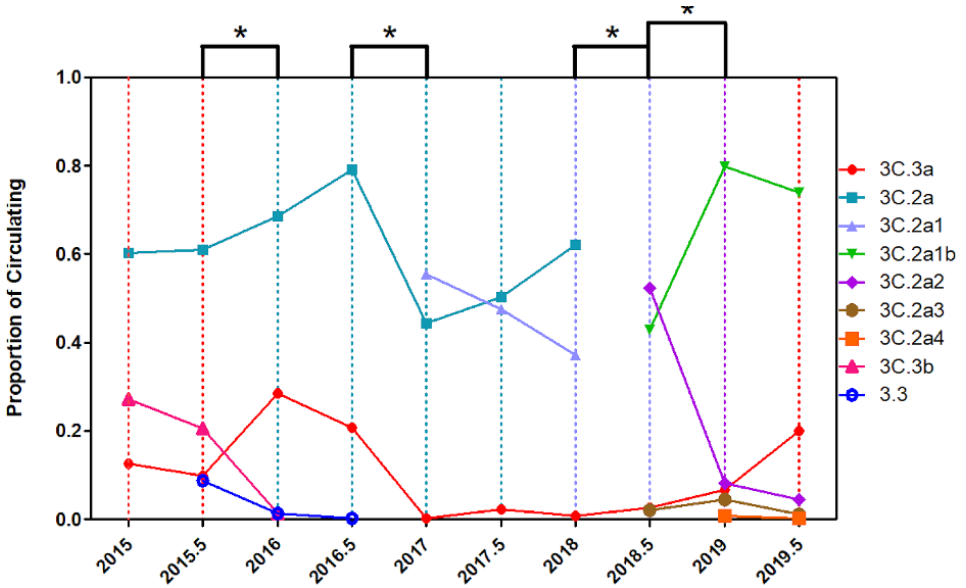


Figure 3: H3N2 strains versus time. Strains were colour coded. Dates on x-axis align with February and September Influenza Reports from the Crick Institute. The y-axis has been normalized to a proportion. The dotted lines correspond to the clade of the vaccine strain for that year. Proportions were analysed by comparing each year to the preceding year via two-tailed Fisher's Exact Test. Significant changes between years are indicated with * for significance at α=0.001

