



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

## Perspective

## Toward a global virus genomic surveillance network

Verity Hill,<sup>1,\*</sup> George Githinji,<sup>2,3</sup> Chantal B.F. Vogels,<sup>1,4</sup> Ana I. Bento,<sup>5,6</sup> Chrispin Chaguza,<sup>1,4</sup> Christine V.F. Carrington,<sup>7</sup> and Nathan D. Grubaugh<sup>1,4,8,9,\*</sup>

<sup>1</sup>Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA

<sup>2</sup>KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya

<sup>3</sup>Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya

<sup>4</sup>Yale Institute for Global Health, Yale University, New Haven, CT, USA

<sup>5</sup>Department of Epidemiology and Biostatistics, Indiana University School of Public Health-Bloomington, Bloomington, IN, USA

<sup>6</sup>The Rockefeller Foundation, New York, NY, USA

<sup>7</sup>Department of Preclinical Sciences, The University of the West Indies, St. Augustine Campus, St. Augustine, Trinidad and Tobago

<sup>8</sup>Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

<sup>9</sup>Public Health Modeling Unit, Yale School of Public Health, New Haven, CT, USA

\*Correspondence: [verity.hill@yale.edu](mailto:verity.hill@yale.edu) (V.H.), [nathan.grubaugh@yale.edu](mailto:nathan.grubaugh@yale.edu) (N.D.G.)

<https://doi.org/10.1016/j.chom.2023.03.003>

## SUMMARY

The COVID-19 pandemic galvanized the field of virus genomic surveillance, demonstrating its utility for public health. Now, we must harness the momentum that led to increased infrastructure, training, and political will to build a sustainable global genomic surveillance network for other epidemic and endemic viruses. We suggest a generalizable modular sequencing framework wherein users can easily switch between virus targets to maximize cost-effectiveness and maintain readiness for new threats. We also highlight challenges associated with genomic surveillance and when global inequalities persist. We propose solutions to mitigate some of these issues, including training and multilateral partnerships. Exploring alternatives to clinical sequencing can also reduce the cost of surveillance programs. Finally, we discuss how establishing genomic surveillance would aid control programs and potentially provide a warning system for outbreaks, using a global respiratory virus (RSV), an arbovirus (dengue virus), and a regional zoonotic virus (Lassa virus) as examples.

## INTRODUCTION

Virus evolution and genetic diversity can impact outbreak dynamics and control efforts. This has been demonstrated to politicians, public health professionals, and the public by the response to the COVID-19 pandemic, in particular with the emergence of variants. Monitoring evolutionary processes can provide answers that are hidden from traditional epidemiology, both on a small scale by investigating the details of transmission chains<sup>1</sup> and by uncovering the pathways of spread on a global scale.<sup>2</sup> Genomic data also has the potential to enhance disease forecasting models<sup>3</sup> and has been critical for development of vaccines, therapeutics, and molecular diagnostic assays.<sup>4,5</sup> Overall, virus sequencing can be used to better design, apply, and evaluate strategies to mitigate transmission and disease.

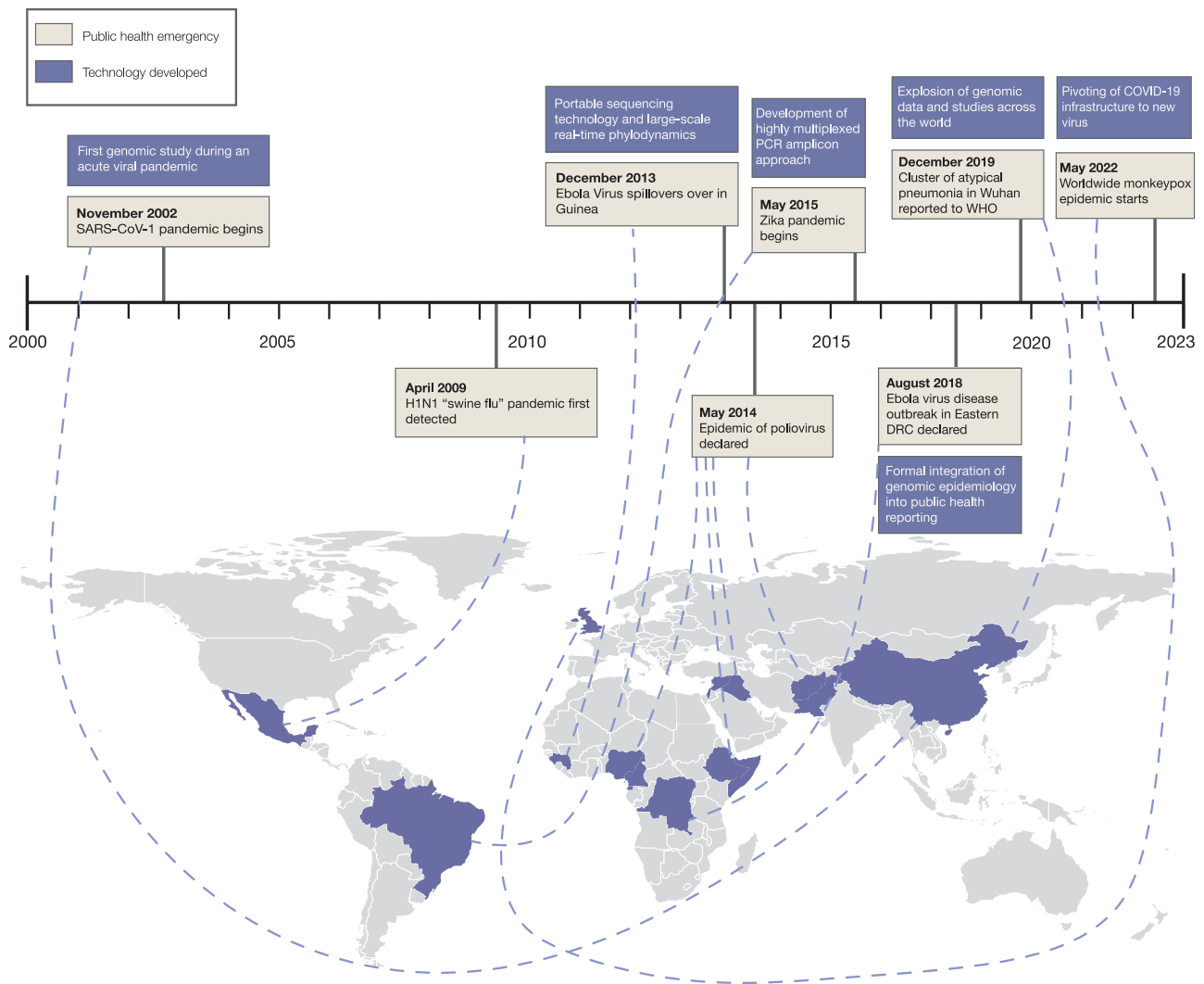
We call for global virus genomic surveillance networks to provide routine data on viral evolution and lineage transmission dynamics during and between outbreaks. These programs do not need to be built from scratch but should utilize the numerous advancements in sequencing technology that facilitated initiatives for influenza virus and HIV, outbreak responses to emerging viruses, and the COVID-19 pandemic. At its core is the implementation of a nimble and adaptable framework, allowing labs to sequence and analyze several different viruses with the same protocol and quickly respond to emerging threats. Further devel-

