# Perspective: The evolutionary dangers of high COVID case counts

C. L. Burch[1]*, D. M. Weinreich[2], Y. Raynes[2], J. M. DeCurzio[1].

[1]Biology Department, University of North Carolina, Chapel Hill, NC 27599-3280.

[2]Department of Ecology and Evolutionary Biology and Center for Computational Molecular Biology, Brown University, Providence, RI

*Correspondence to: cburch@bio.unc.edu

The recent simultaneous appearance of numerous highly contagious variants of SARS-CoV-2 demonstrate that the rate of adaptive evolution in the SARS-CoV-2 population is accelerating. It is no longer appropriate to focus only on epidemiological goals like flattening the curve and vaccinating to achieve herd immunity. We are now in a new phase of the pandemic, in which we must also focus on the evolutionary threats. Here, we use the tools of population genetics, a subfield of evolutionary biology, to look into the pandemic's past and future. We explain why these variants are arising with increasing frequency as the size of the pandemic grows and we explore the likely paths of future SARS-CoV-2 evolution. Our take-home message is that viral evolution during the vaccine rollout will be especially serious, making it essential to bring case counts down immediately.

For those of us who conduct microbial evolution experiments in our labs, the simultaneous appearance of multiple highly-adapted variants is a common and familiar outcome of evolution in our experiments, e.g, (*1*, *2*). Our insight is that the accelerating rate of adaptive evolution in SARS-CoV-2 is a direct result of the steady increase in the supply of viral mutations as the pandemic proceeds. Because the pandemic began from a single virus clone with no genetic variations, the pandemic initially lacked the mutational "fuel" for natural selection to drive adaptive evolution. However, each time the virus reproduced, there was an opportunity for a mutation, as errors occurred when the virus copied its genome. As the daily case counts grew and, with it, the size of the viral population, the rate at which mutations were produced increased: one copy initially provided one opportunity for a mutation, but one million copies provided one million opportunities. Evolution by natural selection accelerated as the growing pandemic fed it more mutational fuel.

Mutations arise randomly with respect to their functional consequences and random changes are much more likely to impair function than to improve it. As a result, adaptive mutations are rare. However, as the population of replicating viruses grows and more mutations are produced, even rare, adaptive mutations will arise. Population geneticists recognize a critical transition in evolutionary dynamics once the population size exceeds the reciprocal of the per-site mutation rate. Roughly speaking, above this threshold, mutations at every site in the genome are produced each generation. Coronaviruses have a per-site mutation rate of approximately $10^{-6}$ per generation (*3*). That is, in each generation and at any given site in the viral genome, there is a one in a million chance that a mutation will occur. Given that active case counts are well above one million, we are well above this critical threshold now. Thus, mutations at every site in the SARS-CoV-2 genome are likely being shed by at least one (and probably tens or hundreds of) infected person(s) each day. With case counts this high, population geneticists predict that

individual adaptive mutations will arise simultaneously in different viral lineages and that multiple adaptive mutations will arise simultaneously within individual viral lineages (*4*). These are exactly the phenomena we observe in laboratory evolution experiments and that we are now witnessing with the B.1.1.7 (United Kingdom), B.1.351 (South Africa), P.1 (Brazil), and other highly transmissible variants (Fig. 1).