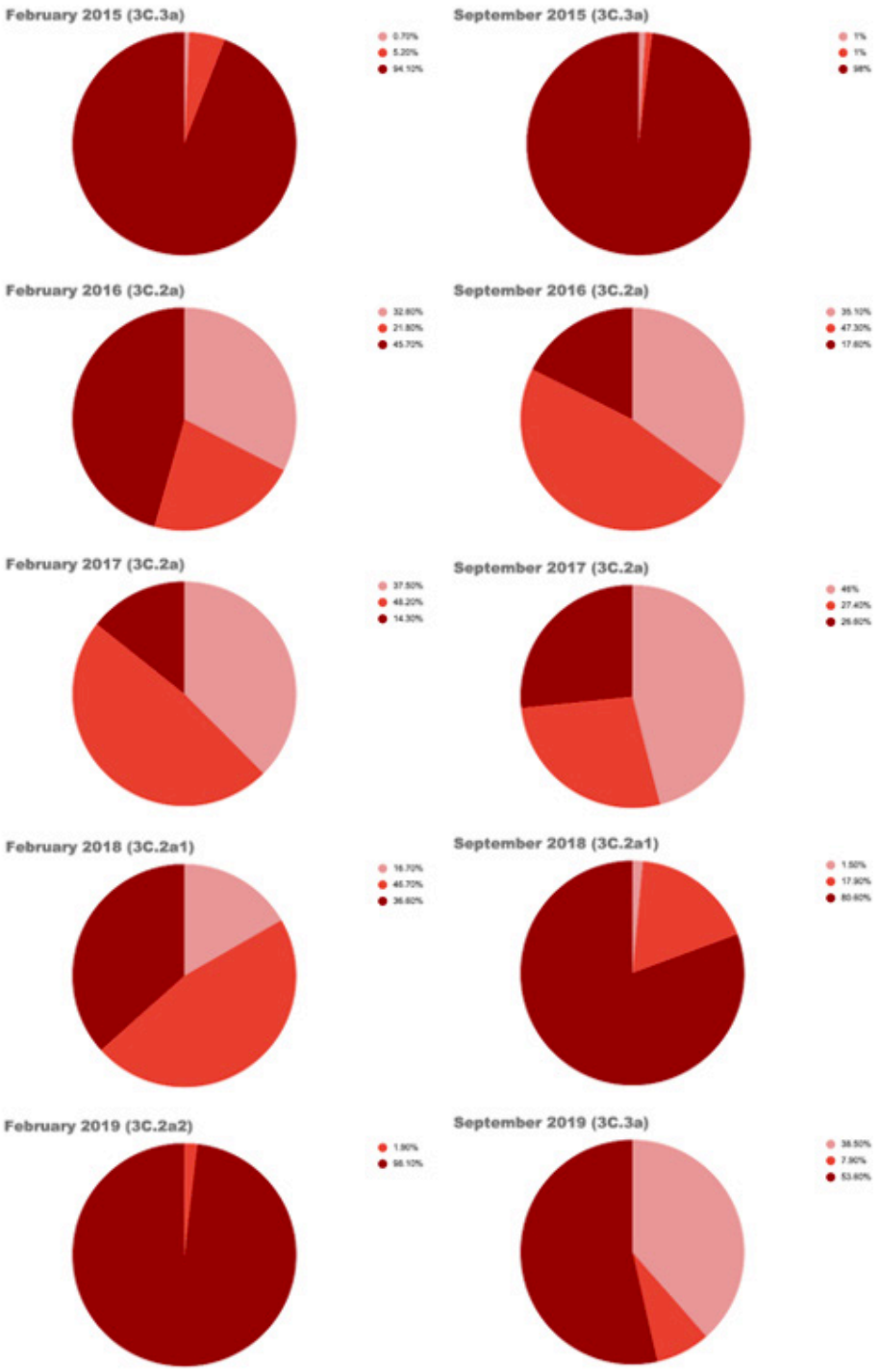


Figure 4: Vaccine efficacy as predicted by hemagglutinin inhibition assay. Colours correspond to <2 (light pink), 4 (medium red) and >8 (dark red) fold decrease in titre as compared to unvaccinated sera. Higher degrees of inhibition(e.g. >8 fold) correspond to more effective vaccines. In brackets is the vaccine strain for each year. Data comes from hemagglutinin inhibition assay using vaccine strain antibodies against all incoming isolates.