opment of globally adopted genotype and lineage classification systems would standardize the viral nomenclature and enhance the usability and comparability of the data, and developing better ways of fair data sharing will ensure that they are broadly usable. Multilateral support from governments, non-governmental organizations, and the biotech industry, which all benefit from genomic surveillance systems, is required to support these efforts. Sufficient long-term investment, buy-in, training, and equity for genomic surveillance can transform the way our global public health community can respond to both epidemic and endemic viruses.

## MOVING SEQUENCING INTO PUBLIC HEALTH

Over the last two decades, the output of virus genomic sequencing has moved from primarily retrospective research toward near real-time analyses. While the former established the field and is important in its own right for examining longer-term dynamics, real-time analyses provide actionable results for public health interventions and a clearer return on investment. This transition to real-time assessments established virus genomic sequencing as a legitimate tool for public health, moving away from its perception as academic “stamp-collecting.” These changes were made possible due to innovations in sequencing,





**Figure 1. Timeline of public health emergencies since 2000 and associated technological advances**

White boxes indicate the timing of the start of epidemics, either based on when the spillover was inferred to happen, or when epidemics were reported or declared. Purple boxes indicate technological advancements. Countries are colored based on where the epidemic was first detected.

computational technology, and methodological advancements, as well as new-found demand for rapid results (Figure 1).

Underlying many of the methodological developments of virus genomic surveillance, as well as much of the physical infrastructure and capacity, is the response to the ongoing HIV pandemic. Due to its high fatality rate when untreated, lack of a vaccine or cure, and socio-political importance, HIV is a priority pathogen across the globe. It also has a relatively complex genome for an RNA virus, with circulating recombinant forms and the ability to evolve resistance to drug treatments. For these reasons, genomic surveillance provides a clear benefit. For example, in Portugal, HIV genomes are sequenced as part of routine clinical care and stored in the HIV drug-resistance database.<sup>6</sup> Phylodynamic analysis of these genomes contributed to the success of Portugal's drug decriminalization policy along with other harm reduction measures for people who inject drugs and an antiretroviral therapy campaign in reducing the HIV burden.<sup>6</sup> In addition to the critical role that vi-

rus sequencing has for HIV programs, these advancements helped to facilitate genomic surveillance for other public health emergencies.

Following HIV, the first epidemic of the 21st century began in November 2002 in rural China. Approximately four months later, the cause was identified as a novel coronavirus (later renamed SARS-CoV-1) and the first whole-genome sequences were submitted to GenBank in April 2003.<sup>7</sup> As the first next-generation sequencing platform (454's GS20) was not introduced until 2005, rapid sequence generation was not possible, and initial analysis was performed on 14 SARS-CoV-1 genomes, produced months after the pandemic began.<sup>8</sup> After the pandemic, a study based on 61 virus sequences identified mutational signatures common to different phases of the epidemic to explore evidence of human adaptation.<sup>9</sup> Therefore, while the SARS pandemic can be viewed as part of the "genomic age," genomic analysis beyond identification was limited and retrospective, largely due to technological limitations.