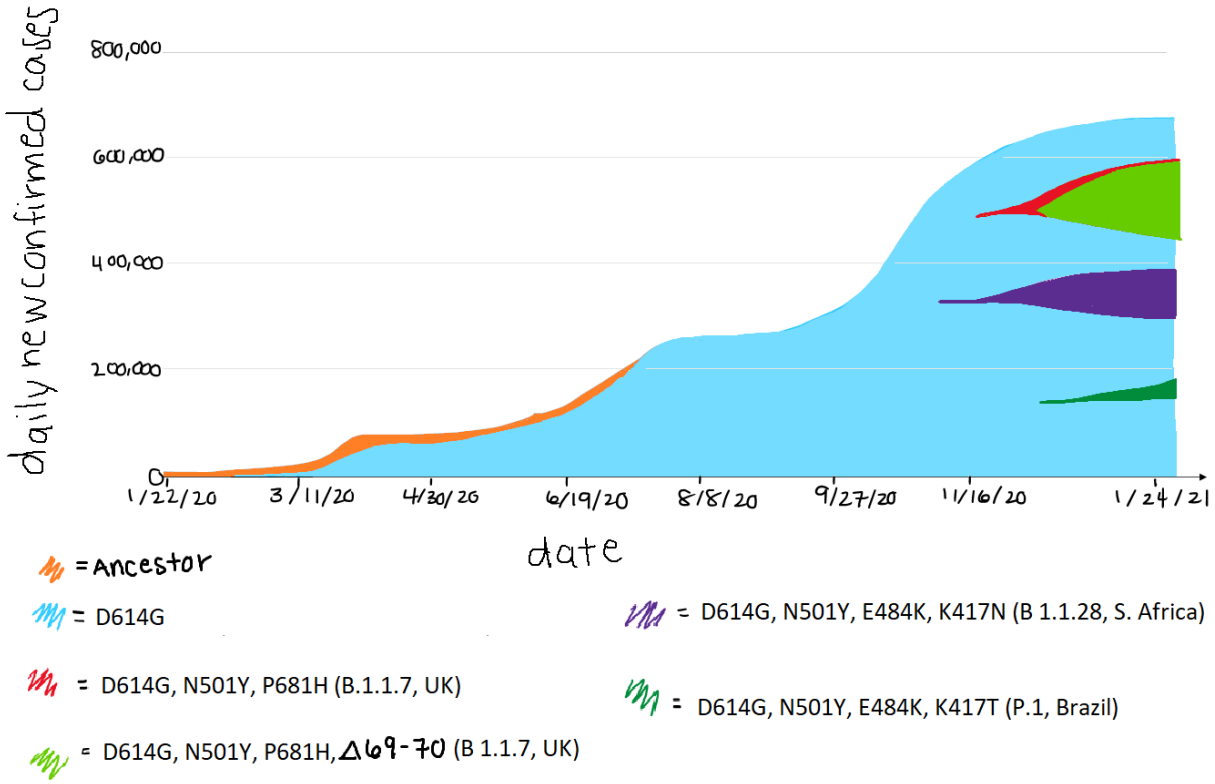


**Fig. 1.** Evolution by natural selection in SARS-CoV-2 is accelerating. The total colored area shows the trend in daily case counts over the course of the pandemic. Case count data were obtained from (*5*) and smoothed for illustrative purposes. Different colors correspond to SARS-CoV-2 variants that arose independently with different sets of mutations (only the amino acid changes in the Spike protein thought to be functionally important are listed). Colors nested within other colors most likely descended from the variant in which they are nested. For a particular date, the height of each color, relative to the total height of the colored area, reflects the approximate frequency of that variant on that date in the SARS-CoV-2 GISAID database (*6*), accessed via nextstrain.org (*7*).

Throughout the pandemic, natural selection has chiefly acted on viral mutations that increase $R_0$, the average number of people that a single infected person will go on to infect, by increasing transmissibility (i.e., contagiousness). That selection favors increasing $R_0$ underlies our understanding of the worldwide spread of the D614G mutation in the spike protein, which has been shown to increase infectivity in cell culture by 5-10 times (*8*). The emergence of highly transmissible variants is dangerous for public health because higher viral transmission will make it more difficult to achieve herd immunity (e.g., we may need vaccination rates nearer to 80% than to 70%).

However, as vaccination rates increase, the strongest selective force acting on the virus will become the vaccine itself. How quickly vaccine resistance evolves will depend on the supply of mutations that specifically contribute to vaccine resistance. Population genetics points to viral population size as the biggest influence on mutational supply under our control. Not only does population size affect the supply of *de novo* (new) mutations, it also affects the amount of pre-existing genetic variation, the probability that vaccine resistance mutations will become established in the population, and evolution of the viral mutation rate itself. We describe each of these effects of population size below.

First, reducing the SARS-CoV-2 population size will decrease the likelihood and rate at which vaccine resistance will evolve from *de novo* mutations — by reducing the opportunity for rare adaptive mutations. Vaccine evasion in Influenza, the virus for which we have the most information, typically requires several mutations in combination (*9*, *10*), an event that has an exceptionally low probability of occurring *de novo* in a single viral reproductive event. However, with case counts as high as they are now, even rare events become likely. Case in point, each of the new highly transmissible variants possess large numbers of mutations, indicating that they resulted from rare and unusual events, such as long-term infections of individuals with compromised immune systems, UK: (*11*), South Africa: (*12*), Brazil: (*13*). The good news is that bringing the case count down will be an exceptionally powerful means of reducing the occurrence of rare events and, thereby, for slowing the evolution of vaccine resistance from *de novo* mutation.