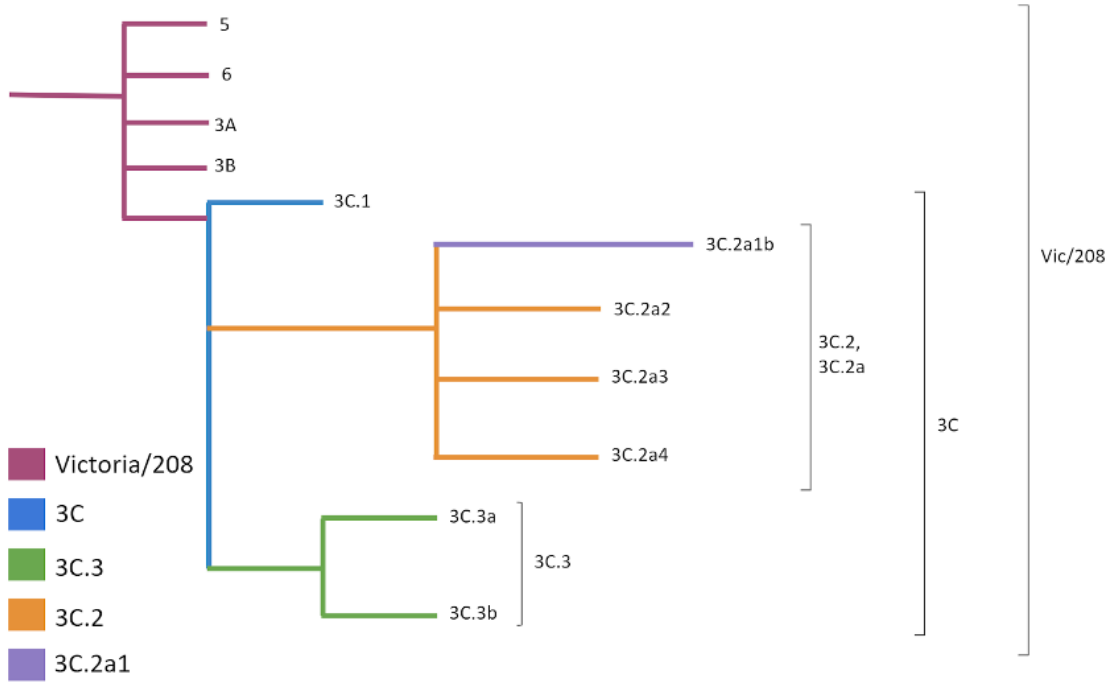


Figure 5: Phylogenetic Tree of A(H3N2) clades and subclades. Information was collected from The Crick Institute's Influenza Reports. Vic/208 refers to the A/Victoria/208/2009 virus, the ancestral clade in which all circulating A(H3N2) viruses stem from. Each clade is defined by specific mutations, as decided during annual WHO flu meetings.

Table 2: Mutations of the virus strain targeted by the recommended vaccine. Data was collected from The Crick Institute's Influenza Reports. Mutations in the table for each clade do not include mutations from the parent strain. The vaccine strains for each season were recommended by the World Health Organization (WHO).

Year	Season (Feb or Sept)	Vaccine strain	Clade	Mutations
2019	Feb	an A/Switzerland/8060/2017 (H3N2)-like virus	3C.2a2	**T131K, R142K and R261Q** in HA1, define subclade 3C.2a2 plus mutations from 3C.2a
	Sept	an A/Kansas/14/2017 (H3N2)-like virus	3C.3a	acquired the substitutions **S91N, N144K** (resulting in the loss of a glycosylation site) **P193S and K326R** in HA1 with many viruses sharing the substitutions **M17L, I149M and A201V** in HA2
2018	Feb	an A/Singapore/INFIMH-16-0019/2016 (H3N2)-like viru	3C.2a1	**R142G, D225G, T160K, L194P, N121K, N171K, I406V, G155E (HA2)**
	Sept	an A/Singapore/INFIMH-16-0019/2016 (H3N2)-like virus	3C.2a1	**N121K, N171K, and R142G** in HA1 and **I77V and G155E** in HA2
2017	Feb	an A/Hong Kong/4801/2014 (H3N2)-like virus	3C.2a	**L3I, N128T** (resulting in the gain of a potential glycosylation site), **N144S** (resulting in the loss of a potential glycosylation site), **N145S, F159Y, K160T** (resulting in the gain of a potential glycosylation site), **V186G, P198S, F219S, N225D and Q311H** in HA1,
	Sept	an A/Hong Kong/4801/2014 (H3N2)-like virus	3C.2a	defined by **L3I, N128T** (resulting in the gain of a potential glycosylation site), **N144S** (resulting in the loss of a potential glycosylation site), **N145S, F159Y, K160T** (resulting in the gain of a potential glycosylation site at residue 158), **P198S, F219S, N225D and Q311H**
2016	Feb	an A/Hong Kong/4801/2014 (H3N2)-like virus	3C.2a	**L3I, N128T** (resulting in the gain of a potential glycosylation site), **N144S** (resulting in the loss of a potential glycosylation site), **N145S, F159Y, K160T** (resulting in the gain of a potential glycosylation site), **V186G, P198S, F219S, N225D and Q311H** in HA1
	Sept	an A/Hong Kong/4801/2014 (H3N2)-like virus	3C.2a	**L3I, N128T** (resulting in the gain of a potential glycosylation site), **N144S** (resulting in the loss of a potential glycosylation site), **N145S, F159Y, K160T** (resulting in the gain of a potential glycosylation site), **V186G, P198S, F219S, N225D and Q311H** in HA1
2015	Feb	an A/Switzerland/9715293/2013 (H3N2)-like virus	3C.3a	**T128A** (resulting in the loss of a potential glycosylation site), **A138S, R142G, N145S, F159S, V186G, and N225D** in HA1
	Sept	an A/Switzerland/9715293/2013 (H3N2)-like virus	3C.3a	**T128A** (resulting in the loss of a potential glycosylation site), **A138S, R142G, N145S, F159S, V186G and N225D** in HA1

Table 1. Mutations that define circulating A(H3N2) strains from February 2012 to September 2019. Data was collected from The Crick Institute's Influenza Reports. After September 2013, reports no longer include mutations from Vic/208 (K62E, K144N) and therefore are not included in the table below. Each strain's mutations are also defined by its parental strain. Certain mutations stop being included for strains after subgroups have been identified within that strain.