Six years later, further technological development enabled rapid sequencing during the 2009 H1N1 influenza A pandemic. Genomic analyses elucidated its animal origins while the pandemic was still spreading,<sup>10</sup> highlighting the need for human and animal influenza surveillance. With endemic viruses that have epidemic potential, the financial and political urgency generated during emergencies must be maintained to monitor viral diversity. Systematic genomic surveillance has been implemented for influenza A virus, and these data are used by the World Health Organization (WHO) to predict antigenic evolution and improve vaccine strain selection.<sup>11</sup> Further, by regularly sequencing human viruses, it is possible to identify and track markers which may impact transmissibility and vaccine effectiveness—for example, in the 2016–2017 influenza season in the Northern Hemisphere, many viruses were detected with a mutation impacting the effectiveness of egg-based vaccines.<sup>12,13</sup>

Genomic surveillance is especially important as part of eradication or elimination programs. In 2014, a polio emergency was declared as multiple countries had circulation of wild-type poliovirus, distinguished from vaccine-derived virus through genomic surveillance.<sup>14</sup> The difference between these two viruses is important, as optimal control strategies vary in terms of the balance between using the injected or oral polio vaccines and are indicative of different failures in the eradication strategy.<sup>14</sup> Genomic surveillance also complements traditional clinical polio surveillance that tracks cases of acute flaccid paralysis, which occurs in less than 1% of infected individuals.<sup>15</sup> Finally, routine sequencing can differentiate between outbreaks caused by new introductions versus cryptic local transmission, identifying high-risk populations that should be targeted for vaccination.<sup>16</sup> Real-time genomic surveillance, therefore, is a crucial element of the poliovirus eradication effort.

The 2013–2016 Ebola virus disease epidemic (EVD) in West Africa marked a watershed moment for near real-time viral genomic surveillance in public health. A total of 1,610 (~5%) cases were sequenced during the epidemic,<sup>17</sup> the largest such dataset until the COVID-19 pandemic. This was in part possible due to advancements in sequencing technology, most notably the Oxford Nanopore Technology (ONT) MinION, released in 2014. The machine's small size makes it the perfect tool as part of a “lab-in-a-suitcase”, able to be transported on a commercial flight.<sup>18</sup> The MinION also does not require a continuous mains-power supply and can be run off a laptop battery, enabling it to be used in contexts with unreliable power supplies. While most of the sequences in this Ebola virus dataset were not produced with an ONT platform, this represented a major development in the democratization of sequencing and increased the speed with which future epidemics could be analyzed.<sup>19</sup> During the epidemic, virus sequencing was used in part to identify sexual transmission of Ebola virus in two clusters.<sup>20,21</sup> Both clusters were small flare-ups after transmission in the area had ceased toward the end of the epidemic, and sequencing revealed that cases in the cluster were related to each other and did not constitute a new zoonosis. Further, by using the background genomic surveillance data that had been sampled throughout the epidemic, the latter cluster was shown to be most closely related to a lineage last sampled two years previously, indicating viral latency.<sup>21</sup> Thus, genomic surveillance of the whole epidemic pro-

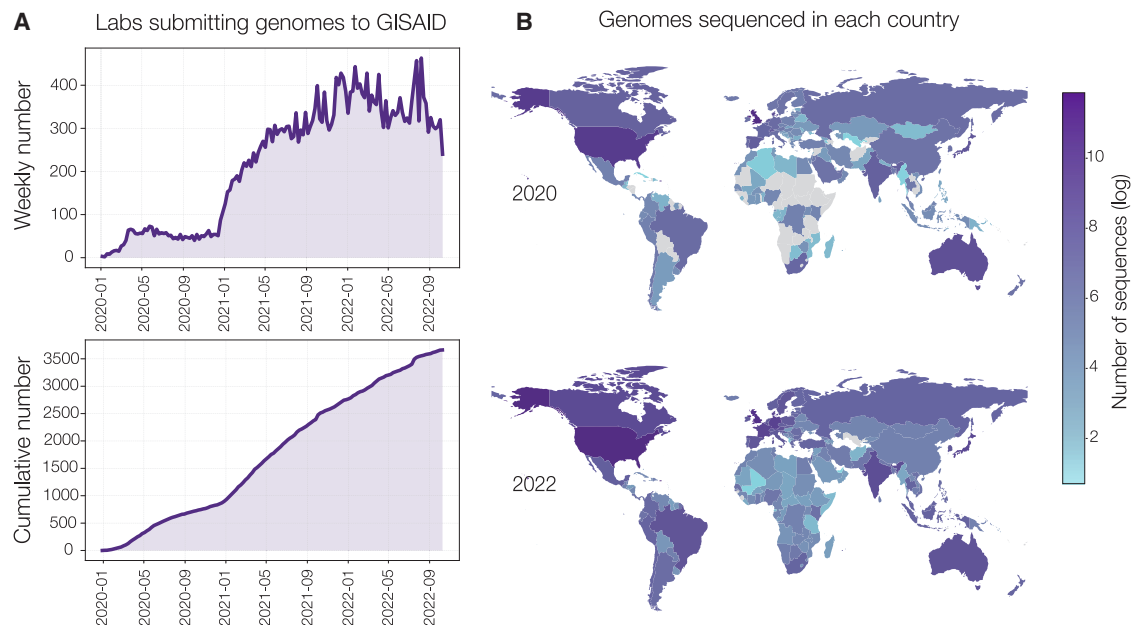
vided the context for these clusters and evidence of a new mode of transmission.

