



The need for linked genomic surveillance of SARS-CoV-2

Caroline Colijn^{1*}, David JD Earn², Jonathan Dushoff³, Nicholas H Ogden⁴, Michael Li⁵, Natalie Knox⁶, Gary Van Domselaar⁶, Kristyn Franklin⁷, Gordon Jolly⁸, Sarah P Otto⁹

Abstract

Genomic surveillance during the coronavirus disease 2019 (COVID-19) pandemic has been key to the timely identification of virus variants with important public health consequences, such as variants that can transmit among and cause severe disease in both vaccinated or recovered individuals. The rapid emergence of the Omicron variant highlighted the speed with which the extent of a threat must be assessed. Rapid sequencing and public health institutions' openness to sharing sequence data internationally give an unprecedented opportunity to do this; however, assessing the epidemiological and clinical properties of any new variant remains challenging. Here we highlight a "band of four" key data sources that can help to detect viral variants that threaten COVID-19 management: 1) genetic (virus sequence) data; 2) epidemiological and geographic data; 3) clinical and demographic data; and 4) immunization data. We emphasize the benefits that can be achieved by linking data from these sources and by combining data from these sources with virus sequence data. The considerable challenges of making genomic data available and linked with virus and patient attributes must be balanced against major consequences of not doing so, especially if new variants of concern emerge and spread without timely detection and action.

Suggested citation: Colijn C, Earn DJD, Dushoff J, Ogden NH, Li M, Knox N, Van Domselaar G, Franklin K, Jolly GW, Otto SP. The need for linked genomic surveillance of SARS-CoV-2. *Can Commun Dis Rep* 2022;48(4):131–9. <https://doi.org/10.14745/ccdr.v48i04a03>

Keywords: genomic surveillance, SARS-CoV-2, viral variants, COVID-19, epidemiology, public health, data sharing

Introduction

Since the start of the pandemic, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has evolved in multiple ways that increase its public health threat, with higher transmissibility (Alpha, Delta, Omicron variants) (1–4), partial immune escape (Beta, Omicron variants) (5,6) and greater severity (Alpha, Delta variants) (7–9). The continued emergence and spread of new variants of interest and variants of concern (VOC) have the potential to undermine our ability to manage the coronavirus disease 2019 (COVID-19) pandemic, with costly consequences to health, healthcare systems and economies. The SARS-CoV-2 virus faces heterogeneous selection: highly vaccinated communities and those with substantial immunity from previous infection are partially protected, while unvaccinated communities and those with waning immune protection are susceptible. With rising immunity levels, selection is expected to favour variants that better escape vaccine or

infection-induced immunity (10). It is particularly crucial to know if a new virus variant emerges with mutations that increase 1) the ability to infect vaccinated or recovered individuals, 2) the transmissibility of the virus and/or 3) the severity of the disease. The rapid spread of the Omicron variant has led to the highest demand yet on hospitals in many areas, despite the disease being less severe on average (11), highlighting the urgency of developing the methods and data processes to answer these questions in time to take appropriate preventive action.

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



Affiliations

¹ Department of Mathematics, Simon Fraser University, Burnaby, BC

² Department of Mathematics & Statistics and M. G. DeGroot Institute for Infectious Disease Research, McMaster University, Hamilton, ON

³ Department of Biology and M. G. DeGroot Institute for Infectious Disease Research, McMaster University, Hamilton, ON

⁴ Public Health Risk Sciences Division, National Microbiology Laboratory, Public Health Agency of Canada, St.-Hyacinthe, QC

⁵ Public Health Risk Sciences Division, National Microbiology Laboratory, Public Health Agency of Canada, Guelph, ON

⁶ National Microbiology Laboratory, Public Health Agency of Canada and Department of Medical Microbiology & Infectious Diseases, University of Manitoba, Winnipeg, MB

⁷ Centre for Immunization and Respiratory Infectious Diseases, Public Health Agency of Canada, Calgary, AB

⁸ Public Health Genomics, Public Health Agency of Canada

⁹ Department of Zoology & Biodiversity Research Centre, University of British Columbia, Vancouver, BC

*Correspondence: ccolijn@sfu.ca



It is to be hoped that SARS-CoV-2 will not evolve higher transmissibility simultaneously with higher severity among vaccinated or recovered individuals. The cellular immune response is strong and complex (12–14), and breakthrough infections have had reduced severity compared to infections in unvaccinated individuals (15). Before Omicron emerged, vaccine-induced antibody responses remained strong across a variety of VOCs (16,17), but Omicron is a stark reminder that variants can emerge that substantially evade our immune responses (1–3,18), at least in terms of neutralizing antibodies (14,18–20), dramatically reducing vaccine-induced protection against infection (21). There is no guarantee that future variants will follow Omicron’s path in terms of severity.

Virus sequencing initiatives and related genomic surveillance systems give a high-resolution and near-real-time view of how SARS-CoV-2 is evolving and spreading and of the mutations that are rising in frequency (22). Establishing surveillance systems that can detect evolving viral characteristics that impact clinical outcomes and effectiveness of control measures is a key aim of viral sequencing efforts (23). For a newly emerging variant with uncertain impact, rapidly assessing the degree of risk to control efforts is paramount and requires multiple sources of data.

Data and linkages that are required

While genomic data alone allow certain inferences (e.g. identifying which cases are related, and identifying which mutations occur in a new variant), substantially greater value can be obtained by combining a “band of four” key data sources: genetic data; epidemiological and geographic data; clinical and demographic data; and immunization (or recovery) data.

