



# Staying Ahead of the Variants: Policy Recommendations to Identify and Manage Current and Future Variants of Concern

February 2021



JOHNS HOPKINS  
BLOOMBERG SCHOOL  
of PUBLIC HEALTH

Center for  
**Health Security**

## Authors

**Kelsey Lane Warmbrod, MS, MPH**

Senior Analyst, Johns Hopkins Center for Health Security

**Rachel West, PhD**

Postdoctoral Fellow, Johns Hopkins Center for Health Security

**Matthew Frieman, PhD**

Associate Professor, University of Maryland School of Medicine

**Dylan George, PhD, MS**

Vice President, In-Q-Tel

**Elena Martin, MPH**

Analyst, Johns Hopkins Center for Health Security

**Caitlin Rivers, PhD, MPH**

Senior Scholar, Johns Hopkins Center for Health Security

## Acknowledgments

We would like to express our gratitude to Anita Cicero, Gigi Kwik Gronvall, Nancy Connell, Marc Trotochaud, and Rick Bright for their invaluable expertise, input and advice in writing this report. The authors would also like to thank Julia Cizek, Kathleen Fox, Margaret Miller, and Prarthana Vasudevan for editing, design, and publication support.

Suggested Citation: Warmbrod KL, West R, Frieman M, George D, Martin E, Rivers C. *Staying Ahead of the Variants: Policy Recommendations to Identify and Manage Current and Future Variants of Concern*. Baltimore, MD: Johns Hopkins Center for Health Security; 2021.

© 2021 The Johns Hopkins University. All rights reserved.

## Executive Summary

As of February 2021, 3 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern with worrisome characteristics have emerged, each on a different continent. The B.1.1.7 variant, first identified in the United Kingdom, is substantially more transmissible than previously circulating variants. The B.1.351 and P.1 variants, first identified in South Africa and Brazil, respectively, both exhibit some degree of immune escape. Each of these variants has precipitated resurgences in the communities where they have become dominant. All 3 have already been identified at low levels in the United States. If they gain a foothold, the same resurgences can be expected here.

Funding for increased genomic surveillance is expected in the next Congressional supplemental, among several investments in SARS-CoV-2 research. Key efforts to expand capacity and improve surveillance systems should be funded with this money. New guidance and policies are also needed to maximize the response. Notably, investments made now to build genomic surveillance infrastructure for coronavirus disease 2019 (COVID-19) will not only help us respond to the pandemic now but will also improve response for outbreaks of other pathogens in the future.

This document explains the current status of SARS-CoV-2 surveillance, sequencing, and variant characterization and provides recommendations for increasing the United States' capacity to respond to new variants.

## Priority Recommendations

1. **Maintain Policies that Slow Transmission:** Variants will continue to emerge as the pandemic unfolds, but the best chance of minimizing their frequency and impact will be to continue public health measures that reduce transmission. This includes mask mandates, social distancing requirements, and limited gatherings.
2. **Prioritize Contact Tracing and Case Investigation for Data Collection:** Cases of variants of concern should be prioritized for contact tracing and case investigation so that public health officials can observe how the new variant behaves compared to previously circulating versions.
3. **Develop a Genomic Surveillance Strategy:** To guide the public health response, maximize resources, and ensure an equitable distribution of benefits, the US Department of Health and Human Services should develop a national strategy for genomic surveillance to implement and direct a robust SARS-CoV-2 genomic surveillance program, drawing on resources and expertise from across the US government.
4. **Improve Coordination for Genomic Surveillance and Characterization:** There are several factors in creating a successful genomic surveillance and characterization network. Clear leadership and coordination will be necessary.

## Introduction

As of February 2021, 3 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern (VOC) with worrisome characteristics have emerged, each on a different continent.<sup>1</sup> The B.1.1.7 variant, first identified in the United Kingdom, is substantially more transmissible than previously circulating variants. The B.1.351 and P.1 variants, first identified in South Africa and Brazil, respectively, both exhibit some degree of immune escape.\*<sup>2,3</sup> Each of these variants has precipitated resurgences in the communities where they have become dominant. All 3 have already been identified at low levels in the United States. If they gain a foothold, the same resurgences can be expected here.

Our dance with the virus will not be limited to these 3 variants. As the coronavirus disease 2019 (COVID-19) pandemic unfolds, selective pressure and genetic drift will drive additional virus variants that are more fit for their survival and spread. This inexorable process is accelerated by high levels of community transmission. As more people become infected, the virus has more opportunities to generate mutations.

Although many mutations in the SARS-CoV-2 virus will be inconsequential, some can be detrimental. During summer 2020, the D614G mutation that is now known to make the virus more transmissible in humans, became the predominant lineage.<sup>4</sup> The B.1.1.7 variant is also substantially more transmissible and has prompted lockdown measures in several countries where healthcare systems were overwhelmed. Medical countermeasures, specifically vaccines, have also been challenged by new variants.<sup>5</sup> Although several authorized and candidate vaccines have shown high levels of efficacy in clinical trials, the target antigens in the vaccine formulation may not reflect emerging variants.<sup>6</sup> To date, preliminary laboratory data suggest that the Pfizer and Moderna vaccines are still effective against the B.1.1.7 variant.<sup>7,8</sup> However, the Moderna vaccine and some monoclonal antibody therapies may be somewhat less effective against the B.1.351 and P.1 variants.<sup>9,10</sup> Hypothetical future variants may further degrade the performance of vaccines and therapeutics, require redesign of diagnostic tests, and render naturally acquired immunity less durable.

Although viral mutation is inevitable, it is possible to anticipate, manage, and mitigate the threat to our collective public health. The key to staying ahead of a rapidly evolving virus is to maintain a continuous, systematic genomic surveillance and functional characterization capability that is able to rapidly detect and evaluate new variants of concern. These data and subsequent analyses can improve risk mitigation efforts throughout an outbreak, including decisions to update medical countermeasures and adapt the public health response—before new variants evade existing control measures. The emergence of the novel, more transmissible variants of SARS-CoV-2 underscores the importance of genomic epidemiology for pandemic response. As coverage and

---

\* Immune escape refers to a virus's ability to persist in a host despite a current or past immune response. For example, this might occur because of viral mutations that render antibodies ineffective.

consistency of viral sequencing increases, so too will our understanding of emerging variants.