Strain	Feb 2012	Sept 2012	Feb 2013	Sept 2013	Feb 2014	Sept 2014	Feb 2015	Sept 2015	Feb 2016	Sept 2016	Feb 2017	Sept 2017	Feb 2018	Sept 2018	Feb 2019	Sept 2019
3A	N144D, N145S, V223I, D487N (HA2) Parent Mutations: K62E, K144N	N144D, N145S, V223I, D487N (HA2) Parent Mutations: K62E, K144N	N144D. N145S, V223I, D158N (HA2) plus parent mutations K26E, K144N	N145S, V223I, D158N (HA2), N144D	No mutation info in report	No mutation info in report	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3B	K26E, K144N, N145S, A198S, V223I, N312S, D158N (HA2)	K26E, K144N, N145S, A198S, V223I, N312S, D158N (HA2)	N/A	N145S, V223I, D158N (HA2), A198S, N312S, new: **T30A**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3C	K26E, K144N, S45N, T48I, A198S, V223I, N312S, some also carry Q33R and N278K	K26E, K144N, S45N, T48I, A198S, V223I, N312S, some also carry Q33R and N278K	K26E, K144N, S45N, T48I, A198S, V223I, N312S, some also carry Q33R and N278K, new: **S145N**, and some carry **N145S, T128A, R142G**	N/A	No mutation info in report	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
5	K26E, K144N, D53N, Y94H, I230V, E280A	K26E, K144N, D53N, Y94H, I230V, E280A	K26E, K144N, D53N, Y94H, I230V, E280A, new: **K2E, N8D**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
6	K26E, K114N, D53N, Y94H, S199A, I230V, E280A	K26E, K114N, D53N, Y94H, S199A, I230V, E280A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3C.1	N/A	N/A	N/A	Group 3: N145S, V223I, D158N (HA2), 3C: S45N, T48I, S145N, A198S, N312S, 3C.1: **Q33R, N278K**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3C.2	N/A	N/A	N/A	Has Group 3, 3C, and 3C.1 mutations plus **N145S**	Has Group 3, 3C, and 3C.1 mutations plus N145S, new: **V186G and D160N** (HA2)	Has Group 3, 3C, and 3C.1 mutations plus N145S, V186G and D160N (HA2)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3C.3	N/A	N/A	N/A	Has Group 3, 3C, 3C.1, 3C.2 mutations plus **T128A, R142G**	Has Group 3, 3C, 3C.1, 3C.2 mutations plus T128A, R142G	Has Group 3, 3C, 3C.1, 3C.2 mutations plus T128A, R142G	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3C.3a	N/A	N/A	N/A	N/A	N/A	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus **A138S, F159S, N225D**	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus A138S, F159S, N225D	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus A138S, F159S, N225D	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus A138S, F159S, N225D new: **P198S, F219S, Y94H**	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus A138S, F159S, N225D, P198S, F219S, Y94H new:**S198P**	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus A138S, F159S, N225D, P198S, F219S, Y94H (some viruses), S198P, **(recent groups fall into two groups with mutations: 1: K92R, Q197K, S198P, S312N 2: L3I, S91N, N144K, F193S, D160 (HA2))**	No mutation info in report	No mutation info in report	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus A138S, F159S, N225D, P198S, F219S, Y94H, S198P, **new: K326R, R142G, N144K, F193S, L3I (does not mention groups)**	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus **A138S, T128A, R142G, L3I, S91N, N144K, F193S, K326R, D160N (HA2) *Some are no longer mentioned**	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus A138S, T128A, R142G, L3I, S91N, N144K, F193S, K326R, D160N (HA2) **new: P193S, M17L (HA2), I149M (HA2), A201V (HA2)**
3C.3b	N/A	N/A	N/A	N/A	N/A	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus **E62K, K83R, N122D, L157S, R261Q, M18K (HA2)**	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus E62K, K83R, N122D, L157S, R261Q, M18K (HA2)	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus E62K, K83R, N122D, L157S, R261Q, M18K (HA2)	Has Group 3, 3C, 3C.1, 3C.2, 3C.3 plus E62K, K83R, N122D, L157S, R261Q, M18K (HA2) new: **P198S, F219S,Q917H, V18K (HA2)**	No mutation info in report	No mutation info in report	N/A	N/A	N/A	N/A	N/A
