VIEWPOINT

Genetic Variants of SARS-CoV-2—What Do They Mean?

Adam S. Lauring, MD, PhD
Division of Infectious Diseases, Department of Internal Medicine and Department of Microbiology and Immunology, University of Michigan, Ann Arbor.

Emma B. Hodcroft, PhD
Institute of Social and Preventive Medicine, University of Bern, Switzerland.

Multimedia

Related article
page 522

Over the course of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic, the clinical, scientific, and public health communities have had to respond to new viral genetic variants. Each one has triggered a flurry of media attention, a range of reactions from the scientific community, and calls from governments to either "stay calm" or pursue immediate countermeasures. While many scientists were initially skeptical about the significance of the D614G alteration, the emergence of the new "UK variant"—lineage B.1.1.7—has raised widespread concern. Understanding which variants are concerning, and why, requires an appreciation of virus evolution and the genomic epidemiology of SARS-CoV-2.

Mutations, Variants, and Spread

Mutations arise as a natural by-product of viral replication.[1] RNA viruses typically have higher mutation rates than DNA viruses. Coronaviruses, however, make fewer mutations than most RNA viruses because they encode an enzyme that corrects some of the errors made during replication. In most cases, the fate of a newly arising mutation is determined by natural selection. Those that confer a competitive advantage with respect to viral replication, transmission, or escape from immunity will increase in frequency, and those that reduce viral fitness tend to be culled from the population of circulating viruses. However, mutations can also increase and decrease in frequency due to chance events. For example, a "founder effect" occurs when a limited number of individual viruses establish a new population during transmission. The mutations present in the genomes of these viral ancestors will dominate the population regardless of their effects on viral fitness. This same interplay of natural selection and chance events shapes virus evolution within hosts, in communities, and across countries.

> Understanding which variants are concerning, and why, requires an appreciation of virus evolution and the genomic epidemiology of SARS-CoV-2.

Although the terms *mutation*, *variant*, and *strain* are often used interchangeably in describing the epidemiology of SARS-CoV-2, the distinctions are important. *Mutation* refers to the actual change in sequence: D614G is an aspartic acid-to-glycine substitution at position 614 of the spike glycoprotein. Genomes that differ in sequence are often called *variants*. This term is somewhat less precise because 2 variants can differ by 1 mutation or many. Strictly speaking, a variant is a *strain* when it has a demonstrably different phenotype (eg, a difference in antigenicity, transmissibility, or virulence).

Evaluation of a new SARS-CoV-2 variant should include assessment of the following questions: Did the variant achieve prominence through natural selection or chance events? If the evidence suggests natural selection, which mutation(s) are being selected? What is the adaptive benefit of these mutations? What effect do these mutations have on transmissibility and spread, antigenicity, or virulence?

Spike D614G

The D614G mutation in the spike glycoprotein of SARS-CoV-2 was first detected at a significant level in early March 2020 and spread to global dominance over the next month.[2] The mutation initially appeared to arise independently and simultaneously sweep across multiple geographic regions. This apparent convergent evolution was suggestive of natural selection and an adaptive benefit of D614G. However, subsequent sequencing efforts identified the D614G mutation in viruses in several Chinese provinces in late January. This raised the possibility that global dispersal of this mutation could have resulted from chance founder events, in which viruses harboring 614G just happened to initiate the majority of early transmission events in multiple locations.

