Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2

O'Toole, Áine

2021

SOFTWARE TOOL ARTICLE

Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 [version 1; peer review: 3 approved]

Áine O'Toole [ID][1]*, Verity Hill[1]*, Oliver G. Pybus[2], Alexander Watts[3,4],
Issac I. Bogoch [ID][5,6], Kamran Khan[3-5], Jane P. Messina[7],
The COVID-19 Genomics UK (COG-UK) consortium,
Network for Genomic Surveillance in South Africa (NGS-SA),
Brazil-UK CADDE Genomic Network, Houriiyah Tegally[8], Richard R. Lessells [ID][8],
Jennifer Giandhari[8], Sureshnee Pillay [ID][8], Kefentse Arnold Tumedi [ID][9],
Gape Nyepetsi[10], Malebogo Kebabonye[11], Maitshwarelo Matsheka [ID][9],
Madisa Mine[10], Sima Tokajian [ID][12], Hamad Hassan[13], Tamara Salloum [ID][12],
Georgi Merhi [ID][12], Jad Koweyes [ID][12], Jemma L. Geoghegan[14,15], Joep de Ligt [ID][15],
Xiaoyun Ren[15], Matthew Storey[15], Nikki E. Freed[16], Chitra Pattabiraman [ID][17],
Pramada Prasad[17], Anita S. Desai[17], Ravi Vasanthapuram[17], Thomas F. Schulz [ID][18],
Lars Steinbrück [ID][18], Tanja Stadler[19], Swiss Viollier Sequencing Consortium,
Antonio Parisi[20], Angelica Bianco[20], Darío García de Viedma[21,22],
Sergio Buenestado-Serrano [ID][21], Vítor Borges [ID][23], Joana Isidro[23], Sílvia Duarte[24],
João Paulo Gomes[23], Neta S. Zuckerman[25], Michal Mandelboim[25], Orna Mor[25],
Torsten Seemann[26], Alicia Arnott[27], Jenny Draper[27], Mailie Gall [ID][27],
William Rawlinson[28], Ira Deveson [ID][29], Sanmarié Schlebusch[30], Jamie McMahon[30],
Lex Leong[31], Chuan Kok Lim[31], Maria Chironna[32], Daniela Loconsole [ID][32],
Antonin Bal[33], Laurence Josset[33], Edward Holmes[34], Kirsten St. George[35],
Erica Lasek-Nesselquist[35], Reina S. Sikkema[36], Bas Oude Munnink[36],
Marion Koopmans[36], Mia Brytting[37], V. Sudha rani [ID][38], S. Pavani[38],
Teemu Smura[39], Albert Heim[18], Satu Kurkela[40], Massab Umair[41],
Muhammad Salman[41], Barbara Bartolini[42], Martina Rueca[42], Christian Drosten[43],
Thorsten Wolff [ID][44], Olin Silander[16], Dirk Eggink[45], Chantal Reusken[45],
Harry Vennema[45], Aekyung Park[46], Christine Carrington [ID][47], Nikita Sahadeo[47],
Michael Carr[48], Gabo Gonzalez[48], SEARCH Alliance San Diego,
National Virus Reference Laboratory, SeqCOVID-Spain,
Danish Covid-19 Genome Consortium (DCGC),

Communicable Diseases Genomic Network (CDGN),
Dutch National SARS-CoV-2 surveillance program,
Division of Emerging Infectious Diseases (KDCA), Tulio de Oliveira[8],
Nuno Faria [ID][2,49], Andrew Rambaut[1], Moritz U. G. Kraemer[2]

[1]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK
[2]Department of Zoology, University of Oxford, Oxford, UK
[3]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Canada
[4]BlueDot, Toronto, Canada
[5]Department of Medicine, University of Toronto, Toronto, Canada
[6]Divisions of General Internal Medicine and Infectious Diseases, University Health Network, Toronto, Canada
[7]Department of Geography, University of Oxford, Oxford, UK
[8]KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa
[9]Botswana Institute for Technology Research and Innovation, Gaborone, Botswana
[10]National Health Laboratory, Gaborone, Botswana
[11]Ministry of Health and Wellness, Gaborone, Botswana
[12]Department of Natural Sciences, Lebanese American University, Beirut, Lebanon
[13]Faculty of Public Health, Lebanese University, Beirut, Lebanon
[14]Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand
[15]Institute of Environmental Science and Research, Wellington, New Zealand
[16]School of Natural and Computational Sciences, Massey University, Auckland, New Zealand
[17]Department of Neurovirology, National Institute of Mental Health and Neurosciences, Bengaluru, India
[18]Institute of Virology, Hannover Medical School, Hannover, Germany
[19]Department of Biosystems Science and Engineering, ETH Zürich, Zurich, Switzerland
[20]Istituto Zooprofilattico sperimentale della Puglia e della Basilicata, Puglia, Italy
[21]Hospital General Universitario Gregorio Marañón; Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain
[22]CIBER Enfermedades Respiratorias CIBERES, Madrid, Spain
[23]Bioinformatics Unit, Department of Infectious Diseases, National Institute of Health Doutor Ricardo Jorge (INSA), Lisbon, Portugal
[24]Innovation and Technology Unit, Department of Human Genetics, National Institute of Health Doutor Ricardo Jorge (INSA), Lisbon, Portugal
[25]Central Virology Laboratory, Israel Ministry of Health, Sheba Medical Center, Ramat Gan, Israel
[26]Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology & Immunology, University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Australia
[27]New South Wales Health Pathology - Institute of Clinical Pathology and Medical Research, Sydney, Australia
[28]New South Wales Health Pathology Randwick, Prince of Wales Hospital, Sydney, Australia
[29]Kinghorn Centre for Clinical Genomics, Sydney, Australia
[30]Queensland Reference Centre for Microbial and Public Health Genomics, Forensic and Scientific Services, Health Support Queensland, Queensland Health South Australia Pathology, Adelaide, Australia
[31]South Australia Pathology, Adelaide, Australia
[32]Department of Biomedical Sciences and Human Oncology, University of Bari, Bari, Italy
[33]Centre National de Référence des virus des infections respiratoires, Hospices Civils de Lyon, Lyon, France
[34]University of Sydney, Sydney, Australia
[35]Wadsworth Center, New York State Department of Health, Albany, New York, USA
[36]ErasmusMC, Department of Viroscience, WHO collaborating centre for arbovirus and viral hemorrhagic fever Reference and Research, Rotterdam, The Netherlands
[37]The Public Health Agency of Sweden, Department of Microbiology, Solna, Sweden
[38]Upgraded Department of Microbiology, Osmania Medical College, Hyderabad, Telangana, India
[39]Department of Virology, University of Helsinki, Helsinki, Finland
[40]HUS Diagnostic Center, HUSLAB, Clinical Microbiology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[41]Department of Virology, National Institute of Health, Islamabad, Pakistan
[42]National Institute for Infectious Diseases "L. Spallanzani", Rome, Italy
[43]Institute for Virology, Charité Universitätsmedizin, Berlin, Germany
[44]Robert Koch-Institut, , Head, Unit 17, Influenza and other Respiratory Viruses, Seestr. 10, Berlin, Germany
[45]WHO COVID-19 reference laboratory, Centre for Infectious Disease Control-National Institute for Public Health and the

