

Introduction and transmission of SARS-CoV-2 B.1.1.7 in Denmark

5 Thomas Y. Michaelsen¹, Marc Bennedbæk², Lasse E. Christiansen³, Mia S. F. Jørgensen⁴, Camilla H. Møller⁵, Emil A. Sørensen¹, Simon Knutsson¹, Jakob Brandt¹, Thomas B. N. Jensen¹, Clarisse Chiche-Lapierre¹, Emilio F. Collados¹, Trine Sørensen¹, Celine Petersen¹, Vang Le-Quy⁶, Mantas Sereika¹, Frederik T. Hansen¹, Morten Rasmussen⁷, Jannik Fonager⁷, Søren M. Karst⁷, Rasmus L. Marvig⁸, Marc Stegger⁹, Raphael N. Sieber⁹, Robert Skov⁵, Rebecca Legarth⁴, Tyra G. Krause⁵, Anders Fomsgaard⁷, The Danish Covid-19 Genome Consortium (DCGC)¹⁰, Mads Albertsen^{1,*}

Affiliations:

10 ¹ Department of Chemistry and Bioscience, Aalborg University; Aalborg, Denmark.

² Centre of Excellence for Health, Immunity and Infection (CHIP), Department of Infectious Diseases, Rigshospitalet, University of Copenhagen; Copenhagen, Denmark.

³ Department of Applied Mathematics and Computer Science, Technical University of Denmark; Lyngby, Denmark.

15 ⁴ Infectious Disease Epidemiology & Prevention, Statens Serum Institut; Copenhagen, Denmark.

⁵ Infectious Disease Preparedness, Statens Serum Institut; Copenhagen, Denmark.

⁶ Unit for Research Data Services (CLAAUDIA), Aalborg University; Aalborg, Denmark.

20 ⁷ Department of Virus & Microbiological Special Diagnostics, Statens Serum Institut; Copenhagen, Denmark.

⁸ Center for Genomic Medicine, Rigshospitalet; Copenhagen, Denmark

⁹ Department of Bacteria, Parasites and Fungi, Statens Serum Institut; Copenhagen, Denmark

¹⁰ <https://www.covid19genomics.dk/about>

*Corresponding author. Email: ma@bio.aau.dk.

Abstract

25 In early 2021, the SARS-CoV-2 lineage B.1.1.7 became dominant across large parts of the world. In Denmark, comprehensive and real-time test, contact-tracing, and sequencing efforts were applied to sustain epidemic control. Here, we use these data to investigate the transmissibility, 30 introduction, and onward transmission of B.1.1.7 in Denmark. In a period with stable restrictions, we estimated an increased B.1.1.7 transmissibility of 58% (95% CI: [56%,60%]) relative to other lineages. Epidemiological and phylogenetic analyses revealed that 37% of B.1.1.7 cases were related to the initial introduction in November 2020. Continuous introductions contributed 35 substantially to case numbers, highlighting the benefit of balanced travel restrictions and self-isolation procedures coupled with comprehensive surveillance efforts, to sustain epidemic control in the face of emerging variants.

A year into the SARS-CoV-2 pandemic it became clear that variants with increased transmissibility and/or reduced vaccine efficacy would have a profound impact on our lives, including prolongation of the pandemic as well as increasing morbidity and mortality. On 14 December 2020, Denmark received a notification from the Early Warning and Response System of the European Union (EWRS) about a novel SARS-CoV-2 lineage (B.1.1.7) spreading fast throughout the United Kingdom (UK). The notification was a direct response to genomic sequences uploaded to the GISAID database by the Danish Covid-19 Genome Consortium (DCGC). On 18 December 2020, the New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG), advisers to the British Health Authorities, released a public statement, reporting a 71% (95% CI: [67%,75%]) increased transmissibility for the B.1.1.7 lineage and elevated it to variant-of-concern (VOC) (1). The B.1.1.7 lineage appeared in the UK in September 2020 and has since spread across the world (2), including Denmark where the first identified case was sampled on 14 November 2020 (3). Several studies have since estimated an increased transmissibility of B.1.1.7, ranging from 29% to 90% (4–7). This is in line with early estimates of 36%–55% from Denmark (8, 9). Hypotheses for increased transmissibility include higher viral load, longer infection period, and better receptor binding (10–12).

The B.1.1.7 lineage contains several genetic changes compared to the specimens first identified in Wuhan, China. Most notably is the spike protein amino acid substitution N501Y which enhances binding to the ACE2 receptor and is thought to increase transmissibility (12). Acquired mutations can be used to construct the phylogenetic relationship between genomes and by coupling with temporal information, transmission timing and dynamics of emerging lineages can be inferred (13). These approaches have been used by several studies to gain insights into the SARS-CoV-2 pandemic on a national level (6, 7, 14–17).

Due to the early warning from the UK and an already extensive sequencing effort in Denmark, it was possible to predict growth dynamics for B.1.1.7 in Denmark, when it was still less than 2% of all circulating SARS-CoV-2 lineages. Hence, on 2 January 2021, B.1.1.7 was predicted to exceed 50% relative abundance nationwide in the middle of February (3, 18), which formed a key argument in the political decision to further extend the lockdown at the time and intensify contact tracing on all B.1.1.7 cases (19). In this study we substantiate early results by providing an updated estimate of increased transmissibility of B.1.1.7. Furthermore, using the detailed and comprehensive nationwide data, we outline the dynamics of B.1.1.7 in Denmark to evaluate the impact of introductions and regional restrictions on the prevalence and spread of B.1.1.7.