## Mitigating Against Circulating Variants of Concern

With hundreds of cases of B.1.1.7 already detected in the United States, as well as a number of cases of B.1.351 and P.1, public health officials should prepare for these variants to impact their communities. The B.1.1.7 variant has caused severe resurgences in countries where they have become dominant, including prompting a lockdown in the United Kingdom.<sup>11</sup> And although the B.1.351 and P.1 variants have not been fully characterized, they are thought to demonstrate some immune escape.<sup>12</sup> Early evidence suggests these variants may reduce the effectiveness of vaccines and therapeutics and may make people with naturally acquired immunity susceptible to reinfection by these novel variants, despite protection against past SARS-CoV-2 circulating virus.<sup>13,14</sup>

Public health officials must be prepared to use broad strategies to prepare and respond, as they have been doing during the pandemic thus far. The goal of the response should be to reduce transmission and maintain low levels of spread through policy controls until sufficient population immunity is achieved through widespread vaccination to prevent severe resurgences. Doing so will also allow therapeutics and vaccines to be updated as needed. The following policy actions will help to manage the current variants of concern:

**Maintain policies that slow transmission.** The number of new cases reported each day has fallen in the United States since in mid-January 2021. However, incidence still remains high nationwide at 28 cases per 100,000 per day.<sup>15</sup> Although there will continue to be pressure for political leaders to relax restrictions, it is important to maintain policy measures that slow transmission of the virus until incidence falls to much lower levels. Mask mandates, capacity limitations in commercial establishments, and limitations on indoor gatherings are vital to sustain until transmission can be controlled through alternate means. When incidence is below 10 cases per 100,000 per day, contact tracing becomes more feasible and can have a greater impact on breaking chains of transmission.<sup>16</sup>

**Prepare for another resurgence.** Although moderate levels of population immunity, combined with increased vaccination coverage, will likely blunt the impact of B.1.1.7 compared to the experiences in the United Kingdom and Denmark, leaders in the United States should prepare for a resurgence with this or other variants, nonetheless. These plans may not be needed, but having them at the ready will improve the effectiveness of the response if they are put to use. It will be important to have clear guidance from the US Centers for Disease Control and Prevention (CDC) for phased actions based on incidence thresholds. State and local leaders should also decide in advance what actions they will take if the epidemiological situation worsens and threatens to overwhelm health systems. Although school closures and further restrictions on businesses are highly undesirable, they must remain options in severe

scenarios, such as if hospitals are on track to become overwhelmed. Similarly, public health and hospital systems should review and update their hospital preparedness plans in case of another surge, including incorporating lessons learned from the immense strain that healthcare systems and staff faced during the winter wave.

**Accelerate widespread vaccination.** Jurisdictions have been working tirelessly to administer SARS-CoV-2 vaccines since they were authorized in December 2020. These efforts must continue through the spring and summer months to reach as many people as possible. Vaccination will not only protect individuals, including those at high risk of severe illness, it will also make it more difficult for the virus to find new hosts, thereby reducing community prevalence and indirectly protecting people who have not been vaccinated. State and local leaders, as well as public health and medical professionals, should continue to adapt their mass vaccination strategies over time, such as by expanding distribution networks to include pharmacies and primary care clinics to expand widespread vaccination coverage.

**Do all that is possible to increase availability of diagnostic and screening testing.**

Although demand for diagnostic testing in people with symptoms or known exposure will fall as community prevalence declines, testing must continue to be a mainstay in the COVID-19 pandemic toolkit. As vaccination campaigns roll out, resources and logistics for testing should continue to expand. Diagnostic testing should continue to be free and easily accessible everywhere, especially in hardest-hit communities. In addition to diagnostic testing, screening tests in group settings like congregate living facilities, institutions of incarceration, and schools can help to identify cases before they become outbreaks. These 2 testing approaches together can identify people who need to isolate, provide public health officials information about hotspots, and protect high-risk settings.

**Encourage the public to use more protective masks.** There is now robust evidence that universal mask wearing slows transmission of SARS-CoV-2, including a recent study from the CDC that finds a 95% reduction in infectious aerosols with a well-fitting mask.<sup>17</sup> Policies that encourage or require mask wearing by people when spending time in public should be maintained until transmission decreases substantially. With the emergence of variants with higher transmissibility and the growing availability of high-quality masks on the commercial market, public health officials should emphasize the benefits of and encourage the use of higher-quality masks, including those with multiple layers and a snug fit.<sup>18</sup> Clarification of mask quality standards and proper use guidance should be clear and actionable.

**Leverage contact tracing and case investigation as tools for data collection.** Although contact tracing and case investigations are most widely understood as a means to break chains of transmission, they are also important tools for collecting data and learning about transmission dynamics, even in high-prevalence settings where tracing of every case is not feasible. In that situation, cases of variants of concern should be prioritized for contact tracing and case investigation so that public health officials can

observe how the new variant behaves compared to previously circulating versions. Data about the secondary attack rate, conditions under which transmission occurred, and epidemiological parameters like serial intervals are all helpful for understanding the public health impact of genetic changes.

## **Augmenting Our National Ability to Identify and Characterize Future Variants**

### **New Investments Are Needed to Bolster US Genomic Surveillance Systems**

Genetic epidemiology is a powerful tool for outbreak detection and response. Outbreak investigations of hepatitis C, tuberculosis, influenza, foodborne outbreaks, and HIV have all been enhanced by genetic epidemiology methods. These approaches have been used to map transmission dynamics, investigate causative sources, and understand how pathogens evolve.<sup>19-28</sup> More recently, it has become clear that genetic and molecular epidemiology will be crucial in the next phase of the COVID-19 response.

One facet of genetic epidemiology, genomic surveillance, is used to monitor changes in the viral landscape over time, including the detection of new variants of concern. Genomic surveillance involves sampling and sequencing viral isolates from people who are infected and comparing these sequences to identify mutations in the genome. Although the United States currently has substantial sequencing capacity, it is primarily used for academic investigation rather than public health surveillance. Significant leadership and investment is urgently needed to implement and ramp up genetic epidemiology capabilities to address the emergence of virus variants.

Countries that have invested heavily in this capability have already seen returns. The United Kingdom, through the COVID-19 Genomics Consortium, currently sequences over 9,900 samples per week, constituting approximately 6% to 8% of their positive cases.<sup>29</sup> This practice enabled scientists to identify the B.1.1.7 variant when it was circulating at very low levels and to track the progression as it became dominant, which has given public health officials more time and information to appropriately modify their response. The United Kingdom is now extending its capacity to help other countries. Denmark, another world leader in genomic surveillance, sequences around 12% of its cases, and has also been able to identify and carefully monitor the B.1.1.7 variant.<sup>30</sup> In contrast, at the end of December, the 51,212 sequences the United States submitted to the widely used database GISAID, represented only 0.3% of cases reported at that time.<sup>31</sup> The percentage of reported cases sequenced has not significantly improved since then.<sup>32</sup>

The United States has several genomic surveillance programs and initiatives underway, but they are currently underresourced for what is now required to track the SARS-CoV-2 variants of concern. The CDC Office of Advanced Molecular Detection (AMD) was

founded in 2014 to integrate genomic surveillance and bioinformatics into the nation's public health response and is the driving force behind current US genomic surveillance efforts for public health. AMD collaborates with health departments to provide the technology needed for sequencing and to train staff in genomic data analysis and interpretation for several types of infectious disease outbreaks.<sup>33</sup>

The CDC has also introduced the first program to coordinate genomic surveillance of SARS-CoV-2 among academic labs, commercial sequencing companies, and the Federal government. This consortium, called SARS-CoV-2 Sequencing for Public Health Emergency Response, Epidemiology and Surveillance (SPHERES), is a forum for stakeholders to exchange best practices and share resources.<sup>34</sup> However, the program has not yet received sufficient funding to support operations at the scale needed for effective pandemic response.