Environment, Bilthoven, The Netherlands

[46]Division of Emerging Infectious Diseases, Bureau of Infectious Disease Diagnosis Control, Korea Disease Control and Prevention Agency, Cheongju-si, Chungcheongbuk-do, South Korea

[47]University of the West Indies, St. Augustine, Trinidad and Tobago

[48]National Virus Reference Laboratory, University College Dublin, Dublin, Ireland

[49]Imperial College London, London, UK

* Equal contributors

Abstract

Late in 2020, two genetically-distinct clusters of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) with mutations of biological concern were reported, one in the United Kingdom and one in South Africa. Using a combination of data from routine surveillance, genomic sequencing and international travel we track the international dispersal of lineages B.1.1.7 and B.1.351 (variant 501Y-V2). We account for potential biases in genomic surveillance efforts by including passenger volumes from location of where the lineage was first reported, London and South Africa respectively. Using the software tool grinch (global report investigating novel coronavirus haplotypes), we track the international spread of lineages of concern with automated daily reports, Further, we have built a custom tracking website (cov-lineages.org/global_report.html) which hosts this daily report and will continue to include novel SARS-CoV-2 lineages of concern as they are detected.

Keywords

genomic surveillance, air travel, SARS-CoV-2, genomics, genome sequencing, virus, surveillance, pandemic, B.1.1.7, B.1.351, N501Y, coronavirus, sequencing, genomic epidemiology
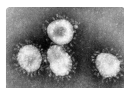
This article is included in the Coronavirus (COVID-19) collection.

Open Peer Review

Reviewer Status ✓ ✓ ✓

	Invited Reviewers		
	1	**2**	**3**
version 2 (revision) 17 Sep 2021			
version 1 19 May 2021	✓ report	✓ report	✓ report

1. **Rob Lanfear** (iD), Australian National University, Canberra, Australia

2. **Anderson F. Brito** (iD), Yale School of Public Health, New Haven, USA

3. **George Githinji** (iD), Kenya Medical Research Institute (KEMRI) - Wellcome Trust Research Programme, Kilifi, Kenya

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding authors: Áine O'Toole (aine.otoole@ed.ac.uk), Moritz U. G. Kraemer (moritz.kraemer@zoo.ox.ac.uk)

Author roles: O'Toole Á: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Hill V**: Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Pybus OG**: Conceptualization, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Watts A**: Data Curation, Resources, Visualization; **Bogoch II**: Data Curation, Resources, Visualization; **Khan K**: Data Curation, Resources, Visualization; **Messina JP**: Data Curation, Resources; **Tegally H**: Data Curation, Formal Analysis, Resources, Writing – Review & Editing; **Lessells RR**: Data Curation, Resources; **Giandhari J**: Data Curation, Formal Analysis, Resources; **Pillay S**: Data Curation, Formal Analysis, Resources; **Tumedi KA**: Data Curation, Formal Analysis, Resources; **Nyepetsi G**: Data Curation, Resources; **Kebabonye M**: Resources; **Matsheka M**: Resources, Supervision; **Mine M**: Resources, Supervision; **Tokajian S**: Data Curation, Formal Analysis, Resources; **Hassan H**

: Data Curation, Formal Analysis, Resources; **Salloum T**: Data Curation, Formal Analysis, Resources; **Merhi G**: Data Curation, Formal Analysis, Resources; **Koweyes J**: Data Curation, Formal Analysis, Resources; **Geoghegan JL**: Data Curation, Formal Analysis, Resources; **de Ligt J**: Data Curation, Formal Analysis, Resources; **Ren X**: Data Curation, Formal Analysis, Resources; **Storey M**: Data Curation, Formal Analysis, Resources; **Freed NE**: Data Curation, Formal Analysis, Resources; **Pattabiraman C**: Data Curation, Formal Analysis, Resources; **Prasad P**: Data Curation, Formal Analysis, Resources; **Desai AS**: Data Curation, Formal Analysis, Resources; **Vasanthapuram R**: Data Curation, Formal Analysis, Resources; **Schulz TF**: Data Curation, Formal Analysis, Resources; **Steinbrück L**: Data Curation, Formal Analysis, Resources; **Stadler T**: Data Curation, Formal Analysis, Resources, Supervision; **Parisi A**: Data Curation, Formal Analysis, Resources; **Bianco A**: Data Curation, Formal Analysis, Resources; **García de Viedma D**: Data Curation, Formal Analysis, Resources; **Buenestado-Serrano S**: Data Curation, Formal Analysis, Resources; **Borges V**: Data Curation, Formal Analysis, Resources; **Isidro J**: Data Curation, Formal Analysis, Resources; **Duarte S**: Data Curation, Formal Analysis, Resources; **Gomes JP**: Data Curation, Formal Analysis, Resources; **Zuckerman NS**: Data Curation, Formal Analysis, Resources; **Mandelboim M**: Data Curation, Formal Analysis, Resources; **Mor O**: Data Curation, Resources; **Seemann T**: Data Curation, Formal Analysis, Resources; **Arnott A**: Data Curation, Formal Analysis, Resources; **Draper J**: Data Curation, Formal Analysis, Resources; **Gall M**: Data Curation, Formal Analysis, Resources; **Rawlinson W**: Data Curation, Formal Analysis, Resources; **Deveson I**: Data Curation, Formal Analysis, Resources; **Schlebusch S**: Data Curation, Formal Analysis, Resources; **McMahon J**: Data Curation, Resources; **Leong L**: Data Curation, Formal Analysis, Resources; **Lim CK**: Data Curation, Formal Analysis, Resources; **Chironna M**: Data Curation, Formal Analysis, Resources; **Loconsole D**: Data Curation, Formal Analysis, Resources; **Bal A**: Data Curation, Formal Analysis, Resources; **Josset L**: Data Curation, Formal Analysis, Resources; **Holmes E**: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **St. George K**: Data Curation, Formal Analysis, Resources; **Lasek-Nesselquist E**: Data Curation, Formal Analysis, Resources; **Sikkema RS**: Data Curation, Formal Analysis, Resources; **Oude Munnink B**: Data Curation, Formal Analysis, Resources; **Koopmans M**: Data Curation, Formal Analysis, Resources; **Brytting M**: Data Curation, Formal Analysis, Resources; **Sudha rani V**: Data Curation, Resources; **Pavani S**: Data Curation, Resources; **Smura T**: Data Curation, Formal Analysis, Resources; **Heim A**: Data Curation, Formal Analysis, Resources; **Kurkela S**: Data Curation, Formal Analysis, Resources; **Umair M**: Data Curation, Formal Analysis, Resources; **Salman M**: Data Curation, Formal Analysis, Resources; **Bartolini B**: Data Curation, Formal Analysis, Resources; **Rueca M**: Data Curation, Formal Analysis, Resources; **Drosten C**: Data Curation, Formal Analysis, Resources; **Wolff T**: Data Curation, Formal Analysis, Resources; **Silander O**: Data Curation, Formal Analysis, Resources; **Eggink D**: Data Curation, Resources; **Reusken C**: Data Curation, Resources; **Vennema H**: Data Curation, Formal Analysis, Resources; **Park A**: Data Curation, Formal Analysis, Resources; **Carrington C**: Data Curation, Formal Analysis, Resources; **Sahadeo N**: Data Curation, Formal Analysis, Resources; **Carr M**: Data Curation, Formal Analysis, Resources; **Gonzalez G**: Data Curation, Formal Analysis, Resources; **de Oliveira T**: Data Curation, Formal Analysis, Resources; **Faria N**: Data Curation, Formal Analysis, Investigation, Resources; **Rambaut A**: Conceptualization, Formal Analysis, Funding Acquisition, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Kraemer MUG**: Conceptualization, Data Curation, Formal Analysis, Investigation, Project Administration, Resources, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Introduction