Genetic data refer to attributes of the virus. Here we focused on SARS-CoV-2 whole genome sequence data, but note that polymerase chain reaction testing can identify specific mutations or deletions without fully sequencing the virus genome and so can provide rapid VOC detection.

Epidemiological and geographic data refer to information about the transmission context, including the geographic location and the reason for testing or sequencing (e.g. whether the individual was part of a known outbreak, was a traveller, was randomly sampled, was a vaccine breakthrough infection, was someone previously infected or was tested for other reasons). Epidemiological data also include information about the source and location of exposure: workplace outbreak; household; travel; community exposure; animal exposure; and health care worker, as well as any other contact investigation information (e.g. indoors vs outdoors, ventilation, community setting).

Clinical and demographic data refer to attributes of individuals infected with SARS-CoV-2, including treatments provided, outcomes (e.g. symptoms, severity) and demographic aspects (e.g. age, comorbidities, exposure risks).

Immunization (or recovery) data refer to attributes of past COVID-19 infection or vaccination, including vaccine type(s), number of doses and dates of doses.

These data are typically gathered by different parts of a health system at different times and are used for a variety of purposes, creating challenges for data linkage. Medical facilities manage the clinical course of disease, contact tracing and other case data are gathered by epidemiological teams in public health, vaccination status may be in medical records or known only to the individual, while sequence information is often collected at specialized sequencing centres. Along the way, information may be lost or remain disconnected. Jurisdictions differ in the extent to which linkages among these data can be made; however, linking these four data sources is the most promising way to rapidly detect variants that have the potential to break through pandemic containment measures.

Opportunities with partial data

It is essential to understand vaccine effectiveness against a variety of outcomes (infection, symptoms, hospitalization, death), as well as intrinsic transmissibility and severity in vaccinated and unvaccinated individuals. These can change rapidly as new variants arise and spread. Links to genetic data can attribute transmissibility, severity and vaccine effectiveness to viral types, and thereby provide a better basis for projecting infections and healthcare burden in the context of vaccination. Viral evolution also causes a continual turn-over in how we classify a virus, as names are given only when a variant has spread and become sufficiently distinct (e.g. by Phylogenetic Assignment of Named Global Outbreak Lineages) (24). Consequently, case data with linked lineage information need to be updated as our classification system changes, and this is only possible if links to sequence data, as opposed to lineage names, are maintained.

With only viral sequences and sample dates, it is possible to identify unusual new variants, bursts of mutations, “mutator” lineages that evolve faster than predicted (25,26) or genetic changes that spread more rapidly than expected; however, rapid growth is difficult to interpret. Rapid growth could be due to viral characteristics, epidemiological fluctuations, travel-associated introductions or sampling artifacts (26). For example, the mutational profile of the Omicron variant was a cause for concern as it includes both new mutations and a number of mutations already seen in other VOC—including mutations known to enable the virus to evade neutralizing antibodies (27). Because of their genetic surveillance system, the Department of Health in South Africa sounded the alarm about Omicron (B.1.1.529; November 25, 2021) after detecting the new subvariant and witnessing its rapid spread in a matter of weeks (first collected on November 11, 2021). The researchers noted key outstanding questions about the effect of Omicron on transmissibility, effectiveness of vaccines and disease severity, which cannot be determined from data on the number of detected Omicron sequences alone (28).



The fields of phylogeography and phylodynamics have enabled the use of virus sequence data to infer the geographic movements of viruses (24,25), identify factors driving transmission across geographic regions (29), estimate the effective reproduction number over time (30,31) and link virus sequences to epidemiological models for a range of applications (32,33); however, there are limitations. Phylogeographic analyses are affected by geographic differences in both sampling rates and strategies. Phylodynamic estimates of reproduction numbers over time tend to be retrospective, apply to large virus populations at the national or international scale, have high degrees of uncertainty and are often not immediately actionable at smaller locations—where public health units need to act. Combining sequence data with the other three bands of data offers more opportunities to use virus sequences to understand transmission, severity and immunity. This combination does not necessarily require individual-level linked data; much could be done with data that are de-identified and even data reported for small groups rather than individuals. Even disaggregating outcomes by VOC status would have very high value, as noted recently for Omicron (34).

If the epidemiological context is known, it is possible to distinguish the emergence of a variant with a high growth rate from growth driven by chance “founder effects” (e.g. superspreader events, social gatherings among unvaccinated individuals, introductions vs transmission in care settings or increased sampling due to a particular outbreak) (35,36). Making this distinction increases the reliability of the inference and the value for both research and public health (36,37). For example, Volz *et al.* combined sequencing and polymerase chain reaction testing data with reason for sequencing (community samples) and geography in estimating transmissibility of the Alpha variant B.1.1.7 (1). Virus sequences can also be linked to travel history to monitor the spread of emerging variants and to inform public health measures aiming to limit importation (24,38,39).

In densely sampled outbreaks, linking virus sequences to epidemiology can offer information of immediate relevance to infection prevention, especially when analysis can be done in real time. Lucey *et al.* used whole genome sequence data to identify previously undetected transmission events in hospital-acquired infections, finding evidence that transmission occurred from both symptomatic and asymptomatic healthcare workers, and occurred disproportionately in patients who required high levels of nursing care, informing better prevention tools (40). In a real-time genomic epidemiology study in Australia, sequencing linked to epidemiological data indicated the probable source of infection and identified previously unknown connections between institutions (37,41). Linking virus sequences to additional host and epidemiological data, such as the location of exposure, would also make it possible to detect mutations that give the virus a context-specific advantage, such as transmitting more efficiently outdoors or among specific age groups.

