





Temporal Dynamics and Evolution of SARS-CoV-2 Demonstrate the Necessity of Ongoing Viral Genome Sequencing in Ontario, Canada

 Calvin P. Sjaarda,^{a,b} Jennifer L. Guthrie,^c Samira Mubareka,^{d,e} Jared T. Simpson,^{f,g} Bettina Hamelin,^h Henry Wong,ⁱ Leanne Mortimer,^{j,k,l} Robert Slinger,^{l,m}  Andrew G. McArthur,^{n,o,p} Marc Desjardins,^{j,q} Allison McGeer,^{r,s} Tony Mazzulli,^{c,e,r} Katya Douchant,^t Danielle Brabant-Kirwan,^u Ramzi Fattouh,^v Aaron Campigotto,^{e,w} Samir N. Patel,^{c,e} Nahuel Fittipaldi,^{c,e,s} Robert I. Colautti,^x Prameet M. Sheth^{i,t,y,z} for the Ontario's COVID-19 Genomics Rapid Response Coalition

^aDepartment of Psychiatry, Queen's University, Kingston, Ontario, Canada

^bQueen's Genomics Lab at Ongwanada (Q-GLO), Ongwanada Resource Centre, Kingston, Ontario, Canada

^cPublic Health Ontario, Toronto, Ontario, Canada

^dSunnybrook Health Science Centre, Division of Infectious Diseases, Toronto, Ontario, Canada

^eLaboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

^fOntario Institute for Cancer Research, Toronto, Ontario, Canada

^gDepartment of Computer Science, University of Toronto, Toronto, Ontario, Canada

^hOntario Genomics Institute, Toronto, Ontario, Canada

ⁱDivision of Microbiology, Kingston Health Sciences Centre, Kingston, Ontario, Canada

^jDivision of Microbiology, Eastern Ontario Regional Laboratories, Ottawa, Ontario, Canada

^kDepartment of Laboratory Medicine and Pathology, CHEO, Ottawa, Ontario, Canada

^lDepartment of Pathology and Laboratory Medicine, University of Ottawa, Ottawa, Ontario, Canada

^mDepartment of Pediatrics, University of Ottawa, Ottawa, Ontario, Canada

ⁿDepartment of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada

^oM.G. DeGroot Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada

^pDavid Braley Centre for Antibiotic Discovery, McMaster University, Hamilton, Ontario, Canada

^qFaculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

^rDepartment of Microbiology, Mount Sinai Hospital, University Health Network, Toronto, Ontario, Canada

^sDepartment of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada

^tDepartment of Biomedical & Molecular Sciences, Queen's University, Kingston, Ontario, Canada

^uDivision of Microbiology, Health Sciences North, Sudbury, Ontario, Canada

^vSt. Michael's Unity Health, Toronto, Ontario, Canada

^wHospital for Sick Children, Toronto, Ontario, Canada

^xBiology Department, Queen's University, Kingston, Ontario, Canada

^yDepartment of Pathology and Molecular Medicine, Queen's University, Kingston, Ontario, Canada

^zGastrointestinal Disease Research Unit, Kingston Health Sciences Center, Kingston, Ontario, Canada

Robert I. Colautti and Prameet M. Sheth contributed equally to this work.

ABSTRACT Genome-wide variation in SARS-CoV-2 reveals evolution and transmission dynamics which are critical considerations for disease control and prevention decisions. Here, we review estimates of the genome-wide viral mutation rates, summarize current COVID-19 case load in the province of Ontario, Canada (5 January 2021), and analyze published SARS-CoV-2 genomes from Ontario (collected prior to 24 November 2020) to test for more infectious genetic variants or lineages. The reported mutation rate ($\sim 10^{-6}$ nucleotide [nt]⁻¹ cycle⁻¹) for SARS-CoV-2 is typical for coronaviruses. Analysis of published SARS-CoV-2 genomes revealed that the G614 spike protein mutation has dominated infections in Ontario and that SARS-CoV-2 lineages present in Ontario have not differed significantly in their rate of spread. These results suggest that the SARS-CoV-2 population circulating in Ontario has not changed significantly to date. However,

Citation Sjaarda CP, Guthrie JL, Mubareka S, Simpson JT, Hamelin B, Wong H, Mortimer L, Slinger R, McArthur AG, Desjardins M, McGeer A, Mazzulli T, Douchant K, Brabant-Kirwan D, Fattouh R, Campigotto A, Patel SN, Fittipaldi N, Colautti RI, Sheth PM, for the Ontario's COVID-19 Genomics Rapid Response Coalition. 2021. Temporal dynamics and evolution of SARS-CoV-2 demonstrate the necessity of ongoing viral genome sequencing in Ontario, Canada. *mSphere* 6:e00011-21. <https://doi.org/10.1128/mSphere.00011-21>.

Editor Matthew B. Frieman, University of Maryland School of Medicine

Copyright © 2021 Sjaarda et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Calvin P. Sjaarda, calvin.sjaarda@queensu.ca.

Published 5 May 2021

ongoing genome monitoring is essential for identification of new variants and lineages that may contribute to increased viral transmission.