Second, reducing the SARS-CoV-2 population size now will eliminate many low frequency mutations from the viral population by chance, making them unavailable to future selection for vaccine resistance. If SARS-CoV-2 resembles other RNA viruses, we expect mutations that confer vaccine resistance to be relatively rare now — and thus most readily lost by chance — as past selection for $R_0$ would not have elevated their frequency. For example, in Influenza, mutations that confer vaccine resistance and mutations that mediate $R_0$ (by changing binding affinity for the host receptor) occur in different regions of the virus' host binding protein (14). Critically, the opportunity to eliminate rare vaccine resistance variants by reducing population size will disappear as soon as the increasing number of vaccinated individuals starts imposing the kind of selection that will elevate the frequency of such variants. The time to act is now. [We note that variants that arose as the result of *past* selection for resistance to convalescent plasma or prior infection (*11*, *13*), will not be eliminated by reductions in population size. We are stuck with these.]

Third, reducing the SARS-CoV-2 population size now will reduce the number of contacts between infected and vaccinated individuals. To increase in frequency in the SARS-CoV-2 population, vaccine resistance mutations will have to benefit from their enhanced ability to infect vaccinated individuals. If infected persons rarely come into contact with vaccinated persons, this kind of selection favoring vaccine resistance will have little opportunity to act. Thus, it is critical that vaccinated individuals continue to wear masks. Of the other strategies that could reduce the number of such contacts, only reducing the SARS-CoV-2 case counts in areas where we are deploying vaccines seems viable. Reducing vaccine deployment in areas where case counts are high is not ethical and altering our vaccination order, e.g., by vaccinating whole households

simultaneously to reduce within-household transmission, would not prioritize the most vulnerable individuals.

Finally, large population size is known to favor the evolution of higher mutation rates (*15*). Evolutionary theory and laboratory evolution experiments show that high mutation variants evolve readily when many adaptive mutations are available and genetic recombination is rare. These conditions may or may not be met in SARS-CoV-2, e.g., (*16*), but the SARS-CoV-2 genome-wide mutation rate is below that of other RNA viruses and mutations that confer higher mutation rates are known to exist in SARS-CoV-1 (*3*). Clearly, the evolution of higher mutation rates in SARS-CoV-2 would be an evolutionary and public health disaster. Our best defense is again to reduce case counts and, with them, the viral population size.

We are heartened that world governments, the press, and the public are all deeply concerned by simultaneous appearance of multiple highly transmissible variants. But as population geneticists, our message is: it's worse than most people realize. We are concerned about the public health consequences of variants that arose in the fall and are being detected now, but we are even more concerned about the variants that will arise this spring as vaccine deployment exerts natural selection in favor of vaccine resistance. Every factor with the potential to increase the rate at which vaccine resistance evolves by natural selection is made worse by the high daily case counts.

We need to bring the daily case counts down and we need to bring them down ***now***. We urge policy makers, epidemiologists, social scientists, and the press to include the evolutionary potential of SARS-CoV-2 in their deliberations. These evolutionary considerations will be particularly critical in communicating to the public the need for continued masking, social distancing and even shut downs, as needed. The rapid advent of vaccines is a triumph of science and a source of hope, but we remain vulnerable to SARS-CoV-2 evolution. The guarantee of further evolution must be met with vigilance and action.

**References and Notes**


1. G. I. Lang, D. P. Rice, M. J. Hickman, E. Sodergren, G. M. Weinstock, D. Botstein, M. M. Desai, Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*. **500**, 571–574 (2013).
2. B. H. Good, M. J. McDonald, J. E. Barrick, R. E. Lenski, M. M. Desai, The dynamics of molecular evolution over 60,000 generations. *Nature*. **551**, 45–50 (2017).
3. L. D. Eckerle, M. M. Becker, R. A. Halpin, K. Li, E. Venter, X. Lu, S. Scherbakova, R. L. Graham, R. S. Baric, T. B. Stockwell, D. J. Spiro, M. R. Denison, Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *PLoS Pathog.* **6**, e1000896 (2010).
4. H. J. Muller, Some Genetic Aspects of Sex. *Am. Nat.* **66**, 118–138 (1932).
5. M. Roser, H. Ritchie, E. Ortiz-Ospina, J. Hasell, Coronavirus Pandemic (COVID-19). *Our World in Data* (2020) (available at https://ourworldindata.org/coronavirus).
6. S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall*. **1**, 33–46 (2017).