In December 2020, routine genomic surveillance in the United Kingdom (UK)[1] reported a new and genetically distinct phylogenetic cluster of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (variant VOC202012/01, lineage B.1.1.7). Preliminary analysis suggests that this lineage carries an unusually large number of genetic changes[2]. The earliest known cases of B.1.1.7 were sampled in southern England in late September 2020, and by December the lineage had spread to most UK regions and was growing rapidly[3]. In October 2020, a separate SARS-CoV-2 cluster (variant 501Y.V2, lineage B.1.351), which carried a different constellation of genetic changes, was detected by the Network for Genomic Surveillance in South Africa[4,5]. Both lineages carry mutations, especially in the virus spike protein, that may affect virus function, and both appear to have grown rapidly in relative frequency since their discovery. Early analyses of the spatial spread of SARS-CoV-2 highlights the potential for rapid virus dissemination through national and international travel[6,7]. Therefore continued genomic monitoring of lineages of concern is required.

Methods

To better characterise the international distribution of lineages B.1.1.7 and B.1.351 we collated SARS-CoV-2 sequences from GISAID[8,9] and assigned lineages using pangolin (v2.1.6, https://github.com/cov-lineages/pangolin), which implements the nomenclature scheme described in Rambaut *et al.*,[10]. Genomes are assigned lineage B.1.1.7 if they exhibit at least 5 of the 17 mutations inferred to have arisen on the phylogenetic branch immediately ancestral to the cluster (Table 1)[2]; or to B.1.351 if they exhibit at least 5 of 9 lineage-associated mutations (Table 1)[5]. Lineage count and frequency data have been calculated daily using grinch. Using Air Transport Association (IATA) travel data from October 2020, available through bluedot.global, we aggregated and collated the passenger volumes from international airports in London and South Africa to international destinations on same booking. Destinations with more than 5,000 passengers from London and more than 300 passengers from South Africa during the month of October are displayed on the cov-lineages.org website and in the underlying data for this publication[11]. grinch, with custom python modules that make use of geopandas v0.9, matplotlib v3.2 and seaborn v0.10, combines this information and produces reports with descriptive tables and figures that can be found at https://cov-lineages.org/global_report.html.

Implementation

All of the code underlying this daily lineage tracking web-report can be found at GitHub and Zenodo[12]. grinch is a python-based tool, the analysis pipeline of which is built on a snakemake backbone[13]. Every 24 hours a scheduled cron[14] task runs on our local servers. We download the latest data from GISAID and deduplicate based on sequence names. The sequences are assigned their most likely lineage using pangolin's latest version and model files. All processed metadata is available and maintained on the cov-lineages.org GitHub repository. To run grinch, the user must have access to a GISAID direct download key and a password and provide these within a configuration file for use. The command used to run grinch is grinch -i grinch_config.yaml, using the config file provided at doi:10.5281/zenodo.4640379[15].

Operation

Most users will not run grinch themselves, instead all information and useful descriptive figures are provided daily on the web report. Users can navigate to cov-lineages.org in a web browser of choice to view the latest daily report.

Results and discussion

As of 7th Jan 2021, 45 countries had reported the presence of B.1.1.7 and 13 countries had reported B.1.351/501Y.V2. B.1.1.7 and B.1.351 genome sequences were available for 28 and 8 countries, respectively (Figure 1a, b, c)[11]. Although some countries report increases in the relative frequency of B.1.1.7, genome sequencing efforts vary considerably. Potential targeting of sequencing towards travelers from the UK could bias frequency estimates upwards (Figure 1b, c) and differing genome sharing policies and delays may also skew reporting estimates. The time between the initial collection date of a new variant sample in a country and the first availability of a corresponding virus genome on GISAID was, on average, 12 days (range 1–71).

The number of B.1.1.7 and B.1.351/501Y.V2 genome sequences reported in each country is a consequence of (i) the intensity of local genomic surveillance; (ii) the level of concern about new variant introductions; (iii) the volume of international travel among affected countries, and (iv) the amount of local transmission following the introduction of lineage from elsewhere. To explore these factors, we analysed the most recent available International Air Transport Association (IATA) travel data (October 2020). We collated the total number of origin-to-destination air journeys between major London international airports and each country. The calculation was repeated for journeys originating in all international South African airports. We focussed on London and South Africa as they are the locations with the first reports and highest reported prevalence of lineages

Table 1. Defining mutations for lineages of interest.