The 2016 Zika epidemic in the Americas also ushered in new technology to improve the scale of virus genomic surveillance. Sequencing total RNA dilutes the viral “signal,” which is especially problematic with clinical samples that often have low viral loads. A highly multiplexed PCR amplicon approach (described in more detail below) was piloted during the EVD epidemic<sup>18</sup> and fully implemented for Zika virus out of necessity—sequencing had limited success with other methods.<sup>22,23</sup> As later demonstrated during the COVID-19 pandemic, the real advantage of the amplicon approach is that it massively scaled sequencing due to its high per-sample success rate. Amplicon sequencing, combined with the portable MinION, facilitated a mobile lab to sequence Zika virus across Brazil.<sup>24</sup> These data were initially used to disentangle the timings of the origin of the virus into South America,<sup>2</sup> as well as later highlighting the importance of northeastern Brazil in seeding multiple locations across Latin America.<sup>22</sup>

In the time since the West African epidemic, Ebola virus genomic surveillance has become more commonplace, in particular in the Democratic Republic of the Congo. Almost 800 cases were sequenced at regular intervals during the course of the 2018–2020 epidemic by the Institut de Recherche Biomedicale in Kinshasa. The team analyzed the sequences mostly within a month of sample collection to produce genomic epidemiology situation reports in French and English.<sup>25</sup> This genomic surveillance identified superspreading events connected to clergy and motorbike taxi drivers, and the vaccination policy was subsequently changed to include these professions as high-risk.<sup>25</sup> Importantly, this team sequenced 24% of recorded cases, all conducted in-country along with most of the bioinformatic analysis with far less reliance on external collaborators than during the 2013–2016 epidemic. This rapid and systematic sequencing program is built on the experience of previous emergencies, and it represents a significant step forward in the inclusion of genomic surveillance for outbreak response and sustainable capacity in viral sequencing.

## MONUMENTAL SEQUENCING EFFORTS DURING THE COVID-19 PANDEMIC

The sequencing effort during the COVID-19 pandemic is unprecedented, both in terms of the number of cases sequenced (over 14 million) and the number of countries producing their own genomic data<sup>26,27</sup> (Figure S1). Within two months of the first SARS-CoV-2 sequence being released by China Centers for Disease Control and Prevention (CDC) on January 11th, 2020,<sup>28</sup> 25 additional countries had submitted at least one sequence to the data repository GISAID. While some countries were prepared to sequence a novel virus, many leveraged the urgency of the pandemic to rapidly build capacity, particularly in Africa and the Caribbean. Substantial investment during 2021 enabled the Africa CDC to set up or support ten African sequencing hubs.<sup>29</sup> The impact of this investment is measurable (Figure 2): in April 2020, 20 of the 55 (36%) African countries had in-country SARS-CoV-2 sequencing capacity, and by mid-2022, this number had increased to 39 (71%); and both the volume of sequences and the turnaround time substantially improved



**Figure 2. Global SARS-CoV-2 sequencing effort during the COVID-19 pandemic**

(A) The number of laboratories worldwide submitting genomes to GISAID, by week and cumulative. (B) The log number of genomes sampled in each country between 2020 and 2022.

between 2020 and 2021.<sup>29,30</sup> A similar trend was observed in the Caribbean, which had limited local capacity for viral genomic surveillance prior to the pandemic. In late 2020, ONT MinION sequencing was implemented at The University of the West Indies in Trinidad and Tobago, which provided capacity for 17 member states of the Caribbean Public Health Agency (CARPHA).<sup>31</sup> This continued until December 2021, when CARPHA developed its own in-house capacity. In addition, at least 8 of the 30 Caribbean islands now have sequencing capacity and produced their own SARS-CoV-2 sequences. Worldwide, by September 2022 over 3,500 labs from 216 countries or territories produced and shared SARS-CoV-2 sequencing data (Figures 2A and S1A).

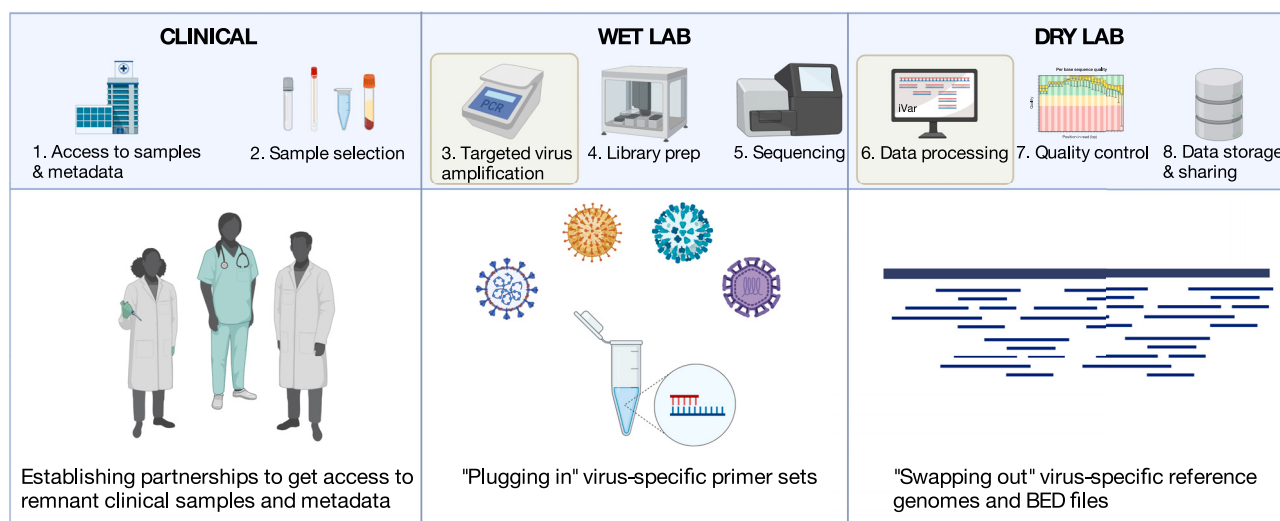
While many sequences were generated within the country from which they were sampled, one of the success stories of the COVID-19 pandemic was the development and strengthening of regional sequencing centers. Some, such as the Center for Epidemic Response and Innovation (CERI) and Kwazulu-Natal Research Innovation and Sequencing Platform (KRISP) in South Africa, the African Center of Excellence for Genomics of Infectious Diseases (ACEGiD) in Nigeria, and Institut Pasteur in Senegal were founded long before the pandemic began. Others, such as the COVID-19 Genomics UK (COG-UK) consortium, were built rapidly as the pandemic accelerated in 2020. This speed was possible due to long-standing academic and public health partnerships and decades of previous investment and training.<sup>32</sup> Clinical and academic laboratories dispersed across the UK, which already contained sequencing equipment and expertise, pivoted rapidly to sequencing SARS-CoV-2. Sequences from the hospital-based and centralized community testing were submitted to the same server, and (along with GISAID data) were passed through the same data cleaning and processing pipeline every day<sup>33</sup> (<https://github.com/COG-UK/>