7. J. Hadfield, C. Megill, S. M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford, R. A. Neher, Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*. **34**, 4121–4123 (2018).
8. B. Korber, W. M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E. E. Giorgi, T. Bhattacharya, B. Foley, K. M. Hastie, M. D. Parker, D. G. Partridge, C. M. Evans, T. M. Freeman, T. I. de Silva, Sheffield COVID-19 Genomics Group, C. McDanal, L. G. Perez, H. Tang, A. Moon-Walker, S. P. Whelan, C. C. LaBranche, E. O. Saphire, D. C. Montefiori, Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*. **182**, 812–827.e19 (2020).
9. R. M. Bush, C. A. Bender, K. Subbarao, N. J. Cox, W. M. Fitch, Predicting the evolution of human influenza A. *Science*. **286**, 1921–1925 (1999).
10. L. I. Gong, M. A. Suchard, J. D. Bloom, Stability-mediated epistasis constrains the evolution of an influenza protein. *Elife*. **2**, e00631 (2013).
11. A. Rambaut, N. Loman, O. Pybus, W. Barclay, J. Barrett, A. Carabelli, T. Connor, T. Peacock, D. L. Robertson, E. Volz, on behalf of COVID-19 Genomics Consortium UK (CoG-UK), Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations (2020), (available at 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).
12. H. Tegally, E. Wilkinson, M. Giovanetti, A. Iranzadeh, V. Fonseca, J. Giandhari, D. Doolabh, S. Pillay, E. J. San, N. Msomi, K. Mlisana, A. von Gottberg, S. Walaza, M. Allam, A. Ismail, T. Mohale, A. J. Glass, S. Engelbrecht, G. Van Zyl, W. Preiser, F. Petruccione, A. Sigal, D. Hardie, G. Marais, M. Hsiao, S. Korsman, M.-A. Davies, L. Tyers, I. Mudau, D. York, C. Maslo, D. Goedhals, S. Abrahams, O. Laguda-Akingba, A. Alisoltani-Dehkordi, A. Godzik, C. K. Wibmer, B. T. Sewell, J. Lourenço, L. C. J. Alcantara, S. L. K. Pond, S. Weaver, D. Martin, R. J. Lessells, J. N. Bhiman, C. Williamson, T. de Oliveira, Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *bioRxiv* (2020), doi:10.1101/2020.12.21.20248640.
13. N. R. Faria, I. M. Claro, D. Candido, L. A. Moyses Franco, P. S. Andrade, T. M. Coletti, C. A. M. Silva, F. C. Sales, E. R. Manuli, R. S. Aguiar, N. Gaburo, C. da C. Camilo, N. A. Fraiji, M. A. Esashika Crispim, M. do P. S. S. Carvalho, A. Rambaut, N. Loman, O. G. Pybus, E. C. Sabino, on behalf of CADDE Genomic Network, Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings (2020), (available at https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586).
14. M. B. Doud, J. D. Bloom, Accurate Measurement of the Effects of All Amino-Acid Mutations on Influenza Hemagglutinin. *Viruses*. **8** (2016), doi:10.3390/v8060155.
15. Y. Raynes, C. S. Wylie, P. D. Sniegowski, D. M. Weinreich, Sign of selection on mutation rate modifiers depends on population size. *Proceedings of the National Academy of Sciences*. **115**, 3422–3427 (2018).
16. A. Ignatieva, J. Hein, P. A. Jenkins, Investigation of ongoing recombination through genealogical reconstruction for SARS-CoV-2. *Cold Spring Harbor Laboratory* (2021), doi:10.1101/2021.01.21.427579.

**Acknowledgments**

The authors thank M. Servedio, J. Kingsolver, D. Pfennig, K. Pfennig, and S. Otto for helpful suggestions. C.L.B. is supported by National Science Foundation grant DEB-2014943. D.M.W. and Y.R. are supported by National Science Foundation grant DEB-1556300. J.D. is supported by National Institutes of Health training grant T32 AI007419.