We modelled the growth rate of B.1.1.7 relative to other lineages in the period from 4 January to 7 February 2021. The study period was chosen as both the PCR testing and sequencing capacity increased substantially until January, reaching a weekly rate of >13,000 tests pr. 100,000 people and >75% of all positive samples having a high-quality genome (fig. 1A, fig. S1). Furthermore, a window with stable restrictions was enforced from January 4th until partial reopening of primary schools from February 8th. A Poisson regression model was fitted on daily counts of B.1.1.7 to estimate increased transmissibility of B.1.1.7 relative to all other lineages. We found a growth rate of 0.021 and -0.062 per day for B.1.1.7 and other lineages respectively, corresponding to an increased transmissibility of 58% (95% CI: [56%,60%]) for B.1.1.7 assuming a generation time of 5.5 days (10, 20). We did not observe significant differences in transmissibility between regions ($\chi^2_{(4)} = 3.17, p = 0.54$).

After the study period ended, case numbers of B.1.1.7 in South Denmark and Capital increased to higher levels than what was predicted by our model (fig. S2A), likely caused by large superspreading events, evident from the rapid increase of unique local haplotypes in South

85 Denmark (H17 and H87) and Capital (H43) regions (fig. S2B). All regions crossed 50% relative
abundance of B.1.1.7 with a few days difference (fig. 1B), except North Denmark, with 1-2 weeks
delay and several incidences of B.1.1.7 decline (fig. 1B).

90 Introductions into Denmark and onward transmission between regions were investigated
using a time-scaled phylogenetic tree of 1976 Danish B.1.1.7 genomes and 3611 representative
international sequences from outside Denmark (fig 2A). We used pastML (21) to perform an
ancestral state reconstruction, to infer if sequences were introduced from outside Denmark or
transmitted from another Danish region. We define an introduction lineage as a phylogenetic group
of Danish sequences that share the same introduction event, similar to the definition of UK
95 transmission lineages in (15). We do not infer the timing of introduction events from the phylogeny
(7, 15, 22), but instead use the first occurrence of an introduction lineage as a proxy justified by
the intense testing and high sequencing rate (fig. S1). An introduction lineage can be further
subdivided into transmission clusters from the ancestral state reconstruction, which is based on the
regional and temporal sub-distributions of an introduction lineage. The inferred connections
between transmission clusters outlines the direction of transmission between regions. The first
100 cases of B.1.1.7 observed in the Capital region of Denmark on 14 November 2020 defines the
introduction lineage CA1, which contained the majority of Danish sequences (736 of 1976, 37%)
across all Danish regions (fig. 2B, fig. S3). We tried to resolve the origin of CA1, but detailed
epidemiological investigations revealed no direct or indirect travel link and the ancestral state
could not be resolved with good confidence. The CA1 introduction lineage is associated with the
105 first reported cases of B.1.1.7 in all Danish regions, including the first North Denmark transmission
cluster CA1-135. This transmission cluster was part of a large but local outbreak verified by
epidemiological investigations and primarily associated with a single local superspreading event
(haplotype H77, fig. S2B). CA1-135 dominated the North Denmark cases until it was last observed
in week 2 (fig. 2B). This contrasts the course of the first transmission clusters in other regions,
110 which showed similar initial growth (CA1-108 in South Denmark and CA1-124 in Zealand) but
remained at high prevalence throughout the study period.

115 Interestingly, all transmission clusters appearing up until week 1 in North Denmark were
not observed in the remaining study period, which were not true for other regions where a constant
level of sustained transmission was present for many larger transmission clusters (fig. 3A). In
North Denmark, 7 of 11 municipalities had elevated restrictive measures until 16 November 2020
due to circulation of mink-related lineages (23). The intensified contact-tracing and genomic
surveillance in that period likely kept the reproductive numbers for North Denmark below that of
other Danish regions, making it difficult for any cluster to develop from local to sustained
transmission. B.1.1.7 was only designated as VOC from late December 2020 onwards (1) and
120 authorities were not able to do sufficient genome-based contact tracing due to low number of
samples sequenced (around 25% of positive cases) and slow turnaround (median 11 days) at that
time. Therefore, the dynamics in North Denmark is likely not attributed to focused contact tracing
of B.1.1.7. This contrasts the current situation in June 2021, with substantially improved
sequencing capacity (>90% of positive cases and median 5 days turnaround) and rapid mutation
125 of concern screening to initiate contact tracing (24, 25).

130 Overall, we identified 103 introductions and 107 transmission events between regions of
B.1.1.7 in the study period (fig. 3). The majority of introductions occurred during the first half of
January 2021 (fig. 3B); 37 introductions showed no onward transmission, 36 introductions had
between 2-5 cases of whom 8 exclusively had transmission within the same household, and only
5 caused >100 cases. To investigate the impact of introductions across time we considered early
cases of the same introduction lineage to be introduction-related, while later cases occur due to
sustained transmission in society. To evaluate this, we incremented the number of follow-up days

for cases to be assigned as introduction-related (fig. 3C). Overall, the relative contribution of introduction-related cases peaked during onset of B.1.1.7 and again in mid-January (fig. 3C).
135 Introductions alone (follow-up =0 days in fig. 3C) accounted for around 5%-10% of all cases with substantial temporal variation. This is comparable to a US study on data from the first half of 2020, where travel restrictions also applied (22). Using 14 days follow-up since the first occurrence of an introduction to define introduction-related cases within the same introduction lineage, there were periods where >50% of cases were due to importation from outside Denmark (fig. 3C) despite
140 travel-restrictions and self-isolation recommendations being in place.

In collaboration with the Danish Patient Safety Authority which administers the contact tracing efforts in Denmark, we performed detailed epidemiological investigations of registry data to identify direct or indirect travel links for all 1976 cases in this study. A total of 62 persons infected with B.1.1.7 (3%) had recorded travel history to 23 different destinations. The most
145 frequent destinations were the United Arab Emirates (n = 21), United Kingdom (n = 5), and Pakistan (n = 4). We were able to associate 32 (31%) introductions directly to travelling (fig. 3A, fig. S5). We performed a separate ancestral state reconstruction with country-level resolution on international sequences, allowing us to compare the phylogenetically inferred origin to the travel history. Of 33 introductions with travel history, 9 (27%) matched by country and 12 (36%) by
150 high-level geographical region. This highlights the limitations of available genomic data, which is still heavily biased both geographically and temporally (26), making high-resolution inference of introduction origin from phylogenetic data alone very difficult. From the ancestral state reconstruction, we inferred the temporal dynamics of B.1.1.7 dispersion between the five major administrative regions of Denmark. The majority of transmission was local and occurred within
155 regions, but would to a large extent cease if not fueled by introductions from other regions or outside Denmark (fig. 3D). The Capital and South Denmark regions with international borders and substantial international travel-activity also received the most introductions, possibly a combined effect of multiple factors such as leisure travel and commuting for work. Interestingly, the Capital region also had the highest number of outgoing transmissions to other regions, acting as an
160 important source for sustaining B.1.1.7 in Denmark, particularly in the Zealand region which has a large number of residents commuting to the Capital region on a regular basis. In many countries the takeover of B.1.1.7 has prolonged or induced further restrictions to sustain epidemic control. In Denmark, genomic surveillance has been key to rapidly predict growth of B.1.1.7, leading to enforced restrictions in effect from 5 January 2021 when B.1.1.7 was still below 2% relative
165 abundance of all SARS-CoV-2 variants (27). The enforced restrictions practically eliminated all other circulating variants within 2 months (fig. 1B) and enabled epidemic control to be maintained as society was gradually reopened from 8 February 2021. We validated the initially predicted increased transmissibility of B.1.1.7 and found that superspreading events, identified through sequencing data, could explain deviations from the model (fig. S2). At the time when B.1.1.7 was
170 recognized as VOC, the initial CA1 introduction lineage had expanded considerably across all regions, which substantially accelerated the initial phase of the B.1.1.7 pandemic in Denmark. However, ongoing introductions and regional transmissions continuously fueled and further accelerated growth of this lineage. Since 22 December 2020, authorities enforced elevated restrictions for travelers from the UK to halt B.1.1.7 expansion (19), similar to the restrictions
175 enforced for inbound travelers from Italy and Austria in the spring of 2020 to stop SARS-CoV-2 from spreading to Denmark (28). Despite these restrictions, our study shows a substantial impact of introductions, suggesting that country-specific restrictions and non-enforced recommendations of self-isolation might be insufficient. Finally, VOC designation of B.1.1.7 and upscaling of capacity in Denmark was too late to hinder the takeover of this lineage, which had to be managed
180 by broad societal restrictions. However, real-time testing, genomic tracking and contact-tracing is

now at a level in Denmark which enables control of emerging VOCs like B.1.617.2 (3) and allow policy makers to make informed decisions of opening the society while maintaining control of the epidemic in the face of emerging variants.

185

References and Notes

1. New and Emerging Respiratory Virus Threats Advisory Group, “NERVTAG meeting on SARS-CoV-2 variant under investigation VUI-202012/01” (2020).
2. European Centre for Disease Prevention and Control, “Risk related to the spread of new SARS-CoV-2 variants of concern in the EU/EEA – first update” (ECDC, 2021).
- 190 3. Statens Serum Institut, SARS-CoV-2-virusvarianter. *SSI’s portal med information om covid-19* (2021), (available at <https://covid19.ssi.dk/virusvarianter>).
4. N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. Munday, C. A. B. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne, Others, Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. *MedRxiv*, 2020–2012 (2021).
- 195 5. H. S. Vohringer, T. Sanderson, M. Sinnott, N. De Maio, T. Nguyen, R. Goater, F. Schwach, I. Harrison, J. Hellewell, C. Ariani, S. Goncalves, D. Jackson, I. Johnston, A. W. Jung, C. Saint, J. Sillitoe, M. Suci, N. Goldman, E. Birney, S. Funk, E. Volz, D. Kwiatkowski, M. Chand, I. Martincorena, J. C. Barrett, M. Gerstung, The Wellcome Sanger Institute Covid-19 Surveillance Team, The COVID-19 Genomics UK (COG-UK) Consortium, Genomic reconstruction of the SARS-CoV-2 epidemic across England from September 2020 to May 200 2021, , doi:10.1101/2021.05.22.21257633.
6. T. Alpert, A. F. Brito, E. Lasek-Nesselquist, J. Rothman, A. L. Valesano, M. J. MacKay, M. E. Petrone, M. I. Breban, A. E. Watkins, C. B. F. Vogels, C. C. Kalinich, S. Dellicour, A. Russell, J. P. Kelly, M. Shudt, J. Plitnick, E. Schneider, W. J. Fitzsimmons, G. Khullar, J. Metti, J. T. Dudley, M. Nash, N. Beaubier, J. Wang, C. Liu, P. Hui, A. Muyombwe, R. Downing, J. Razeq, S. M. Bart, A. Grills, S. M. Morrison, S. Murphy, C. Neal, E. Laszlo, H. Rennert, M. Cushing, L. Westblade, P. Velu, A. Craney, L. Cong, D. R. Peaper, M. L. Landry, P. W. Cook, J. R. Fauver, C. E. Mason, A. S. Luring, K. St George, D. R. MacCannell, N. D. Grubaugh, Early introductions and transmission of SARS-CoV-2 variant 210 B.1.1.7 in the United States. *Cell* (2021), doi:10.1016/j.cell.2021.03.061.
7. N. L. Washington, K. Gangavarapu, M. Zeller, A. Bolze, E. T. Cirulli, K. M. Schiabor Barrett, B. B. Larsen, C. Anderson, S. White, T. Cassens, S. Jacobs, G. Levan, J. Nguyen, J. M. Ramirez 3rd, C. Rivera-Garcia, E. Sandoval, X. Wang, D. Wong, E. Spencer, R. Robles-Sikisaka, E. Kurzban, L. D. Hughes, X. Deng, C. Wang, V. Servellita, H. Valentine, P. De Hoff, P. Seaver, S. Sathe, K. Gietzen, B. Sickler, J. Antico, K. Hoon, J. Liu, A. Harding, O. Bakhtar, T. Basler, B. Austin, D. MacCannell, M. Isaksson, P. G. Febbo, D. Becker, M. Laurent, E. McDonald, G. W. Yeo, R. Knight, L. C. Laurent, E. de Feo, M. Worobey, C. Y. 215

- 220 Chiu, M. A. Suchard, J. T. Lu, W. Lee, K. G. Andersen, Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell* (2021), doi:10.1016/j.cell.2021.03.052.
8. Ekspertgruppen for matematisk modellering, “Prognoser for smittetal med fokus på udviklingen i cluster B.1.1.7” (Statens Serum Institut, 2021).
9. Ekspertgruppen for matematisk modellering, “Notat om prognoser for smittetal og indlæggelser ved scenarier for genåbning af 0.-4. klasse i grundskolen” (2021).
- 225 10. N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. D. Munday, C. A. B. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne, T. Wenseleers, A. Gimma, W. Waites, K. L. M. Wong, K. van Zandvoort, J. D. Silverman, CMMID COVID-19 Working Group, COVID-19 Genomics UK (COG-UK) Consortium, K. Diaz-Ordaz, R. Keogh, R. M. Eggo, S. Funk, M. Jit, K. E. Atkins, W. J. Edmunds, Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. **372** (2021),
230 doi:10.1126/science.abg3055.
11. P. Calistri, L. Amato, I. Puglia, F. Cito, A. Di Giuseppe, M. L. Danzetta, D. Morelli, M. Di Domenico, M. Caporale, S. Scialabba, O. Portanti, V. Curini, F. Perletta, C. Cammà, M. Ancora, G. Savini, G. Migliorati, N. D’Alterio, A. Lorusso, Infection sustained by lineage
235 B.1.1.7 of SARS-CoV-2 is characterised by longer persistence and higher viral RNA loads in nasopharyngeal swabs. *Int. J. Infect. Dis.* **105**, 753–755 (2021).
12. R. Andrew, L. Nick, P. Oliver, B. Wendy, B. Jeff, C. Alesandro, C. Tom, P. Tom, L. R. David, V. Erik, Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *virological.org* (2020), (available at
240 <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563>).
13. B. Morel, P. Barbera, L. Czech, B. Bettisworth, L. Hübner, S. Lutteropp, D. Serdari, E.-G. Kostaki, I. Mamais, A. M. Kozlov, P. Pavlidis, D. Paraskevis, A. Stamatakis, Phylogenetic analysis of SARS-CoV-2 data is difficult. *Mol. Biol. Evol.* (2020),
245 doi:10.1093/molbev/msaa314.
14. S. Dellicour, K. Durkin, S. L. Hong, B. Vanmechelen, J. Martí-Carreras, M. S. Gill, C. Meex, S. Bontems, E. André, M. Gilbert, C. Walker, N. De Maio, N. R. Faria, J. Hadfield, M.-P. Hayette, V. Bours, T. Wawina-Bokalanga, M. Artesi, G. Baele, P. Maes, A
250 Phylodynamic Workflow to Rapidly Gain Insights into the Dispersal History and Dynamics of SARS-CoV-2 Lineages. *Mol. Biol. Evol.* (2020), doi:10.1093/molbev/msaa284.
15. L. du Plessis, J. T. McCrone, A. E. Zarebski, V. Hill, C. Ruis, B. Gutierrez, J. Raghwani, J. Ashworth, R. Colquhoun, T. R. Connor, N. R. Faria, B. Jackson, N. J. Loman, Á. O’Toole, S. M. Nicholls, K. V. Parag, E. Scher, T. I. Vasylyeva, E. M. Volz, A. Watts, I. I. Bogoch, K. Khan, COVID-19 Genomics UK (COG-UK) Consortium, D. M. Aanensen, M. U. G.
255 Kraemer, A. Rambaut, O. G. Pybus, Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science*. **371**, 708–712 (2021).
16. H. Tegally, E. Wilkinson, R. J. Lessells, J. Giandhari, S. Pillay, N. Msomi, K. Mlisana, J. N. Bhiman, A. von Gottberg, S. Walaza, V. Fonseca, M. Allam, A. Ismail, A. J. Glass, S.

- 260 Engelbrecht, G. Van Zyl, W. Preiser, C. Williamson, F. Petruccione, A. Sigal, I. Gazy, D.
Hardie, N.-Y. Hsiao, D. Martin, D. York, D. Goedhals, E. J. San, M. Giovanetti, J.
Lourenço, L. C. J. Alcantara, T. de Oliveira, Sixteen novel lineages of SARS-CoV-2 in
South Africa. *Nat. Med.* **27**, 440–446 (2021).
- 265 17. A. Popa, J.-W. Genger, M. D. Nicholson, T. Penz, D. Schmid, S. W. Aberle, B. Agerer, A.
Lercher, L. Endler, H. Colaço, M. Smyth, M. Schuster, M. L. Grau, F. Martínez-Jiménez,
O. Pich, W. Borena, E. Pawelka, Z. Keszei, M. Senekowitsch, J. Laine, J. H. Aberle, M.
Redlberger-Fritz, M. Karolyi, A. Zoufaly, S. Maritschnik, M. Borkovec, P. Hufnagl, M.
Nairz, G. Weiss, M. T. Wolfinger, D. von Laer, G. Superti-Furga, N. Lopez-Bigas, E.
270 Puchhammer-Stöckl, F. Allerberger, F. Michor, C. Bock, A. Bergthaler, Genomic
epidemiology of superspreading events in Austria reveals mutational dynamics and
transmission properties of SARS-CoV-2. *Sci. Transl. Med.* **12** (2020),
doi:10.1126/scitranslmed.abe2555.
18. Ekspertgruppen for matematisk modellering, “Scenarier for udviklingen i den engelske
virusvariant af SARS-COV-2 (cluster B.1.1.7)” (Statens Serum Institut, 2021).
- 275 19. Nationalt Kommunikationscenter – COVID-19 (NKC), “Skærpelse af rejsevejledninger og
indrejserestriktioner” (2021).
20. S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker,
H. Zhu, T. Berah, J. W. Eaton, M. Monod, Imperial College COVID-19 Response Team, A.
C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, S.
Bhatt, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe.
280 *Nature*. **584**, 257–261 (2020).
21. S. A. Ishikawa, A. Zhukova, W. Iwasaki, O. Gascuel, A Fast Likelihood Method to
Reconstruct and Visualize Ancestral Scenarios. *Mol. Biol. Evol.* **36**, 2069–2085 (2019).
- 285 22. N. F. Müller, C. Wagner, C. D. Frazar, P. Roychoudhury, J. Lee, L. H. Moncla, B. Pelle, M.
Richardson, E. Ryke, H. Xie, L. Shrestha, A. Addetia, V. M. Rachleff, N. A. P. Lieberman,
M.-L. Huang, R. Gautom, G. Melly, B. Hiatt, P. Dykema, A. Adler, E. Brandstetter, P. D.
Han, K. Fay, M. Ilcisin, K. Lacombe, T. R. Sibley, M. Truong, C. R. Wolf, M. Boeckh, J.
A. Englund, M. Famulare, B. R. Lutz, M. J. Rieder, M. Thompson, J. S. Duchin, L. M.
Starita, H. Y. Chu, J. Shendure, K. R. Jerome, S. Lindquist, A. L. Greninger, D. A.
290 Nickerson, T. Bedford, Viral genomes reveal patterns of the SARS-CoV-2 outbreak in
Washington State. *Sci. Transl. Med.* (2021), doi:10.1126/scitranslmed.abf0202.
23. Statsministeriet, Pressemøde den 5. november 2020. *stm.dk* (2020), (available at
<https://www.stm.dk/presse/pressemoedearkiv/pressemoede-den-5-november-2020/>).
24. T. S. Jørgensen, K. Blin, F. Kuntke, H. K. Salling, T. Y. Michaelsen, M. Albertsen, H.
Larsen, A rapid, cost efficient and simple method to identify current SARS-CoV-2 variants

- 295 of concern by Sanger sequencing part of the spike protein gene. *medRxiv* (2021), ,
doi:10.1101/2021.03.27.21252266.
25. Statens Serum Institut, Delta-PCR-testen. *SSI's portal med information om covid-19* (19
March, 2021), (available at <https://covid19.ssi.dk/diagnostik/delta-pcr-testen>).
26. GISAID initiative, Submission Tracker Global. *gisaid.org* (2021), (available at
300 <https://www.gisaid.org/index.php?id=208>).
27. Nationalt Kommunikationscenter – COVID-19 (NKC), “Skærpet forsamlingsforbud og nye
tiltag for at imødegå cluster B.1.1.7” (2021).
28. Statsministeriet, Pressemøde d. 11 marts 2020. *stm.dk* (2020), (available at
305 <https://www.stm.dk/presse/pressemoedearkiv/pressemoede-om-covid-19-den-11-marts-2020/>).
29. Statens Serum Institut, Download fil med overvågningsdata. *SSI's portal med information
om covid-19* (2021), (available at [https://covid19.ssi.dk/overvagningsdata/download-fil-
med-overvaagningdata](https://covid19.ssi.dk/overvagningsdata/download-fil-med-overvaagningdata)).
30. F. P. Lyngse, K. Mølbak, R. L. Skov, L. E. Christiansen, L. H. Mortensen, M. Albertsen, C.
310 H. Møller, T. G. Krause, M. Rasmussen, T. Y. Michaelsen, M. Voldstedlund, J. Fonager, N.
Steenhard, C. Kirkeby, The Danish Covid-19 Genome Consortium, Increased
Transmissibility of SARS-CoV-2 Lineage B.1.1.7 by Age and Viral Load: Evidence from
Danish Households, , doi:10.1101/2021.04.16.21255459.
31. J. Quick, nCov-2019 sequencing protocol V.1. *protocols.io* (2020), (available at
315 [https://www.protocols.io/view/ncov-2019-sequencing-protocol-
bbmuik6w?version_warning=no](https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w?version_warning=no)).
32. E. A. Sørensen, S. M. Karst, S. Knutsson, AAU-nCoV-2019 Tailed Long Amplicon
Sequencing. *protocols.io* (28 April, 2020), doi:10.17504/protocols.io.bfc3jiyn.
33. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads.
320 *EMBnet.journal*. **17**, 10–12 (2011).
34. N. Loman, W. Rowe, A. Rambaut, nCoV-2019 novel coronavirus bioinformatics protocol.
ARTIC network (2020), (available at [https://artic.network/ncov-2019/ncov2019-
bioinformatics-sop.html](https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html)).
35. M. E. Brooks, K. Kristensen, K. J. Van Benthem, A. Magnusson, C. W. Berg, A. Nielsen,
325 H. J. Skaug, M. Machler, B. M. Bolker, glmmTMB balances speed and flexibility among
packages for zero-inflated generalized linear mixed modeling. *R J.* **9**, 378–400 (2017).
36. Á. O’toole, E. Scher, A. Underwood, B. Jackson, V. Hill, J. Mccrone, C. Ruis, K. Abu-
Dahab, B. Taylor, C. Yeats, L. du Plessis, D. Aanensen, E. Holmes, O. Pybus, A. Rambaut,

- 330 pangolin: lineage assignment in an emerging pandemic as an epidemiological tool. *github*
(2021), (available at <https://github.com/cov-lineages/pangolin>).
37. A. Rambaut, E. C. Holmes, Á. O’Toole, V. Hill, J. T. McCrone, C. Ruis, L. du Plessis, O. G. Pybus, A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* **5**, 1403–1407 (2020).
38. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
335
39. S. Tavaré, Others, Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on mathematics in the life sciences.* **17**, 57–86 (1986).
40. Z. Yang, A space-time process model for the evolution of DNA sequences. *Genetics.* **139**, 993–1005 (1995).
41. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermini, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods.* **14**, 587–589 (2017).
340
42. A. Rambaut, T. T. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007
345 (2016).
43. B. Q. Minh, M. A. T. Nguyen, A. von Haeseler, Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195 (2013).
44. T.-H. To, M. Jung, S. Lycett, O. Gascuel, Fast Dating Using Least-Squares Criteria and Algorithms. *Syst. Biol.* **65**, 82–97 (2016).
45. A. Zhukova, L. Blassel, F. Lemoine, M. Morel, J. Voznica, O. Gascuel, Origin, evolution and global spread of SARS-CoV-2. *C. R. Biol.* (2020), doi:10.5802/crbiol.29.
350
46. J. Felsenstein, Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376 (1981).

355 **Acknowledgments**

We gratefully acknowledge the Danish hospitals for local sequencing and data sharing. Thanks to all data contributors from outside Denmark, including the authors and their originating laboratories responsible for obtaining the specimens, as well as their submitting laboratories for generating the genetic sequence and metadata and sharing it via the GISAID Initiative. We would like to
360 acknowledge the people at Data Integration and Analysis (DIAS), Statens Serum Institute, for managing and maintaining the national MiBa database which much of this work is based on. Thanks to the people at the Danish Patient Safety Authority for providing us with detailed travel records of SARS-CoV-2 positive cases. Thanks to Anna Zhukova, Institut Pasteur for modifying the pastML software to better fit our needs upon request. Finally, a sincere thanks to Edyth Parker,
365 Scripps Research, for insightful phylogenetic discussions.

Funding

VILLUM FONDEN grant (15510) (MA), Poul Due Jensen Foundation (Corona Danica) (MA), Styrelsen for Forskning og Uddannelse (0238-00002B) (MA)

Author contributions

370 TYM analyzed all phylogenetic and epidemiological data and prepared all figures. TYM and MA wrote the paper. TYM, MB, MR, JF, and RLM designed the phylogenetic analysis. LEC and CHM designed the Poisson regression models. MSFJ, RL, and TYM performed epidemiological investigations. EAS, SK, JB, SMK, TBNJ, CCL, EFC, TS, CP, VL, MR, JF, MSe, FTH, MSt, RNS, TYM and MA generated the sequencing data. MA, RS, TGK, AF oversaw the study. All
375 authors reviewed and edited the manuscript.

Competing interests

All authors declare no competing interests.

Data and materials availability

This study was conducted on administrative register data. According to Danish law, informed
380 consent is not needed for such research as the publication only contains aggregated results and no personal data is shown. The epidemiological data used in this work is person-sensitive and we are prohibited to make it publicly available according to Danish legislation. SARS-CoV-2 consensus genome sequences associated with this work have been uploaded to the GISAID database in accordance with Danish law (no. 285 vers. 2021-02-27). As consequence, dates are binned by
385 week and maximum spatial resolution is at regional level. Accession numbers are available in the Supplementary Materials (data file S1). All code used to run the phylogenetic analysis, Poisson regression models, and generate the visualizations used in this work is available at <https://github.com/TYMichaelson/B117-DK-introduction>.

Supplementary Materials

390 Materials and Methods
List of Danish Covid-19 Genome Consortium members
Figs. S1 to S4
Data file S1
References (12, 21, 29-46)

395

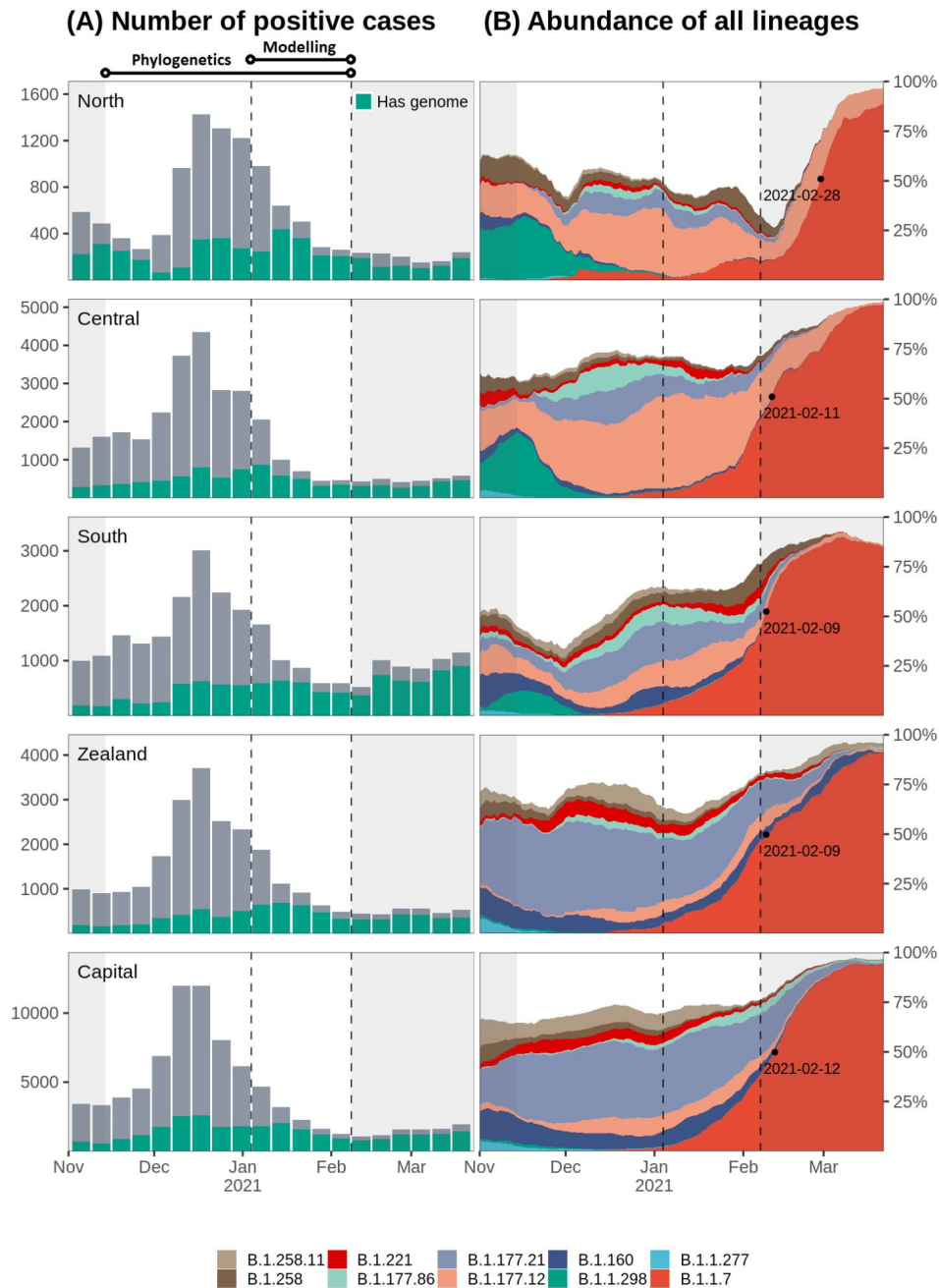


Fig. 1. Sequencing rate and lineage dynamics. Each row corresponds to the region highlighted to the far left. The two vertical dashed lines indicate the beginning and end of study period used to infer B.1.1.7 transmissibility, while the non-shaded area shows the period used for phylogenetic analysis. The time outside the study are shaded in grey. **(A)** sequencing rate across time. Total height corresponds to the total number of cases for each week. Green bars indicate the number of cases for which a high-quality genome was available (less than 3000 ambiguous bases). **(B)** relative abundance of ten most abundant PANGO lineages across Denmark in the study period. Computed from rolling averages of daily cases with genome available over a 14-days window. Less abundant lineages are not shown but included in the unfilled area of the plots. The black dot and date corresponds to the point where B.1.1.7 crosses 50% relative abundance.

400

405

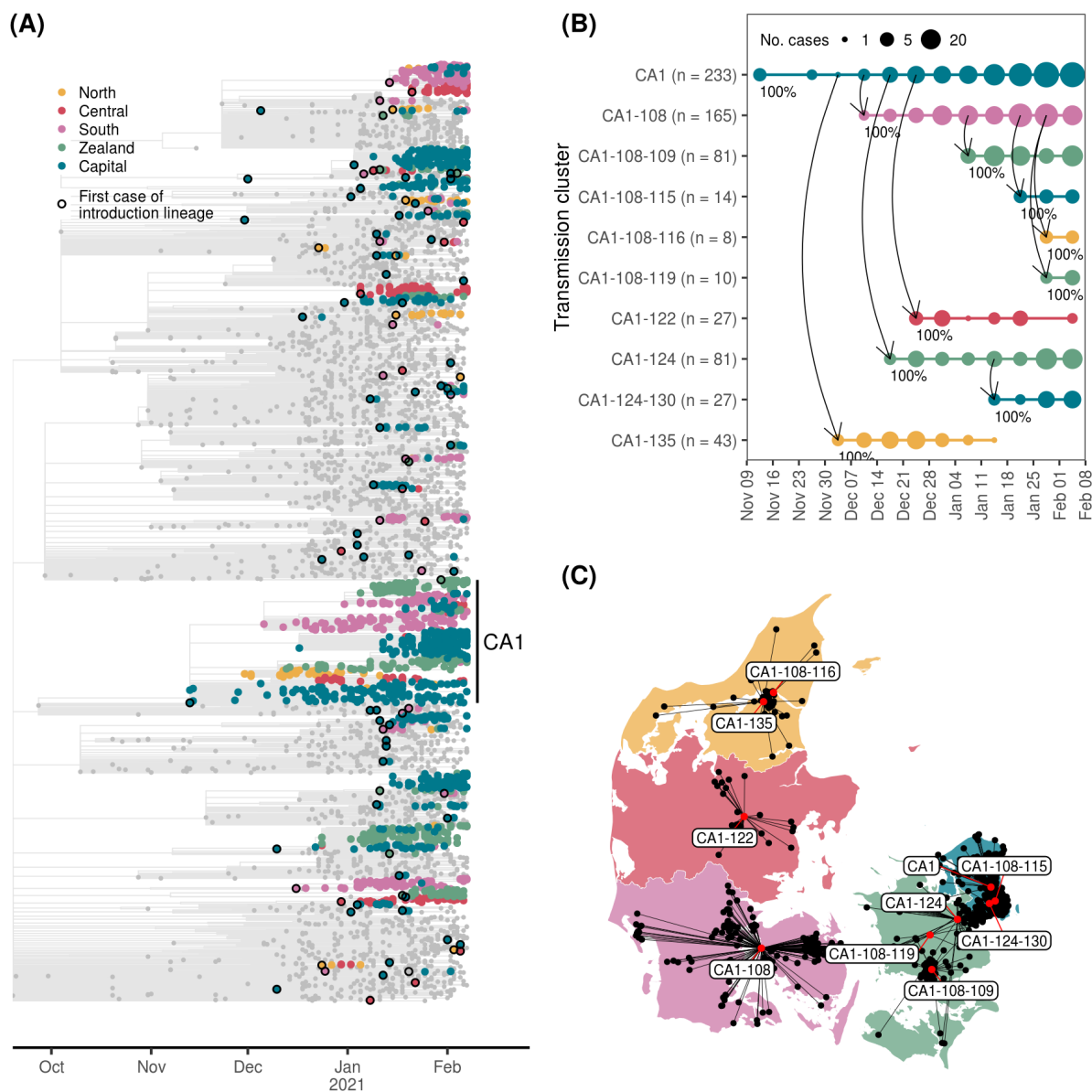


Fig. 2. Early introduction of B.1.1.7 and onwards transmission between Danish regions. (A) time-adjusted phylogenetic tree. Tree branches and non-Danish sequences are coloured grey. First sequence in a transmission lineage is used to infer time of importation and highlighted with a black circle. **(B)** and **(C)** shows a focused temporal analysis of the CA1 introduction lineage, focusing on descending clusters with more than 5 cases. **(B)** the temporal span and dynamics of onwards transmission across regions. Each descendent transmission cluster of CA1 is indicated with a solid horizontal line, with points at each week sized according to the number of cases. The vertical curved arrows indicate direction of transmission. The marginal probability of the observed transmission link is indicated at each arrow. In **(C)** the corresponding cases are placed on a map of Denmark using black points, with up to 5 km of random jitter added. Cases from the same transmission cluster are linked by black lines to their centroid location indicated by a red dot. Labels indicate cluster id's.

410

415

420

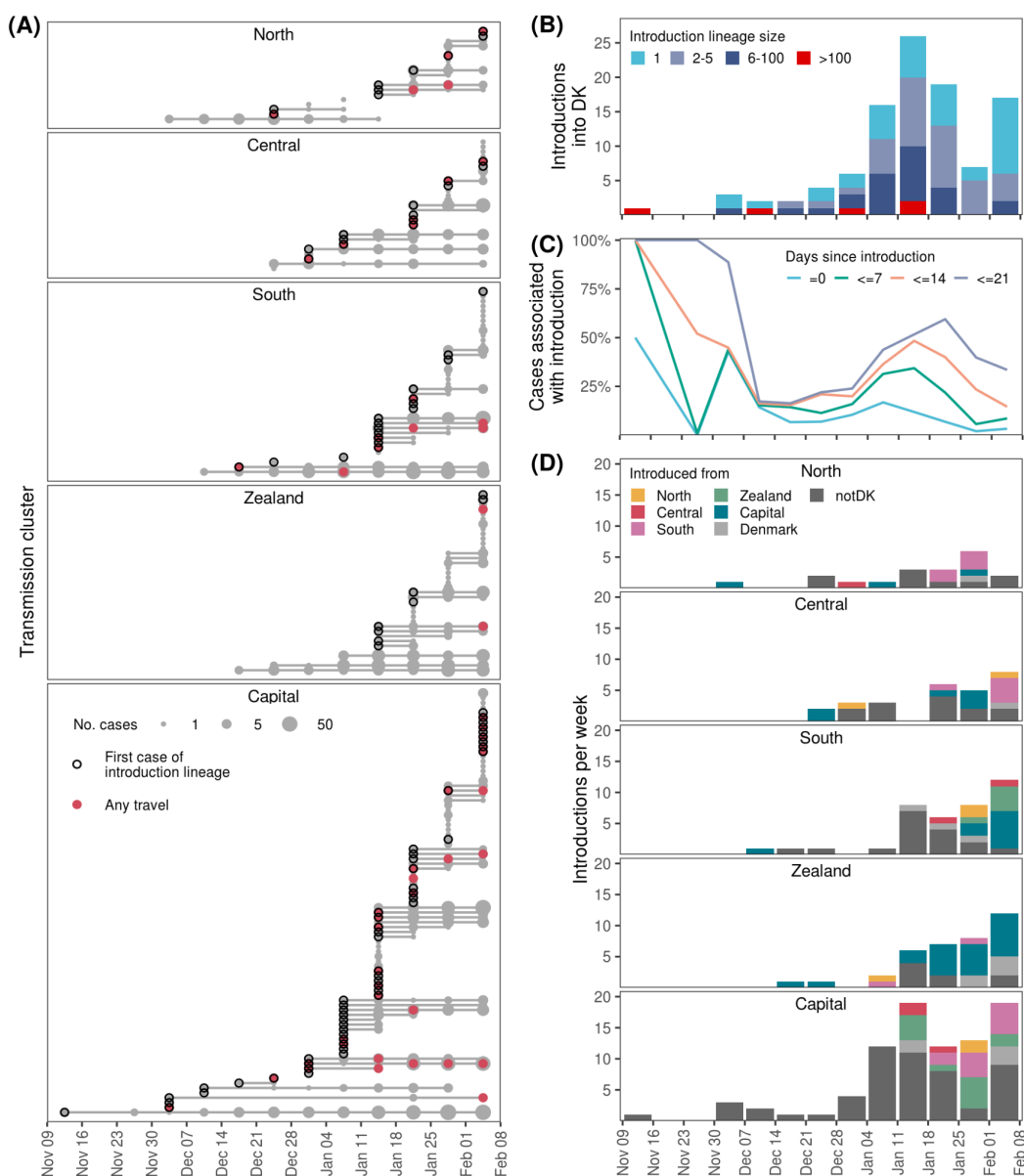


Fig. 3. Analysis of importation events into Denmark and onwards transmission between regions. (A) the temporal span and development of each transmission cluster, grouped by Danish region. Each transmission cluster is indicated with a solid horizontal line, with points at each week sized according to the number of cases. Circles indicate first observed case of an introduction lineage imported from outside Denmark. Red points indicate presence of cases with travel history in the given week. (B) the number of introductions into Denmark across time. Colours indicate the total number of offspring cases associated with each introduction lineage. (C) the relative contribution of introductions versus ongoing existing transmission. Introduction-related cases were defined with various cutoffs for the maximum number of days between the case and the first observed case for each introduction lineage, as indicated by different colours. (D) the origin of introductions across time for each region. If there was equal support for multiple regions as origin for an introduction Denmark was used as the origin. Label notDK indicates an introduction from outside Denmark. Only ancestral state changes with a marginal probability >95% were included in the analysis.

425

430

435