KEYWORDS COVID-19, epidemiology, G614D, genetics, infectious disease, PANGOLIN lineage, public health, SARS-CoV-2

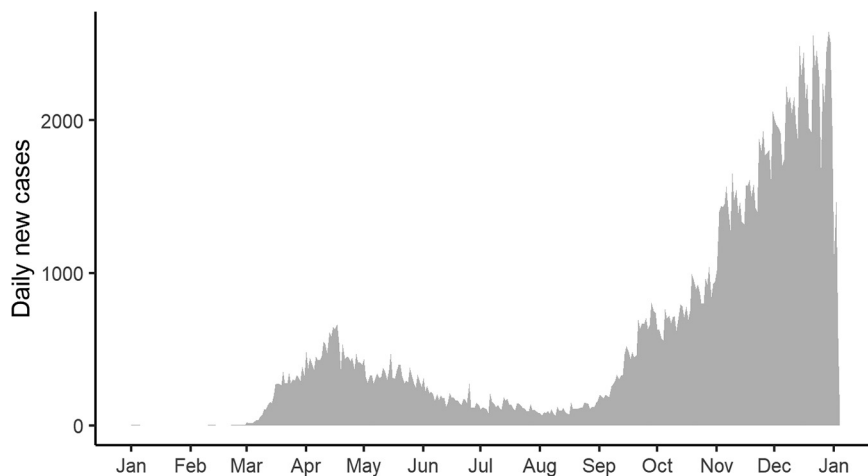
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus disease 2019 (COVID-19), has had devastating consequences on human health and wellbeing (1, 2), health care systems (3), and the global economy (4). Numerous studies have demonstrated the immense value of time-resolved SARS-CoV-2 genome sequencing for tracing viral origin (5, 6), mutational dynamics (7, 8), and transmission properties (9, 10) to inform public health decision-making (11).

The first COVID-19 case in Ontario, Canada, was reported in January 2020, followed thereafter by a rapid increase in reported cases beginning in early March (Fig. 1). The number of new infections spiked in April followed by a significant decrease in new infections by mid-August, which coincided with heavy restrictions on everyday human activity, including physical distancing, limited social events, face mask requirements in public spaces, reduction of nonessential travel, and closing of schools and workplaces. However, only a month later the province saw a sharp rise in COVID-19 cases, signifying a second wave that has since surpassed the first wave in daily case counts (Fig. 1).

Potential contributors to the surge of infections in Ontario include changing behaviors of the host permitting the virus to be transmitted more easily or evolution of the pathogen enabling it to overcome barriers that had previously slowed its spread. To address concerns regarding the mutation of SARS-CoV-2 since its introduction into Ontario and the impact that these mutations may have on transmission efficiency of the virus, we have addressed the following questions. What is the published mutation rate of SARS-CoV-2 and how does it compare with other RNA viruses? Are there SARS-CoV-2 mutations (spike D614G) or lineages circulating in Ontario that may contribute to increased transmission of the virus which correlates with increases in infection rates?

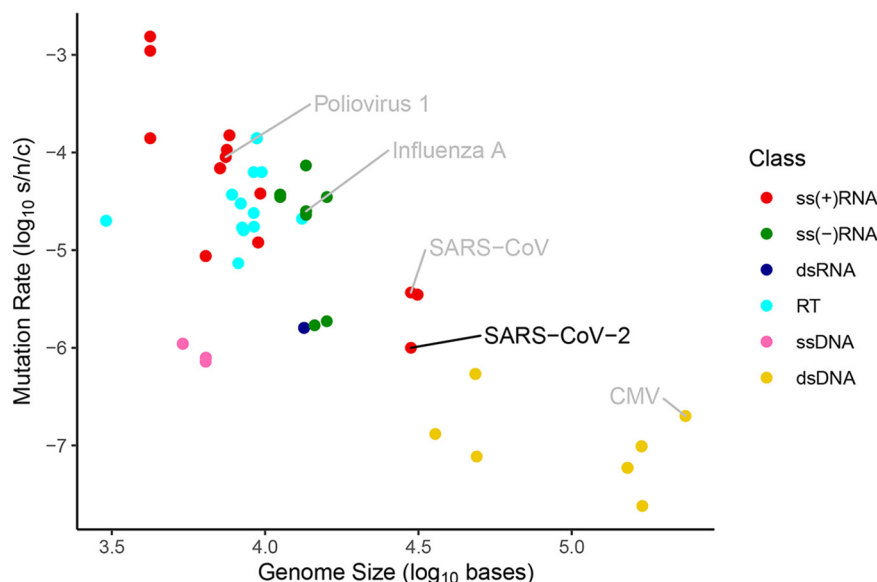
SARS-CoV-2 MUTATION RATE

Viral evolution through genetic mutations enables viruses to increase virulence and transmissibility, escape host defenses, and infect new host species. There are several general trends that determine virus mutation rates (12–17) (Fig. 2). First, RNA viruses have mutation rates that are between 10 and 100 times higher than DNA viruses; a mutation



Source: <https://data.ontario.ca/dataset/>

FIG 1 Daily COVID-19 case counts in Ontario, Canada, from 1 January 2020 until 5 January 2021 show surges of infections beginning in March and September 2020.