Linking viral sequence data with host data on age, sex, race, occupation, dwelling type, comorbidities and other clinical/demographic data permits virus and host factors contributing to severe disease to be identified. For example, Bager *et al.* used linked data for virus sequences, hospitalization outcome and a large number of host covariates to demonstrate a higher adjusted risk ratio of hospitalization for the Alpha variant (42). Similarly, Fisman and Tuite estimated the increase in risk of hospitalization, intensive care unit admission and death from N501Y-containing variants and the Delta variant (43). Further resolution could be achieved with whole genome sequence in place of VOC screening data.

Linked immunization and sequence data are essential to determine whether newly emerging types and/or variants reduce vaccine effectiveness and to what extent. For example, Skowronski *et al.* linked VOC typing with vaccine status and testing information to show that a single dose of messenger ribonucleic acid (mRNA) vaccines was similarly effective against the Alpha and Gamma variants and non-VOC SARS-CoV-2 (44). Examining clusters or sets of closely related virus sequences together with immunization status informs us about potential transmission. If a cluster consists mainly of vaccinated individuals, this suggests considerable transmission among these individuals; however, if breakthrough infections are preferentially sequenced, an apparent cluster of breakthrough cases could be missing many unvaccinated individuals who comprised most of the transmission. Distinguishing between these requires linking sequences, vaccination status and reason for sequencing, which may include contact tracing or household information.

The entire band of four is needed to determine whether a virus variant can be transmitted by vaccinated individuals and cause severe disease among them: sequence data can tell us whether this is a new variant; epidemiological data and vaccination data can tell us whether it is being transmitted among vaccinated individuals and clinical data will indicate whether the variant is causing severe disease. Without these four linked pieces—shared sufficiently rapidly and over a large enough area to have strong statistical power—there will be gaps that substantially weaken our ability to monitor the virus’ changing phenotype. Small-scale but aggregated and de-identified data may be sufficient for early warnings and help to avert concerns over privacy.

Data sharing and statistical power

Many jurisdictions may gather virus sequences and clinical, epidemiological and immunization data, but may not permit linkage among them due to structural or other barriers. Even where timely joint analysis of these data is possible, however, there is an additional challenge that an emerging variant or type is necessarily rare when it is first emerging. Sharing data across jurisdictions results in greatly improved statistical power by increasing the total amount of data available. Data delays are an additional problem. Even for countries sharing virus genomic data through the Global Initiative on Sharing All Influenza Data database, lags can span months (45). These extensive time



lags hamper international efforts to track variants and their mutations, determine which are rising in frequency and where, track variants' epidemiological and biological consequences and develop effective public health policy (45). Furthermore, even where sequences are shared in a timely manner to the Global Initiative on Sharing All Influenza Data database, they are typically not shared alongside epidemiological, clinical/demographic and immunization data. Indeed, the barriers to public health data sharing are extensive: van Panhuis *et al.* described technical, motivational, economic, political, legal and ethical barriers (46). Many of these are of daily relevance in the COVID-19 pandemic.

Timeliness matters

To make an immediate practical difference, these data linkages and analyses need to be conducted with as little delay as possible. The sooner a new VOC can be characterized, the more warning decision-makers have about the risk. Identifying the spread of a VOC requires strong real-time genomic surveillance with sampling that reflects community transmission, and it requires regular reporting on the makeup of the virus population.

There are significant challenges to developing timely surveillance for emerging VOC, and these challenges differ according to whether the concern is an increase in severity, immune escape, transmissibility or a combination. It takes many infections before we can estimate a difference in severity, yet changes in severity will shape the impact on the healthcare burden. But only a minority of individuals experience severe disease, and there are inherent delays between infection and eventual outcomes. By the time the risks of hospital and acute care needs can be estimated, many hundreds or thousands of infections will have occurred. To stratify severity estimates by viral factors requires even more hospital records and therefore more infections (potentially thousands). This can be ameliorated slightly by focusing on measures with minimal time lags (for example hospital admissions rather than occupancy) and with timely reporting.

Differences in transmissibility are likely to be apparent earlier than differences in severity, because transmission occurs for all infections (whereas severe outcomes occur for a small minority). Indeed, with both the Alpha and Delta variants, increases in transmissibility were detected well ahead of increases in severity (1,7). Differences in immune evasion may or may not be apparent soon after the relevant variants arise, depending on the genomic surveillance system (e.g. prioritization of breakthrough infections, extent of surveillance) and whether the new type causes severe disease among vaccinated individuals.

An effective surveillance system also requires linking timely detection with timely action. Public health and policy makers need to assess when to take action in the face of the uncertainty that is inherent in early assessments of variants that might increase transmission, severity or immune escape. Early localized actions that prevent a VOC from spreading widely, while costly

in the short-term, reduce the risk of prolonged and global challenges to effective COVID-19 control.