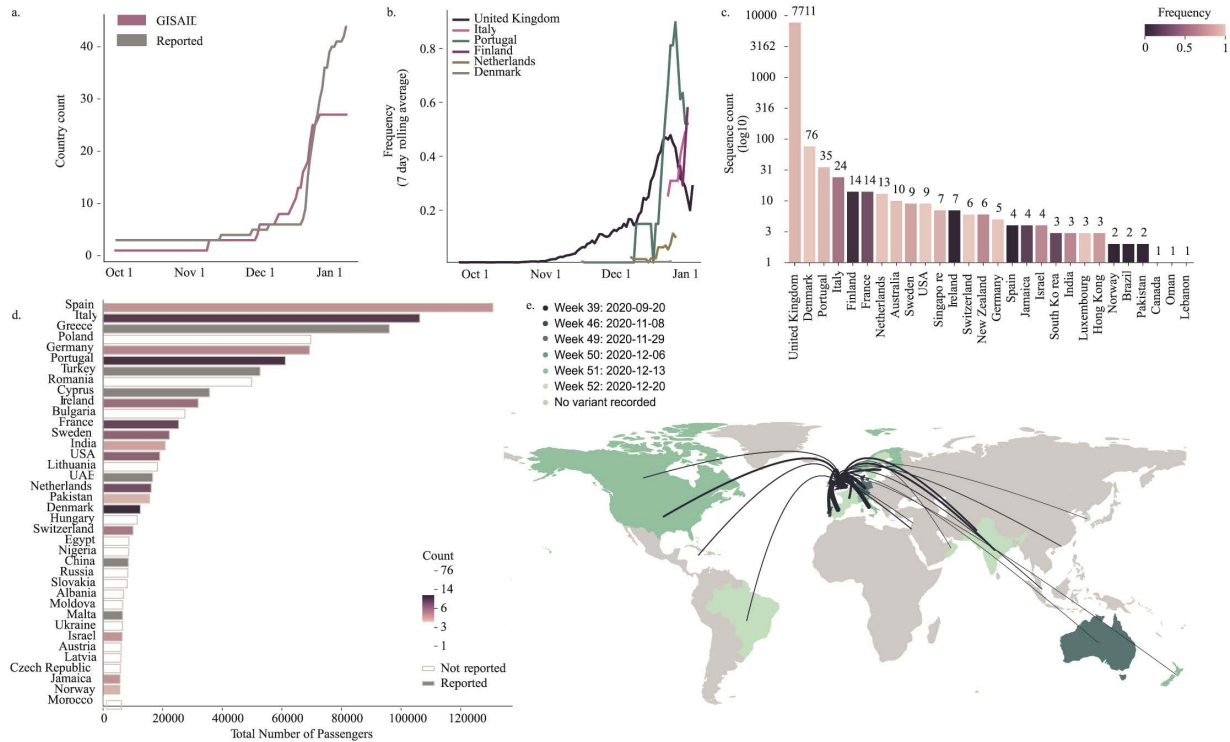| Lineage	Defining mutations
B.1.1.7	orf1ab:T1001I; orf1ab:A1708D; orf1ab:I2230T; del:11288:9; del:21765:6; del:21991:3; S:N501Y; S:A570D; S:P681H; S:T716I; S:S982A; S:D1118H; Orf8:Q27*; Orf8:R52I; Orf8:Y73C; N:D3L; N:S235F
B.1.351/501Y-V2	E:P71L; N:T205I; orf1a:K1655N; S:D80A; S:D215G; S:K417N; S:E484K; S:N501Y; S:E484K

Figure 1. a) The cumulative number of countries with reports of lineage B.1.1.7 (grey line) and cumulative number of genomes of B.1.1.7 deposited in GISAID. **b**) Rolling seven-day average of the proportion of B.1.1.7 genomes in countries with more than ten sequences of the variant, and with more than ten days between the first B.1.1.7 sequence and the most recent one compared to all sampled genomes in that country. **c**) Number of sequences (log10) per country. Colour indicates the proportion of sequences that are classified as lineage B.1.1.7. **d**) Number of air travellers from major international London airports (Heathrow, Gatwick, Luton, City, Stansted, Southend) during October 2020. Colour indicates the number of sampled genomes of lineage B.1.1.7. **e**) Map of international flights from major international London airports to countries with B.1.1.7 sequences. Colours indicate the date of earliest detection of B.1.1.7. in each country. The width of the lines indicates the number of flights. International Air Transport Association data used here account for ~90% of passenger travel itineraries on commercial flights, excluding transportation via unscheduled charter flights (the remainder is modelled using market intelligence). Data shown represents origin-destination journeys during October 2020. Routes to countries that have not yet detected B.1.1.7 and deposited data on GISAID are not included.

B.1.1.7 and B.1.351 respectively[2,5]. However, due to low SARS-CoV-2 genomic surveillance in many locations, we cannot reject the hypotheses that these lineages initially originated elsewhere. Figure 1d shows destinations receiving

>5,000 travellers in October 2020 from the UK (Figure 2 shows destinations receiving

>300 travellers from South Africa).

Of the countries that receive >5,000 travellers from London, 16 have sequenced B.1.1.7. Of the 45 countries that have identified B.1.1.7 (32 in travellers and 13 with local onward transmission), only 6 perform real-time routine genomic surveillance (Denmark, UK, Iceland, The Netherlands, Australia,