sequence from China on 24 January 2020 followed by Germany on 28 January 2020 (29). This variant was found in 10% of globally published genomes by 1 March 2020, and became the most dominant form within a month (8). In the months that followed, hundreds of new lineages containing the D614G variant have been described. Countries that avoided a first wave of SARS-CoV-2 in January and February 2020, including most African and South American countries, report SARS-CoV-2 genotypes that are almost exclusively the G614 variant (29). The D614G amino acid change occurs in the spike protein but is outside the observed spike/ACE2 binding domain (amino acids 387 to 516) (30). The substitution from D to G reportedly enhances SARS-CoV-2 infectivity, competitive fitness, and transmission in primary human cells and animal models (31), which may contribute to the increasing variant frequency at multiple geographic locations, even those with an established D614 SARS-CoV-2 population (8, 32). In addition, viral loads are higher (demonstrated by lower reverse transcription-quantitative PCR (RT-qPCR) threshold cycle [C_T] values) in COVID-19 patients infected with the G614 variant compared with the D614 variant (8, 32). Recently, at least one study has suggested that the second wave of COVID-19 cases may be a result of increased transmissibility of SARS-CoV-2 with the G614 variant (33). In contrast, another study found that D614G does not associate with significantly increased viral transmission and suggests that increases in D614 frequency are a demographic artifact due to a founder effect (34). To our knowledge, no study has shown that spike protein variants differ in virulence within human populations, and the potential impacts, if any, of D614G on the COVID-19 pandemic remain unclear (35).

Research, public health, and clinical laboratories in Ontario have been sequencing SARS-CoV-2 genomes since January 2020 and by 5 January 2021 shared a cumulative 1,743 genomes on the GISAID database (www.gisaid.org) (36). As this study investigated the temporal distribution of SARS-CoV-2 variants and lineages, the 78 samples in the GISAID database reporting a collection date with year only were removed from the analysis. In addition, samples collected on or after 24 November 2020, the first point prevalence analysis (to be published elsewhere), were also excluded, resulting in 1,466 viral genomes being included in this analysis. The SARS-CoV-2 genomes that were classified by GISAID as part of the G, GR, GH, or GV clade had the G614 variant, while all other clade classifications had the D614 variant (www.gisaid.org/references/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/). The earliest sequences (January and February) reported in Ontario have the wild-type sequence (D614), but most of the sequences reported during the first wave (March to June) are the mutant type (G614) (Fig. 3). In fact, the G614 variant accounts for 95.2% of all SARS-CoV-2 genomes reported in Ontario. Given that the G614 variant was already prevalent in Ontario at the beginning of the pandemic and that the D614 variant has not been detected since May, it is likely that the second wave has likewise been dominated by the G614 variant of SARS-CoV-2, suggesting that increased case numbers corresponding to the second wave of COVID-19 are likely unrelated to the spike protein variant. Our understanding of the circulating viral variants in Ontario, especially in rural areas, is limited by the available SARS-CoV-2 genome data on the GISAID server as the data demonstrate temporal bias (Fig. 3) and are geographically weighted to the greater Toronto area and eastern Ontario due to the laboratories sequencing SARS-CoV-2 being in these regions. As more viral genomes are sequenced and shared, this sampling bias may be reduced.

FREQUENCY OF CIRCULATING LINEAGES

A dynamic nomenclature system based on a phylogenetic framework was proposed by Rambaut et al. (37) in July 2020 to help track the global spread of SARS-CoV-2 lineages. This nomenclature was quickly adopted by the international community, and a tool known as PANGOLIN (Phylogenetic Assignment of Named Global Outbreak Lineages) (<https://github.com/hCoV-2019/pangolin>) was developed to assign lineages of newly sequenced genomes. At the root of the SARS-CoV-2 pandemic is lineage A,

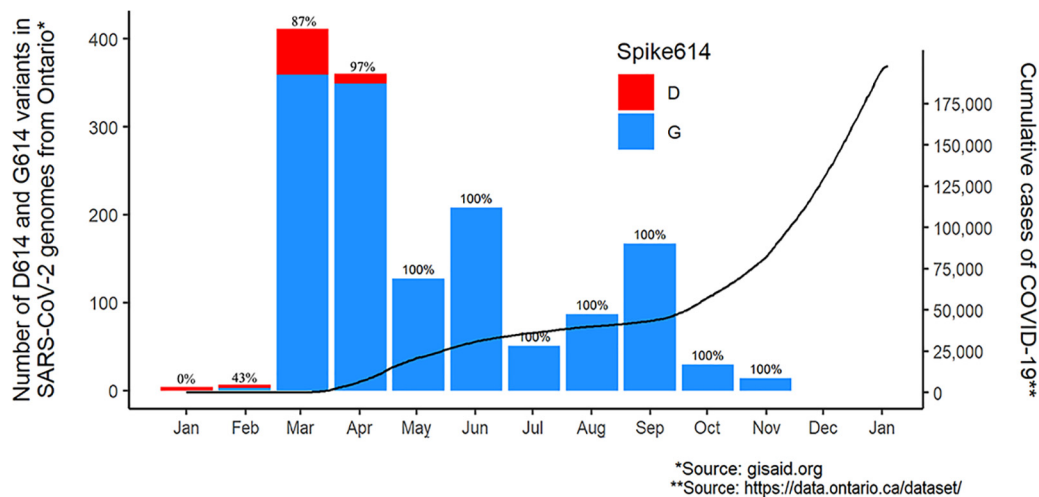


FIG 3 Total COVID-19 cases (line, right axis) and D614G variant distribution in published SARS-CoV-2 genomes (bars, left axis) from COVID-19 cases in Ontario from January 2020 to January 2021. Column labels represent the percentage of total sequences containing the G614 variant in each month.