Discussion

Timely and accurate surveillance requires a range of expertise spanning infectious disease epidemiology, statistics, virus evolution, genomics and public health. Benefits are gained not just from combining data but from conducting joint analyses, bringing together a sufficient range of expertise to increase the chance of early detection of an emerging threat. Many standard approaches used to estimate transmissibility, vaccine effectiveness and severity (e.g. attack rates, test negative study designs) are only possible after community transmission is well established. Designing systems to warn of possible elevated transmission, immune evasion and severity when there are still few cases requires integrating many sources of information and expertise and developing and using analytical methods designed to combine these data streams. Furthermore, progress in establishing linked surveillance for SARS-CoV-2 is likely to benefit surveillance for other respiratory pathogens, including newly emerging zoonotic viruses and high-burden pathogens such as influenza and respiratory syncytial virus. Improvements in sequencing technology also allow sequencing multiple viral pathogens sampled from patients or the environment, improving the ability to respond rapidly to any newly emerging virus (47).

There are precedents for strong genomic-based surveillance systems with linkage to clinical and epidemiological data. PulseNet Canada (48) is a virtual electronic network that delivers systemic surveillance for enteric disease and ensures that genomes of causal bacteria are rapidly sequenced. The presence of clusters of cases triggers coordinated outbreak investigations in which data are collected and linked to sequences to assess the full extent of the outbreak and identify the source. For SARS-CoV-2 surveillance, the Canadian COVID-19 Genomics Network (16) aims to establish large-scale virus and host sequencing at a national scale to inform decision-making and track the evolution and spread of the virus. Such national platforms can enable data linkage, either with public access or with privileged access given to approved researchers. Although to date such goals have been hampered in Canada, in part by limited or delayed access to virus sequences and limited linkage.

Throughout the SARS-CoV-2 pandemic, the United Kingdom has led the world in data linking, analyses and public communication in its efforts to understand SARS-CoV-2 evolution and impact on public health. The COVID-19 UK Genomics Consortium (49) performs and coordinates sequencing, with over 1.5 M publicly available viral genomes as of February 17, 2022 (50). Sequences are linked with clinical and epidemiological information and are stored securely. Public health agencies use genomic data linked to clinical, demographic and epidemiological data in the public health response and can provide de-identified COVID-19 patient information into the Cloud Infrastructure for Microbial



Bioinformatics (CLIMB-COVID-19) (51) database. There are systems in place for researchers to access the data.

A recent briefing (SARS-CoV-2 VOC and variants under investigation in England: technical briefing 36) from the UK Health Security Agency (21) provides an excellent example of the impact of research enabled by data linkage in the United Kingdom. This report summarizes research linking Phylogenetic Assignment of Named Global Outbreak lineage information to contact tracing data, permitting the discovery that the BA.2 sublineage of Omicron has shorter serial intervals than the BA.1 sublineage, which in turn impacts the interpretation of selection (higher rate of spread is in part due to faster transmission rather than more overall transmission). Linking to vaccination data, age profiles and severity permitted estimates of protection against severe disease and the likely health care burden of BA.2. Sequence and screen-based characterization of the rise of BA.2 allowed estimates of its rate of spread, which is needed to project the future burden of infection and disease. The report is a collaboration of teams that combine expertise in genomics, outbreak surveillance, contact tracing, epidemiology and data analytics, linking and analyzing emerging data with very rapid turn-around and thereby benefitting the global community.

Beyond national-level analyses, linking data at a local level can provide important insight into transmission routes and outbreak risks; for example, genomic epidemiology tools have been used to examine transmission at the scale of outbreaks (52–56). By linking sequences, clinical outcome, epidemiological data and vaccination status, such local analyses can alert public health to the emergence of a concerning cluster. If there was a growing cluster with transmission among vaccinated individuals and high severity, this could be detected early. Both national and local-scale analyses require linkage among disparate data systems through unique identifiers, collaboration across multiple disciplines, and a process by which researchers can access linked data to develop and validate methods.

Conclusion

The SARS-CoV-2 virus will continue to evolve. We cannot predict where new variants of concern will arise, nor rely on them being detected early in locations that have strong genomic surveillance. The more we build strong surveillance systems worldwide, with high-quality data and linkages, the earlier we will be able to detect new variants and act accordingly. Many wealthy countries have high rates of vaccination, which leads to selection of variants with the ability to transmit among vaccinated individuals. With extensive international travel, emerging variants will be able to rapidly migrate around the world, and any that evade immunity will not be as impacted by vaccination requirements. In the worst case, viral evolution could undermine the potential for vaccination to mitigate the pandemic, even in countries that have not yet reached high vaccination rates. Countries with the resources to conduct high volumes of sequencing and to develop

strongly linked surveillance programs are also the ones that have most benefited from early and extensive vaccination programs. Developing and supporting strong genomic surveillance that enables monitoring the virus' phenotypes is important to help ensure that the vaccines remain effective for the rest of the world.

Authors' statement

CC — Conceived the project, led discussion with all authors, wrote the first draft

SO — Literature overview

NO — Literature overview

GJ — Literature overview

GvD — Literature overview

All authors performed writing-review and editing. All authors contributed text and approved the final manuscript.

Competing interests

None.

Acknowledgments

None.

Funding

We would like to acknowledge funding support from the Natural Sciences and Engineering Research Council (CANMOD;RGPIN/06624-2019). DJDE and JD are grateful for support from The Michael G. DeGroot Institute for Infectious Disease Research (IIDR) at McMaster University. SO is supported by NSERC RGPIN-2016-03711. The sponsors of the funding sources were not involved in this work.