Sweden), 3 have prioritised sequencing based on S-gene target failure tests[16], 30 primarily targeted sequencing towards arriving travellers from the UK, and there was no information available for 10 (details at https://github.com/cov-lineages/lineages-website/blob/master/_data/). Of the 13 countries that have identified B.1.351 (four with local onward transmission including South Africa), 4 perform routine sequencing (South Africa, UK, Botswana, Australia), 6 target sequencing of travellers, and there was no information available for 3. Consequently, the number of sequences reported does not correlate with flight numbers, but rather reflects current genomic surveillance effort. For example, in September, the UK sequenced ~13% of its reported cases and Denmark sequenced ~21%. In comparison, Israel sequenced ~0.002% of its cases during the same period[17,18].
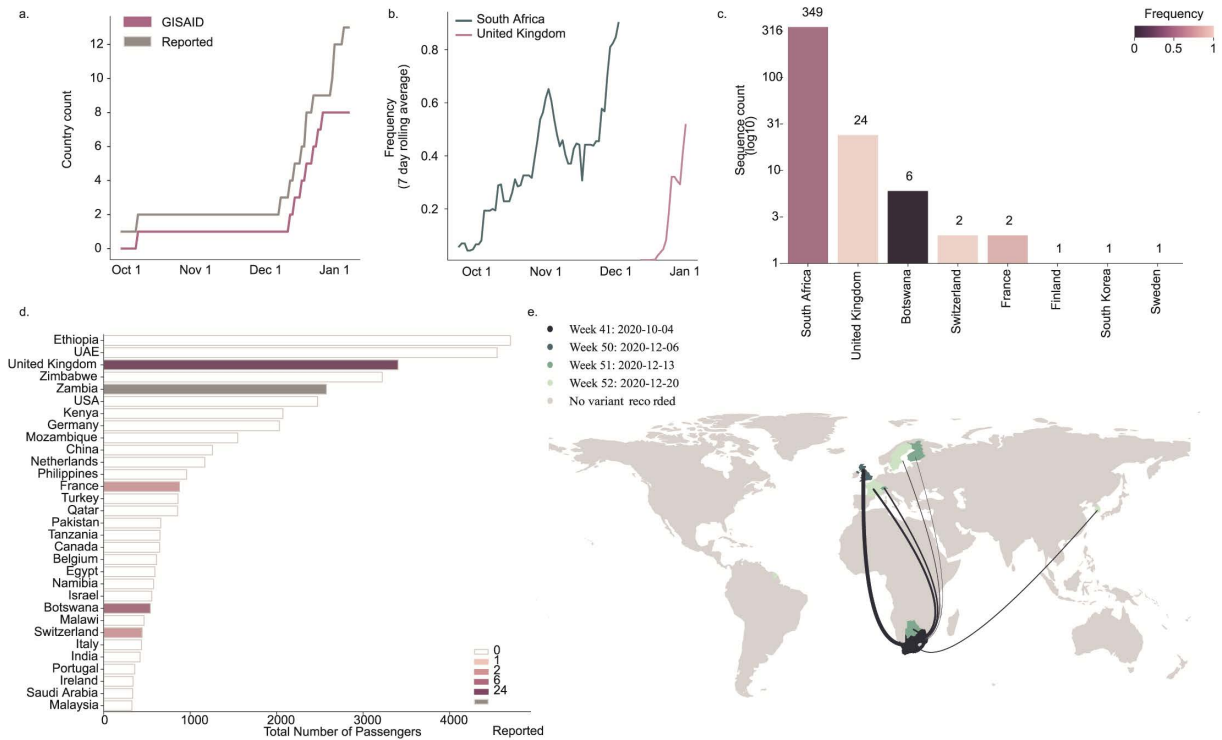
Figure 2. a) Shows the cumulative number of countries with reports of lineage B.1.351 (black line) and cumulative number of genomes of B.1.351 deposited in GISAID. **b**) Rolling seven-day average of the proportion of B.1.1.7 genomes in countries with more than ten sequences of the variant, and with more than ten days between the first B.1.1.7 sequence and the most recent one compared to all sampled genomes in that country. **c**) Number of sequences (log10) per country. Colour indicates the proportion of sequences that are classified as lineage B.1.351 **d**) Number of air travellers from South Africa during October 2020. Colour indicates the number of sampled genomes of lineage B.1.351. **e**) Map of international flights to countries with B.1.351 sequences. Colours indicate the date of earliest detection of B.1.351 in each country. The width of the lines indicates the number of flights. International Air Transport Association data used here account for ~90% of passenger travel itineraries on commercial flights, excluding transportation via unscheduled charter flights (the remainder is modelled using market intelligence). Data shown represents origin-destination journeys during October 2020. Routes to countries that have not yet detected B.1.351 and deposited data on GISAID are not included.

Our study has several limitations. The passenger flight data do not include recent changes to holiday travel, and recent restrictions on travel from the UK and South Africa is not reflected in the mobility data. Further, flight data may not accurately reflect the final destination if multiple tickets are purchased.

The discovery and rapid spread of B.1.1.7 and B.1.351/ 501Y.V2 highlights the importance of real-time and open data for tracking the spread of SARS-CoV-2 and for informing future public health interventions and travel advice.

Data availability
Underlying data
Zenodo: Accession IDs included in publication Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2. https://doi.org/10.5281/zenodo.4642401[9].

This project contains the following underlying data:
- Accession IDs of B.1.1.7 and B.1.351 genome sequences included in report up until January 7th, 2021. All accession IDs link to data on the GISAID repository, http://doi.org/10.17616/R3Q59F. These data are available under the terms of the GISAID EpiFlu™ Database Access Agreement.

Zenodo: cov-lineages.org website. https://doi.org/10.5281/zenodo.4640140[11].

This project contains the following underlying data:
- Website data archived at time of publication

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Extended data

Zenodo: Supplementary materials with group affiliations for Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2. https://doi.org/10.5281/zenodo.4704471[19].

This project contains the following extended data:
- Supplementary materials with group authorship affiliations and full acknowledgements.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Software availability

- **Software available from:** https://cov-lineages.org/global_report.html

- **Source code available from:** https://github.com/cov-lineages/grinch

- **Archived source code at time of publication:** https://doi.org/10.5281/zenodo.4640037[12]; https://doi.org/10.5281/zenodo.4640379[15]

- **Licenses:** GNU General Public License v3.0; Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Authors contributions

A.OT., O.G.P., J.P.M., N.R.F., A.R., M.U.G.K. conceived the study. A.OT, V.H., M.U.G.K. O.G.P., A.R. wrote the first draft. A.OT, V.H., M.U.G.K., A.R., AW, conducted data analysis. S.T., T.Salloum, G.M., J.K., J.G., J.d.L., X.R., M.S, N.Freed, C.P., P.P., A.D., R.V., T.F.S., L.S., T.Stadler, A.P., A.B., D.G.d.V., S.B-S., V.B., J.I., S.D., J.P.G., N.Z., M.M., O.M., T.Seemann, N.S., B.H., M.Sait, A.A., J.D., M.G., W.R., I.D., S.S., J.M., L.L., C.K.L., M.C., D.L., A.B., L.J., K.S.G., E.L-N., R.S., B.M., M.Koopmans, M.B., V.S.R., S.P. , T.Smura, A.H., S.K., M.U., M.Salman, B.B., M.R., C.D., T.W., O.S., D.E., C.R., H.V., A.P. contributed to the genomic dataset and facilitated data and sample availability. All authors interpreted the data and contributed to writing.