[grapevine/](https://github.com/COG-UK/)). The outputs of this pipeline included metadata, alignments, and a phylogeny, which were then passed to programs such as CIVET<sup>34</sup> for rapid cluster investigations and close to real-time phylodynamic analysis and reporting. This had the dual benefit of enabling rapid reporting to governmental funders, as well as allowing data producers to see real benefit of their work on a local level in terms of infection control.

A key technological advancement during the SARS-CoV-2 pandemic was the development of a lineage classification system, providing a relatively universal, high-resolution picture of the genetic diversity of SARS-CoV-2.<sup>35</sup> Along with its associated tool (“pangolin”) to assign those lineages,<sup>26</sup> which also has a user-friendly web browser version (<https://pangolin.cog-uk.io/>), researchers, clinicians, and policy-makers have a common language with which to discuss SARS-CoV-2 genetic diversity in their region, without having to either generate or interpret phylogenies. There were also other nomenclature systems developed in 2020, such as the Nextstrain clade system, which aimed to describe the phylogeny in a broader and more stable way, by having criteria for minimum size and persistence; as compared to pango lineages which are finer scale to capture outbreaks and ongoing transmission in close to real-time.<sup>36</sup> However, the clade system became less commonly used as the pandemic progressed, and users coalesced more around pango lineages. In general, having a single nomenclature system in place prior to a major outbreak of a disease, preferably maintained by an independent organization like the WHO rather than academics, would improve communication early on. Along these lines, the WHO set up a system of naming variants of concern with Greek letters.<sup>37</sup> This aimed to make phenotypically important lineages easier to discuss and to avoid the use of stigmatizing names.

The enormous amount of sequencing data generated during the pandemic contributed to thousands of high-impact papers





**Figure 3. Modular amplicon-based virus sequencing framework**

Steps where swapping of key elements make the framework generalizable to other viruses are highlighted in yellow. Created with [BioRender.com](https://www.biorender.com).

on national,<sup>38</sup> regional,<sup>39</sup> and transcontinental virus dynamics and evolution.<sup>40</sup> Importantly, this includes the discovery of novel variants with estimates of their transmissibility and immune evasion properties.<sup>41</sup> It is notable that sequencing efforts increased around the time of waves caused by new variants (Figure 2A; see January 2021 with the rise of Alpha, Beta, and Gamma variants), indicating when it was the easiest to justify the expenditure on genome sequencing. On the flip side, as much of the world is scaling back pandemic control efforts, fewer countries are submitting sequences to GISAID in 2022 (Figure S1). Now, the goal globally must be to not lose the financial and training investment developed during the pandemic and to invest some of it into a rapid response system for the next pandemic, as well as applying it to endemic viruses.

## MODULAR VIRUS SEQUENCING FRAMEWORK

The new and expanded sequencing capacity implemented in response to the COVID-19 pandemic provides an opportunity to create global genomic surveillance systems for other medically important viruses. As the volume of samples needing to be sequenced for SARS-CoV-2 declines, those pipelines should be diverted rather than closed—it is much easier to dial complex programs up or down rather than on or off. This helps to maintain access to diagnostic samples, employment of trained personnel, equipment functionality, supply stocks, database management, and, perhaps most importantly, rationale for sustained funding.

Transitioning between different sequencing approaches, however, can be harsh. Many SARS-CoV-2 sequencing labs, both large and small, invested heavily in developing optimal procedures and well-trained staff. If the change in approach is major, including needing to purchase new reagents and/or equipment and learn/validate new protocols, the lab may deem the onboarding costs too high to justify, and the pipeline turns off. Thus, the key to diverting sequencing capacity is protocol continuity.

Most labs routinely sequencing SARS-CoV-2 include these core components (Figure 3).

- (1) Access to remnant diagnostic samples and associated metadata
- (2) Sample selection
- (3) Targeted virus amplification or enrichment
- (4) Preparation of sequence libraries
- (5) Sequencing
- (6) Data processing and consensus sequence generation
- (7) Data quality control (QC) checks
- (8) Data upload to public databases and storage

Each step could change when switching to a new virus, but some of the most drastic differences are often the protocols to prepare the samples for sequencing (steps 3 & 4) and the downstream data processing (step 6). Most labs sequencing SARS-CoV-2 (~30-kb RNA virus) utilize a highly multiplexed PCR reaction to amplify the virus (step 3), a library preparation optimized for amplicons (step 4), and a bioinformatics pipeline that removes primer sequences and aligns the reads to the SARS-CoV-2 reference genome (step 6). In comparison, the standard approach for sequencing human monkeypox virus (~200-kb DNA virus) was an untargeted DNA metagenomic protocol that sequences most DNA in a sample.<sup>42</sup> This approach, however, uses completely different reagents and protocols for steps 3 and 4, and a different bioinformatic pipeline for step 6, creating an entry barrier for labs without this expertise. Additionally, while metagenomic sequencing is critical for detecting unknown pathogens, it is far less efficient than targeted amplicon sequencing for viruses. A simpler solution is to adapt the existing SARS-CoV-2 pipeline for monkeypox virus, as has been demonstrated with minimal protocol changes.<sup>3</sup> By swapping out the PCR primer sets (step 3), the same library prep kits can be used (step 4) with the lab's preference for sequencing platform (e.g., ONT, Illumina, Ion Torrent, BGI; step 5). Then the bioinformatic pipelines only need to be updated with the new

reference files (step 6). This modular approach allows labs to use a single base set of wet and dry lab protocols for almost any virus of interest (Figure 3).

Many virus sequencing protocols are already the product of adaptation. The multiplexed PCR-based sequencing approach used for SARS-CoV-2 is a derivative of a protocol originally designed for Zika virus<sup>23</sup> and is similar to approaches for sequencing Ebola virus and historical HIV samples.<sup>18,43</sup> These protocols all use dozens, sometimes over a hundred, of overlapping primer pairs to amplify the target genome in two PCR reactions. A primer design tool called “PrimalScheme” ([primalscheme.com](http://primalscheme.com)<sup>18,23</sup>) makes it easy to design amplicon schemes for other viruses—for example, sets for chikungunya, Yellow fever, and West Nile,<sup>44,45</sup> in addition to the widely used SARS-CoV-2 schemes. Many other renditions of the base virus amplicon protocol have also been developed,<sup>46,47</sup> and primer schemes are relatively easy to design with adequate genomic information. Then, regardless of what amplicon library prep kits or sequencing platforms a lab may use, custom or validated primer schemes can be easily exchanged to sequence different viruses.

A generalizable modular sequencing approach also simplifies the bioinformatic pipelines used to process the data. For example, software packages like iVar,<sup>44</sup> which utilizes standard programs for analyzing sequencing data (e.g., BWA and SAMtools<sup>48,49</sup>), are designed for amplicon sequencing pipelines and are virus-agnostic: the user provides the primer scheme genome coordinates (i.e., BED file for primer sequence trimming) and the reference genome (i.e. FASTA file for aligning reads to create a consensus genome). There are likely magnitudes more custom bioinformatic pipelines than there are wet lab protocols. The key is to design them to be “plug and play” with validation checks to ensure that the pipelines can process different virus genome structures (e.g., lengths and segments) and alignment depths. As amplicon-based sequencing data are notoriously prone to contamination, these modular strategies can also help to standardize the quality control steps to ensure accurate data releases and reporting.

Modular protocols can make it feasible and more economical to establish routine sequencing of multiple endemic viruses that are important to the region. Importantly for when a new virus arrives, it also provides the framework to rapidly respond.

### OPPORTUNITIES TO OVERCOME EXISTING CHALLENGES

Even with available sequencing tools, establishing virus genomic surveillance networks requires addressing several inequities. Despite the monumental efforts to sequence >14 million COVID-19 cases, SARS-CoV-2 sequencing was still unbalanced.<sup>27</sup> This is because investment in public health is often lacking, especially in low-resource settings, where genomic surveillance efforts compete with the more fundamental needs that support testing and medical countermeasures.<sup>50</sup> Thus, countries must first and foremost improve their disease surveillance systems, which would then create the networks and systems to integrate virus genomics into public health.

At the forefront of the challenges is funding. Even in well-resourced countries like the US, public health is woefully underfunded despite its significant return on investment.<sup>51</sup> Typically,

public health funding follows emergencies, leaving gaps in routine activities as priorities shift. Continuously restarting technical systems like genome sequencing—requiring hiring and/or training, purchasing expensive equipment, and validating pipelines—is an inefficient way to spend money. Rather, long-term support is needed to maximize the returns. In well-resourced countries, the major obstacle is political buy-in. SARS-CoV-2 sequencing during the COVID-19 pandemic demonstrated its value, with even some influential politicians calling for its increased usage. The real challenge is convincing leaders of the value of routinely sequencing medically important endemic viruses as a standard component of surveillance, which could provide the biggest benefit to public health. Not only could it enhance forecasting<sup>3</sup> and the precision of control methods (see Case Studies section below), but it would also create the infrastructure to rapidly and efficiently respond to outbreaks. Thus, we urgently need more studies that highlight the financial return on investment that routine virus sequencing provides to help justify its permanence in public health.

For global virus genomic surveillance systems to be successful, strengthening sustainable efforts in low and middle-income countries (LMICs) is key. Here, convincing political leaders for access to funding may not be the primary barrier if those resources do not exist. One solution is for new grants focused on long-term programs that encourage local sequencing to be specifically awarded to principal investigators in LMICs. These investments, if made to be sustainable, would thereby build capacity and shift leadership opportunities from high income countries to LMICs. In addition, the biotech industry should directly support the system from which they profit. For example, Pfizer is predicted to make \$34 billion in 2022 from the bivalent COVID-19 vaccine that was only possible through global variant surveillance,<sup>52</sup> but the company did not provide any financial support (or bivalent vaccine doses) to South Africa where Omicron was first identified. Without sustained support, LMICs are forced to find ways to make available resources go further. The modular sequencing framework that we describe above would thus be advantageous as the same reagents could be used for many different viruses. Portable sequencers like the ONT MinION can also reduce the entry price, and they can be scaled for bulk sequencing. This also eliminates the annual license and service costs that come with large benchtop sequencing platforms. Finally, smart sampling designs can help to ensure the most value per genome, as routinely sequencing a large proportion of the cases (e.g., >10%) may not provide much new information for surveillance.<sup>27,53</sup>

Exacerbating the funding issues is that the costs of obtaining equipment and reagents in many LMICs can be 10× higher than in Europe and North America. Further, significant delays in customs can lead to reagent degradation,<sup>25</sup> and materials from some vendors are not even possible to obtain (e.g., ONT in parts of South Asia). We desperately need a call to action to limit price inflation from secondary vendors and expand distribution centers across the Global South to ensure access to materials at a lower cost. Hopefully, if routine sequencing increases in LMICs, vendors will make a concerted effort to target these emerging markets for their businesses. These suggestions, however, do not address all of the challenges of balancing genomic surveillance.

Access to training programs or a trained workforce can impede the establishment of routine sequencing in many environments, but this is an even greater concern in LMICs. In the wet lab, learning the sequencing protocols is relatively straightforward as it mostly uses standard techniques like PCR and purification. The main focuses for the wet lab training should be on workflows and QC to prevent and monitor contamination, protocol troubleshooting, and operating the sequencing platforms. The dry lab is a greater challenge as there are fewer public health professionals with bioinformatics training. It may not be difficult for someone with basic knowledge of the command line to run a data processing pipeline, but this should not be a “black box” in case of errors or the need to adapt. Having remote collaborators is a temporary but useful solution, especially for skill transfer in the early stages of setting up a genomic surveillance system in LMICs. However, autonomy through local expertise is best for the long-term success of a program. Hence, expanding training programs like those listed in [Table S1](#) are crucial for genomic surveillance systems.

Centralized sequencing, on a national or regional level, presents a solution to many of the above issues. Its main advantage is the focusing of funding; instead of many labs competing for limited resources, larger centers can apply for larger grants from international organizations. This greater level of funding can then translate into several key elements. Higher throughput sequencing equipment is expensive to buy but is ultimately cheaper per sample,<sup>54</sup> making sequencing more cost-effective. Higher-powered computational equipment means that more complex analyses can be done faster and coherent datasets can be generated in real-time, as in COG-UK (see above). Larger, well-funded centers also attract the top talent in the region, focusing expertise in a single area and enabling successful collaborations. This expertise can then be used to advocate for a region, both geopolitically (as with campaigning for cheaper reagents) and in building wider capacity through training. Finally, the efficiency of sequencing centers carries through into sample selection. With a top-down overview, it becomes possible to select representative clinical samples based on location and time (Box 1 in Hill et al.<sup>55</sup>), allowing more viral diversity in the region to be sampled with fewer resources.

There are specific concerns that must be addressed if a centralized model is adopted. First, specimen transport is not a trivial issue, and networks must be set up to ensure rapid and safe transport, especially important when dealing with clinical samples of high biosafety-level pathogens.<sup>25</sup> This may be mitigated partially by building on existing non-genomic laboratory networks, such as those set up for diagnosing existing priority pathogens. Connected to this, laboratory information management systems (LIMS) that keep metadata connected to specific samples are integral. Successful LIMS can rely on simple measures like consistent serial numbers, or more complex ones including scanning sample barcodes at each stage of its transport. Keeping accurate labels and dates is vital for correct analysis further downstream, as found when the evolutionary rate of Ebola virus was underestimated due to taxon label misassignment.<sup>56</sup>

Maintaining local ownership of the data (sovereignty) and associated buy-in by sample collectors can also be tricky with a centralized approach. This can be mitigated by the fast turn-

around of sequences into readable reports or dashboards like Nextstrain,<sup>57</sup> allowing local sample collectors to act on their data where necessary, and to see where their region fits into the wider picture. A major drawback of a centralized system is a lack of agility: it is much harder to focus on locally important questions, such as a hospital outbreak, and results can be slower. Finally, centralization can lead to a lack of local capacity development, possibly focusing human and financial resources on locations or countries that already have better capacity, exacerbating existing inequities. Therefore, the ideal would be a balance of centralized and localized capacity to maximize efficiency without losing the flexibility and speed of small-scale sequencing and maintaining capacity at a local level.

The above does not solve all of the issues with creating balanced global surveillance systems. The ultimate goal, one that is easier said than done, is to (re)distribute the resources needed for routine sequencing to create an equitable system. As demonstrated by the emergence of SARS-CoV-2 variants, international surveillance is as important for local public health actions as local surveillance.

## ALTERNATIVE APPROACHES TO SUPPLEMENT CLINICAL SEQUENCING

Routinely sequencing from local outbreaks is not the only way to conduct virus genomic surveillance. Gaps in resources, lab capacity, and access to clinical diagnostic samples can delay the implementation and routine availability of bulk sequencing. Thus, it is important to have alternative strategies to supplement the sequencing of diagnosed cases to still obtain a population-level view of circulating lineages. Moreover, sequencing environmental and (in)vertebrate animal samples can help to reveal epidemiological patterns that would otherwise be hidden by only using human samples. The exact approaches and sample types rely heavily on the viruses of interest.