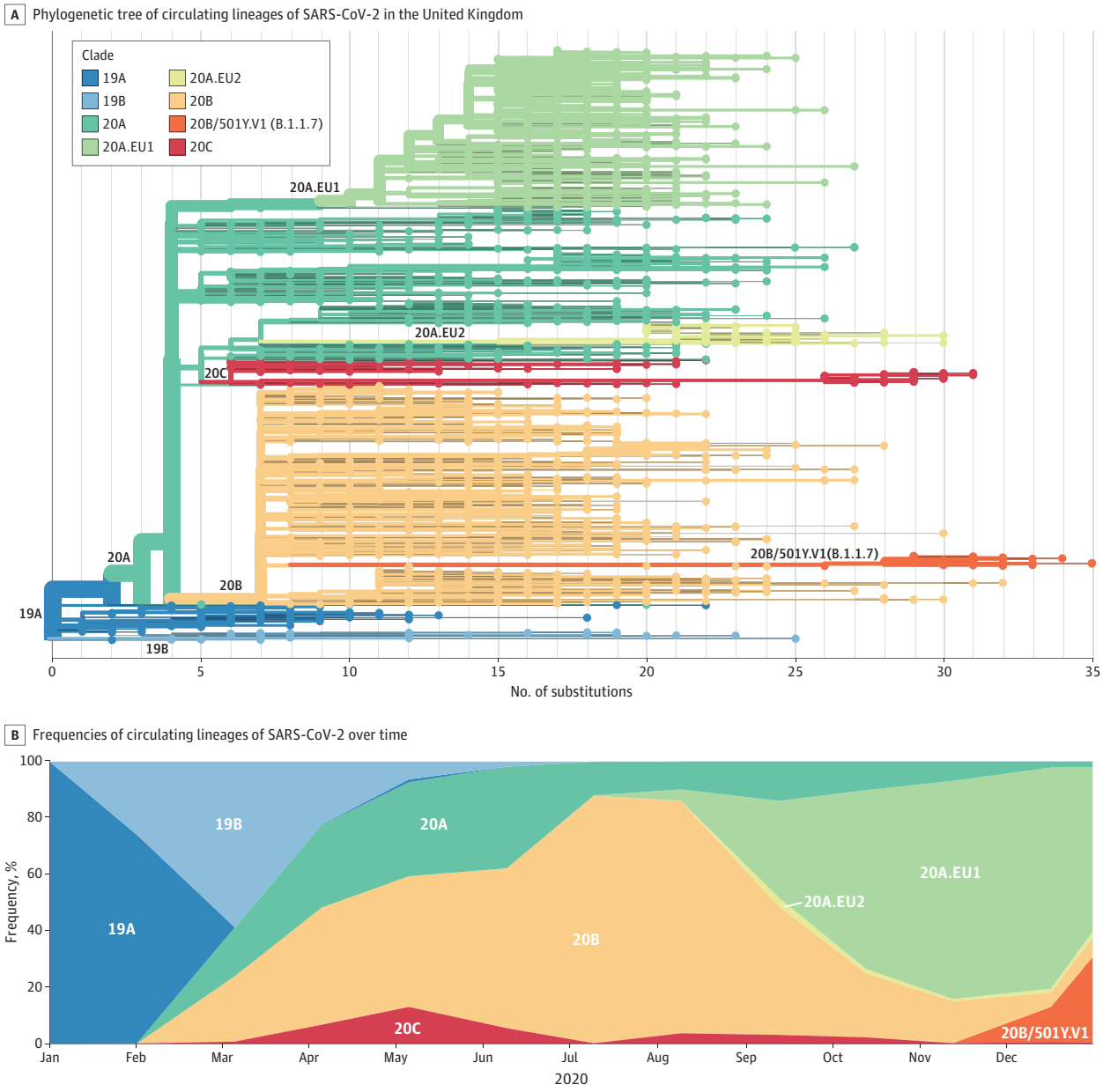
This plausible null hypothesis led many in the evolution community to doubt that the D614G mutation was adaptive, despite in vitro data showing its effects on receptor binding. A recent population genetic and phylodynamic analysis of more than 25 000 sequences from the UK found that viruses bearing 614G did appear to spread faster and seed larger phylogenetic clusters than viruses with 614D.[3] The effect size was modest, and the varying models did not always achieve statistical significance. More recently, complementary work in animal models indicates that 614G viruses transmit more efficiently.[4,5]

Spike N453Y and Mink

Concerning outbreaks of SARS-CoV-2 began to emerge on mink farms in the Netherlands and Denmark in late spring and early summer 2020.[6] Genomic and epidemiologic investigation of an early outbreak in the Netherlands demonstrated human to mink, mink to mink, and mink to human transmission.[7] In early November 2020 Danish authorities reported 214 cases of human coronavirus disease 2019 (COVID-19) associated with mink farms. Many SARS-CoV-2 sequences from the Netherlands and Danish outbreaks had a Y453F mutation in the receptor binding domain of

Corresponding Author: Adam S. Lauring, MD, PhD, Division of Infectious Diseases, Department of Internal Medicine and Department of Microbiology and Immunology, University of Michigan, 1150 W Medical Center Dr, MSRB1 5510B, Ann Arbor, MI 48109-5680 (alauring@med.umich.edu).

Figure. Spread of a New SARS-CoV-2 Variant



A, Phylogenetic tree showing the relationship of lineage B.1.1.7 (20B/501Y.V1, orange branch and tips) to other circulating lineages. The long branch length for this lineage reflects the fact that it accumulated a significant number of mutations prior to being discovered. B, Frequencies of circulating lineages over time. Lineages are colored as in the tree, with lineage B.1.1.7 (20B/501Y.V1) shown in orange.

spike, which might mediate increased binding affinity for mink ACE2 (angiotensin-converting enzyme 2). Eleven individuals from the Danish outbreak had a variant termed cluster 5, which had 3 additional mutations in spike (del69_70, I692V, and M1229I). An initial investigation of 9 human convalescent serum samples suggested a modest and variably statistically significant reduction in neutralization activity against cluster 5 viruses (mean, 3.58 fold; range, 0-13.5). The apparent adaptation of SARS-CoV-2 to mink was nevertheless concerning because continued evolution of the virus in an animal reservoir could potentially lead to recurrent spillover events of novel SARS-CoV-2 from mink to humans and other mammals. For this reason, many countries have increased surveillance efforts and in some cases implemented large-scale culls (ie, selective slaughter) of mink on farms.