Acknowledgements

References

1. COVID-19 Genomics UK (COG-UK) consortiumcontact@cogconsortium.uk: **An integrated national scale SARS-CoV-2 genomic surveillance network.** *Lancet Microbe.* 2020; **1**(3): e99–100.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. 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**. 2020; published online Dec 18. (accessed Jan 8, 2021).
   Reference Source

3. 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**. *bioRxiv.* 2021.
   Publisher Full Text

4. Msomi N, Mlisana K, de Oliveira T: **A genomics network established to respond rapidly to public health threats in South Africa.** *Lancet Microbe.* 2020; **1**(6): e229–30.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Tegally H, Wilkinson E, Giovanetti M, *et al.*: **Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa**. *bioRxiv.* 2020.
   Publisher Full Text

6. du Plessis L, McCrone JT, Zarebski AE, *et al.*: **Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK.** *Science.* 2021; **371**(6530): 708–712.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Lu J, du Plessis L, Liu Z, *et al.*: **Genomic Epidemiology of SARS-CoV-2 in Guangdong Province, China.** *Cell.* 2020; **181**(5): 997–1003.e9.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Elbe S, Buckland-Merrett G: **Data, disease and diplomacy: GISAID's innovative contribution to global health.** *Glob Chall.* 2017; **1**(1): 33–46.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. O'Toole A: **Accession IDs included in publication Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 [Data set].** *Zenodo.* 2021.
   http://www.doi.org/10.5281/zenodo.4642401

10. Rambaut A, Holmes EC, O'Toole Á,, *et al.*: **A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology.** *Nat Microbiol.* 2020; **5**(11): 1403–7.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. O'Toole A: **cov-lineages.org website.** *Zenodo.* 2021.
    http://www.doi.org/10.5281/zenodo.4640140

12. O Toole A, Hill V: **grinch.** *Zenodo.* 2021.
    http://www.doi.org/10.5281/zenodo.4640037

13. Mölder F, Jablonski KP, Letcher B, *et al.*: **Sustainable data analysis with Snakemake [version 1; peer review: 1 approved, 1 approved with reservations].** *F1000Res.* 2021; **10**: 33.
    Publisher Full Text

14. Reznick L: **Using cron and crontab.** *Sys Admin.* 1993; **2**(4): 29–32.

15. O'Toole A: **grinch_config.yaml [Data set].** *Zenodo.* 2021.
    http://www.doi.org/10.5281/zenodo.4640379

16. Bal A, Destras G, Gaymard A, *et al.*: **Two-step strategy for the identification of SARS-CoV-2 variants co-occurring with spike deletion H69-V70, Lyon, France, August to December 2020**. *bioRxiv.* 2020.
    Publisher Full Text

17. Hasell J, Mathieu E, Beltekian D, *et al.*: **A cross-country database of COVID-19 testing.** *Sci Data.* 2020; **7**(1): 345.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Dong E, Du H, Gardner L: **An interactive web-based dashboard to track COVID-19 in real time.** *Lancet Infect Dis.* 2020; **20**(5): 533–4.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. O'Toole A: **Supplementary materials with group affiliations for Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2.** *Zenodo.* 2021.
    http://www.doi.org/10.5281/zenodo.4704471

Open Peer Review

Current Peer Review Status: ✓ ✓ ✓

Version 1

Reviewer Report 10 June 2021

https://doi.org/10.21956/wellcomeopenres.18372.r43967

✓ **George Githinji** (iD)

Epidemiology and Demography Department, Kenya Medical Research Institute (KEMRI) - Wellcome Trust Research Programme, Kilifi, Kenya

The article by O'Toole 2021 *et al.* describes a bioinformatics tool for the analysis of SARS-CoV-2 sequence data. The article is concise, and the relevant details have been considered. For example, the software and source code is available and well documented. The tool has shown great utility in public health based on its application in tracking and describing two SARS-CoV-2 variants of global concern.

Some minor comments below:
1. The transition from an article of public health importance to a software tool is abrupt. I think a paragraph or a link aimed at orientating the audience would be useful.

2. It would be useful to outline the special niche that the tool occupies or the gaps it fills relative to similar utilities and webpages such as covariants.org and outbreak.info.

3. The readme file at https://github.com/cov-lineages/grinch lacks full installation documentation. An introductory paragraph of the tool and its utility would also be useful. The scripts directory could be better organised by separating the snakemake files from the regular python files. I would image a workflows dir and scripts dir

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: bioinformatics, molecular epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 10 Sep 2021

Áine Niamh O'Toole, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

Thank you for your review and time, and apologies for taking so long to respond!

The transition from an article of public health importance to a software tool is abrupt. I think a paragraph or a link aimed at orientating the audience would be useful.
We have added in a paragraph that bridges the public health information and the resource information.

It would be useful to outline the special niche that the tool occupies or the gaps it fills relative to similar utilities and webpages such as covariants.org and outbreak.info.
The linking paragraph also discusses other resources such as outbreak.info and covariants.org- we now also provide a non-exhaustive list of a several other resources on the cov-lineages.org website

The readme file at https://github.com/cov-lineages/grinch lacks full installation documentation. An introductory paragraph of the tool and its utility would also be useful. The scripts directory could be better organised by separating the snakemake files from the regular python files. I would image a workflows dir and scripts dir
We have updated the readme on the grinch repository and updated usage in that it can be run in full analysis pipeline mode or in report only mode. We have also added in a brief description of the grinch setup and where the macro data is hosted.

Competing Interests: No competing interests were disclosed.

Reviewer Report 03 June 2021

https://doi.org/10.21956/wellcomeopenres.18372.r43964

✔️   **Anderson F. Brito** 🆔

Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA

Thank you for developing these tools for daily tracking of SARS-CoV-2 lineages.

Some comments about the tool and its functionalities:
- In the manuscript, the authors mention that users with access to GISAID direct download could run this pipeline locally, and generate their own reports. Can this tool be adapted to display genomic results for tracking national spread, or even state-level spread of lineages?

- If this pipeline is intended to be constantly executed locally by users, it would be helpful to provide more information about how to install and run the pipeline, including reference to example input and output files. I have tried to run the pipeline using my GISAID data provision credentials, but that was not successful, as I ran into errors for which I could not find a solution online (GitHub and Zenodo).