which originated from China and disseminated globally. Arising from lineage A within China is lineage B, defined by two single nucleotide polymorphisms (SNPs), T8782C and C28144T (37). A number of other early lineages have been associated with specific geographies, including A.1 in the Washington State, USA, outbreak, B.1 in the Italian outbreak, and B.1.1, the major European lineage which was spread throughout the world (37). The variant causing the D614G mutation is characteristic of all SARS-CoV-2 genomes in B.1 and its descendant lineages (38). Many of these early lineages disseminated globally and are now present in most countries, with most local epidemics seeded by many independent introductions of the virus (39). Several lineages, including B.1.147 and B.1.114, have been characterized as Canadian SARS-CoV-2 lineages (<https://cov-lineages.org/descriptions.html>).

There are a total of 65 lineages identified in the 1,466 SARS-CoV-2 genomes published from COVID-19 cases in Ontario; however, only 22 lineages were observed ≥ 10 times and account for 91.9% of the published samples (Fig. 4A). No lineage dominates the more recent published genomes (Fig. 4A), and we detect no significant change in the monthly proportion of sequenced genomic lineages (Fig. 4B; $F_{20,52} = 0.3098$; $P = >0.9$; $R^2 = 0.106$). Instead, lineage prevalence varies widely from month to month, consistent with stochastic effects of human activity such as a superspreader event that increases prevalence of a particular strain in 1 month only to be stamped out with contact tracing the next. However, this analysis is limited by the temporal and geographic biases discussed above. The novel lineage B.1.177 reported across Europe during the summer of 2020 was described in a single case in Ontario in September 2020, but no further cases with this lineage were found to date (40). Furthermore, lineage B.1.1.7 was recently described in the United Kingdom and preliminary characterization suggests that the new strain is significantly more transmissible (41). The presence of this lineage was confirmed in Ontario on 26 December 2020. Therefore, there is a clear need to sequence more SARS-CoV-2 genomes, more broadly sampled, especially from the second wave of infections, to increase statistical power for detecting more subtle variation among lineages in their contribution to the second and future waves of infections in Ontario.

CONCLUSION

The COVID-19 pandemic has had a profound effect on the social and economic welfare of people and governments worldwide as they struggle to reduce the spread of the virus through implementation of comprehensive and aggressive public health measures. Our results contribute to our understanding of SARS-CoV-2, its reported

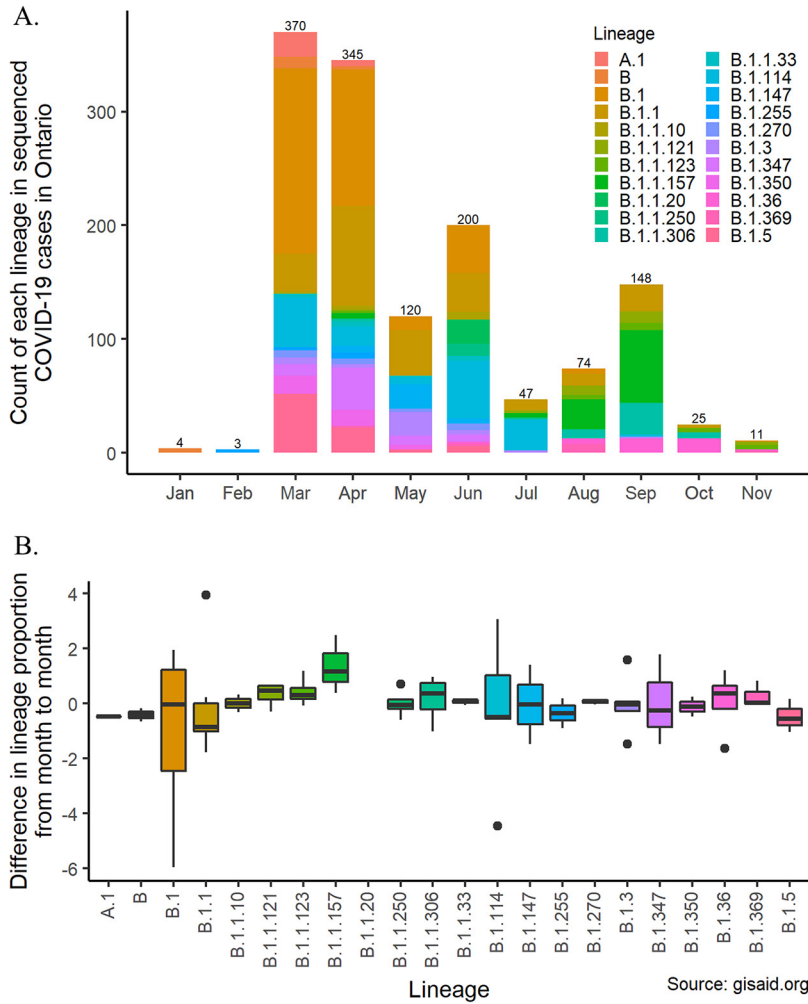


FIG 4 Distribution of major SARS-CoV-2 lineages sequenced from COVID-19 cases in Ontario. (A) Temporal distribution of major SARS-CoV-2 lineages of published genomes from January to December 2020. Column labels represent cumulative number of viral genomes for that month. (B) Box plot representing the monthly change in lineage prevalence (i.e., proportion of genomes) of each of 22 major SARS-CoV-2 lineages circulating in Ontario.

mutation rate, and the variants and lineages that have been circulating in Ontario, Canada. SARS-CoV-2 has a genome size and mutation rate that are typical for coronaviruses and is reported to have a similar or lower mutation rate than SARS-CoV. Furthermore, Ontario’s circulating SARS-CoV-2 represents a relatively mixed genetic population resulting from multiple introductions and within-province spread. As discussed, preliminary evidence reportedly implicates the SARS-CoV-2 G614 variant as a more transmissible variant that may contribute to a second wave of the pandemic. However, this variant has been present in Ontario since the beginning of the pandemic, has been the dominant form of the virus in Ontario and in many other global regions, and is present in most sublineages that are currently circulating globally. We also detect no increased prevalence of a particular lineage that can account for the increase in cases in Ontario. It is possible that changes in human behavior are more likely explanations for the current wave of infections, perhaps due to relaxed control measures or declining compliance with existing measures. To slow transmission of COVID-19 and preserve health system capacity, the Province of Ontario implemented a second provincewide shutdown effective 26 December 2020. Sparse data limit our ability to detect differences in infection rate and highlight the need for Public Health

Ontario Laboratory and academic research groups across Ontario to sequence and archive SARS-CoV-2 genomes from COVID-19 cases, particularly over the course of the second wave. Ongoing SARS-CoV-2 genomic surveillance, like the Canadian COVID Genomics Network (CanCOGeN) initiative, is essential to identify mutations that allow reconstruction of transmission pathways and detection of variants that affect transmissibility, virulence, or host mortality.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TABLE S1, PDF file, 0.01 MB.

ACKNOWLEDGMENTS

We gratefully acknowledge Ongwanada Resource Centre and its Board of Directors for their support of Q-GLO and this project. We thank the following additional members of Public Health Ontario: Sarah Teatero, Sandra Zittermann, and Yao Chen. We acknowledge the Ontario COVID-19 Genomics Coalition (oncov.ca) and CanCoGen (cancogen.ca) for coordinating and funding SARS-CoV-2 sequencing in Ontario. We thank Finlay Maguire (Dalhousie University) and Natalie Knox (National Microbiology Laboratory, Canada) for collaboration on the development of the SIGNAL bioinformatics workflow that generated many of the SARS-CoV-2 genome sequences utilized. We also thank Ilinca Lungu, Bernard Lam, Jeremy Johns, Paul Krzyzanowski, Michael Laszloffy, Lawrence E. Heisler, Richard de Borja, Felicia Vincelli, and Philip Zuzarte for data produced at OICR and released to GISAID. We gratefully acknowledge the originating laboratories and submitting laboratories for their contributions to the GISAID database (see Table S1 in the supplemental material). All contributors of data may be contacted directly via the GISAID website (www.gisaid.org).

Additional members of the Ontario's COVID-19 (ONCoV) Genomics Rapid Response Coalition: Hooman Derakhshani (McMaster University), Michael G. Surette (McMaster University), Amogelang R. Raphenya (McMaster University), Jalees A. Nasir (McMaster University), Emily M. Panousis (McMaster University), Ahmed N. Draia (McMaster University), Hamza Mbareche (Sunnybrook Research Institute), Kuganya Nirmalarajah (Sunnybrook Research Institute), Patryk Aftanas (Sunnybrook Research Institute).

C.P.S., P.M.S., and R.I.C. contributed to the conception and design of the project and performed analysis of data; C.P.S. drafted the manuscript; all authors contributed data, reviewed the manuscript, and approved the submitted version.

REFERENCES

1. Ammar A, Brach M, Trabelsi K, Chtourou H, Boukhris O, Masmoudi L, Bouaziz B, Bentley E, How D, Ahmed M, Muller P, Muller N, Aloui A, Hammouda O, Paineiras-Domingos LL, Braakman-Jansen A, Wrede C, Bastoni S, Pernambuco CS, Mataruna L, Taheri M, Irandoust K, Khacharem A, Bragazzi NL, Chamari K, Glenn JM, Bott NT, Gargouri F, Chaari L, Batatia H, Ali GM, Abdelkarim O, Jarraya M, Abed KE, Souissi N, Van Gemert-Pijnen L, Riemann BL, Riemann L, Moalla W, Gomez-Raja J, Epstein M, Sanderman R, Schulz SV, Jerg A, Al-Horani R, Mansi T, Jmail M, Barbosa F, Ferreira-Santos F, Simunic B, Pisot R, Gaggioli A, Bailey SJ, Steinacker JM, Driss T, Hoekelmann A. 2020. Effects of COVID-19 home confinement on eating behaviour and physical activity: results of the ECLB-COVID19 International Online Survey. *Nutrients* 12:1583. <https://doi.org/10.3390/nu12061583>.
2. Madjid M, Safavi-Naeini P, Solomon SD, Vardeny O. 2020. Potential effects of coronaviruses on the cardiovascular system: a review. *JAMA Cardiol* 5:831–840. <https://doi.org/10.1001/jamacardio.2020.1286>.
3. Iyengar K, Mabrouk A, Jain VK, Venkatesan A, Vaishya R. 2020. Learning opportunities from COVID-19 and future effects on health care system. *Diabetes Metab Syndr* 14:943–946. <https://doi.org/10.1016/j.dsx.2020.06.036>.
4. Ludvigson SC, Ma S, Ng S. 2020. COVID-19 and the macroeconomic effects of costly disasters. National Bureau of Economic Research Working Paper Series no. 26987. National Bureau of Economic Research, Cambridge, MA.
5. Deng X, Gu W, Federman S, Du Plessis L, Pybus OG, Faria NR, Wang C, Yu G, Bushnell B, Pan CY, Guevara H, Sotomayor-Gonzalez A, Zorn K, Gopez A, Servellita V, Hsu E, Miller S, Bedford T, Greninger AL, Roychoudhury P, Starita LM, Famulare M, Chu HY, Shendure J, Jerome KR, Anderson C, Gangavarapu K, Zeller M, Spencer E, Andersen KG, MacCannell D, Paden CR, Li Y, Zhang J, Tong S, Armstrong G, Morrow S, Willis M, Matyas BT, Mase S, Kasirye O, Park M, Masinde G, Chan C, Yu AT, Chai SJ, Villarino E, Bonin B, Wadford DA, Chiu CY. 2020. Genomic surveillance reveals multiple introductions of SARS-CoV-2 into Northern California. *Science* 369:582–587. <https://doi.org/10.1126/science.abb9263>.
6. Sjaarda CP, Rustom N, Evans GA, Huang D, Perez-Patrigeon S, Hudson ML, Wong H, Sun Z, Guan TH, Ayub M, Soares CN, Colautti RI, Sheth PM. 2021. Phylogenomics reveals viral sources, transmission, and potential superinfection in early-stage COVID-19 patients in Ontario. *Sci Rep* 11:3697. <https://doi.org/10.1038/s41598-021-83355-1>.
7. Dearlove B, Lewitus E, Bai H, Li Y, Reeves DB, Joyce MG, Scott PT, Amare MF, Vasani S, Michael NL, Modjarrad K, Rolland M. 2020. A SARS-CoV-2 vaccine candidate would likely match all currently circulating variants. *Proc Natl Acad Sci U S A* 117:1111–1116. <https://doi.org/10.1073/pnas.2005111117>.

- Natl Acad Sci U S A 117:23652–23662. <https://doi.org/10.1073/pnas.2008281117>.
8. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N, Giorgi EE, Bhattacharya T, Foley B, Hastie KM, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI, Angyal A, Brown RL, Carrilero L, Green LR, Groves DC, Johnson KJ, Keeley AJ, Lindsey BB, Parsons PJ, Raza M, Rowland-Jones S, Smith N, Tucker RM, Wang D, Wyles MD, McDanal C, Perez LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC, Saphire EO, Montefiori DC. 2020. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* 182:812–827.e19. <https://doi.org/10.1016/j.cell.2020.06.043>.
 9. Popa A, Genger J-W, Nicholson MD, Penz T, Schmid D, Aberle SW, Agerer B, Lercher A, Endler L, Colaço H, Smyth M, Schuster M, Grau ML, Martínez-Jiménez F, Pich O, Borena W, Pawelka E, Keszei Z, Senekowitsch M, Laine J, Aberle JH, Redlberger-Fritz M, Karolyi M, Zoufaly A, Maritschnik S, Borkovec M, Hufnagl P, Nairz M, Weiss G, Wolfinger MT, von Laer D, Superti-Furga G, Lopez-Bigas N, Puchhammer-Stöckl E, Allerberger F, Michor F, Bock C, Bergthaler A. 2020. Genomic epidemiology of super-spreading events in Austria reveals mutational dynamics and transmission properties of SARS-CoV-2. *Sci Transl Med* 12:eabe2555. <https://doi.org/10.1126/scitranslmed.abe2555>.
 10. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, Taylor J, Spicer K, Bardossy AC, Oakley LP, Tanwar S, Dyal JW, Harney J, Chisty Z, Bell JM, Methner M, Paul P, Carlson CM, McLaughlin HP, Thornburg N, Tong S, Tamin A, Tao Y, Uehara A, Harcourt J, Clark S, Brostrom-Smith C, Page LC, Kay M, Lewis J, Montgomery P, Stone ND, Clark TA, Honein MA, Duchin JS, Jernigan JA. 2020. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N Engl J Med* 382:2081–2090. <https://doi.org/10.1056/NEJMoa2008457>.
 11. Oude Munnink BB, Nieuwenhuijse DF, Stein M, O'Toole Á, Haverkate M, Mollers M, Kamga SK, Schapendonk C, Pronk M, Lexmond P, van der Linden A, Bestebroer T, Chestakova I, Overmars RJ, van Nieuwkoop S, Molenkamp R, van der Eijk AA, GeurtsvanKessel C, Vennema H, Meijer A, Rambaut A, van Dissel J, Sikkema RS, Timen A, Koopmans M, Dutch-Covid-19 response team. 2020. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nat Med* 26:1405–1410. <https://doi.org/10.1038/s41591-020-0997-y>.
 12. Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. 2010. Viral mutation rates. *J Virol* 84:9733–9748. <https://doi.org/10.1128/JVI.00694-10>.
 13. Sanjuan R, Domingo-Calap P. 2016. Mechanisms of viral mutation. *Cell Mol Life Sci* 73:4433–4448. <https://doi.org/10.1007/s00018-016-2299-6>.
 14. Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. 2000. The evolution of mutation rates: separating causes from consequences. *Bioessays* 22:1057–1066. [https://doi.org/10.1002/1521-1878\(200012\)22:12<1057::AID-BIES3>3.0.CO;2-W](https://doi.org/10.1002/1521-1878(200012)22:12<1057::AID-BIES3>3.0.CO;2-W).
 15. Peck KM, Lauring AS. 2018. Complexities of viral mutation rates. *J Virol* 92:e01031-17. <https://doi.org/10.1128/JVI.01031-17>.
 16. Bar-On YM, Flamholz A, Phillips R, Milo R. 2020. SARS-CoV-2 (COVID-19) by the numbers. *Elife* 9:e57309. <https://doi.org/10.7554/eLife.57309>.
 17. Yeh S-H, Wang H-Y, Tsai C-Y, Kao C-L, Yang J-Y, Liu H-W, Su I-J, Tsai S-F, Chen D-S, Chen P-J, Chen D-S, Lee Y-T, Teng C-M, Yang P-C, Ho H-N, Chen P-J, Chang M-F, Wang J-T, Chang S-C, Kao C-L, Wang W-K, Hsiao C-H, Hsueh P-R, National Taiwan University SARS Research Team. 2004. Characterization of severe acute respiratory syndrome coronavirus genomes in Taiwan: molecular epidemiology and genome evolution. *Proc Natl Acad Sci U S A* 101:2542–2547. <https://doi.org/10.1073/pnas.0307904100>.
 18. Drake JW, Charlesworth B, Charlesworth D, Crow JF. 1998. Rates of spontaneous mutation. *Genetics* 148:1667–1686. <https://doi.org/10.1093/genetics/148.4.1667>.
 19. Lynch M, Ackerman MS, Gout J-F, Long H, Sung W, Thomas WK, Foster PL. 2016. Genetic drift, selection and the evolution of the mutation rate. *Nat Rev Genet* 17:704–714. <https://doi.org/10.1038/nrg.2016.104>.
 20. Robson F, Khan KS, Le TK, Paris C, Demirbag S, Barfuss P, Rocchi P, Ng WL. 2020. Coronavirus RNA proofreading: molecular basis and therapeutic targeting. *Mol Cell* 79:710–727. <https://doi.org/10.1016/j.molcel.2020.07.027>.
 21. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res* 117:17–37. <https://doi.org/10.1016/j.virusres.2006.01.017>.
 22. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ. 2020. A new coronavirus associated with human respiratory disease in China. *Nature* 579:265–269. <https://doi.org/10.1038/s41586-020-2008-3>.
 23. Taiaroa G, Rawlinson D, Featherstone L, Pitt M, Caly L, Druce J, Purcell D, Harty L, Tran T, Roberts J, Scott N, Catton M, Williamson D, Coin L, Duchene S. 2020. Direct RNA sequencing and early evolution of SARS-CoV-2. *bioRxiv* <https://doi.org/10.1101/2020.03.05.976167>.
 24. Volz E, Baguelin M, Bhatia S, Boonyasiri A, Cori A, Cucunubá Z, Cuomo-Dannenburg G, Donnelly CA, Dorigatti I, FitzJohn R. 2020. Report 5: phylogenetic analysis of SARS-CoV-2. Imperial College London COVID-19 Response Team, London, United Kingdom.
 25. Jia Y, Shen G, Zhang Y, Huang K-S, Ho H-Y, Hor W-S, Yang C-H, Li C, Wang W-L. 2020. Analysis of the mutation dynamics of SARS-CoV-2 reveals the spread history and emergence of RBD mutant with lower ACE2 binding affinity. *bioRxiv* <https://doi.org/10.1101/2020.04.09.034942>.
 26. Rausch JW, Capoferri AA, Katusiime MG, Patro SC, Kearney MF. 2020. Low genetic diversity may be an Achilles heel of SARS-CoV-2. *Proc Natl Acad Sci U S A* 117:24614–24616. <https://doi.org/10.1073/pnas.2017726117>.
 27. Islam MR, Hoque MN, Rahman MS, Alam A, Akther M, Puspo JA, Akter S, Sultana M, Crandall KA, Hossain MA. 2020. Genome-wide analysis of SARS-CoV-2 virus strains circulating worldwide implicates heterogeneity. *Sci Rep* 10:14004. <https://doi.org/10.1038/s41598-020-70812-6>.
 28. Hernandez-Reiche AS, Hernandez MM, Sullivan MJ, Ciferri B, Alshammari H, Obla A, Fabre S, Kleiner G, Polanco J, Khan Z, Albuquerque B, van de Guchte A, Dutta J, Francoeur N, Melo BS, Oussenko I, Deikus G, Soto J, Sridhar SH, Wang YC, Twyman K, Kasarskis A, Altman DR, Smith M, Sebra R, Aberg J, Krammer F, Garcia-Sastre A, Luksza M, Patel G, Paniz-Mondolfi A, Gitman M, Sordillo EM, Simon V, van Bakel H. 2020. Introductions and early spread of SARS-CoV-2 in the New York City area. *Science* 369:297–301. <https://doi.org/10.1126/science.abc1917>.
 29. Alouane T, Laamarti M, Essabbar A, Hakmi M, Bouricha EM, Chemao-Elfihri MW, Kartti S, Boumajdi N, Bendani H, Laamarti R, Ghriif F, Allam L, Aaniz T, Ouadghiri M, El Hafidi N, El Jaoudi R, Benrahma H, El Attar J, Mentag R, Sbabou L, Nejari C, Amzazi S, Belyamani L, Ibrahim A. 2020. Genomic diversity and hotspot mutations in 30,983 SARS-CoV-2 genomes: moving toward a universal vaccine for the “confined virus”? *Pathogens (Basel)* 9:829. <https://doi.org/10.3390/pathogens9100829>.
 30. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X. 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581:215–220. <https://doi.org/10.1038/s41586-020-2180-5>.
 31. Hou YJ, Chiba S, Halfmann P, Ehre C, Kuroda M, Dinno KH, Leist SR, Schäfer A, Nakajima N, Takahashi K, Lee RE, Mascenik TM, Graham R, Edwards CE, Tse LV, Okuda K, Markmann AJ, Bartelt L, de Silva A, Margolis DM, Boucher RC, Randell SH, Suzuki T, Gralinski LE, Kawaoka Y, Baric RS. 2020. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* 370:1464–1468. <https://doi.org/10.1126/science.abe8499>.
 32. Volz E, Hill V, McCrone JT, Price A, Jorgensen D, O'Toole Á, Southgate J, Johnson R, Jackson B, Nascimento FF, Rey SM, Nicholls SM, Colquhoun RM, da Silva Filipe A, Shepherd J, Pascall DJ, Shah R, Jesudason N, Li K, Jarrett R, Pacchiarini N, Bull M, Geidelberg L, Siveroni I, Koshy C, Wise E, Cortes N, Lynch J, Kidd S, Mori M, Fairley DJ, Curran T, McKenna JP, Adams H, Fraser C, Golubchik T, Bonsall D, Moore C, Caddy SL, Khokhar FA, Wantoch M, Reynolds N, Warne B, Maksimovic J, Spellman K, McCluggage K, John M, Beer R, Affi S, Morgan S. 2021. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell* 184:64–75.e11. <https://doi.org/10.1016/j.cell.2020.11.020>.
 33. Long SW, Olsen RJ, Christensen PA, Bernard DW, Davis JJ, Shukla M, Nguyen M, Saavedra MO, Yerramilli P, Pruitt L, Subedi S, Kuo H-C, Hendrickson H, Eskandari G, Nguyen HAT, Long JH, Kumaraswami M, Goike J, Boutz D, Gollihar J, McLellan JS, Chou C-W, Javanmardi K, Finkelstein IJ, Musser JM. 2020. Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan area. *mBio* 11:e02707-20. <https://doi.org/10.1128/mBio.02707-20>.
 34. van Dorp L, Richard D, Tan CCS, Shaw LP, Acman M, Balloux F. 2020. Evidence for increased transmissibility from recurrent mutations in SARS-CoV-2. *Nat Commun* 11:5986. <https://doi.org/10.1038/s41467-020-19818-2>.
 35. Grubaugh ND, Hanage WP, Rasmussen AL. 2020. Making sense of mutation: what D614G means for the COVID-19 pandemic remains unclear. *Cell* 182:794–795. <https://doi.org/10.1016/j.cell.2020.06.040>.
 36. Shu Y, McCauley J. 2017. GISAID: global initiative on sharing all influenza data - from vision to reality. *Euro Surveill* 22:30494. <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494>.
 37. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, du Plessis L, Pybus OG. 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 5:1403–1407. <https://doi.org/10.1038/s41564-020-0770-5>.