Lineage B.1.1.7 and N501Y

Lineage B.1.1.7 (also called 501Y.V1) is a phylogenetic cluster that is rapidly spreading in southeastern England[8] (**Figure**). It had accumulated 17 lineage-defining mutations prior to its detection in early September, which suggests a significant amount of prior evolution, possibly in a chronically infected host. As of December 28, 2020, this variant accounted for approximately 28% of cases of

SARS-CoV-2 infection in England, and population genetic models suggest that it is spreading 56% more quickly than other lineages.[9] Unlike D614G, which could plausibly have benefited from early chance events, lineage B.1.1.7 expanded when SARS-CoV-2 cases were widespread and has seemingly achieved dominance by outcompeting an existing population of circulating variants. This is strongly suggestive of natural selection of a virus that is more transmissible at a population level. While public health interventions like masks, physical distancing, and limitations on large gatherings should remain effective, control of a more transmissible variant would likely require more stringent application and widespread adoption of these measures.

Eight of the lineage B.1.1.7 mutations are in the spike glycoprotein, including N501Y in the receptor binding domain, deletion 69_70, and P681H in the furin cleavage site. All of these mutations could plausibly influence ACE2 binding and viral replication. The 501Y spike variants are predicted to have a higher affinity for human ACE2, and a different variant, also with an N501Y mutation, is rapidly spreading in South Africa. The effects of these mutations on antigenicity are currently unclear.

Antigenicity and Vaccine Effectiveness

Genomic surveillance of SARS-CoV-2 variants has largely focused on mutations in the spike glycoprotein, which mediates attachment to cells and is a major target of neutralizing antibodies. There is intense interest in whether mutations in the spike glycoprotein mediate escape from host antibodies and could potentially compromise vaccine effectiveness, since spike is the major viral antigen in the current vaccines. At this point, strong selection of a variant at the population level is probably not driven by host antibody because there are not sufficient numbers of immune individuals to systematically push the virus in a given direction. In contrast, if a variant has one or more mutations in spike that increase transmissibility, it could quickly outcompete and replace other circulating variants. Because current vaccines provoke an immune response to the entire spike protein, it is hoped that effective protection may still occur despite a few changes at antigenic sites in SARS-CoV-2 variants.

Separating cause from consequence is important in evaluating data on antibody neutralization of spike variants. Regardless of why the mutations were selected, it is reasonable to expect that many mutations in spike might affect neutralization by convalescent sera. It is therefore important to consider both the magnitude of the change in neutralization and the number of serum samples evaluated. Another issue is that viral glycoproteins are subject to evolutionary trade-offs. Sometimes a mutation that enhances one viral property, such as binding to a receptor, can reduce another property, such as escaping host antibody. Indeed, recent evidence suggests this could be the case for D614G.[10] It is possible that mutations in spike that are "good" for the virus right now could also make it less fit in the context of population-level immunity in the future. Defining these dynamics, and their potential influence on vaccine effectiveness, will require large-scale monitoring of SARS-CoV-2 evolution and host immunity for a long time to come.

ARTICLE INFORMATION

Published Online: January 6, 2021. doi:10.1001/jama.2020.27124

Conflict of Interest Disclosures: Dr Lauring reported receiving personal fees from Sanofi as a consultant on oseltamivir and influenza and personal fees from Roche as a member of a steering committee for a clinical trial of baloxavir and influenza, outside the submitted work. No other disclosures were reported.

REFERENCES

1. Grubaugh ND, Petrone ME, Holmes EC. We shouldn't worry when a virus mutates during disease outbreaks. *Nat Microbiol*. 2020;5(4):529-530. doi:10.1038/s41564-020-0690-4

2. Korber B, Fischer WM, Gnanakaran S, et al; Sheffield COVID-19 Genomics Group. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. 2020;182(4):812-827.e19. doi:10.1016/j.cell.2020.06.043

3. Volz E, Hill V, McCrone JT, et al; COG-UK Consortium. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell*. 2020;S0092867420315373.

4. Plante JA, Liu Y, Liu J, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature*. Published online December 23, 2020. doi:10.1038/s41586-020-2895-3

5. Hou YJ, Chiba S, Halfmann P, et al. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science*. 2020;370(6523):1464-1468. doi:10.1126/science.abe8499

6. European Centre for Disease Prevention and Control. Detection of new SARS-CoV-2 variants related to mink. Posted November 12, 2020. Accessed January 3, 2021. https://www.ecdc.europa.eu/sites/default/files/documents/RRA-SARS-CoV-2-in-mink-12-nov-2020.pdf

7. Oude Munnink BB, Sikkema RS, Nieuwenhuijse DF, et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science*. 2020;eabe5901. doi:10.1126/science.abe5901

8. Rambaut A, Loman N, Pybus O, et al; COVID-19 Genomics Consortium UK. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. Virological.org. Posted December 16, 2020. Accessed January 3, 2021. https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563

9. Davies NG, Barnard RC, Jarvis CI, et al Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. *CMMID*. Preprint published online December 23, 2020. Updated December 31, 2020. doi:10.1101/2020.12.24.20248822

10. Weissman D, Alameh M-G, de Silva T, et al. D614G spike mutation increases SARS CoV-2 susceptibility to neutralization. *Cell Host Microbe*. 2020;S193131282030634X.