A second program, the National SARS-CoV-2 Strain Surveillance (NS3) program, requests state health laboratories to send at least 10 samples and their associated metadata to the CDC every 2 weeks for whole genome sequencing.<sup>35</sup> Announced in October 2020, this program is meant to increase the number of cases sequenced to build a public, population-level genetic epidemiology monitoring system with geographic diversity to detect new variants if they arise.<sup>36</sup> The metadata attached to the sequences is intended to allow public health officials to better understand the epidemiology and improve outbreak investigations. However, the NS3 program only requests about 1,503 samples biweekly, which is insufficient coverage for a robust genetic epidemiology monitoring or surveillance system.<sup>37</sup> Illumina, a commercial sequencing company, estimates that identifying variants of concern at 0.01% to 1.00% prevalence would require sequencing around 5% of polymerase chain reaction-positive cases reported.<sup>38,39</sup> As incidence falls, the target throughput would fall as well. At 100,000 reported cases per day, 5,000 sequences per day would be needed to meet the 5% benchmark. Although that target may be refined through further efforts to optimize an adequate genomic surveillance system, it provides a reasonable benchmark for allocating resources.

There have been several barriers to improving the genetic and molecular epidemiology system in the United States. One of the primary challenges is funding: sequencing and the subsequent bioinformatic analysis is resource intensive. The cost of SARS-CoV-2 sequencing is highly variable and dependent on the study objectives and protocol, with some estimates ranging from about \$25 to \$400 per sample when multiplexing.<sup>40,41</sup> Although the cost of sequencing has dropped dramatically,<sup>42</sup> there are still significant resource requirements for sequencing of clinical samples, including the costs of specialized training, equipment, and reagents.

One of the major hurdles to efficiently using genomic surveillance to inform public health decision making is the complexity of managing sequencing data.<sup>43</sup> Sequencing generates large quantities of data that must be appropriately processed, curated, stored, and queried to be useful in answering questions relevant to outbreak response.<sup>44</sup> Putting genomic sequence data in context of the clinical and epidemiological data

yields more robust results and information. However, interoperability between clinical, epidemiological, and genomic sequence data remains challenging. Furthermore, the data must be analyzed by bioinformatics experts who are equipped to conduct analyses and manage data. Most health departments do not have the data science, computational resources, or bioinformatics expertise necessary to manage a large genetic and/or molecular epidemiology program.

## **Additional Laboratory Resources and Increased Coordination Across Labs Are Needed to Characterize Circulating Variants**

Not all variants identified through genomic surveillance will pose an increased threat to public health. Mutations in the genome do not always translate to changes in protein structure or even function. In some cases, epidemiologists or clinicians might see a change in transmission patterns or clinical presentation and request that sequencing be conducted to investigate whether a genetic component is driving the change. In other cases, a variant could become more dominant or show other worrying signals from genomic surveillance data, which prompts epidemiologists to begin investigating if the variant is causing altered transmission.

In either case, studies are needed to better understand the impacts of mutations in the viral genome. Epidemiologists must investigate if the variant has generated greater numbers of cases or more clinically severe cases. Scientists must conduct phenotypic studies to understand changes in the virus.<sup>45,46</sup> If transmission dynamics are affected or if existing medical countermeasures are a poor match, the public health response should be adapted accordingly before the variant becomes more prevalent.<sup>47</sup>

A loose network of laboratories is involved in variant characterization. Currently, many of these agencies and labs are working with the SPHERES consortium.<sup>34</sup> The CDC NS3 sequencing project is also partnering with commercial diagnostic labs and providing \$9 million to academic research centers, although this is primarily focused on sequence generation.<sup>48</sup> Larger facilities such as those found at the CDC, National Institutes of Health (NIH), National Cancer Institute, Rocky Mountain Labs, and United States Army Medical Research Institute of Infectious Diseases (USAMRIID) can also contribute to characterization, and many of these labs have participated in SARS-CoV-2 characterization already.<sup>49-54</sup> Additional scientific resources should be brought to bear to expand capabilities, including Department of Energy (DOE) laboratories and the National Emerging Infectious Disease Laboratory, that can provide *in silico* and *in vitro* support.<sup>55-60</sup> Additional capacity is also available in public health labs and academic institutions.<sup>61</sup>

Importantly, many of these institutes have animal and high biosafety-level facilities that will ensure the studies are performed safely and securely.<sup>62-64</sup> Despite this diverse pool of participating labs, there is currently little coordination of characterization studies. Data flow, data storage, and translation of this data into public health interventions will require a clear hierarchy of reporting and collaboration.

Of note, as variants emerge and are characterized, any that are suspected of posing a significant public health threat should require careful biosecurity and biosafety measures. Proper training, compensation, and facility management will be essential to ensure the safety and integrity of these studies.<sup>50,65</sup> Many facilities, such as those at USAMRIID, already have these systems in place.<sup>61</sup> Clear guidance from agencies such as the CDC will help inform local and state facilities about sample preparation, storage, and processing, so that these studies can be completed with the proper controls.

Epidemiological and clinical investigation can provide further evidence for how variants behave in the human population by tying clinical outcomes and more traditional epidemiological data to sequencing data. For example, case-control studies comparing transmission and clinical outcomes between 2 variants can reveal increased transmissibility, severity, or incubation period. Prioritizing cases of variants of concern for contact tracing and case investigation is an important source of information to inform characterization.

Critically, genomic surveillance must be maintained to understand trends. Variants of concern are only identified in the context of changing patterns over time; sequencing for 1 day or month will only provide data for a snapshot of that single point or period in time. When establishing a genomic surveillance program, sustainability must be at the foundation. Building an infrastructure for COVID-19 response will also benefit routine outbreak response for a wide variety of diseases and improve preparedness for future outbreaks of novel pathogens.

## Recommendations

### Strategy and Coordination

**Develop a Genomic Surveillance Strategy.** Going forward, it is clear that the scale of genomic surveillance in the United States will need to move closer to the example set by the United Kingdom and Denmark. To maximize resources and ensure an equitable distribution of benefits, the US Department of Health and Human Services should develop a national strategy for genomic surveillance to implement and direct a robust SARS-CoV-2 genomic surveillance program, drawing on resources and expertise from across the US government. The strategy should include plans or directives for resource allocation, workforce requirements, logistics for samples, material transfer agreements, data use agreements, standardized protocols, communication, data management guidelines, data system requirements, data interoperability guidance, and guidance for ensuring patient privacy. The strategy should also enumerate how resources, such as sequencing instruments and reagents and technical expertise, will be equitably distributed across the country and accessible to all states and territories. The National Academies of Science, Engineering, and Medicine report on *Genomic Epidemiology Data Infrastructure Needs for SARS-CoV-2*<sup>66</sup> and the World Health Organization's *Genomic Sequencing of SARS-CoV-2: A Guide to Implementation for Maximum Impact on Public Health*<sup>67</sup> could serve as the basis for the strategy.

**Establish an Academic Center of Excellence for Outbreak Bioinformatics.** An effective genomic surveillance ecosystem extends beyond sequencing into intricate data analysis. Many United States public health agencies lack the robust analytical capabilities to interpret genomic sequencing, especially in an ongoing outbreak. Establishing data technologies, maintaining skilled data scientists, improving bioinformatics models, and providing timely results have been challenging. Funding an academic center of excellence for outbreak bioinformatics would provide ongoing capabilities not only for infectious disease pandemics but also foodborne disease and antimicrobial resistant bacteria outbreaks. This center of excellence could be structured around the goals of improving bioinformatics models, supporting public health decision making, and refining data architectures needed to respond quickly during public health emergencies.