38. Mercatelli D, Giorgi FM. 2020. Geographic and genomic distribution of SARS-CoV-2 mutations. *Front Microbiol* 11:1800. <https://doi.org/10.3389/fmicb.2020.01800>.
39. van Dorp L, Acman M, Richard D, Shaw LP, Ford CE, Ormond L, Owen CJ, Pang J, Tan CCS, Boshier FAT, Ortiz AT, Balloux F. 2020. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infect Genet Evol* 83:104351. <https://doi.org/10.1016/j.meegid.2020.104351>.
40. Guthrie JL, Teatero S, Zittermann S, Chen Y, Sullivan A, Rilkoff H, Joshi E, Sivaraman K, de Borja R, Sundaravadanam Y, Laszloffy M, Heisler L, Allen VG, Simpson JT, Fittipaldi N. 2020. Detection of the novel SARS-CoV-2 European lineage B.1.177 in Ontario, Canada. medRxiv <https://doi.org/10.1101/2020.11.30.20241265>.
41. 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 N, McCrone JT, Gonçalves S, Jorgensen D, Myers R, Hill V, Jackson DK, Gaythorpe K, Groves N, Sillitoe J, Kwiatkowski DP, Flaxman S, Ratman O, Bhatt S, Hopkins S, Gandy A, Rambaut A, Ferguson NM. 2021. Transmission of SARS-CoV-2 lineage B.1.1.7 in England: insights from linking epidemiological and genetic data. medRxiv <https://doi.org/10.1101/2020.12.30.20249034>.