References

1. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, Hinsley WR, Laydon DJ, Dabrera G, O'Toole Á, Amato R, Ragonnet-Cronin M, Harrison I, Jackson B, Ariani CV, Boyd O, Loman NJ, McCrone JT, Gonçalves S, Jorgensen D, Myers R, Hill V, Jackson DK, Gaythorpe K, Groves N, Sillitoe J, Kwiatkowski DP, Flaxman S, Ratmann O, Bhatt S, Hopkins S, Gandy A, Rambaut A, Ferguson NM; COVID-19 Genomics UK (COG-UK) consortium. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* 2021;593(7858):266–9. DOI PubMed



2. Public Health England. SARS-CoV-2 variants of concern and variants under investigation in England. Technical briefing 23. London (UK): Public Health England; 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1018547/Technical_Briefing_23_21_09_16.pdf
3. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, Pearson CA, Russell TW, Tully DC, Washburne AD, Wenseleers T, Gimma A, Waites W, Wong KL, van Zandvoort K, Silverman JD, Diaz-Ordaz K, Keogh R, Eggo RM, Funk S, Jit M, Atkins KE, Edmunds WJ; CMMID COVID-19 Working Group; COVID-19 Genomics United Kingdom (COG-UK) Consortium. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* 2021;372(6538):eabg3055. [DOI PubMed](#)
4. Liu Y, Rocklöv J. The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. *J Travel Med* 2021;28(7):taab124. [DOI](#)
5. Lefèvre B, Tondeur L, Madec Y, Grant R, Lina B, van der Werf S, Rabaud C, Fontanet A. Beta SARS-CoV-2 variant and BNT162b2 vaccine effectiveness in long-term care facilities in France. *Lancet Healthy Longev* 2021;2(11):e685–7. [DOI PubMed](#)
6. National Collaborating Centre for Infectious Diseases. Updates on COVID-19 Variants of Concern (VOC). Winnipeg (MB): NCCID; 2022 (accessed 2022-02-18). <https://nccid.ca/covid-19-variants/>
7. Scientific Advisory Group for Emergencies. NERVTAG: Note on B.1.1.7 severity, 20 January 2021. Government of the United Kingdom; (updated 2021-02-12; access 2021-10-12). <https://www.gov.uk/government/publications/nervtag-note-on-b117-severity-20-january-2021>
8. Frampton D, Rampling T, Cross A, Bailey H, Heaney J, Byott M, Scott R, Sconza R, Price J, Margaritis M, Bergstrom M, Spyer MJ, Miralhes PB, Grant P, Kirk S, Valerio C, Mangera Z, Prabhakar T, Moreno-Cuesta J, Arulkumaran N, Singer M, Shin GY, Sanchez E, Paraskevopoulou SM, Pillay D, McKendry RA, Mirfenderesky M, Houlihan CF, Nastouli E. Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect Dis* 2021;21(9):1246–56. [DOI PubMed](#)
9. Fisman DN, Tuite AR. Evaluation of the relative virulence of novel SARS-CoV-2 variants: a retrospective cohort study in Ontario, Canada. *CMAJ* 2021;193(42):E1619–25. [DOI PubMed](#)
10. Otto SP, Day T, Arino J, Colijn C, Dushoff J, Li M, Mechai S, Van Domselaar G, Wu J, Earn DJ, Ogden NH. The origins and potential future of SARS-CoV-2 variants of concern in the evolving COVID-19 pandemic. *Curr Biol* 2021;31(14):R918–29. [DOI PubMed](#)
11. Ulloa AC, Buchan SA, Daneman N, Brown KA. Early estimates of SARS-CoV-2 Omicron variant severity based on a matched cohort study, Ontario, Canada. *medRxiv* 2021.12.24.21268382. [DOI](#)
12. Le Bert N, Tan AT, Kunasegaran K, Tham CY, Hafezi M, Chia A, Chng MH, Lin M, Tan N, Linster M, Chia WN, Chen MI, Wang LF, Ooi EE, Kalimuddin S, Tambyah PA, Low JG, Tan YJ, Bertoletti A. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* 2020;584(7821):457–62. [DOI PubMed](#)
13. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, Rawlings SA, Sutherland A, Premkumar L, Jadi RS, Marrama D, de Silva AM, Frazier A, Carlin AF, Greenbaum JA, Peters B, Krammer F, Smith DM, Crotty S, Sette A. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* 2020;181(7):1489–1501.e15. [DOI PubMed](#)
14. Wu K, Werner AP, Koch M, Choi A, Narayanan E, Stewart-Jones GB, Colpitts T, Bennett H, Boyoglu-Barnum S, Shi W, Moliva JI, Sullivan NJ, Graham BS, Carfi A, Corbett KS, Seder RA, Edwards DK. Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine. *N Engl J Med* 2021;384(15):1468–70. [DOI PubMed](#)
15. Chia PY, Ong SWX, Chiew CJ, Ang LW, Chavatte JM, Mak TM, Cui L, Kalimuddin S, Chia WN, Tan CW, Chai LYA, Tan SY, Zheng S, Lin RTP, Wang L, Leo YS, Lee VJ, Lye DC, Young BE. Virological and serological kinetics of SARS-CoV-2 Delta variant vaccine-breakthrough infections: a multi-center cohort study. *Clin Microbiol Infect* 2021:S1198-743X(21)00638-8. [DOI](#)
16. Skelly DT, Harding AC, Gilbert-Jaramillo J, Knight ML, Longet S, Brown A, Adele S, Adland E, Brown H; Medawar laboratory Team, Tipton T, Stafford L, Johnson SA, Amini A; OPTIC Clinical Group, Tan TK, Schimanski L, Huang K-YA, Rijal P; PITCH Study Group, CMORE/PHOSP-C Group, Frater J, Goulder P, Conion CP, Jeffery K, Dold C, Pollard AJ, Sigal A, de Oliveira T, Townsend AR, Klenerman P, Dunachie SJ, Barnes E, Carroll MW, James WS. Two doses of SARS-CoV-2 vaccination induce more robust immune responses to emerging SARS-CoV2 variants of concern than does natural infection. *Research Square*. 2021. <https://www.researchsquare.com/article/rs-226857/latest>
17. Tarke A, Sidney J, Methot N, Zhang Y, Dan JM, Goodwin B, Rubiro P, Sutherland A, da Silva Antunes R, Frazier A, Rawlings SA, Smith DM, Peters B, Scheuermann RH, Weiskopf D, Crotty S, Grifoni A, Sette A. Negligible impact of SARS-CoV-2 variants on CD4+ and CD8+ T cell reactivity in COVID-19 exposed donors and vaccinees. *bioRxiv* 2021.02.27.433180. [DOI](#)



18. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DD, Mishra S, Crispim MA, Sales FC, Hawryluk I, McCrone JT, Hulswit RJ, Franco LA, Ramundo MS, de Jesus JG, Andrade PS, Coletti TM, Ferreira GM, Silva CA, Manuli ER, Pereira RH, Peixoto PS, Kraemer MU, Gaburo N Jr, Camilo CD, Hoeltgebaum H, Souza WM, Rocha EC, de Souza LM, de Pinho MC, Araujo LJ, Malta FS, de Lima AB, Silva JD, Zauli DA, Ferreira AC, Schnekenberg RP, Laydon DJ, Walker PG, Schlüter HM, Dos Santos AL, Vidal MS, Del Caro VS, Filho RM, Dos Santos HM, Aguiar RS, Proença-Modena JL, Nelson B, Hay JA, Monod M, Miscouridou X, Coupland H, Sonabend R, Vollmer M, Gandy A, Prete CA Jr, Nascimento VH, Suchard MA, Bowden TA, Pond SL, Wu CH, Ratmann O, Ferguson NM, Dye C, Loman NJ, Lemey P, Rambaut A, Fraiji NA, Carvalho MD, Pybus OG, Flaxman S, Bhatt S, Sabino EC. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* 2021;372(6544):815–21. DOI PubMed
19. Abu-Raddad LJ, Chemaitelly H, Butt AA; National Study Group for COVID-19 Vaccination. Effectiveness of the BNT162b2 Covid-19 vaccine against the B.1.1.7 and B.1.351 variants. *N Engl J Med* 2021;385(2):187–9. DOI PubMed
20. Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, Stowe J, Tessier E, Groves N, Dabrera G, Myers R, Campbell CN, Amirthalingam G, Edmunds M, Zambon M, Brown KE, Hopkins S, Chand M, Ramsay M. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *N Engl J Med* 2021;385(7):585–94. DOI PubMed
21. UK Health Security Agency. Investigation of SARS-CoV-2 variants: technical briefings. London (UK): Government of the United Kingdom; 2021 (accessed 2022-02-17). <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings>
22. Robishaw JD, Alter SM, Solano JJ, Shih RD, DeMets DL, Maki DG, Hennekens CH. Genomic surveillance to combat COVID-19: challenges and opportunities. *Lancet Microbe* 2021;2(9):e481–4. DOI PubMed
23. Genome Canada. CanCOGeN VirusSeq. Ottawa (ON): Genome Canada; 2022 (accessed 2021-07-28). <https://www.genomecanada.ca/en/cancogen/cancogen-virusseq>
24. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, Russell CA, Smith DJ, Pybus OG, Brockmann D, Suchard MA. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog* 2014;10(2):e1003932. DOI PubMed
25. McLaughlin A, Montoya V, Miller RL, Mordecai GJ, Worobey M, Poon AFY, Joy JB. Early and ongoing importations of SARS-CoV-2 in Canada. *medRxiv* 2021.04.09.21255131. DOI
26. Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, Pavlin B, Vandemaële K, Van Kerkhove MD, Jombart T, Morgan O, le Polain de Waroux O. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Euro Surveill* 2021;26(24):2100509. DOI PubMed
27. Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, Huang W, Li Q, Wang P, An R, Wang J, Wang Y, Niu X, Yang S, Liang H, Sun H, Li T, Yu Y, Cui Q, Liu S, Yang X, Du S, Zhang Z, Hao X, Shao F, Jin R, Wang X, Xiao J, Wang Y, Xie XS. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* 2022;602(7898):657–63. DOI PubMed
28. News24Video. Urgent briefing on latest developments around the Covid-19 vaccination programme. YouTube; 25 Nov 2021. <https://www.youtube.com/watch?v=Vh4XMueP1zQ>
29. Dudas G, Carvalho LM, Bedford T, Tatem AJ, Baele G, Faria NR, Park DJ, Ladner JT, Arias A, Asogun D, Bielejec F, Caddy SL, Cotten M, D'Ambrozio J, Dellicour S, Di Caro A, DiCarlo JW, Duraffour S, Elmore MJ, Fakoli LS, Faye O, Gilbert ML, Gevao SM, Gire S, Gladden-Young A, Gnirke A, Goba A, Grant DS, Haagmans BL, Hiscox JA, Jah U, Kugelman JR, Liu D, Lu J, Malboeuf CM, Mate S, Matthews DA, Matranga CB, Meredith LW, Qu J, Quick J, Pas SD, Phan MV, Pollakis G, Reusken CB, Sanchez-Lockhart M, Schaffner SF, Schieffelin JS, Sealfon RS, Simon-Loriere E, Smits SL, Stoecker K, Thorne L, Tobin EA, Vandt MA, Watson SJ, West K, Whitmer S, Wiley MR, Winnicki SM, Wohl S, Wölfel R, Yozwiak NL, Andersen KG, Blyden SO, Bolay F, Carroll MW, Dahn B, Diallo B, Formenty P, Fraser C, Gao GF, Garry RF, Goodfellow I, Günther S, Hapfi CT, Holmes EC, Kargbo B, Keita S, Kellam P, Koopmans MP, Kuhn JH, Loman NJ, Magassouba N, Naidoo D, Nichol ST, Nyenswah T, Palacios G, Pybus OG, Sabeti PC, Sall A, Ströher U, Wurie I, Suchard MA, Lemey P, Rambaut A. Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* 2017;544(7650):309–15. DOI PubMed
30. Stadler T, Kühnert D, Bonhoeffer S, Drummond AJ. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc Natl Acad Sci USA* 2013;110(1):228–33. DOI PubMed
31. Stadler T, Bonhoeffer S. Uncovering epidemiological dynamics in heterogeneous host populations using phylogenetic methods. *Philos Trans R Soc Lond B Biol Sci* 2013;368(1614):20120198. DOI PubMed
32. Volz EM, Siveroni I. Bayesian phylodynamic inference with complex models. *PLOS Comput Biol* 2018;14(11):e1006546. DOI PubMed
33. Ratmann O, Hodcroft EB, Pickles M, Cori A, Hall M, Lycett S, Colijn C, Dearlove B, Didelot X, Frost S, Hossain AS, Joy JB, Kendall M, Kühnert D, Leventhal GE, Liang R, Plazzotta G, Poon AF, Rasmussen DA, Stadler T, Volz E, Weis C, Leigh Brown AJ, Fraser C; PANGEA-HIV Consortium. Phylogenetic Tools for Generalized HIV-1 Epidemics: Findings from the PANGEA-HIV Methods Comparison. *Mol Biol Evol* 2017;34(1):185–203. DOI PubMed
34. Scott L, Hsiao NY, Moyo S, Singh L, Tegally H, Dor G, Maes P, Pybus OG, Kraemer MU, Semenova E, Bhatt S, Flaxman S, Faria NR, de Oliveira T. Track Omicron's spread with molecular data. *Science* 2021;374(6574):1454–5. DOI PubMed



