

the fact that the Moon is generally depleted in volatile elements and compounds, which are chemical species that can readily vaporize. If the Moon formed from material ejected during a giant impact with Earth, such species would have been vaporized and lost at the high temperatures involved in this event¹³.

The concentrations of incompatible elements in the Chang'e-5 basalt samples are comparable to those typically found in KREEP, which is consistent with data from orbiting spacecraft. Tian *et al.* obtained an unexpected result, however, when they looked at the isotopic composition of strontium and neodymium – the isotopic abundances of these two elements are affected by radioactive decay. Their analysis suggests that the Chang'e-5 basalts derive from the melting of ancient sources in the lunar interior that are depleted in incompatible elements. These elements became enriched in the basalts because the magmas from which they crystallized came from minimal melting of the lunar interior and underwent large amounts of cooling and crystallization during their transit from source to eruption. Both the partial melting and the crystallization enriched the remaining magma in incompatible elements, but neither process affected the isotopic composition of the magma's source, which is depleted in these elements.

The age and compositional characteristics of the Chang'e-5 samples have at least two implications for our understanding of lunar structure and evolution. First, even though regions that have high concentrations of incompatible elements are seen across most of Oceanus Procellarum (Fig. 1), the Chang'e-5 data show that not all of these are composed of KREEP. This means that KREEP might form a much smaller component of the lunar interior than was suspected from Apollo samples, and data taken from orbit. Second, the melting that led to lunar volcanic eruptions two billion years ago did not involve heating due to high concentrations of radioactive elements in the magma source regions. Nor did it occur as a result of high water concentrations that reduce the melting temperature of rock in the same way that salt reduces the melting temperature of ice. An alternative explanation must now be sought for how the lunar interior was hot enough to drive volcanic eruptions two billion years ago.

The Chang'e-5 results demonstrate that samples returned from previously unvisited regions of the lunar surface can prompt a revision of models of lunar evolution that were developed on the basis of the Apollo and Luna samples. This is not surprising, given that the combined Apollo and Luna missions sampled only a restricted portion of the Moon's Earth-facing side. Much like Earth's surface, the lunar surface is a mosaic of materials created over the past 4.5 billion years. Each piece

of that mosaic provides different information about the history of the Moon. The Chang'e-5 results show that sample-return missions to previously unexplored portions of the lunar surface will help models of the evolution of our nearest planetary neighbour to eventually converge on reality.

Richard W. Carlson is in the Earth and Planets Laboratory, Carnegie Institution for Science, Washington DC 20015, USA.
e-mail: rcarlson@carnegiescience.edu

1. Spudis, P. *Nature Geosci.* **2**, 234–236 (2009).
2. Tian, H.-C. *et al. Nature* **600**, 59–63 (2021).
3. Hu, S. *et al. Nature* **600**, 49–53 (2021).

4. Li, Q.-L. *et al. Nature* **600**, 54–58 (2021).
5. Che, X. *et al. Science* **374**, 887–890 (2021).
6. Qian, Y. Q. *et al. J. Geophys. Res. Planets* **123**, 1407–1430 (2018).
7. Jolliff, B. L., Gillis, J. J., Haskin, L. A., Korotev, R. L. & Wieczorek, M. A. *J. Geophys. Res.* **105**, 4197–4216 (2000).
8. Warren, P. H. *Annu. Rev. Earth Planet. Sci.* **13**, 201–240 (1985).
9. Borg, L. E., Gaffney, A. M. & Shearer, C. K. *Meteorit. Planet. Sci.* **50**, 715–732 (2015).
10. Miljković, K. *et al. Science* **342**, 724–726 (2013).
11. Laneville, M., Wieczorek, M. A., Breuer, D. & Tosi, N. *J. Geophys. Res. Planets* **118**, 1435–1452 (2013).
12. Saal, A. E., Hauri, E. H., Langmuir, C. H. & Perfit, M. R. *Nature* **419**, 451–455 (2002).
13. Stevenson, D. J. *Annu. Rev. Earth Planet. Sci.* **15**, 271–315 (1987).

The author declares no competing interests.

Coronavirus

A reconstruction of early cryptic COVID spread

Simon Cauchemez & Paolo Bosetti

To respond better to future pandemics, we must understand how the SARS-CoV-2 virus dispersed so rapidly. A model of COVID-19 spread sheds light on cryptic transmission, undetected by surveillance efforts, in early 2020. **See p.127**

Over the past 20 months, the COVID-19 pandemic has caused more than 4.9 million reported deaths (<https://coronavirus.jhu.edu>), and measures to limit the spread of the SARS-CoV-2 virus have affected the lives of people around the world. Although modelling has helped to reconstruct the early dynamics of the epidemic in some countries, we still lack a coherent picture of how the pandemic unfolded globally. On page 127, Davis *et al.*¹ use a worldwide model to assess early, cryptic transmission of SARS-CoV-2 – the spread of the virus that was not detected by initial surveillance efforts – in the United States and Europe.