- About the online reports, increasing the font size in the plots being displayed (bar, curves, etc) would make labels and legends more intelligible, and improving the readability of their content.

- About the flight data, why only flight counts from October are shown? Are these data only used for tracking the potential spread in early stages of viral emergence, or do you see other uses for such data?

Concerning the manuscript, a few minor points:
- The colour gradient in the legend of Figure 1 is incomplete and does not go from 1 to 76. I think it must be just a formatting issue.

- How was the "reported" cases shown in Figures 1 and 2 detected? By differential PCR? I know that applies to B.1.1.7, but what about B.1.351?

- The legend in Figure 2 refers to "B.1.1.7" sequences, while the figure shows "B.1.351" sequences. It must be a typo.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Partly

Are sufficient details of the code, methods and analysis (if applicable) provided to allow

replication of the software development and its use by others?
Partly

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Partly

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Virology, Bioinformatics, Evolution, Epidemiology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

> Author Response 10 Sep 2021
> **Áine Niamh O'Toole**, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK
>
> - **In the manuscript, the authors mention that users with access to GISAID direct download could run this pipeline locally, and generate their own reports. Can this tool be adapted to display genomic results for tracking national spread, or even state-level spread of lineages?**
>   We've added in an option to grinch with the --analysis flag, that can either run the whole GISAID-processing pipeline, or just generate a set of reports from a metadata file. It currently would need resolution to remain at the country level, but we like the idea of doing a more local report and will work towards implementing something like this in the future.
>
> - **If this pipeline is intended to be constantly executed locally by users, it would be helpful to provide more information about how to install and run the pipeline, including reference to example input and output files. I have tried to run the pipeline using my GISAID data provision credentials, but that was not successful, as I ran into errors for which I could not find a solution online (GitHub and Zenodo).**
>   We have supplied some more information on how to install on the GitHub repository and run the pipeline, however the tool isn't necessarily intended for users to run locally themselves as we process the data on this end and share the macro count data on GitHub.
>
> - **About the online reports, increasing the font size in the plots being displayed (bar, curves, etc) would make labels and legends more intelligible, and improving the readability of their content.**

We have made the axes longer to account for more countries, however are working towards re-implementing these reports in javascript so they can be interactive and more responsive to the browser size.

- **About the flight data, why only flight counts from October are shown? Are these data only used for tracking the potential spread in early stages of viral emergence, or do you see other uses for such data?**
  October related to the date of early spread of both lineages described in the text and was due to limitations of access to data. We hope to continue to develop this resource and supply more recent dates that track over time.

Concerning the manuscript, a few minor points:
- **The colour gradient in the legend of Figure 1 is incomplete and does not go from 1 to 76. I think it must be just a formatting issue.**
  This was a proofs issue and we've hopefully rectified this now.

- **How was the "reported" cases shown in Figures 1 and 2 detected? By differential PCR? I know that applies to B.1.1.7, but what about B.1.351?**
  This was the set of manually curated media reports that we were tracking at the time.

- **The legend in Figure 2 refers to "B.1.1.7" sequences, while the figure shows "B.1.351" sequences. It must be a typo.**
  Thank you we have now fixed this in the figure.

Competing Interests: No competing interests were disclosed.

Author Response 10 Sep 2021

Áine Niamh O'Toole, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

Also- thank you for your time reviewing this manuscript (we do really appreciate it!) and apologies for taking so long to respond.

Competing Interests: No competing interests were disclosed.

Reviewer Report 27 May 2021

https://doi.org/10.21956/wellcomeopenres.18372.r43966

✔

Rob Lanfear [iD]

Department of Ecology and Evolution, Research School of Biology, Australian National University, Canberra, ACT, Australia

This article describes a software tool, grinch, that can be used to produce automated reports on SARS-CoV-2 lineages. The authors apply it to two lineages of concern in the article, and also highlight that the main utility of grinch is not in static one-off reports, but in regularly updated reports available at https://cov-lineages.org/global_report.html.

The paper clearly describes the software and demonstrates its utility. I'd like to commend the authors for putting this tool and the associated website together so quickly, for maintaining both to a very high standard, for making sure that all of the work is open and reproducible, and for the huge amount of work and enormous collaborative effort that has gone into this clear and concise report.

I have no serious reservations about the software tool or the data, analyses, or conclusions presented in the manuscript. The software is clear, open-source, sufficiently documented, and almost all of the proposed utility is presented on a clear and regularly updated website. The manuscript is clearly written, well researched, concise, and the conclusions are well justified by the analyses.

Of course, I do have a few comments, some of which I hope might be useful in improving the paper and/or the website.

Minor comments on the manuscript:

1. I felt there was some tension in this article about whether it's a software note or a public health report. The title suggests the latter, but much of the article (and the article type of "Software Tool Article") suggests the former. Most of this tension for me as a reader came from looking at the title, which has no mention of software, so I think sets up expectations that differ from what is then provided (quite reasonably) in the paper. A very simple way to address this would be to start the title with "Using grinch to track..." or to end it with "... using grinch".

2. Similar to point 1, the abstract doesn't actually mention 'grinch' or https://cov-lineages.org/global_report.html. It would seem clearer to me to incorporate in the abstract the framing that this article presents a generally applicable software tool, demonstrated on two lineages of concern.

3. I would like to see some mention of related efforts somewhere in the report. A full detailed comparison is neither warranted nor useful here because all such websites can and should change regularly, but a couple of sentences comparing cov-lineages.org to sites like outbreak.info and covariants.org would be very useful. At a minimum, it seems useful to list the similar sites the authors are aware of, if only because the fact one can see similar patterns presented on those sites serves as a useful validation of the software presented in this paper.

4. Given the situation, this is a desirable, not a requirement, but I'd love to see some unit tests on the GitHub repo. It seems potentially important to have this when the intention is to

produce daily updates for public health. (Though I note that getting the same end result from completely independent implementations on other sites is probably worth more than a lot of unit tests).

5. I struggled with Figure 1D. It wasn't clear to me what 'reported' and 'not reported' mean. And the legend makes it really hard to figure out how colours map to counts.

6. It's stated that there is no correlation between the numbers of sequences and flight numbers. It would be nice to see the scatter plot for this (maybe as an inset to figure 1D?), as well as the effect size and p-value of a suitable model.

7. Following from the previous point, the explanation for the lack of a correlation with absolute numbers seems reasonable. But it still seems to me that flight numbers could correlate with the frequency of B117 at a fixed time interval from the first detected case in a given locality (thus somewhat factoring out sequencing effort in the locality). Is it possible to add this analysis?

8. Please add installation instructions to the GitHub repo

Minor comments on https://cov-lineages.org/global_report.html:
1. Figure 3 for each lineage is a map of sequence counts by region. I find the legend here completely baffling. All it states is grey=No variant (that makes sense), pink = 1 sequence (that makes sense too), and purple = 'Max sequences'. I have no idea what to make of this. How many is 'Max', and how am I supposed to quantitatively interpret intermediate colours to pink and purple? It's so obvious I'm certain there are good reasons why this isn't already done, but it does seem like a continuous colour scale is what *should* be used here. Similar to the scale in Figure 2 (grey for no data, shades of green nicely spaced and annotated for different values of a continuous variable).

2. For the widespread lineages like B.1.1.7, there's a lot of overplotting on Figures 4 and 5, which make the counts and the country names very difficult to read. This could be addressed by just making the figures larger.

3. The table of links to news reports is absolutely wonderful. Would it be possible to include a button here to allow users to suggest additional news links? (I assume there's an existing mechanism for doing this, but I couldn't find one, so if not maybe just a link to a github issue with (potentially) a pre-filled title and required information would help?)

Really minor comments about the manuscript:
1. The first use of IATA (first para of the methods) is missing "International", i.e. it says "Using Air Transport...".

2. The second use of IATA (second para of results) does not need to be spelled out.

3. Figure 1A seems like it is missing a second Y axis for the number of GISAID genomes reported.

4. In the PDF version and the HTML version it seems that new lines were added wherever

there was a '>', e.g. '>5,000 travellers in October' and '>300 travellers from South Africa' both (erroneously?) start on new lines.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: I am a paid consultant to GISAID, the database on which much of the data analysed in this article is hosted.

Reviewer Expertise: Phylogenetics, molecular evolution, bioinformatics. I have a passing familiarity with SARS-CoV-2 data analysis.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 10 Sep 2021
Áine Niamh O'Toole, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

Thank you so much for your detailed review and our apologies for not responding sooner. All your comments were well justified and fair. We've responded to and believe have addressed your comments and concerns below.

I felt there was some tension in this article about whether it's a software note or a public health report. The title suggests the latter, but much of the article (and the article type of "Software Tool Article") suggests the former. Most of this tension for me as a reader came from looking at the title, which has no mention of software, so I think sets up expectations that differ from what is then provided (quite reasonably) in the paper. A very simple way to address this would be to start the title with "Using grinch to track..." or to end it with "... using grinch".

I completely agree, it's a very fair point! We had originally submitted the article as a public

health report but this didn't fit with the Wellcome Open Research journal remit, so resubmitted under Software Tool. I have changed the title as suggested to with with "using grinch".

Similar to point 1, the abstract doesn't actually mention 'grinch' or https://cov-lineages.org/global_report.html. It would seem clearer to me to incorporate in the abstract the framing that this article presents a generally applicable software tool, demonstrated on two lineages of concern.

Our abstract and introduction now both contain reference to grinch and the reports at cov-lineages.org

I would like to see some mention of related efforts somewhere in the report. A full detailed comparison is neither warranted nor useful here because all such websites can and should change regularly, but a couple of sentences comparing cov-lineages.org to sites like outbreak.info and covariants.org would be very useful. At a minimum, it seems useful to list the similar sites the authors are aware of, if only because the fact one can see similar patterns presented on those sites serves as a useful validation of the software presented in this paper.

We have added in a short paragraph about these resources and a link to a more extensive (but definitely non-exhaustive) list of resources.

Given the situation, this is a desirable, not a requirement, but I'd love to see some unit tests on the GitHub repo. It seems potentially important to have this when the intention is to produce daily updates for public health. (Though I note that getting the same end result from completely independent implementations on other sites is probably worth more than a lot of unit tests).

Since publication, we've re-worked the back-end analysis pipeline and the GISAID data is now processed with the datapipe pipeline written by Rachel Colquhoun (https://github.com/COG-UK/datapipe). The reports and webpages are still generated with grinch, however the main data processing steps are now done with the robust datapipe pipeline.

It's stated that there is no correlation between the numbers of sequences and flight numbers. It would be nice to see the scatter plot for this (maybe as an inset to figure 1D?), as well as the effect size and p-value of a suitable model. Following from the previous point, the explanation for the lack of a correlation with absolute numbers seems reasonable. But it still seems to me that flight numbers could correlate with the frequency of B117 at a fixed time interval from the first detected case in a given locality (thus somewhat factoring out sequencing effort in the locality). Is it possible to add this analysis?

We have amended to state there is no clear relationship, rather than correlation.

Please add installation instructions to the GitHub repo.

We have added updated usage and install instructions, and a description of the behaviour to the repository, at https://github.com/cov-lineages/grinch/blob/main/README.md

Minor comments on https://cov-lineages.org/global_report.html:

1. **Figure 3 for each lineage is a map of sequence counts by region. I find the legend here completely baffling. All it states is grey=No variant (that makes sense), pink = 1 sequence (that makes sense too), and purple = 'Max sequences'. I have no idea what to make of this. How many is 'Max', and how am I supposed to quantitatively interpret intermediate colours to pink and purple? It's so obvious I'm certain there are good reasons why this isn't already done, but it does seem like a continuous colour scale is what should be used here. Similar to the scale in Figure 2 (grey for no data, shades of green nicely spaced and annotated for different values of a continuous variable).**

   We have since amended the legend for the report.

2. **For the widespread lineages like B.1.1.7, there's a lot of overplotting on Figures 4 and 5, which make the counts and the country names very difficult to read. This could be addressed by just making the figures larger.**

   We have made these figures larger, but recognise that these plots may be better displayed by having a top 20 country barchart (like in https://cov-lineages.org/lineage.html?lineage=B.1.1.7) and intend to adopt an interactive, more attractive global report in the future.

3. **The table of links to news reports is absolutely wonderful. Would it be possible to include a button here to allow users to suggest additional news links? (I assume there's an existing mechanism for doing this, but I couldn't find one, so if not maybe just a link to a github issue with (potentially) a pre-filled title and required information would help?)**
   We really like this idea and will work towards implementing it!

Competing Interests: No competing interests were disclosed.