**Coordinating Computational Resources Across US Government Agencies and Other Key Stakeholders.** The DOE and NIH have the computational resources needed to support a robust genetic and molecular epidemiology program, including several that already participate in SPHERES. The CDC, DOE, and NIH should expand partnerships to meet computational needs across stakeholders. DOE national laboratories could also provide computational expertise to states to assist them with collecting and managing their data so that public health authorities at all levels can access critical information. Plans to include and leverage academic, private sector, and philanthropic organization resources should also be prioritized.

**Strengthen Genomic Surveillance and Variant Characterization Coordination.** Significantly expanding the volume of sequences, genotypic, and phenotypic analysis in the United States will require extensive coordination of various data inputs, curation, analysis, and reporting outputs. Currently, AMD is the primary genomic surveillance office at the CDC. While annual influenza surveillance is managed by the CDC, there has never been an event requiring coordination for genomic surveillance between all US states and territories at this scale in the past. Given current needs, AMD should be bolstered to become the central coordinator for genomic surveillance nationwide. In addition to coordinating with each state and territory health department and other stakeholders, AMD will need to manage a large data reporting system and disseminate interpretable results to the public, policy leaders, and public health officials. As SPHERES is already managed by AMD, existing SPHERES partnerships can be leveraged to direct and support the expansion of AMD into the larger coordinating body needed. Investments into AMD and the resulting infrastructure can be used to respond to other diseases in the future.

The current reporting structure of academic, industry, public health, and national labs performing viral characterization studies is not well defined and there is no overarching coordination effort between all entities. The NIH has extensive resources at *in silico*, *in vitro*, and *in vivo* research levels across multiple institutions across the country. Further, they provide grants to many academic institutions that are already engaged in characterization studies. The NIH could host data storage and reporting pipelines on variant characterization studies, for example by expanding their current National

COVID Cohort Collaborative, which serves as a central database for storage and analysis of COVID-19 clinical data.<sup>68,69</sup> Compiling all characterization data in one location will improve communication between researchers and provide a better resource for clinicians and public health practitioners to quickly find information as needed.

## Legislative Action and Funding

**Expand the Public Health Bioinformatics Workforce.** The United States does not have sufficient bioinformatics expertise in state and local health departments to support robust genomic surveillance. Most public health training and education programs do not teach bioinformatics, and there are a limited number of fellowship or training programs to bolster that expertise.<sup>70</sup> The public health bioinformatic workforce can be expanded and sustained by increasing training opportunities, funding more permanent positions within public health departments, and providing incentives for bioinformaticians to enter and stay in public health.

**Fund Research for Viral Characterization.** Current efforts to characterize SARS-CoV-2 variants are relatively independent and are often located at well-established academic and national labs. Beyond the AMD budget, additional funds are needed to enrich characterization studies and create a strong laboratory network across the country. Funding could support current labs working to characterize virus variants and also create new labs and partnerships at local public health centers and academic institutions. This effort could be similar to the CDC Influenza Surveillance Network, which has successfully engaged public health departments, academic labs, clinical labs, healthcare providers, and other partners throughout the country.<sup>71</sup> Another potential model is the NIH Centers of Excellence for Influenza Research and Surveillance, which currently support 6 centers of research excellence that characterize and monitor influenza.<sup>72</sup> A similar program to coordinate characterization research, provide guidance and standardized protocols to member labs as needed, and compile results from efforts across the network for interpretation and dissemination to public health practitioners. A panel of experts from across disciplines could be convened by this coordinating office to help create prioritizations for characterization research.

## Policy, Guidance, and Protocols

**Develop a Risk Characterization Framework.** Assessing variants of concern is common for other pathogens, and risk characterization frameworks have been established for influenza viruses. The influenza risk assessment tool enables a large group of experts to systematically assess the available evidence on new influenza variants in order to assess the risk to public health.<sup>73,74</sup> Similar risk assessment frameworks should be established for coronaviruses and used to inform the response to COVID-19 and prepare for future pandemics.

**Ensure Diversity and Proportional Representation in Samples Sent for Sequencing.** In addition to geographical representation, samples sent for sequencing must represent the diversity of the United States. If molecular epidemiology capacity is limited only to

high-resource settings or only some geographic regions, sampling bias will occur due to a failure to capture the true representation of variants in a population. This bias can be especially problematic considering the disproportionate impacts of the pandemic on Black, Indigenous, Latino/Latinx, immigrant, and other communities in low-resource jurisdictions. The CDC should create clear guidelines for jurisdictions on an appropriate sampling strategy for sequencing that will ensure the collection of samples used to inform response is representative.

**Create Standardized Protocols and Guidance for COVID-19 Genetic and Molecular Epidemiology.** The CDC and National Center for Biotechnology Information should release standardized protocols for sequencing assays and bioinformatic protocols to avoid biases in data from differences in how jurisdictions are conducting sequencing or the subsequent analysis. The CDC should provide guidance on how to ensure no human DNA is accidentally made public or misused, who owns samples and data, how to appropriately handle and protect data, and if and when informed consent is needed. These protocols, guidelines, and standards may be set by an independent advisory committee to ensure that all labs performing characterization studies may have similar study design and data analysis. This will improve the ease of reporting and comparing data between districts and across variants.

## Conclusion

The emergence of new variants of SARS-CoV-2 will bring new challenges to efforts to control transmission. At this stage of the pandemic, public health systems have been stretched and depleted of energy and resources. Renewed investment to ensure the capacity to detect and respond to new variants of SARS-CoV-2 is desperately needed. Investments in infrastructure made now to identify, characterize, and sequence new variants will have tremendous impacts not only on our response to COVID-19 but also on future infectious disease threats that will emerge. Our future preparedness depends on our ability to take the lessons learned during the current response and reinforce the necessary systems to ensure a public health catastrophe like this one never happens again.