3C.2a	N/A	N/A	N/A	N/A	N/A	Has Group 3, 3C, 3C.1, 3C.2 plus **L3I, N144S, F159Y, K160T, N225D, Q311H**	Has Group 3, 3C, 3C.1, 3C.2 plus L3I, N144S, F159Y, K160T, N225D, Q311H	Has Group 3, 3C, 3C.1, 3C.2 plus L3I, N144S, F159Y, K160T, N225D, Q311H	Has Group 3, 3C, 3C.1, 3C.2 plus L3I, N144S, F159Y, K160T, N225D, Q311H new: **N128T, P198S, F219S**	Has Group 3, 3C, 3C.1, 3C.2 plus L3I, N144S, F159Y, K160T, N225D, Q311H, N128T, P198S, F219S	Has Group 3, 3C, 3C.1, 3C.2 plus L3I, N144S, F159Y, K160T, N225D, Q311H, N128T, P198S, F219S, **(there are subgroups within 3C.2a but do not have their own clade; these mutations are N121K, S144K, S219Y, N122D, Y94H, I140T)**	Has Group 3, 3C, 3C.1, 3C.2 plus L3I, N144S, F159Y, K160T, N225D, Q311H, N128T, P198S, F219S, (there are subgroups within 3C.2a but do not have their own clade; these mutations are N121K, S144K, S219Y, N122D, Y94H, I140T **new: S262N, R261Q, T131K, R142K, N31S, D53N, R142G, S144R, N171K, I192T, Q197H)**	Has Group 3, 3C, 3C.1, 3C.2 plus L3I, N144S, F159Y, K160T, N225D, Q311H, N128T, P198S, F219S, (there are subgroups within 3C.2a but do not have their own clade; these mutations are N121K, S144K, S219Y, N122D, Y94H, I140T, S262N, R261Q, T131K, R142K, N31S, D53N, R142G, S144R, N171K, I192T, Q197H) ***subgroups start to get named but are not included in Table 6**	***Mutations from before are put into subgroups** and they now only define it as having Group 3, 3C, 3C.1, 3C.2 plusL3I, N144S, F159Y, K160T, N225D, Q311H, N128T, P198S, F219S	*Mutations from before are put into subgroups and they now only define it as having Group 3, 3C, 3C.1, 3C.2 plusL3I, N144S, F159Y, K160T, N225D, Q311H, N128T, P198S, F219S	Group 3, 3C, 3C.1, 3C.2 **plus N128T, S138A, N144S, S159Y, K160T, Q311H, G142R, D160N (HA2)**
3.3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	No mutation info in report	No mutation info in report	No mutation info in report	No mutation info in report	N/A	N/A	N/A	N/A	N/A
3C.2a1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a plus **N171K, I77V (HA2), N121K, R142G, G150E (HA2)**	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a plus N171K, I77V (HA2), N121K, R142G, G150E (HA2) **(there are additional subgroups within 3C.2a1)**	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a plus N171K, I77V (HA2), N121K, R142G, G150E (HA2) (there are additional subgroups within 3C.2a1)	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a plus N171K, I77V (HA2), N121K, R142G, G150E, ***Subgroups are named**	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a plus N171K, I77V (HA2), N121K, R142G, G150E, *Subgroups are named	N/A
3C.2a1b	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a, 3C.2a1 plus **K92R, H311Q, I58V, R142G, E62G, T135K**	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a, 3C.2a1 plus K92R, H311Q, R142G, E62G, T135K, **new: T131K, K135T V200I (HA2)**	Has Group 3, 3C, 3C.1, 3C.2 , 3C.2a, 3C.2a1 plus **E62G, K92R, N121K, R142G, N171K, H311Q, I77V (HA2), G155E (HA2)**
3C.2a2	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Has 3C.2a which includes parents from before plus **T131K, R142K, R261Q**	Has 3C.2a which includes parents from before plus T131K, R142K, R261Q	Has 3C.2a which includes parents from before plus T131K, R142K, R261Q
3C.2a3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Has 3C.2a which includes parents plus **N121K, S144K, N122D, S262N, R261Q**	Has 3C.2a which includes parents plus **N121K, S144K, T135K, R142G, R261Q, T128A**	Has 3C.2a which includes parents plus **N121K, S144K,**
3C.2a4	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Has 3C.2a which includes parents plus **N31S, D52N, R142G, S144R, N171K, I192T, Q197H**	Has 3C.2a which includes parents plus **N31S, D52N, S144R, N171K, I192T, Q197H**