35. Hamilton WL, Tonkin-Hill G, Smith ER, Aggarwal D, Houldcroft CJ, Warne B, Meredith LW, Hosmillo M, Jahun AS, Curran MD, Parmar S, Caller LG, Caddy SL, Khokhar FA, Yakovleva A, Hall G, Feltwell T, Pinckert ML, Georgana I, Chaudhry Y, Brown CS, Gonçalves S, Amato R, Harrison EM, Brown NM, Beale MA, Spencer Chapman M, Jackson DK, Johnston I, Alderton A, Sillitoe J, Langford C, Dougan G, Peacock SJ, Kwiatowski DP, Goodfellow IG, Torok ME; COVID-19 Genomics Consortium UK. Genomic epidemiology of COVID-19 in care homes in the east of England. *eLife* 2021;10:e64618. [DOI PubMed](#)
36. Meredith LW, Hamilton WL, Warne B, Houldcroft CJ, Hosmillo M, Jahun AS, Curran MD, Parmar S, Caller LG, Caddy SL, Khokhar FA, Yakovleva A, Hall G, Feltwell T, Forrest S, Sridhar S, Weekes MP, Baker S, Brown N, Moore E, Popay A, Roddick I, Reacher M, Gouliouris T, Peacock SJ, Dougan G, Török ME, Goodfellow I. Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic surveillance study. *Lancet Infect Dis* 2020;20(11):1263–72. [DOI PubMed](#)
37. Seemann T, Lane CR, Sherry NL, Duchene S, Gonçalves da Silva A, Caly L, Sait M, Ballard SA, Horan K, Schultz MB, Hoang T, Easton M, Dougall S, Stinear TP, Druce J, Catton M, Sutton B, van Diemen A, Alprent C, Williamson DA, Howden BP. Tracking the COVID-19 pandemic in Australia using genomics. *Nat Commun* 2020;11(1):4376. [DOI PubMed](#)
38. Health Canada. Priority strategies to optimize testing and quarantine at Canada's borders. Ottawa (ON): HC; 2021 (accessed 2021-10-21). <https://www.canada.ca/en/health-canada/services/drugs-health-products/covid19-industry/medical-devices/testing-screening-advisory-panel/reports-summaries/priority-strategies-quarantine-borders.html>
39. UK Health Security Agency. Data informing international travel traffic-light risk assessments: 7 October 2021. London (UK): Government of the United Kingdom; 2021 (accessed 2021-10-21). <https://www.gov.uk/government/publications/data-informing-international-travel-traffic-light-risk-assessments-7-october-2021>
40. Lucey M, Macori G, Mullane N, Sutton-Fitzpatrick U, Gonzalez G, Coughlan S, Purcell A, Fenelon L, Fanning S, Schaffer K. Whole-genome Sequencing to Track Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Transmission in Nosocomial Outbreaks. *Clin Infect Dis* 2021;72(11):e727–35. [DOI PubMed](#)
41. Rockett RJ, Arnott A, Lam C, Sadsad R, Timms V, Gray KA, Eden JS, Chang S, Gall M, Draper J, Sim EM, Bachmann NL, Carter I, Basile K, Byun R, O'Sullivan MV, Chen SC, Maddocks S, Sorrell TC, Dwyer DE, Holmes EC, Kok J, Prokopenko M, Sintchenko V. Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling. *Nat Med* 2020;26(9):1398–404. [DOI PubMed](#)
42. Bager P, Wohlfahrt J, Fonager J, Rasmussen M, Albertsen M, Michaelsen TY, Møller CH, Ethelberg S, Legarth R, Button MS, Gubbels S, Voldstedlund M, Mølbak K, Skov RL, Fomsgaard A, Krause TG; Danish Covid-19 Genome Consortium. Risk of hospitalisation associated with infection with SARS-CoV-2 lineage B.1.1.7 in Denmark: an observational cohort study. *Lancet Infect Dis* 2021;21(11):1507–17. [DOI PubMed](#)
43. Fisman DN, Tuite AR. Progressive increase in virulence of novel SARS-CoV-2 variants in Ontario, Canada. *medRxiv* 2021.07.05.21260050. [DOI](#)
44. Skowronski DM, Setayeshgar S, Zou M, Prystajecky N, Tyson JR, Galanis E, Naus M, Patrick DM, Sbihi H, El Adam S, Henry B, Hoang LMN, Sadarangani M, Jassem AN, Krajden M. Single-dose mRNA vaccine effectiveness against SARS-CoV-2, including Alpha and Gamma variants: a test-negative design in adults 70 years and older in British Columbia, Canada. *Clin Infect Dis* 2021:ciab616. [DOI](#)
45. Kalia K, Saberwal G, Sharma G. The lag in SARS-CoV-2 genome submissions to GISAID. *Nat Biotechnol* 2021;39(9):1058–60. [DOI PubMed](#)
46. van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst AJ, Heymann D, Burke DS. A systematic review of barriers to data sharing in public health. *BMC Public Health* 2014;14:1144. [DOI PubMed](#)
47. Briese T, Kapoor A, Mishra N, Jain K, Kumar A, Jabado OJ, Lipkin WI. Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. *MBio* 2015;6(5):e01491–15. [DOI PubMed](#)
48. Public Health Agency of Canada. PulseNet Canada. Ottawa (ON): PHAC; (updated 2019; accessed 2021-09-14). <https://www.canada.ca/en/public-health/programs/pulsenet-canada.html>
49. COVID-19 Genomics UK Consortium (COG-UK); 2021 (accessed 2021-09-14). <https://www.cogconsortium.uk/>
50. COVID-19 Genomics UK Consortium. Impact. COVID-19 Genomics UK Consortium; 2022 (accessed 2022-02-17). <https://www.cogconsortium.uk/priority-areas/research/impact/>
51. Nicholls SM, Poplawski R, Bull MJ, Underwood A, Chapman M, Abu-Dahab K, Taylor B, Colquhoun RM, Rowe WP, Jackson B, Hill V, O'Toole Á, Rey S, Southgate J, Amato R, Livett R, Gonçalves S, Harrison EM, Peacock SJ, Aanensen DM, Rambaut A, Connor TR, Loman NJ; COVID-19 Genomics UK (COG-UK) Consortium. CLIMB-COVID: continuous integration supporting decentralised sequencing for SARS-CoV-2 genomic surveillance. *Genome Biol* 2021;22(1):196. [DOI PubMed](#)