Looking back at the chronology of the first months of the pandemic, it is concerning how fast the virus spread around the world, leading to a massive shutdown of people's social and economic lives. On 10 January 2020, 41 cases of COVID-19 were reported in Wuhan, Hubei province, China. The first reports of infection outside China were made on 13 January (in Thailand) and 16 January (in Japan). Wuhan was locked down on 23 January, followed by lockdowns in Italy (11 March), Spain (14 March), Austria (16 March) and France (17 March). Many countries, caught off-guard by the rapidly changing situation, reported a large death toll. How can we do better next time? To answer this question and improve our preparedness in the face of

future pandemics, it is crucial to build a clearer picture of the initial spread of the virus. This is difficult, because the limited capacity to test for virus infections at the time meant that, in many locations, SARS-CoV-2 transmission might have been undetected.

Davis *et al.* used the Global Epidemic and Mobility (GLEAM) model, which has both stochastic (incorporating elements of randomness) and mechanistic (including defined principles about the biological and social mechanisms associated with viral infection and transmission) components to simulate virus spread on a global scale². The model relies on various types of data to capture the multifactorial nature of the epidemic process. This information includes data describing: the populations in which the virus spread, such as country-specific demographics; the movement of people on international and local scales (for example, airline transport networks and commuting flows); and behaviours, such as information documenting how individuals of different ages mingle with each other. The model also captures biological aspects of SARS-CoV-2 transmission, clinical features (such as lethality for each age group analysed) and the timing of non-pharmaceutical containment measures, such as lockdowns.

Using the model, the authors shed light on how the virus propagated around the world. For example, the model confirms that, at the

start (between January and March 2020), many introductions of the virus went unnoticed in both the United States and Europe (Fig. 1). In the United States, California was the first state to be affected by local transmission (week of 26 January, according to the model), even though evidence of local transmission there was reported only a month later (26 February). In Europe, modelling indicated that local transmission started at the end of January in Italy, the United Kingdom, Germany and France, again with substantial delays before local cases were reported. Overall, Davis and colleagues estimate that, by 8 March 2020, only 1–3 out of every 100 SARS-CoV-2 infections were detected in the United States and Europe.

Remarkably, in all 30 European countries studied and in all US states, local transmission started in a relatively narrow time window of about 2 months between mid-January and mid-March 2020. These variations in starting dates – as well as differences in the timing and strength of interventions – led to highly variable patterns of disease spread in different regions. By 4 July 2020, infection attack rates (the proportions of a region's population that contracted the virus) ranged from around 0.2% to about 15% in the analysed regions.

The modelling can help to assess the performance of testing strategies put in place at the time, which mostly affected passengers arriving from China. Although a large proportion of individuals with COVID-19 entering the United States and Europe indeed came from China in January 2020, neighbouring European countries and US states quickly became major contributors to infection rates in the areas analysed.

The lack of robust epidemic information on international locations affected by transmissions means that the testing policies in most countries were too narrow in scope and, probably, ineffective. Davis and co-workers provide a counterfactual scenario: if testing policies had been broader, enabling the detection of 50% of all imported and local infections in the United States and Europe, the start of local transmission might have been delayed by at least a month in many places. This would have given governments more time to prepare – for example, by increasing health-care capacities and acquiring protective equipment. However, considering the optimistic underlying assumption of this counterfactual scenario, it remains difficult to imagine that countries could have scaled up their testing capacity sufficiently fast to detect 50% of all initial SARS-CoV-2 infections.

There are other methods for reconstructing the spread of SARS-CoV-2 – for example, studies that sequence the RNA of the virus to determine the 'family tree' of circulating strains^{3,4}. In these approaches, the reconstructed history of viral circulation is informed by and consistent with all data available up to the

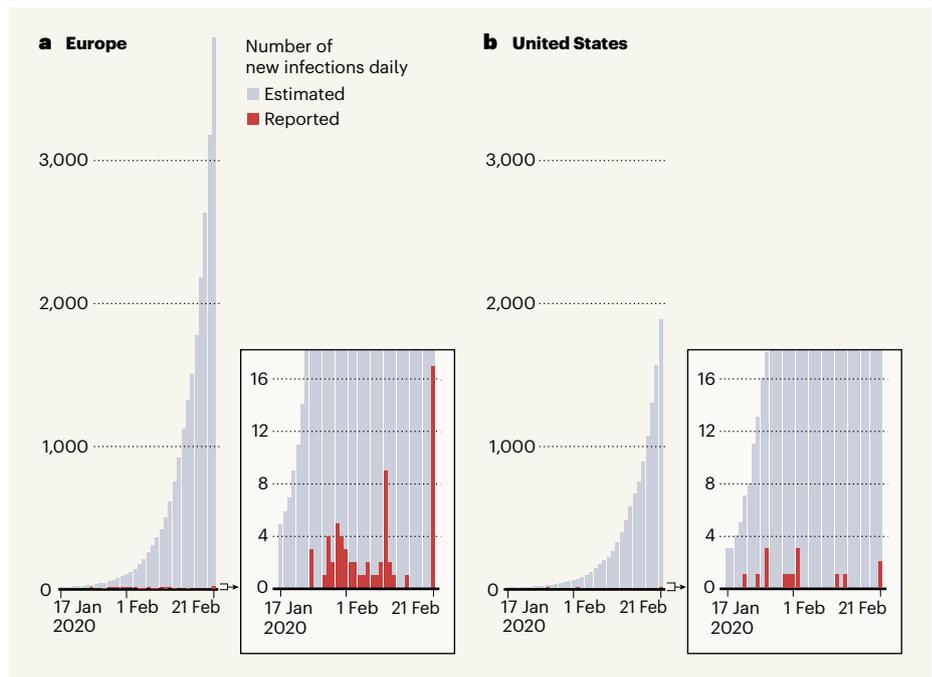


Figure 1 | Reported and estimated numbers of cases of COVID-19 in early 2020. Davis *et al.*¹ used a sophisticated epidemiological model to reconstruct the spread of the SARS-CoV-2 virus in 30 European countries and the United States at the beginning of the COVID-19 pandemic. They estimate that, between 17 January and 21 February 2020, only a tiny proportion of cases were detected by surveillance efforts, and that the rest of the estimated cases resulted from cryptic transmission.

time of analysis. By contrast, the simulations presented by Davis *et al.* are initialized to match data from the start of the pandemic, but are not constrained by what was observed after 21 January 2020. The authors therefore explore a broader range of possible trajectories of viral spread than do the phylogenetic studies, and some of these trajectories might not perfectly match what happened in reality. Despite this feature of the simulations, there is a relatively good correlation between reconstructed incidences of infections and observed numbers of cases.

Compared with other approaches, a key strength of Davis and co-workers' approach is that their model provides a mechanistic understanding of the spread of a pandemic. This makes it possible to model scenarios that document how the pandemic might have unfolded given different policy options, such as the counterfactual scenarios mentioned above that included more-intensive testing from the start of the pandemic. This capacity of models to capture and anticipate complex nonlinear dynamics and assess the potential effects of different policy options explains in large part why such modelling has increasingly been used to support policymaking, and this trend has been amplified by the pandemic. The creation of a Center for Disease Forecasting by the US Centers for Disease Control and Prevention is only the most recent example of this trend⁵.

Several exciting avenues should be considered for improving these models for future

use. Methods that calibrate these models to both epidemiological and virus-sequencing data could improve their performance for 'now-casting' and forecasting viral spread, as well as for assessing the effects of policies in real time. A more detailed understanding of the factors affecting spread (such as mobility, mixing patterns, behavioural changes, climate and population immunity) is necessary to improve model assumptions. Models are only as good as the data they rely on. Although COVID-19 surveillance and related data sets have considerably improved during the pandemic, these efforts must be maintained beyond COVID-19, and stable information systems^{6,7} (for example, <https://coronavirus.jhu.edu>) must continue to be made available to assess global disease dynamics in a robust manner.

Simon Cauchemez and **Paolo Bosetti** are in the Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, University of Paris, UMR 2000, Paris 75015, France. e-mail: simon.cauchemez@pasteur.fr

1. Davis, J. T. *et al.* *Nature* **600**, 127–132 (2021).
2. Balcan, D. *et al.* *J. Comput. Sci.* **1**, 132–145 (2010).
3. Fauver, J. R. *et al.* *Cell* **181**, 990–996 (2020).
4. du Plessis, L. *et al.* *Science* **371**, 708–712 (2021).
5. CDC Stands Up New Disease Forecasting Center. <https://www.cdc.gov/media/releases/2021/p0818-disease-forecasting-center.html> (CDC, 2021).
6. Hale, T. *et al.* *Nature Hum. Behav.* **5**, 529–538 (2021).
7. Hadfield, J. *et al.* *Bioinformatics* **34**, 4121–4123 (2018).

P.B. declares competing interests. See [go.nature.com/3mm00ax](https://www.nature.com/3mm00ax) for details.

This article was published online on 1 November 2021.