## References

1. US Centers for Disease Control and Prevention. Emerging SARS-CoV-2 variants. Updated January 28, 2021. Accessed February 11, 2021. <https://www.cdc.gov/coronavirus/2019-ncov/more/science-and-research/scientific-brief-emerging-variants.html>
2. Volz E, Mishra S, Chand M, et al. Transmission of SARS-CoV-2 lineage B.1.1.7 in England: insights from linking epidemiological and genetic data. Preprint. *medRxiv*. Posted January 4, 2021. Accessed February 11, 2021. <http://doi.org/10.1101/2020.12.30.20249034>
3. Callaway E. Fast-spreading COVID variant can elude immune responses. *Nature*. 2021;589(7843):500-501.
4. 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.
5. Collier DA, Marco AD, Ferreira IATM, et al. SARS-CoV-2 B.1.1.7 escape from mRNA vaccine-elicited neutralizing antibodies. Preprint. *medRxiv*. Posted February 2, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.19.21249840>
6. US Centers for Disease Control and Prevention. COVID-19 and your health. Published February 11, 2020. Accessed February 9, 2021. <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines.html>
7. Moderna. Statement on variants of the SARS-CoV-2 virus. Published December 23, 2020. Accessed February 9, 2021. <https://investors.modernatx.com/news-releases/news-release-details/statement-variants-sars-cov-2-virus>
8. Joseph A. Pfizer-BioNTech Covid-19 vaccine works just as well against variant first detected in U.K., study indicates. *STAT*. January 20, 2021. Accessed February 11, 2021. <https://www.statnews.com/2021/01/20/pfizer-biontech-covid-19-vaccine-works-just-as-well-against-variant-first-detected-in-u-k-study-indicates/>
9. Grady D, Mandavilli A, Thomas K. As virus grows stealthier, vaccine makers reconsider battle plans. *New York Times*. Published January 25, 2021. Updated February 3, 2021. Accessed February 11, 2021. <https://www.nytimes.com/2021/01/25/health/coronavirus-moderna-vaccine-variant.html>
10. Wang P, Liu L, Iketani S, et al. Increased resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7 to antibody neutralization. Preprint. *bioRxiv*. Posted January 26, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.25.428137>
11. Schnirring L. COVID variant-fueled surge, health system pressure trigger new UK lockdown. Center for Infectious Disease Research and Policy website. Published January 4, 2021. Accessed February 11, 2021. <https://www.cidrap.umn.edu/news-perspective/2021/01/covid-variant-fueled-surge-health-system-pressure-trigger-new-uk-lockdown>
12. Wibmer CK, Ayres F, Hermanus T, et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Preprint. *bioRxiv*. Posted January 19, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.18.427166>
13. NOVAVAX. Novavax COVID-19 vaccine demonstrates 89.3% efficacy in UK Phase 3 trial. Published January 28, 2021. Accessed February 11, 2021. <https://ir.novavax.com/news-releases/news-release-details/novavax-covid-19-vaccine-demonstrates-893-efficacy-uk-phase-3>
14. Liu Z, VanBlargan LA, Rothlauf PW, et al. Landscape analysis of escape variants identifies SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. Preprint. *bioRxiv*. Posted November 8, 2020. Accessed February 11, 2021. <https://doi.org/10.1101/2020.11.06.372037>

15. US Centers for Disease Control and Prevention. COVID data tracker. Accessed February 11, 2021. <https://covid.cdc.gov/covid-data-tracker>
16. Johns Hopkins Coronavirus Resource Center. Center for Health Security/NPR/Partners in Health state survey: December 15, 2020. Accessed February 11, 2021. <https://coronavirus.jhu.edu/contact-tracing/state-survey-results>
17. Brooks JT, Beezhold DH, Noti JD, et al. Maximizing fit for cloth and medical procedure masks to improve performance and reduce SARS-CoV-2 transmission and exposure, 2021. *MMWR Morb Mortal Wkly Rep*. February 10, 2021. doi:10.15585/mmwr.mm7007e1
18. Horby P, Huntley C, Davies N, et al. *NERVTAG Note on B.1.1.7 Severity*. Scientific Advisory Groups for Emergencies (SAGE) meeting paper. London: SAGE; 2021. Accessed February 11, 2021. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/955239/NERVTAG\\_paper\\_on\\_variant\\_of\\_concern\\_VOC\\_B.1.1.7.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955239/NERVTAG_paper_on_variant_of_concern_VOC_B.1.1.7.pdf)
19. US Centers for Disease Control and Prevention. Influenza virus genome sequencing and genetic characterization. Page last reviewed October 15, 2019. Accessed February 11, 2021. <https://www.cdc.gov/flu/about/professionals/genetic-characterization.htm>
20. Khoury MJ. Integrating genomics into public health surveillance: ushering in a new era of precision public health. Published July 19, 2017. Accessed February 11, 2021. <https://blogs.cdc.gov/genomics/2017/07/19/integrating-genomics/>
21. US Centers for Disease Control and Prevention. HCV molecular epidemiology. Page last reviewed January 23, 2019. Accessed February 11, 2021. <https://www.cdc.gov/hepatitis/outbreaks/HCV-MolecularEpidemiology.htm>
22. Gire SK, Goba A, Andersen KG, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*. 2014;345(6202):1369-1372.
23. Wohl S, Metsky HC, Schaffner SF, et al. Combining genomics and epidemiology to track mumps virus transmission in the United States. *PLoS Biol*. 2020;18(2):e3000611.
24. Grubaugh ND, Ladner JT, Kraemer MUG, et al. Genomic epidemiology reveals multiple introductions of Zika virus into the United States. *Nature*. 2017;546(7658):401-405.
25. Wohl S, Schaffner SF, Sabeti PC. Genomic analysis of viral outbreaks. *Annu Rev Virol*. 2016;3(1):173-195.
26. Walker TM, Merker M, Kohl TA, Crook DW, Niemann S, Peto TEA. Whole genome sequencing for M/XDR tuberculosis surveillance and for resistance testing. *Clin Microbiol Infect*. 2017;23(3):161-166.
27. Bonsall D, Golubchik T, de Cesare M, et al. A comprehensive genomics solution for HIV surveillance and clinical monitoring in low-income settings. *J Clin Microbiol*. 2020;58(10):e00382-20.
28. Deng X, den Bakker HC, Hendriksen RS. Genomic epidemiology: whole-genome-sequencing-powered surveillance and outbreak investigation of foodborne bacterial pathogens. *Annu Rev Food Sci Technol*. 2016;7:353-374.
29. COVID-19 Genomics UK Consortium (COG-UK). Summary report: COG-UK geographic coverage of SARS-CoV-2 sample sequencing. Published February 2, 2021. Accessed February 11, 2021. [https://www.cogconsortium.uk/wp-content/uploads/2021/02/COG-UK-geo-coverage\\_2021-01-25\\_summary.pdf](https://www.cogconsortium.uk/wp-content/uploads/2021/02/COG-UK-geo-coverage_2021-01-25_summary.pdf)
30. Cyranoski D. Alarming COVID variants show vital role of genomic surveillance. *Nature*. 2021;589(7842):337-338.

31. Stevens H, Berger M. U.S. ranks 43rd worldwide in sequencing to check for coronavirus variants like the one found in the U.K. *Washington Post*. December 23, 2020. Accessed February 11, 2021. <https://www.washingtonpost.com/world/2020/12/23/us-leads-world-coronavirus-cases-ranks-43rd-sequencing-check-variants/>
32. GISAID. Phylogeny. Accessed February 10, 2021. <https://www.gisaid.org/phylogenetics/global/nextstrain/>
33. Centers for Disease Control and Prevention Office of Advanced Molecular Detection. Who we are. Page last reviewed January 8, 2019. Accessed February 11, 2021. <https://www.cdc.gov/amd/who-we-are/index.html>
34. US Centers for Disease Control and Prevention. SPHERES. Updated July 27, 2020. Accessed February 11, 2021. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/spheres.html>
35. US Centers for Disease Control and Prevention. National SARS-CoV-2 Strain Surveillance (NS3) frequently asked questions (FAQs). Published February 5, 2021. Accessed February 11, 2021. [https://www.cdc.gov/programs/preparedness/Crisis-Management/Documents/FEB%202021%20Revised%20NS3%20FAQ\\_02052021%20FNL.pdf](https://www.cdc.gov/programs/preparedness/Crisis-Management/Documents/FEB%202021%20Revised%20NS3%20FAQ_02052021%20FNL.pdf)
36. Walke H. CDC guidance for specimen submission for sequence-based surveillance of SARS-CoV-2. Published October 1, 2020. Accessed February 11, 2021. <https://www.cdc.gov/programs/preparedness/Crisis-Management/Documents/SARS-CoV-2-Strain-Surveillance-Letter-10-01-2020.pdf>
37. US Centers for Disease Control and Prevention (CDC). *Appendix 1: National SARS-CoV-2 Strain Surveillance (NS3) Submissions to CDC for SARS-CoV-2 Positive Samples*. Atlanta: CDC; 2020. Accessed February 11, 2021. <https://www.cdc.gov/programs/preparedness/Crisis-Management/Documents/NS3-Submission-Guidance.pdf>
38. Gottlieb S, Rivers C. Send out the search party for new Covid strains. *Wall Street Journal*. January 24, 2021. Accessed February 9, 2021. <https://www.wsj.com/articles/send-out-the-search-party-for-new-covid-strains-11611525878>
39. Vavrek D, Speroni L, Curnow KJ, Oberholzer M, Moeder V, Febbo PG. Genomic surveillance at scale is required to detect newly emerging strains at an early timepoint. Preprint. *medRxiv*. Posted January 15, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.12.21249613>
40. Paden CR, Tao Y, Queen K, et al. Rapid, sensitive, full-genome sequencing of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis*. 2020;26(10):2401-2405.
41. Peacock S. Reflections on the achievements of COG-UK. COVID-19 Genomics UK website. Published January 22, 2021. Accessed February 11, 2021. [https://www.cogconsortium.uk/news\\_item/reflections-on-the-achievements-of-cog-uk/](https://www.cogconsortium.uk/news_item/reflections-on-the-achievements-of-cog-uk/)
42. National Human Genome Research Institute. DNA sequencing costs: data. Last updated December 7, 2020. Accessed February 11, 2021. <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>
43. Black A, MacCannell DR, Sibley TR, Bedford T. Ten recommendations for supporting open pathogen genomic analysis in public health settings. Preprint. *preprints*. Posted January 11, 2020. Accessed February 11, 2021. <https://www.preprints.org/manuscript/202001.0107/v1>
44. Aarestrup FM, Brown EW, Detter C, et al. Integrating genome-based informatics to modernize global disease monitoring, information sharing, and response. *Emerg Infect Dis*. 2012;18(11):e1.

45. Rambaut A, Loman N, Pybus O, et al. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. Preprint. *Virological*. Posted December 18, 2020. Accessed February 9, 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>
46. Ahmed W, Philip AM, Biswas KH. Stable interaction of the UK B.1.1.7 lineage SARS-CoV-2 S1 spike N501Y mutant with ACE2 revealed by molecular dynamics simulation. Preprint. *bioRxiv*. Posted January 7, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.07.425307>
47. Luring AS, Hodcroft EB. Genetic variants of SARS-CoV-2—what do they mean? *JAMA*. January 6, 2021. doi:10.1001/jama.2020.27124
48. US Centers for Disease Control and Prevention. Genomic surveillance for SARS-CoV-2 variants. Updated February 2, 2021. Accessed February 11, 2021. <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance.html>
49. US Centers for Disease Control and Prevention. *CDC Strategy for Global Response to COVID-19 (2020-2023)*. Atlanta: CDC; 2021. [https://www.cdc.gov/coronavirus/2019-ncov/images/CDCGlobalCOVIDStrategy\\_2020-2023.pdf](https://www.cdc.gov/coronavirus/2019-ncov/images/CDCGlobalCOVIDStrategy_2020-2023.pdf)
50. US Centers for Disease Control and Prevention. Interim laboratory biosafety guidelines for handling and processing specimens associated with coronavirus disease 2019 (COVID-19). Updated January 6, 2021. Accessed February 11, 2021. <https://www.cdc.gov/coronavirus/2019-ncov/lab/lab-biosafety-guidelines.html>
51. National Institutes of Health. NIAID strategic plan details COVID-19 research priorities. Published April 23, 2020. Accessed February 11, 2021. <https://www.nih.gov/news-events/news-releases/niaid-strategic-plan-details-covid-19-research-priorities>
52. National Cancer Institute. NCI part of federal effort to evaluate antibody tests for novel coronavirus. Published May 5, 2020. Accessed February 11, 2021. <https://www.cancer.gov/news-events/cancer-currents-blog/2020/covid-19-nci-antibody-testing-review>
53. National Institute of Allergy and Infectious Diseases. Rocky Mountain Labs overview. Last reviewed September 5, 2019. Accessed February 11, 2021. <https://www.niaid.nih.gov/about/rocky-mountain-overview>
54. Niiler E. The US Army's virus research lab gears up to fight Covid-19. *WIRED*. April 1, 2020. Accessed February 11, 2021. [https://www.wired.com/story/the-us-armys-virus-research-lab-gears-up-to-fight-covid-19/?utm\\_medium=website&utm\\_source=globalbiodefense.com&utm\\_campaign=headlines&utm\\_content=sarscov2](https://www.wired.com/story/the-us-armys-virus-research-lab-gears-up-to-fight-covid-19/?utm_medium=website&utm_source=globalbiodefense.com&utm_campaign=headlines&utm_content=sarscov2)
55. National Emerging Infectious Diseases Laboratories (NEIDL). Current research at the NEIDL. Accessed February 11, 2021. <https://www.bu.edu/neidl/research/current-research/>
56. Sandia National Laboratories. COVID-19 research. Updated July 2, 2020. Accessed February 11, 2021. [https://www.sandia.gov/news/publications/covid-19\\_research/index.html](https://www.sandia.gov/news/publications/covid-19_research/index.html)
57. Pacific Northwest National Laboratory (PNNL). PNNL researchers confront COVID-19 challenges. Accessed February 11, 2021. <https://www.pnnl.gov/covid>
58. Los Alamos National Laboratory. Answering the “what-if” questions on COVID interventions with MEDIAN. Published December 10, 2020. Accessed February 11, 2021. <https://www.lanl.gov/discover/news-stories/2020/December/1210-covid-interventions.php>
59. Lawrence Livermore National Laboratory. New partnership to unleash U.S. supercomputing resources in the fight against COVID-19. Published March 26, 2020. Accessed February 11, 2021. <https://www.llnl.gov/news/new-partnership-unleash-us-supercomputing-resources-fight-against-covid-19>

60. US Department of Energy. National laboratories. Accessed February 11, 2021. <https://www.energy.gov/national-laboratories>
61. Federation of American Scientists. US BSL laboratories. Accessed February 11, 2021. <https://fas.org/programs/bio/research.html>
62. National Institute of Allergy and Infectious Diseases. The need for biosafety labs. Page last reviewed May 10, 2018. Accessed February 11, 2021. <https://www.niaid.nih.gov/research/biosafety-labs-needed>
63. HDR. USAMRIID replacement facility. Accessed February 11, 2021. <https://www.hdrinc.com/portfolio/usamriid-replacement-facility>
64. WSP. WSP USA provide expertise for world's largest high-containment research lab. Published September 25, 2019. Accessed February 11, 2021. <https://www.wsp.com/en-US/insights/usamriid-high-containment-research-lab>
65. Association of Public Health Laboratories. Strengthening lab biosafety & biosecurity. Accessed February 11, 2021. <https://www.aphl.org/programs/preparedness/Pages/Strengthening-Lab-Biosafety-Biosecurity.aspx>
66. National Academies of Sciences, Engineering, and Medicine. *Genomic Epidemiology Data Infrastructure Needs for SARS-CoV-2: Modernizing Pandemic Response Strategies*. Washington, DC: The National Academies Press; 2020. <https://doi.org/10.17226/25879>
67. World Health Organization (WHO). *Genomic Sequencing of SARS-CoV-2: A Guide to Implementation for Maximum Impact on Public Health*. Geneva: WHO; 2021. Accessed February 11, 2021. <https://www.who.int/publications-detail-redirect/9789240018440>
68. National Institute of Nursing Research. Building a COVID-19 analytics platform to turn clinical data into knowledge. Published June 22, 2020. Accessed February 11, 2021. <https://www.ninr.nih.gov/newsandinformation/newsandnotes/nih-n3c>
69. National Center for Advancing Translational Sciences. National COVID Cohort Collaborative (N3C). Last updated January 27, 2021. Accessed February 11, 2021. <https://ncats.nih.gov/n3c>
70. Association of Public Health Laboratories (APHL). APHL-CDC Infectious Diseases Bioinformatics Fellowship. Accessed February 11, 2021. <http://www.aphl.org>
71. US Centers for Disease Control and Prevention. U.S. influenza surveillance system: purpose and methods. Page last reviewed October 6, 2020. Accessed February 11, 2021. <https://www.cdc.gov/flu/weekly/overview.htm>
72. Centers of Excellence for Influenza Research and Surveillance (CEIRS). Accessed February 11, 2021. <http://www.niaidceirs.org/>
73. Russell CA, Kasson PM, Donis RO, et al. Improving pandemic influenza risk assessment. *eLife*. 2014;3:e03883.
74. US Centers for Disease Control and Prevention. Influenza Risk Assessment Tool (IRAT). Page last reviewed November 27, 2020. Accessed February 11, 2021. <https://www.cdc.gov/flu/pandemic-resources/national-strategy/risk-assessment.htm>

## Additional Resources

### Understanding Elements of Genomic Sequencing

An optimal genomic surveillance strategy should include both sentinel and active surveillance components. Sentinel surveillance is a framework for monitoring the occurrence of specific conditions by collecting data on a subset of cases. At regular intervals (eg, weekly) samples are collected from designated providers located across the country. These samples are sequenced and analyzed by bioinformaticians. This process can identify new variants emerging in the population and inform public health authorities about transmission dynamics in the jurisdiction. The primary goal of genetic sequencing for sentinel surveillance is to monitor the landscape of lineages or variants of concern.

The design of a sentinel surveillance system for genomic surveillance is dependent on a number of factors, including available resources, characteristics of the viral genome, and the mutation rate.<sup>1</sup> The influenza sentinel surveillance network in the United States works by collecting about 7,000 samples a year from sites registered in the network, sending those samples to regional labs to sequence the samples, and reporting sequences to a common repository. Each geographic region has a baseline percentage of samples to sequence each week to ensure geographic diversity. A key decision when designing a program is the number of samples to collect for sequences. As the number of samples increases, so does the ability to find newly emerging lineages earlier. However, programs are necessarily limited by financial and personnel constraints. Sentinel surveillance programs must balance those constraints while still maintaining a scale that is capable of identifying new variants or other concerning developments before they become prevalent.

In addition to monitoring for the presence of variants of concern, sentinel surveillance samples can be used to compare sequences from people with severe illness to those with mild illness, for example by comparing hospitalized cases with community cases. These data can be used to inform response and risk assessments. Other information that can be assessed from sentinel surveillance programs include estimates of the size of the outbreak, the reproductive number, introduction and reintroduction events, and spillover events.

The active surveillance component is more directed and is deployed to monitor specific situations where evolutionary pressures are high. Candidate environments for active surveillance include long-term care facilities, universities, correctional facilities, and other high-density living situations. Active surveillance can also be useful in populations with unique viral or disease states. The sequencing of samples from patients who have chronic COVID-19 infections, are undergoing extended antibody-based treatments, have suspected reinfection cases, or are suspected of being infected by an animal source or person who has already been vaccinated against COVID-19 should be targeted. Active surveillance is especially important in cases of reinfection in

people who were previously vaccinated. Sequencing breakthrough cases can be used to identify variants that are able to evade the vaccine. If the reinfection is not attributable to a variant of concern, a better understanding of the epidemiology of reinfection could inform vaccination dosing schedules.

Surveillance in human populations will not cover all risks to human health. SARS-CoV-2 can infect animals as well as humans. When a virus expands its host range, or the types of hosts it can infect, the viral population has more potential hosts and transmission increases. In Denmark, officials observed COVID-19 circulation in mink populations and eventually found evidence of mink to human.<sup>2</sup> To mitigate this risk of expanded host range and transmission from humans to animals and back to humans, public health officials can partner with animal health and wildlife officials to create a strategy to conduct surveillance in wild, farmed, and domesticated animals and determine when positive tests should be sequenced. This sequencing data is most helpful when reported into the same system as human sampled sequences are reported. Additionally, preventing spillover into animal populations, or from animal populations back into human populations, are priorities for pandemic response. As with humans, the more transmission that occurs in animals, the bigger the risk that a variant of concern may arise, so genomic surveillance strategies should consider animal health as well for zoonotic pathogens.

## Understanding Variant Characterization

The characterization process has 3 main levels: *in silico*, *in vitro*, and *in vivo*. *In silico* studies use bioinformatic methods to study changes within the virus based on changes in the genome. *In silico* characterization can involve sequence comparison and phylogenetic tree building, both of which are useful to understand how the virus is changing and how variants relate to one another.<sup>3-5</sup> It can also be used to understand if there are common selection pressures on the virus based on the types of evolution identified and to evaluate whether antigen and molecular tests will still accurately detect the variant.<sup>6</sup> *In silico* work can also include modeling of viral proteins and structure based on sequences alone.<sup>7</sup> While only a prediction, this process can quickly identify suspicious mutations that warrant further investigation.<sup>8</sup> The main resources required for *in silico* studies are bioinformatics specialists and computational resources for sequence analysis. One of the benefits of *in silico* work is that space, facilities, and reagents are lower cost, because most of the work can be performed on a computer. Consequently, the network of *in silico* studies and personnel can be more distributed across the country.

After variants have been characterized at the sequence level, mutations of interest can be further characterized at the *in vitro* level in the lab. Mutations of concern may be introduced to existing constructs of SARS-CoV-2 one at a time, in targeted mutagenesis, so that their impact can be assessed in cell culture.<sup>9</sup> Previous studies, in other viruses or in other SARS-CoV-2 variants, may give clues as to the impact of a mutation.<sup>10,11</sup> For instance, the N501Y mutation is shared between the B.1.1.7 and B.1.351 variants. If the phenotype had been characterized in 1 variant, this could inform the experimental

plan to study that mutation further or to shift focus onto entirely novel mutations. The novel variant may also be studied in its entirety, if samples and high biosafety-level environments are available.<sup>12</sup> This could help characterize the virus as a whole but would not shed light on the roles of individual mutations. It is important to help characterize the behavior of the virus in the method of transmission, infection, propagation, and release from the host cell.

Scientists can parse out the molecular mechanics of mutations, typically using methods such as site-directed mutagenesis or genomic engineering, by introducing mutations one at a time to a SARS-CoV-2 backbone. In doing so, it is possible to assess how each change contributes to a phenotype. However, viruses are often more than the sum of their parts, and combined mutations can lead to phenotypes that may not be entirely predicted by single-point mutation studies. Characterization studies using the actual variant, with all mutations incorporated, is most robust. Experiments producing infectious SARS-CoV-2 virus, especially experiments with novel variants, require high biosafety level facilities and trained personnel.

*In vitro* studies allow translation of sequence-based (*in silico*) identification of variants to a phenotype in different cell types and environments. These studies assess whether variants are influencing the efficacy of therapeutics and vaccines by investigating how mutations in the genome translate to the phenotype of the virus.<sup>13,14</sup> *In vitro* characterization involves studying the virus in cell culture in a laboratory environment, often using molecular studies of infection mechanisms and viral replication in different cell types.

These *in vitro* studies are useful for understanding whether therapeutics, like monoclonal antibodies and convalescent plasma, are impacted by a novel variant. For example, if inhibitory abilities of antibodies or plasma are less effective *in vitro*, that suggests a significant phenotypic change of the virus. Similarly, vaccine efficacy can be evaluated by studying the impact of viral mutations on inhibition by vaccinee convalescent plasma.

*In vitro* studies will require greater investments in reagents, facilities, and trained personnel than *in silico* studies. The majority of *in vitro* studies on SARS-CoV-2 are undertaken in large academic or national-level laboratories, with many such laboratories participating in studies of countermeasures such as convalescent plasma and validating diagnostic tests. Academic and industry partnerships should be leveraged to expand the network of laboratories performing these studies. The main requirements for *in vitro* studies are reagents and equipment, biosafety facilities, and trained research personnel. While this requires greater investment than *in silico* studies, these experiments are essential to bridge the sequence to the function of the viral mutations.

The most complex level of characterization involves live animal (*in vivo*) studies that explore how the virus behaves within an individual or a small group of individuals. Because they involve live animals, *in vivo* studies better capture the tissue environments and physiology that the virus encounters in a natural infection. These studies can further detail the phenotype of viral infectivity, severity of infection, and transmissibility (in specialized models).

During *in vivo* studies, animals are exposed to a virus and then carefully monitored and studied to understand the course of illness. Sometimes the animals are treated or vaccinated with a product to determine whether it is effective. Animal transmission studies are also possible, with the right model that can mimic coughing, respiratory droplet transmission, and socialization.<sup>15,16</sup> The results from these studies, while not directly translatable to humans, can indicate changes to viral phenotype such as immune escape or viral replication changes.

Animal models are complex; they require careful Institutional Animal Care and Use Committee protocol approval, veterinarian involvement, and further reagents. The financial investment is significant at this stage. Many of the facilities performing *in vivo* studies may also perform *in vitro* studies, but the animal models, advanced biosafety facilities, and highly trained research personnel require significant funding.

## References

1. Association of Public Health Laboratories (APHL). *Influenza Virologic Surveillance Right Size Roadmap*. Silver Spring, MD: APHL; 2013. [https://www.aphl.org/AboutAPHL/publications/Documents/ID\\_July2013\\_Influenza-Virologic-Surveillance-Right-Size-Roadmap.pdf](https://www.aphl.org/AboutAPHL/publications/Documents/ID_July2013_Influenza-Virologic-Surveillance-Right-Size-Roadmap.pdf)
2. Hammer AS, Quaade ML, Rasmussen TB, et al. SARS-CoV-2 Transmission between Mink (*Neovison vison*) and Humans, Denmark. *Emerg Infect Dis*. 2021;27(2):547-551.
3. Koyama T, Platt D, Parida L. Variant analysis of SARS-CoV-2 genomes. *Bull World Health Organ*. 2020;98(7):495-504.
4. Isabel S, Graña-Miraglia L, Gutierrez JM, et al. Evolutionary and structural analyses of SARS-CoV-2 D614G spike protein mutation now documented worldwide. *Sci Rep*. 2020;10(1):14031.
5. Batut B, Parsons DP, Fischer S, Beslon G, Knibbe C. In silico experimental evolution: a tool to test evolutionary scenarios. *BMC Bioinformatics*. 2013;14(suppl 15):S11.
6. Wang P, Liu L, Iketani S, et al. Increased resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7 to antibody neutralization. Preprint. *bioRxiv*. Posted January 26, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.25.428137>
7. Yurkovetskiy L, Wang X, Pascal KE, et al. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell*. 2020;183(3):739-751.e8.
8. Korber B, Fischer WM, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. 2020;182(4):812-827.e19.
9. Li Q, Wu J, Nie J, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell*. 2020;182(5):1284-1294.e9.
10. Volz E, Hill V, McCrone JT, et al. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell*. 2021;184(1):64-75.e11.

11. Forni D, Filippi G, Cagliani R, et al. The heptad repeat region is a major selection target in MERS-CoV and related coronaviruses. *Sci Rep*. 2015;5(1):14480.
12. Yao W, Wang Y, Ma D, et al. Circulating SARS-CoV-2 variants B.1.1.7, 501YV2, and P.1 have gained ability to utilize rat and mouse Ace2 and altered in vitro sensitivity to neutralizing antibodies and ACE2-Ig. Preprint. *bioRxiv*. Posted January 30, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.27.428353>
13. Xie X, Liu Y, Liu J, et al. Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. *Nat Med*. February 8, 2021. doi:10.1038/s41591-021-01270-4
14. Hu J, Peng P, Wang K, et al. Emerging SARS-CoV-2 variants reduce neutralization sensitivity to convalescent sera and monoclonal antibodies. Preprint. *bioRxiv*. Posted January 22, 2021. Accessed February 11, 2021. <https://doi.org/10.1101/2021.01.22.427749>
15. Kim Y-I, Kim S-G, Kim S-M, et al. Infection and Rapid Transmission of SARS-CoV-2 in Ferrets. *Cell Host & Microbe*. 2020;27(5):704-709.e2. doi:10.1016/j.chom.2020.03.023
16. 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.

**Johns Hopkins  
Center for Health Security**

621 E. Pratt Street, Suite 210  
Baltimore, MD 21202

Tel: 443-573-3304

Fax: 443-573-3305

[centerhealthsecurity@jhu.edu](mailto:centerhealthsecurity@jhu.edu)  
[centerforhealthsecurity.org](http://centerforhealthsecurity.org)



**JOHNS HOPKINS**  
BLOOMBERG SCHOOL  
of PUBLIC HEALTH

**Center for  
Health Security**