- 52. Jombart T, Cori A, Didelot X, Cauchemez S, Fraser C, Ferguson N. Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. *PLOS Comput Biol* 2014;10(1):e1003457. [DOI PubMed](#)
- 53. Didelot X, Fraser C, Gardy J, Colijn C. Genomic Infectious Disease Epidemiology in Partially Sampled and Ongoing Outbreaks. *Mol Biol Evol* 2017;34(4):997–1007. [DOI PubMed](#)
- 54. Klinkenberg D, Backer JA, Didelot X, Colijn C, Wallinga J. Simultaneous inference of phylogenetic and transmission trees in infectious disease outbreaks. *PLOS Comput Biol* 2017;13(5):e1005495. [DOI PubMed](#)
- 55. Hall M, Woolhouse M, Rambaut A. Epidemic Reconstruction in a Phylogenetics Framework: Transmission Trees as Partitions of the Node Set. *PLOS Comput Biol* 2015;11(12):e1004613. [DOI PubMed](#)
- 56. Hall MD, Colijn C. Transmission Trees on a Known Pathogen Phylogeny: enumeration and Sampling. *Mol Biol Evol* 2019;36(6):1333–43. [DOI PubMed](#)



Get **CCDR** delivered to your inbox

- Know the trends
- Get the testing guidelines
- Stay current on new vaccines
- Learn about emerging infections
- Get the table of contents straight to your inbox

SUBSCRIBE TODAY

Web search: