

# Phylodynamics reveals the role of human travel and contact tracing in controlling the first wave of COVID-19 in four island nations

Jordan Douglas,<sup>1,2,†,‡</sup> Fábio K. Mendes,<sup>1,3,†</sup> Remco Bouckaert,<sup>1,2</sup> Dong Xie,<sup>1,2</sup> Cinthy L. Jiménez-Silva,<sup>1,3</sup> Christiaan Swanepoel,<sup>1,2</sup> Joep de Ligt,<sup>4</sup> Xiaoyun Ren,<sup>4</sup> Matt Storey,<sup>4</sup> James Hadfield,<sup>5</sup> Colin R. Simpson,<sup>6</sup> Jemma L. Geoghegan,<sup>4,7,§</sup> Alexei J. Drummond,<sup>1,2,3</sup> and David Welch<sup>1,2</sup>

<sup>1</sup>Centre for Computational Evolution, The University of Auckland, Auckland 1010, New Zealand, <sup>2</sup>School of Computer Science, The University of Auckland, Auckland 1010, New Zealand, <sup>3</sup>School of Biological Sciences, The University of Auckland, Auckland 1010, New Zealand, <sup>4</sup>Institute of Environmental Science and Research Limited (ESR), Porirua 5420, New Zealand, <sup>5</sup>Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington WA 98109-1024, USA, <sup>6</sup>School of Health, Victoria University of Wellington, Wellington 6012, New Zealand and <sup>7</sup>Department of Microbiology and Immunology, University of Otago, Dunedin 9016, New Zealand

† Authors contributed equally to this work.

‡ <http://orcid.org/0000-0003-0371-9961>

§ <http://orcid.org/0000-0003-0970-0153>

\*Corresponding authors: E-mail: [jordan.douglas@auckland.ac.nz](mailto:jordan.douglas@auckland.ac.nz); [f.mendes@auckland.ac.nz](mailto:f.mendes@auckland.ac.nz)

## Abstract

New Zealand, Australia, Iceland, and Taiwan all saw success in controlling their first waves of Coronavirus Disease 2019 (COVID-19). As islands, they make excellent case studies for exploring the effects of international travel and human movement on the spread of COVID-19. We employed a range of robust phylodynamic methods and genome subsampling strategies to infer the epidemiological history of Severe acute respiratory syndrome coronavirus 2 in these four countries. We compared these results to transmission clusters identified by the New Zealand Ministry of Health by contact tracing strategies. We estimated the effective reproduction number of COVID-19 as 1–1.4 during early stages of the pandemic and show that it declined below 1 as human movement was restricted. We also showed that this disease was introduced many times into each country and that introductions slowed down markedly following the reduction of international travel in mid-March 2020. Finally, we confirmed that New Zealand transmission clusters identified via standard health surveillance strategies largely agree with those defined by genomic data. We have demonstrated how the use of genomic data and computational biology methods can assist health officials in characterising the epidemiology of viral epidemics and for contact tracing.

**Key words:** COVID-19; coronavirus; phylogenomics; human movement; contact tracing; New Zealand; Australia; Taiwan; Iceland

## 1. Introduction

Many continued hoping that the epidemic would soon die out and they and their families be spared. Thus they felt under no obligation to make any change in their habits, as yet. Plague was an unwelcome visitant, bound to take its leave one day as unexpectedly as it had come.

Albert Camus ('The Plague', 1947)

The respiratory tract illness caused by Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) known as Coronavirus Disease 2019 (COVID-19; (Rothan and Byrareddy 2020)) has caused substantial global health and economic impact, and the global epidemic appears far from over. COVID-19 was first detected in the city of Wuhan, Hubei province (China) in late 2019 (Zhu et al. 2020). By 11 March 2020 there were 100,000 cases across 114 countries leading the World Health Organization (WHO) to declare a

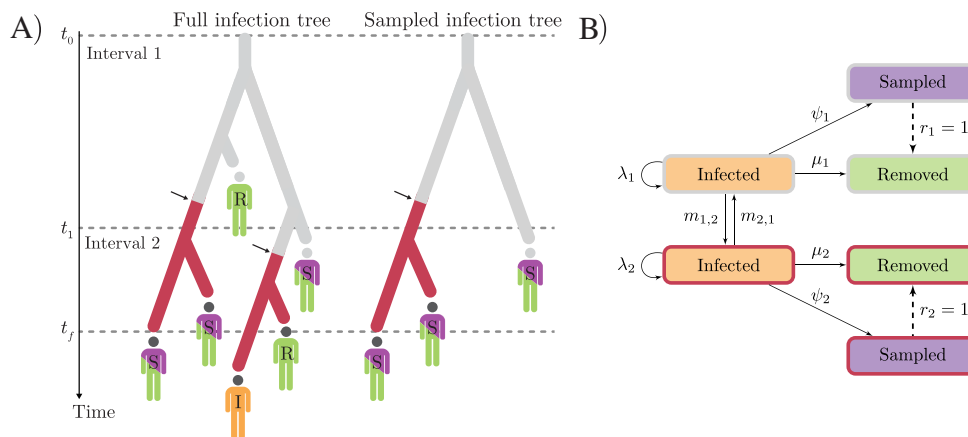
pandemic (World Health Organization 2020a). By early 2021, over 2 million deaths had been reported and over half of the world's population had received stay-at-home orders (Alasdair Sandford 2020).

COVID-19 has now been characterised extensively. The average number of people each case goes on to infect (known as the effective reproduction number,  $R_e$ ), for example, has been estimated as 1–6 during the early stages of the outbreak (Liu et al. 2020; Binny et al. 2020a; Alimohamadi, Taghdir, and Sepandi 2020). Because COVID-19 often causes mild symptoms Day (2020a,b) and surveillance efforts are variable, the difference between the true and reported number of cases could be of an order of magnitude or more (Ruiyun et al. 2020a; Lu et al. 2020). Fortunately, many locations have demonstrated that it is possible to reduce  $R_e$  to less than 1 as a result of mitigation efforts including contact tracing, case isolation, and travel restrictions (Binny et al. 2020a; Chinazzi et al.).

**Table 1.** Summary of the COVID-19 pandemic in the four target islands. Total passenger arrivals into the island in the January–March period are counted. All dates are in the Year 2020. The number of confirmed cases and the percentage of cases with recent overseas travel are reported for the January–April period.

Island	Population (millions)	Arrival (millions)	First case	Border close	Confirmed cases	Overseas travel
New Zealand	5	1.6 (Stats NZ 2020)	February 28	March 19 (RNZ 2020)	1,129 (World Health Organization 2020b)	39% (New Zealand Ministry of Health, Manatu Hauora 2020)
Australia	25	4.3 (Australian Bureau of Statistics, 2020)	January 25	March 20 (Burke 2020)	6,746 (World Health Organization 2020b)	64% (Department of Health 2020)
Iceland	0.34	0.33 <sup>a</sup> (Icelandic Tourist Board 2020)	February 28	March 20 (Schengen Visa Info 2020)	1,797 (World Health Organization 2020b)	19% (The Directorate of Health, The Department of Civil Protection, and Iceland 2020)
Taiwan	23	2.6 (Strong 2020)	January 21	March 19 (Aspinwall 2020)	429 (World Health Organization 2020b)	90% (Taiwan Centers for Disease Control 2020a)

<sup>a</sup>Visitor arrivals only.



**Figure 1.** Depiction of the MTBD model. (A) A full infection tree and a sampled infection tree ( $T$ ), with two time intervals (epochs) and two demes (indicated in grey and red). All lineages in the trees represent infected individuals ( $I$ ). Branching corresponds to transmission and colour change (arrows) corresponds to migration; infected individuals may be removed ( $R$ ) or sampled and then removed ( $S$ ). (B) Transition diagram: solid and dashed arrows labelled with rates and probabilities, respectively.

As we reflect on the early months of the pandemic (until the end of April 2020), different parts of the world tell contrasting stories about their struggle against COVID-19 (Table 1).

Aotearoa New Zealand, Australia, Iceland, and Taiwan, for example, fared comparatively well against their ‘first wave’, with smaller infection peaks and lower excess mortalities compared to many mainland countries (World Health Organization 2020b; FT Visual 2020), with New Zealand and Taiwan achieving elimination or near-elimination of their first waves in that period (Cousins, 2020; Summers et al. 2020). Over the following year, all four locations experienced ‘second waves’ of community transmission (Aljazeera 2020a, 2020b; Iceland Review 2020; The Guardian 2021).

Irrespective of their current COVID-19 situation, these four countries share in common the effectiveness of their response to the first outbreak, which contributed positively to their swift recovery at the time. In March 2020 amidst the first wave, all four nations closed their borders and required isolation or quarantine of those who were allowed through (Hale et al. 2020), either restricting or discouraging human movement (Apple 2020). Stay-at-home orders were issued in New Zealand and Australia but not in Iceland or Taiwan (Hale et al. 2020). Stringent contact tracing, coupled with isolation and/or quarantine of cases and their close contacts, proved to be effective strategies in these four locations

(Baker et al. 2020; Draper et al. 2020; Gudbjartsson et al. 2020a; Jason Wang, Ng, and Brook 2020). These islands thus provide excellent case studies to characterise the effects of human travel on epidemic spread during the early months of the pandemic, a critical endeavour when facing the threat of global endemicity.

Mathematical methods for studying epidemics are as diverse as the diseases they target, making varied use of surveillance and genetic data and ranging from models of mathematical epidemiology (Grassly and Fraser 2008; Binny et al. 2020a; Ferretti et al. 2020; Hethcote 2000) to phylodynamic models (Volz, Koelle, and Bedford 2013). The latter take advantage of the comparable timescales of epidemiological and evolutionary processes (Grenfell et al. 2004), which result from short viral generation times and high viral mutation rates (20–40 mutations per year for SARS-CoV-2; (Lai et al. 2020; Xingguang et al. 2020b; Rambaut 2020)). By modelling the transmission process as phylogenetic trees (Fig. 1), phylodynamic and molecular evolution models can be coupled to reconstruct the mutational history of viral genomes (Xingguang et al. 2020b; Rambaut 2020; Lai et al. 2020), thereby elucidating the epidemiological processes that shaped them.

In this study, we considered epidemic dynamics in New Zealand, Australia, Iceland, and Taiwan over the first few months

of the pandemic up until the end of April. The goal was to determine how SARS-CoV-2 case numbers varied as human mobility changed in response to government restrictions. We incorporated human movement into our analyses through the modelling of mobile phone data. More specifically, we hypothesised that introductions of SARS-CoV-2 would sharply decline following international travel and that viral transmission ( $R_e$ ) would decline following human movement.

While the notion that restricting human movement constrains the spread of infectious diseases may appear exhaustively addressed, our study is unique in its integration of human mobility data with genomic and epidemiological data, in geographically isolated island settings. There is a great benefit in studying the spread of disease in these locations, as their geographic isolation reduces the effect of unregulated and undocumented land border crossing that is common to mainland countries. We do so by analysing viral genomic data under a range of phylogenetic methods, while at the same time controlling for data subsampling strategies. Lastly, we compare the phylodynamic history of New Zealand's first outbreak with that inferred by the New Zealand Ministry of Health (NZMH) in order to assess the ability of contact tracing at identifying transmission pathways.

## 2. Methods

In the study of disease phylodynamics, a complete epidemic can be modelled as a rooted, binary phylogenetic tree whose  $N$  tips are all infected individuals (Fig. 1) and whose characteristics and drivers we wish to infer. In a rapidly growing epidemic, the age of this tree serves as a good proxy for the duration of the epidemic. Because it is seldom possible to sample the entire infected population, one instead reconstructs a subtree  $T$  from a sample of cases of size  $n$ , where  $n < N$ . This sample provides  $n$  viral genome sequences we refer to as  $D$ , as well as each sequence's sampling date, collectively represented by vector  $y$ .  $D$  and  $y$  thus constitute the input data and are herein referred to as 'samples'. In the case of structured phylodynamic models, samples in  $y$  are further annotated by 'deme'—a geographical unit such as a city, a country, or a continent.

A Bayesian phylodynamic model is characterised by the following unnormalised posterior probability density function:

$$f(\mathbf{T}, \mu_c, \theta_s, \theta_\tau | D, y) \propto P(D | \mathbf{T}, \mu_c, \theta_s) f(\mathbf{T} | \theta_\tau) f_{pr}(\mu_c, \theta_s, \theta_\tau), \quad (1)$$

where  $P(D | \mathbf{T}, \mu_c, \theta_s)$  is the phylogenetic likelihood of tree  $T$  given alignment  $D$ , whose sequences accumulate substitutions at rate  $\mu_c$  under a substitution model with parameters  $\theta_s$ .  $f(\mathbf{T} | \theta_\tau)$  is the probability density function of the tree prior sampling distribution, given tree prior parameters  $\theta_\tau$  (in this study we consider a range of phylodynamic models that differ in their tree prior; see sections 2.1 and 2.2 and Supplementary materials). Finally,  $f_{pr}(\mu_c, \theta_s, \theta_\tau)$  is the prior distribution on the remaining parameters.

### 2.1 Multitype birth–death model

The multitype birth–death (MTBD) model (Kuhnert et al. 2016; Sciré et al. 2020) is a structured phylodynamic model that assumes infected ('I'; Fig. 1) hosts of viral lineages are organised into a set of demes ('types') and migrate between them at rate  $m$  (Fig. 1). Infected individuals transmit the infection at birth rate  $\lambda$  and can be sampled ('S'; Fig. 1) and sequenced at sampling rate  $\psi$ ; after which they are removed ('R'; Fig. 1) with probability  $r$ . We set  $r = 1$  by assuming (sampled) confirmed cases are isolated and are unlikely to cause further infections. Removal

without sampling happens at death rate  $\mu$  (Fig. 1). Here, 'death' refers to anything causing hosts to no longer be infectious, including recovery with immunity, host death, or behavioural changes (e.g. self-isolation).

The probability density function of  $T$  under the MTBD sampling distribution (Kuhnert et al. 2016) and its parameters  $\theta_\tau$  is given by:

$$f(\mathbf{T} | \theta_\tau) = f(\mathbf{T} | \lambda, \mu, \psi, m, r, \pi_G, t, \mathcal{O}), \quad (2)$$

where  $\pi_G$  represents the frequency of lineages in each deme at the root and  $\mathcal{O}$  is tree  $T$ 's age.

The MTBD model is reparameterised into a computationally and epidemiologically convenient form, with parameters  $R_e = \frac{\lambda}{\mu + \psi r}$ ,  $b = \mu + \psi r$ , and  $S = \frac{\psi}{\psi + \mu}$ .  $R_e$  is the disease's effective reproductive number (Stadler et al. 2013).  $b$  is the total rate at which individuals become non-infectious, and  $s$  is the probability of an individual being a sample (out of all removed individuals). The phylodynamic tree prior under this parameterisation is

$$f(\mathbf{T} | \theta_\tau) = f(\mathbf{T} | R_e, b, s, m, r, \pi_G, t, \mathcal{O}). \quad (3)$$

### 2.2 Alternate phylogenetic models

As a sensitivity analysis, we invoke three additional phylogenetic models on top of MTBD and compare the results between the four models. These models are (1) a discrete phylogeography model (DPG, Lemey et al. (2009)), (2) a discrete phylogeography model with informed effective population sizes and epochs (DPG+), and (3) the structured coalescent model with informed effective population sizes and epochs (SC, Muller, Rasmussen, and Stadler (2018)). Details of these models are in Supplementary information.

### 2.3 Demes and infection time intervals

We consider four independent geographical models, each characterised by two demes: a target island deme (New Zealand, Australia, Iceland, or Taiwan; referred to as the 'IS' deme) and a mutually exclusive 'rest-of-the-world' deme ('RW') that contributes samples from other countries in the world. All RW-specific parameters are considered nuisance parameters.

Under our MTBD model,  $R_e$ ,  $b$ ,  $s$ , and  $m$  are estimated across a series of time intervals (and held constant within intervals) defined by time boundaries  $t$ , with  $t_0 < t_1 < \dots < t_f$ , where  $t$  is fixed at informed dates.  $t_0$  is the beginning of the infection, and  $t_f$  is both the last sampling time and the end of the last interval (Fig. 1). Time intervals are both parameter- and deme-specific.

$R_e$  and  $b$  are modelled with three epochs  $t_0 < t_1 < t_2 < t_f$ , where  $t_1$  is the first reported case in each deme and  $t_2$  is the point of human mobility decline, according to a Bayesian hierarchical sigmoidal model trained on Apple mobility data (see Results and Supplementary information). These mobility data consist of human driving, walking, and public transit use as tracked through Apple mobile phone user movement. Epoch dates for  $s$  and  $m$  can be found in Table S3.

### 2.4 SARS-CoV-2 genomic data and subsampling

We sequenced and assembled 217 New Zealand SARS-CoV-2 genomes (representing 19 per cent of confirmed cases during the relevant time period) that are now available on Global Initiative on Sharing All Influenza Data (GISAID; Shu and John 2017), which provided the remaining 3,000+ genome sequences analysed here. Due to computational limitations, we subsampled GISAID sequences as follows: (1) two different subsample sizes

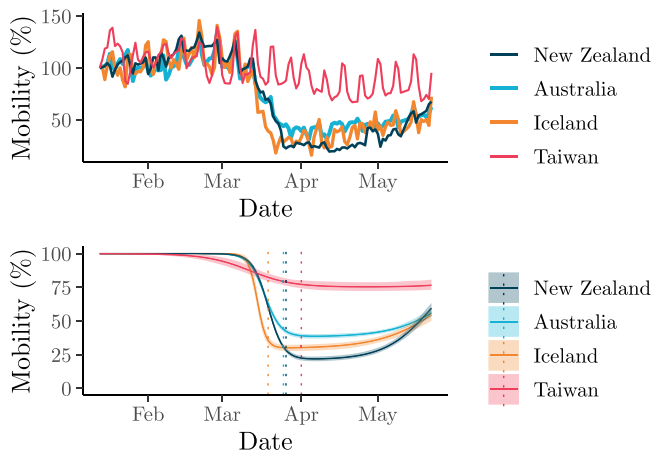
(‘small’ vs. ‘large’), with 200 and 500 RW (‘rest of the world’ deme) sequences, and (2) two sample selection methods (‘time’ vs. ‘active’) that chose sequences uniformly across time or proportional to the number of active cases over time. Up to 250 sequences from the IS (island) deme are included in each alignment (where more than 250 were available, we subsampled them using the respective sampling method). We restricted our analyses to the early months of the pandemic (until the end of April 2020) because this time window encompasses largely resolved (i.e. from onset to end) outbreaks in all four countries, while also making data set sizes more manageable and precluding further subsampling and thinning of data points.

Permuting the above two options yields four subsampling protocols, which were applied to the four IS demes, generating a total of sixteen SARS-CoV-2 alignments (30-kb long, 0.3 per cent ambiguous sites on average). Each alignment consists of sequences from one of the four islands (IS) and from the rest of the world (RW). Moreover, each alignment was analysed with each phylodynamic model (MTBD, SC, DPG, and DPG+), coming to a total of forty-eight Bayesian phylogenetic analyses.

### 3. Results

#### 3.1 Modelling human movement as tracked by mobile phone data

With the goal of characterising the decrease in human movement over the weeks our SARS-CoV-2 samples were collected; we employed a novel Bayesian hierarchical sigmoidal model for the analysis of mobile phone data (Apple 2020; Fig. 2). This analysis revealed that human movement generally decreased in all four islands, with Taiwan being the first to show a declining trend among the islands, but also the least pronounced one. New Zealand, Australia, and Iceland underwent a marked decline in human movement during March 2020, with a slight recovery starting around mid-April. Weekly travel oscillations can be observed in the original mobility data, but are averaged out by the sigmoidal model.



**Figure 2.** Human walking movement relative to Day 1 (100 per cent). Data were obtained by tracking the mobility of Apple mobile phone users (Apple 2020). Top: movement data from Apple. Bottom: fitted sigmoid model (solid lines show the median and shaded curves show the 2.5–97.5 per cent quantile interval); vertical dotted lines correspond to the estimated mean date of mobility change, which defines the  $t_2$  epoch date for each deme.

The dates of mobility reduction provided by this mobility model informed the start of the final time interval  $t_2$  in our three-epoch models of  $R_e$  and  $b$ .

#### 3.2 SARS-CoV-2 phylodynamics

We tested four genome subsampling regimes under four phylodynamic models. This sensitivity analysis suggests that inference under the MTBD and SC models is fairly robust to subsampling protocols, as opposed to DPG and DPG+, which yielded unrealistically small estimates of infection duration and of viral introduction counts (see Iceland; Fig. 5), particularly from the ‘active’ method. Such model behaviour is unsurprising and has been reported elsewhere (Nicola et al. 2015), and it was only slightly alleviated when ‘large’ subsamples were used.

Analyses under MTBD and SC were prohibitively time-demanding using the larger data sets. For consistency, we only present results from the ‘small-time’ data set in the main article, but further results can be found in Supplementary information. Four ‘small-time’ data sets consist of up to 250 sequences from the target deme IS and 200 sequences sampled uniformly through time from the rest-of-the-world deme RW.

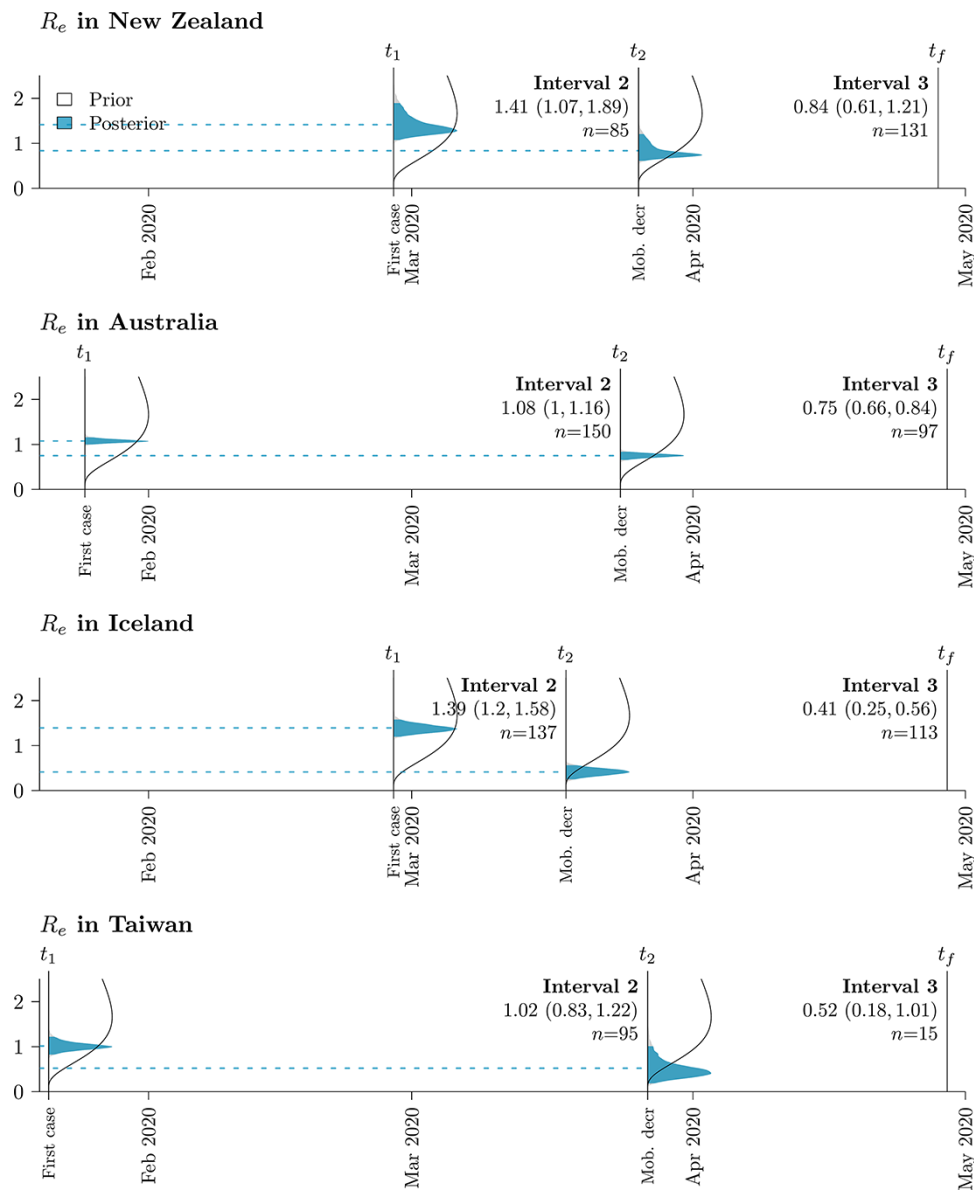
We estimated  $R_e$  independently for each epoch using the MTBD model (Fig. 3). This analysis estimates  $R_e$  as 1.02–1.41 during the early infection of each respective country (following the first case) and later at less than 1 (following human mobility reduction). These estimates were not sensitive to sampling strategies. Each island IS saw multiple introductions of SARS-CoV-2 from the rest of the world (RW; Figs. 4 and 5). Under MTBD, we estimated the mean number of sample sequence introductions to be 63, 87, 49, and 65 for New Zealand, Australia, Iceland, and Taiwan, respectively. Conversely, we estimated just one virus export event from each country to the rest of the world, across the time period and sample considered. Similar results were observed using the SC model, with neither SC nor MTBD producing drastically different estimates between subsampling strategies.

The larger estimates of  $R_e$  for New Zealand and Iceland ( $\sim 1.4$  before mobility decline) reflect their samples having larger proportions of community transmission, as opposed to viral introduction events, in comparison to Australia and Taiwan, which have smaller  $R_e$  estimates ( $\sim 1$ ). The genome samples themselves are a product of the public health surveillance policy within each country, and non-random sampling of cases could lead to biased estimates of  $R_e$ ; for example, intensive sampling of large outbreaks would inflate  $R_e$  and increased sequencing of overseas returnees is likely to deflate  $R_e$ . All countries studied here practised widespread sampling and sequencing of cases, which would produce relatively unbiased samples.

#### 3.3 Comparison with contact tracing in New Zealand

The NZMH defines a significant COVID-19 cluster to comprise ten or more cases connected through transmission and who are not all part of the same household (New Zealand Ministry of Health 2020). Five of these clusters are represented by five or more of our samples (Fig. 4A and C), with a sixth non-significant cluster (based in the city of Auckland).

A clade is defined as a set of samples (within a phylogenetic tree) that share an exclusive common ancestor. Here we define a phylogenetic cluster as a clade whose samples share a cluster label and whose origin can be explained by invoking a single introduction into a focal deme. Under this definition (1) case exports are allowed, (2) isolated labelled tips outside of the clade are considered type-1 contact tracing mismatches (i.e. contact tracers



**Figure 3.**  $R_e$  prior (white) and posterior (blue) probability distributions across islands and time intervals ( $n$  samples in each interval). Posterior means are indicated by dashed lines. Mean estimates (and 95 per cent highest posterior density intervals) are indicated under each time interval:  $t_1$  is the first case in the target deme,  $t_2$  the start of mobility reduction, and  $t_f$  the final sample in the data set.

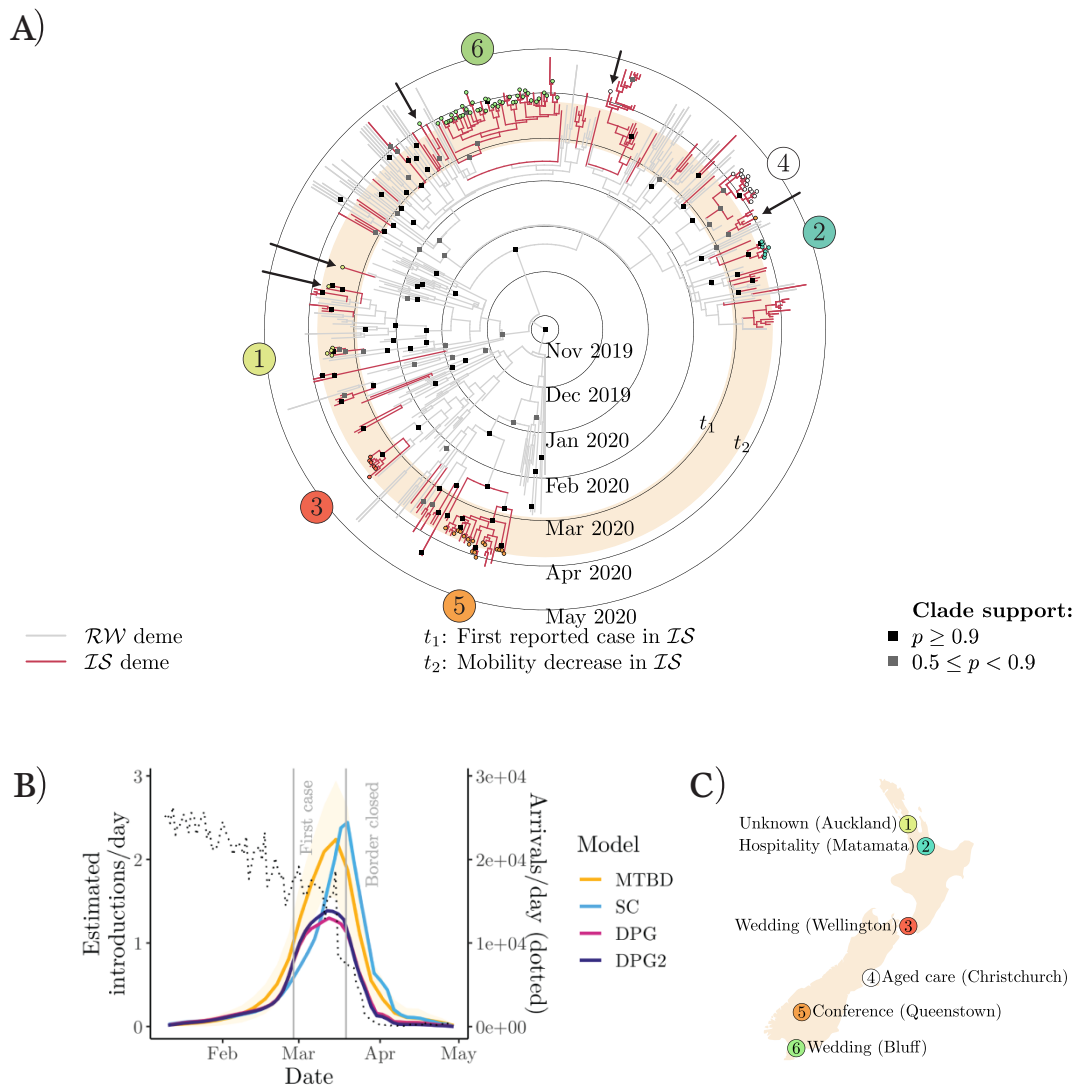
assigned a case to a cluster they were not genomically linked to), and (3) unlabelled tips that are both from the focal deme and placed within the clade are considered type-2 mismatches (i.e. genomics assigned a case to a cluster when contact tracers could not).

Using this definition, all clusters identified by NZMH were recovered (among those represented by our samples). However, there were five type-1 mismatches: two from an Auckland-based cluster, one from Christchurch, one from Queenstown, and one from Bluff (shown by arrows in Fig. 4). We estimated eighteen type-2 mismatches. Our alternative subsampling schemes produced similar results: estimating 5, 4, and 7 type-1 mismatches and 18, 19, and 18 type-2 mismatches for ‘small-active’, ‘large-time’, and ‘large-active’, respectively. This suggests that the method is fairly robust to subsampling strategy.

## 4. Discussion

In this article, we explored the introduction and transmission history of COVID-19 in New Zealand, Australia, Iceland, and Taiwan—geographically isolated countries that so far have been successful in preventing this disease from wreaking havoc as extensively as it did in other parts of the world. We also delved deeper into the outcome of contact tracing within New Zealand specifically. This study is novel in its integration of genomic, epidemiological, and mobile phone mobility data through explicit statistical models. The methods employed are robust and transparent in that they (1) explicitly justify and incorporate prior knowledge about epidemiological parameters (see Supplementary information), (2) report uncertainty in parameter estimation, and (3) explore the effect that different data subsampling protocols have on parameter estimation.





**Figure 4.** (A) New Zealand COVID-19 maximum-clade-credibility tree (from ‘small-time’ alignment). Samples are labelled according to outbreaks identified by NZMH. Arrows indicate cases assigned to a cluster by NZMH, which are not monophyletic with the rest of the cluster. (B) Number of SARS-CoV-2 introductions in New Zealand over time. (C) New Zealand cluster locations (with greater than five cases).

#### 4.1 Phylodynamic modelling complements contact tracing

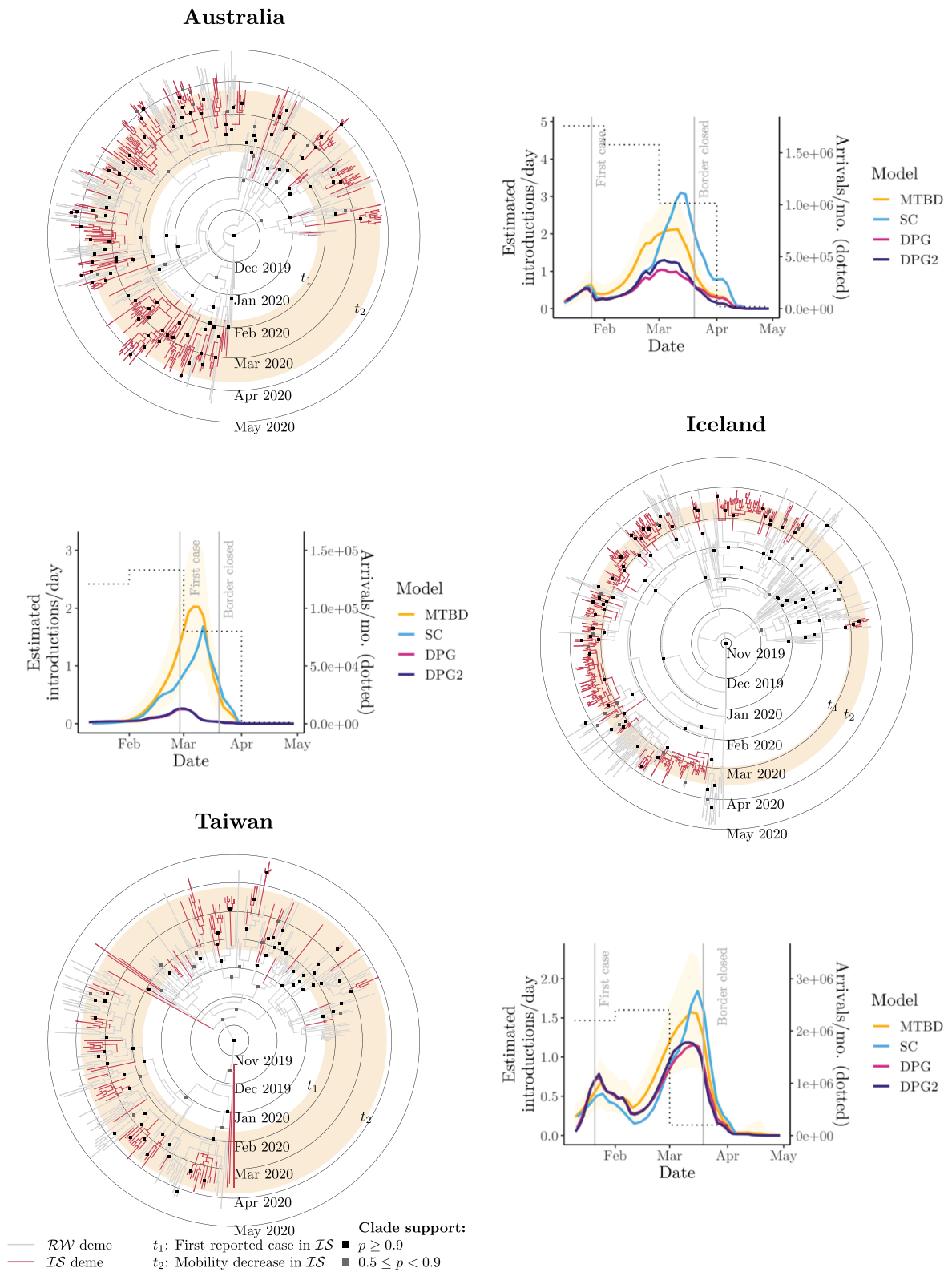
Identifying cases and the sources of their infections is an essential exercise in slowing infectious diseases such as COVID-19 (Hellewell et al. 2020), as recognised by the Director-General of the WHO in his pronouncement: ‘You cannot fight a fire blindfolded. And we cannot stop this pandemic if we don’t know who is infected’. Modern health surveillance strategies include medical databases containing travel history (Jason Wang, Ng, and Brook 2020), the development of mobile phone applications (Ferretti et al. 2020), and the identification of outbreak subgroups known as ‘clusters’, whose cases share the same origin (New Zealand Ministry of Health 2020). At the core of health surveillance practice lays contact tracing—the goal of which is to pinpoint the origin of infections by reconstructing the travel history of confirmed cases and identifying all individuals with whom cases have come in contact.

Nonetheless, contact tracing has its limitations. Establishing the link between cases is restricted, for example, by the ability of cases to recall their recent contacts and travel history,

which can lead to type-2 mismatches between epidemiological and genetic clusters. Furthermore, a case may have attended an event causally linked to a cluster but acquired their infection elsewhere and did not infect anyone at the event (type-1 mismatches). The implications of these types of error are manifold. If cluster sizes are underestimated, then so too is the rate of disease spread. If the extent of import-related cases is overestimated, then the impact of international air travel and the extent of community transmission cannot be fully accounted for.

After considering 217 genomes (representing 19 per cent of the cases in New Zealand at the time), our results suggest that there are additional eighteen unclassified cases in New Zealand that could have been linked to known clusters but were not (type-2 mismatch) and five cases that were linked to a cluster where they did not acquire or transmit the infection (type-1). However, by and large, our phylodynamic analysis is in agreement with conclusions reached by NZMH.

Overall, we have shown that contact tracing has been accurate among the cases considered and succeeded to a large degree in identifying individuals belonging to the same infection cluster.



**Figure 5.** Maximum-clade-credibility COVID-19 tree (from ‘small-time’ alignment) for Australia (top), Iceland (middle), and Taiwan (bottom). Plots show the estimated number of SARS-CoV-2 introductions over time.

We have demonstrated how the rapid real-time availability and assessment of viral genomic data can complement and augment health surveillance strategies and thus improve the accuracy of contact tracing methods.

As the pandemic has progressed, real-time whole-genome sequence data have been integrated into the health response of the four countries discussed here (Geoghegan et al. 2021;

Seemann et al. 2020; Gudbjartsson et al. 2020b; Taiwan Centers for Disease Control 2020b).

Genome sequencing is increasingly affordable and shared global surveillance resources such as NextStrain (Hadfield, et al. 2018) and GISAID are becoming more complete and useful. We advocate widely incorporating modern methods of genetic epidemiology into standard contact tracing and epidemic surveillance in order to ensure that COVID-19 and future epidemics leave as promptly as they come.

## 4.2 Restricting human movement constrains COVID-19 spread

As we have demonstrated for all four countries, genomic data can be used in the estimation of disease introductions over time, a goal particularly hard to achieve with contact tracing alone. Our results suggest that SARS-CoV-2 was introduced into each island on many independent occasions, corroborating the available epidemiological data from health officials (Table 1). Viral imports rapidly declined around the time of border closure in mid-late March 2020 as a result of decreased international travel (Figs. 4 and 5). We believe this in turn played a significant role in limiting the spread of COVID-19.

In contrast, even with plentiful genomic data, it is non-trivial to reliably estimate the number of exportation events. Such an estimate would depend on a large, unbiased global sample that is challenging to obtain and whose analysis is prohibitively slow.

Mobile phone data revealed idiosyncratic human movement patterns (Fig. 2). During the early stages of the pandemic, human mobility rapidly declined in both New Zealand and Australia as a result of stay-at-home orders (Hale et al. 2020), as well as in Iceland despite no such official restraints. Taiwan, on the other hand, did not exhibit as much of a decline in human movement nor did it enforce stay-at-home orders on the general population. A goal of this study was to characterise the extent to which this decline in human movement could impact the spread of COVID-19.

We estimated effective reproduction numbers,  $R_e$ , as approximately 1–1.4 at the onset of the pandemic in all four islands (Fig. 3). These estimates are not as high as in other countries where the pandemic was more widespread and where  $R_e$  was typically estimated in the range of 2–6 (Liu et al. 2020; Binny et al. 2020a; Alimohamadi, Taghdir, and Sepandi 2020). Our low estimates reflect the fact that a large fraction of cases in these countries were travel-related imports and that community transmission rates were low.  $R_e$  declined as human movement decreased in each of the countries studied, reaching values below 1.0 (Fig. 3). Estimates of  $R_e$  based only on epidemiological data are also consistently below 1 during these time periods (Binny et al. 2020a,b; Price et al. 2020; Karmakov et al. 2020). For New Zealand, Australia, and Iceland, the decline in  $R_e$  by 30–70 per cent coincided with a decline in human mobility of 50–80 per cent.

Our findings suggest that the first wave of COVID-19 in the four islands was controlled by greatly reducing international travel coupled with strong reductions in domestic movement. While the human movement data tell us nothing about the role contact tracing played in the four islands, it is noted that traditional contact tracing methods are insufficient in themselves (Ferretti et al. 2020) and that contact tracing systems were not well established across all territories early in the pandemic (Verrall 2020). The exception to this may be Taiwan, where an extreme restriction of domestic human movement was not required for suppressing a pandemic like COVID-19. In the first three months, Taiwan saw fewer

cases-per-capita than New Zealand by an order of magnitude, for instance (Table 1), despite the latter enforcing a strict nationwide lockdown while having lower population density. Taiwan's success has been attributed to the Taiwan Centers for Disease Control (Taiwan CDC) who were well prepared for such an epidemic following their experience with the SARS-CoV-1 outbreak in 2003—the early implementation of mass masking in public spaces and border control (Summers et al. 2020; Jason Wang, Ng, and Brook 2020).

Nevertheless, a stringent mobility-reduction response against COVID-19, likely contributed to SARS-CoV-2 elimination in New Zealand by May 9 (Cousins, 2020), thus being the first country with over 1,000 confirmed cases to return to zero active cases, allowing it to re-open its economy, albeit with a restricted border. Most importantly, human morbidity and mortality were limited. Conversely, while Australia also saw a decline in human mobility and initially enjoyed some measure of success against COVID-19, its more lenient approach, compared with either New Zealand's or Taiwan's, might underlie Australia's longer path to elimination (Hale et al. 2020; Aljazeera 2020b).

In conclusion, we have shown that the introduction of SARS-CoV-2 was significantly inhibited as a result of international travel reduction in March 2020—notably the closing of borders in the four islands examined. We have also shown that the spread of SARS-CoV-2 within these communities, while initially lower than other parts of the world, saw a further reduction following the decline of human mobility within the countries. However, positive outcomes against this and similar pandemics will be dependent on a multitude of other factors including contact tracing, case isolation, and hygienic practices (MacIntyre 2020). As such, none of these should be seen as a silver bullet.

## Data availability

For reproducibility, BEAST2 input files for the four phylogenomic models (with sequences removed) are available at <https://github.com/CompEvol/Covid4Island>.

## Supplementary data

Supplementary data is available at *Virus Evolution* online.

## Acknowledgements

We thank T. Vaughan and J. Sciré for discussions about the MTBD model, and all of the authors who have shared their genomic data on GISAID. Finally, we wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high-performance computing facilities, funded jointly by NeSI's collaborator institutions and through the MBIE's Research Infrastructure programme.

## Funding

This research was supported by a contract from the Health Research Council of New Zealand (20/1018), a COVID-19 Innovation Acceleration Fund (grant CIAF-0470) from the Ministry of Business, Innovation & Employment's Research Infrastructure (MBIE), and Marsden grants 16-UOA-277 and 18-UOA-096 from the Royal Society of New Zealand. Genome sequencing was partially funded by the NZMH.

**Conflict of interest:** The authors have no conflicting interests to declare.



## Ethics and permissions

Nasopharyngeal samples testing positive for SARS-CoV-2 were obtained from public health medical diagnostics laboratories located throughout New Zealand. All samples were de-identified before receipt by the researchers. Under contract for the Ministry of Health, ESR has the approval to conduct genomic sequencing for surveillance of notifiable diseases.

## References

- Alimohamadi, Y., Taghdir, M., and Sepandi, M. (2020) 'The Estimate of the Basic Reproduction Number for Novel Coronavirus Disease (COVID-19): A Systematic Review and Meta-Analysis', *Journal of Preventive Medicine and Public Health*, 53: 151–7.
- Aljazeera (2020a) We Beat the Virus Again: COVID Restrictions in Auckland Lifted. <<https://www.aljazeera.com/news/2020/10/5/new-zealandsardern-lifts-coronavirus-restrictions-in-auckland>> accessed 5 Oct 2020.
- (2020b) Australia Records Zero New COVID-19 Cases. <<https://www.aljazeera.com/news/2020/11/1/national-donutaustralia-records-zero-new-community-covid-case>> accessed 1 Nov 2020.
- Apple (2020) Mobility Trends Reports. <<https://www.apple.com/covid19/mobility>> accessed 25 May 2020.
- Aspinwall, N. (2020) Taiwan Closes Borders in Preparation for Possible Second Wave of the Coronavirus. *The Diplomat*. <<https://thediplomat.com/2020/03/taiwan-closes-borders-in-preparation-for-possible-secondwave-of-the-coronavirus/>> accessed 20 Mar 2020.
- Australian Bureau of Statistics (2020) 3401.0 - Overseas Arrivals and Departures (Australia, Mar 2020) <<https://www.abs.gov.au/ausstats/abs@nsf/mf/3401.0>> accessed 12 May 2020.
- Baker, M. G. et al. (2020) 'New Zealand's Elimination Strategy for the Covid-19 Pandemic and What Is Required to Make It Work', *The New Zealand Medical Journal*, 133: 10–4.
- Binny, R. N. et al. (2020a) 'Effect of Alert Level 4 on Effective Reproduction Number: Review of International COVID-19 Cases', *medRxiv*.
- (2020b) 'Effective Reproduction Number for COVID-19 in Aotearoa New Zealand', *medRxiv*.
- Board, I. T. (2020) *Tourism in Iceland in Figures*. <<https://www.ferdamalastofa.is/static/files/ferdamalastofa/Frettamyndir/2020/mai/april-2020-e.pdf>> accessed May 2020.
- Bouckaert, R. et al. (2019) 'BEAST 2.5: An Advanced Software Platform for Bayesian Evolutionary Analysis', *PLOS Computational Biology*, 15: e1006650.
- Burke, K. (2020) Australia Closes Borders to Stop Coronavirus. Seven News. <<https://7news.com.au/lifestyle/health-wellbeing/australia-closes-borders-to-stopcoronavirus-c-752927>> accessed 19 Mar 2020.
- Chinazzi, M. et al. (2020) 'The Effect of Travel Restrictions on the Spread of the 2019 Novel Coronavirus (COVID-19) Outbreak', *Science*, 368: 395–400.
- Day, M. (2020a) 'COVID-19: Identifying and Isolating Asymptomatic People Helped Eliminate Virus in Italian Village', *BMJ*, 368: m1165.
- (2020b) 'COVID-19: Four Fifths of Cases are Asymptomatic, China Figures Indicate', *BMJ*, 369: m1375.
- The Directorate of Health, The Department of Civil Protection, and Iceland (2020) *Emergency Management. COVID-19 in Iceland Statistics, by Origin of Infection*. <<https://www.covid.is/data>> accessed 1 May 2020.
- Draper, A. D. K. et al. (2020) 'The First 2 Months of Covid-19 Contact Tracing in the Northern Territory of Australia, March-April 2020', *Communicable Diseases Intelligence*, 44: 10–44.
- Ferretti, L. et al. (2020) 'Quantifying SARS-CoV-2 Transmission Suggests Epidemic Control with Digital Contact Tracing', *Science*, 368: 6491.
- Geoghegan, J. L. et al. (2021) 'Use of Genomics to Track Coronavirus Disease Outbreaks, New Zealand', *Emerging Infectious Diseases*, 27: 1317.
- Grassly, N. C., and Fraser, C. (2008) 'Mathematical Models of Infectious Disease Transmission', *Nature Reviews Microbiology*, 6: 477–87.
- Grenfell, B. T. et al. (2004) 'Unifying the Epidemiological and Evolutionary Dynamics of Pathogens', *Science*, 303: 327–32.
- The Guardian (2021) Taiwan Records 180 New Cases in Island's Worst Covid Outbreak of Pandemic. <<https://www.theguardian.com/world/2021/may/15/taiwan-records-180-new-cases-in-islands-worst-covidoutbreak-of-pandemic>> accessed 15 May 2021.
- Gudbjartsson, D. F. et al. (2020a) 'Spread of SARS-CoV-2 in the Icelandic Population', *New England Journal of Medicine*, 382: 2302–2315.
- (2020b) 'Spread of Sars-cov-2 in the Icelandic Population', *New England Journal of Medicine*, 382: 2302–15.
- Hadfield, J. et al. (2018) 'Nextstrain: Real-time Tracking of Pathogen Evolution', *Bioinformatics*, 34: 4121–3.
- Hale, T. et al. (2020) *Oxford COVID-19 Government Response Tracker*. Blavatnik School of Government. <<https://github.com/OxCGRT/covid-policy-tracker>> accessed 4 Jun 2020.
- Hellewell, J. et al. (2020) 'Feasibility of Controlling COVID-19 Outbreaks by Isolation of Cases and Contacts', *The Lancet Global Health*, 8: e488–96.
- Hethcote, H. W. (2000) 'The Mathematics of Infectious Diseases', *SIAM Review*, 42: 599–653.
- Iceland Review (2020) COVID-19 in Iceland: Too Early to Declare Victory against Third Wave. <<https://www.icelandreview.com/society/covid-19in-iceland-too-early-to-declare-victory-against-third-wave/>> accessed 19 Oct 2020.
- Jason Wang., C. Ng, C. Y., and Brook, R. H. (2020) 'Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing', *JAMA*, 323: 1341–2.
- Karnakov, Petr, et al. "Data-driven inference of the reproduction number for COVID-19 before and after interventions for 51 European countries". *Swiss medical weekly* 150 (2020): w20313.
- Kuhnert, D. et al. (2016) 'Phylogenomics with Migration: A Computational Framework to Quantify Population Structure from Genomic Data', *Molecular Biology and Evolution*, 33: 2102–16.
- Lai, A. et al. (2020) 'Early Phylogenetic Estimate of the Effective Reproduction Number of SARS-CoV-2', *Journal of Medical Virology*, 92: 675–9.
- Lemey, P. et al. (2009) 'Bayesian Phylogeography Finds Its Roots', *PLoS Computational Biology*, 5: e1000520.
- Liu, Y. et al. (2020) 'The Reproductive Number of COVID-19 Is Higher Compared to SARS Coronavirus', *Journal of Travel Medicine*, 27: 1–4.
- Lu, F. S. et al. (2020) 'Estimating the Prevalence of COVID-19 in the United States: Three Complementary Approaches', *medRxiv*.
- MacIntyre, C. R. (2020) 'Case Isolation, Contact Tracing, and Physical Distancing are Pillars of COVID-19 Pandemic Control, Not Optional Choices', *The Lancet Infectious Diseases*, 20: 1105–6.

- Muller, N. F., Rasmussen, D., and Stadler, T. (2018) 'MASCOT: Parameter and State Inference' under the Marginal Structured Coalescent Approximation', *Bioinformatics*, 34: 3843–8.
- New Zealand Ministry of Health (2020) COVID-19 – Significant Clusters. <<https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-current-situation/covid-19-current-cases/covid-19-significant-clusters>> accessed 19 Jun 2020.
- New Zealand Ministry of Health, Manatu Hauora (2020) Information about Confirmed and Probable Cases of COVID-19 in New Zealand. <<https://www.health.govt.nz/our-work/diseases-and-conditions/covid19-novel-coronavirus/covid-19-current-situation/covid-19-current-cases>> accessed 20 May 2020.
- Nicola, D. M. et al. (2015) 'New Routes to Phylogeography: A Bayesian Structured Coalescent Approximation', *PLoS Genetics*, 11: e1005421.
- Price, David J., et al. "Early Analysis of the Australian COVID-19 epidemic", *Elife* 9 (2020): e58785.
- Rambaut, A. (2020) Phylodynamic Analysis — 176 Genomes — 6 Mar 2020. <<http://virological.org/t/phylodynamic-analysis-176-genomes6-mar-2020/356>> accessed 6 Mar 2020.
- RNZ (2020) NZ to Close Its Borders to Anyone Not a Citizen or Permanent Resident, PM Confirms. <<https://www.rnz.co.nz/news/national/412162/nzto-close-its-borders-to-anyone-not-a-citizen-or-permanentresident-pm-confirms>> accessed 19 Mar 2020.
- Rothan, H. A., and Byrareddy, S. N. (2020) 'The Epidemiology and Pathogenesis of Coronavirus Disease (COVID-19) Outbreak', *Journal of Autoimmunity*, 109: 102433.
- Ruiyun, L. et al. (2020a) 'Substantial Undocumented Infection Facilitates the Rapid Dissemination of Novel Coronavirus (SARS-CoV-2)', *Science*, 368: 489–93.
- Sandford, A. (2020) Coronavirus: Half of Humanity Now on Lockdown as 90 Countries Call for Confinement. Euronews. <<https://www.euronews.com/2020/04/02/coronavirus-in-europe-spain-s-death-toll-hits-10-000after-record-950-new-deaths-in-24-hou>> accessed 3 Apr 2020.
- Schengen Visa Info (2020) Iceland Starts Implementing Schengen and EU Travel Ban. <<https://www.schengenvisainfo.com/news/iceland-startsimplementing-schengen-and-eu-travel-ban/>> accessed 21 Mar 2020.
- Sciré, J. et al. (2020) 'Robust Phylodynamic Analysis of Genetic Sequencing Data from Structured Populations', *bioRxiv*.
- Seemann, T. et al. (2020) 'Tracking the Covid-19 Pandemic in Australia Using Genomics', *Nature Communications*, 11: 1–9.
- Shu, Y., and John, M. (2017) 'GISAID: Global Initiative on Sharing All Influenza Data – from Vision to Reality', *Eurosurveillance*, 22: 30494.
- Stadler, T. et al. (2013) 'Birth-death Skyline Plot Reveals Temporal Changes of Epidemic Spread in HIV and Hepatitis C Virus (HCV)', *Proceedings of the National Academy of Sciences of the United States of America*, 110: 228–33.
- Department of Health (2020) Coronavirus (COVID19) at a Glance 29 April 2020. States and Territories Report. Cases by Source of Infection. <<https://www.health.gov.au/sites/default/files/documents/2020/04/coronavirus-covid-19-at-a-glance-30-april-2020.pdf>> accessed 30 Apr 2020.
- Stats NZ (2020) Provisional International Travel Statistics. <[http://archive.stats.govt.nz/browse\\_for\\_stats/population/Migration/provisional-international-travel-statistics.aspx](http://archive.stats.govt.nz/browse_for_stats/population/Migration/provisional-international-travel-statistics.aspx)> accessed Jan–Mar 2020.
- Strong, M. (2020) Arrivals and Departures in Taiwan Plunge by 3 Million during First Quarter of 2020. Taiwan News. <<https://www.taiwannews.com.tw/en/news/3931135>> accessed 9 May 2020.
- Summers, J. et al. (2020) 'Potential Lessons from the Taiwan and New Zealand Health Responses to the Covid-19 Pandemic', *The Lancet Regional Health-Western Pacific*, 100044.
- Taiwan Centers for Disease Control (2020a) CECC Reports No New Confirmed Cases; 322 Patients Released From Isolation. <<https://www.cdc.gov.tw/En/Bulletin/Detail/shdoyOUaEjWMbkToD9DZEQ?typeid=158>> accessed 30 Apr 2020.
- (2020b) Virus Genome Sequence Disclosed; Case #39 Confirmed to Be Infected with COVID-19 in Egypt. <<https://www.cdc.gov.tw/En/Bulletin/Detail/4Xu2VavNqpCl2VwScBa0NA?typeid=158>> accessed 8 Mar 2020.
- Verrall, A. (2020) Rapid Audit of Contact Tracing for COVID-19 in New Zealand. <[https://www.health.govt.nz/system/files/documents/publications/contact\\_tracing\\_report\\_verrall.pdf](https://www.health.govt.nz/system/files/documents/publications/contact_tracing_report_verrall.pdf)> accessed 20 Apr 2020.
- Visual, F. T. and Data Journalism team. (2020) Coronavirus Tracked: The Latest Figures as Countries Fight to Contain the Pandemic. Financial Times. <<https://www.ft.com/content/a26fbf7e-48f8-11ea-aeb3-955839e06441>> accessed 28 May 2020.
- Cousins, S. (2020) 'New Zealand eliminates COVID-19', *Lancet*, 395: P1474.
- Volz, E. M., Koelle, K., and Bedford, T. (2013) 'Viral Phylodynamics', *PLoS Computational Biology*, 9: e1002947.
- World Health Organization (2020a) WHO Director-general's Opening Remarks at the Media Briefing on COVID-19 11 March 2020. <<https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020>> accessed 11 Mar 2020.
- (2020b) Coronavirus Disease (COVID-19) Situation Report – 101. <[https://www.who.int/docs/defaultsource/coronavirus/situation-reports/20200430-sitrep-101-covid-19.pdf?sfvrsn=2ba4e093\\_2](https://www.who.int/docs/defaultsource/coronavirus/situation-reports/20200430-sitrep-101-covid-19.pdf?sfvrsn=2ba4e093_2)> accessed 30 April 2020.
- Xingguang, L. et al. (2020b) 'Evolutionary History, Potential Intermediate Animal Host, and Cross-species Analyses of SARS-CoV-2', *Journal of Medical Virology*, 92: 602–11.
- Zhu, N. et al. (2020) 'A Novel Coronavirus from Patients with Pneumonia in China, 2019', *New England Journal of Medicine*, 382: 727–33.