International travel surveillance can supplement regional surveillance by essentially using humans as sentinels. It can be used to sample neighboring countries at border crossings or even the whole world at transport hubs. This is especially important when the country itself is not willing or able to perform its own genomic surveillance. For example, diagnostic testing and sequencing of travelers entering Florida, US, uncovered an unreported Zika outbreak in Cuba in 2017, a year after the epidemic was thought to be over.<sup>58</sup> Combined with new phylodynamic methods to explicitly include travel data in the reconstruction of ancestral node locations,<sup>59</sup> this information can be powerful. Travel sampling can be integrated into existing points of entry control measures, performed at borders to help prevent the spread of an infectious disease of interest, and so is relatively cost-efficient to set up.<sup>55,60</sup> Without local surveillance, however, simply sampling travelers may only provide a limited picture of locally circulating lineages and transmission dynamics.

Wastewater genomic surveillance can be useful in monitoring the levels of virus in a given community when a sufficient amount is known about how case counts correspond to the virus concentration in sewage. It has been used extensively in detecting poliovirus because only 1% of those infected have recognizable symptoms. Recently, poliovirus was unexpectedly detected in wastewater in London, leading to a vaccination campaign in



young children. After subsequent sequencing, this outbreak was connected to poliovirus detected in wastewater and New York City and Israel.<sup>61</sup> This approach has been extensively validated for monitoring SARS-CoV-2 transmission and variant detection.<sup>62,63</sup> Wastewater sampling is also an avenue for sampling other animal reservoirs of disease, for example, the possibility of cryptic SARS-CoV-2 lineages originating from rats in the sewers of New York City.<sup>64</sup> While the use of wastewater for virus genomic surveillance will likely dramatically rise in years to come, it is important to highlight that the non-linear relationship between clinical cases and the amount of virus in sewage is pathogen specific and thus not trivial to properly interpret wastewater data signals. Furthermore, wastewater is not a universal approach as it can only be used to detect viruses that are shed in urine and/or feces.

For respiratory viruses, air filter sampling is a new method of non-targeted surveillance.<sup>65</sup> Like wastewater sampling, viable virus can be amplified and then sequenced to examine what is present in an area. Further, given that transmission risk is associated with the amount of virus in the air, sampling air filters from public places could help understand the relationships between exposures and infections and evaluation control measures (such as mask-wearing). Similar to other indirect forms of surveillance, the relationship between prevalence and virus nucleic acid in the air needs to be investigated, the approach validated and sampling technology improved; but it remains a promising avenue for surveillance.<sup>66</sup>

Genomic surveillance for zoonotic and arthropod-borne viruses should also include sampling from their animal reservoirs or vectors to better understand the context of outbreaks. Viruses like Lassa, Ebola, and monkeypox are primarily maintained in animal reservoirs and human outbreaks are the result of spillovers.<sup>67–69</sup> Reservoir surveillance and sequencing, when possible, can help detect new spillover or spillback events and monitor for changes in transmission modalities (see Lassa virus section below). For arthropod-borne viruses, sequencing from invertebrate vectors is often crucial for revealing ecological factors that drive outbreaks.<sup>70</sup> This is especially true for viruses like West Nile (mosquito-borne) and Powassan (tick-borne) in which (1) humans are “dead-end” hosts, (2) human cases are not often detected, and (3) produce low viremias in humans, making it nearly impossible to routinely sequence only from clinical samples. Thus, effective genomic surveillance programs for zoonotic and arthropod-borne viruses must also establish relationships with the groups that are able to routinely sample (in)vertebrate animals to complement sequencing from human cases.

Sequencing viruses from human clinical samples may not always be the most efficient or informative method for genomic surveillance. Creative ways to supplement clinical sequencing, like with travelers, wastewater, or air, are not only critical when resources are low, they may provide early outbreak signals. The key to implementation is partnerships across disciplines.

## BALANCING DATA SOVEREIGNTY AND SHARING

An effective virus genomic surveillance network requires fair and open data sharing within hours to days of generation. This can be complicated by many factors, from technical issues regarding the ease of data submissions to more political issues about

data ownership. The primary challenge is that sequencing data are highly valuable. They can lead to publications in high impact journals and reveal information that can negatively impact a nation's economy. For example, with the SARS-CoV-2 Omicron variant discovered in South Africa, and to a lesser extent the Alpha variant discovered in the UK, sharing data warned the world of the next major COVID-19 wave, but resulting in many countries instituting severe travel restrictions.<sup>71,72</sup> Therefore, rapid data sharing needs to be delicately balanced with protections for the data generators. Some will ultimately choose to not share their data to ensure protection and sovereignty, but the public health benefits will be maximized when everyone has access.

Repositories with fair data usage policies are instrumental in the success of large-scale genomic surveillance and sharing of information among countries. GISAID emerged as one of the most important platforms, developed for influenza A virus and then adding SARS-CoV-2 and now human monkeypox virus.<sup>73</sup> GISAID aims to strike the balance between data protection and accessibility by requiring user agreements to recognize data producers as authors or in the acknowledgments where necessary. The question of where future data should be stored remains open: GISAID has many advantages, but with restrictions on how much data can be downloaded at once, it can be difficult to use for broad surveillance. Further, it is owned by a private individual, which raises issues for its governance and sustainability. NCBI GenBank, run by the US National Institutes of Health, is widely used for other virus genomic data, but it has no protections for data producers. While these and other data sharing platforms continue to be important for genomic surveillance, new systems developed directly from the input of diverse stakeholders are still needed.

The WHO has recently released a new set of guidelines for pathogen genomic data sharing. They emphasize the importance of pragmatism in data generation, including that the first sequences from a region have more impact on the dataset than further sequences from a well-represented region, and (sometimes) that speed matters more than absolute correctness. They also recommend building on the FAIR (findable, accessible, interoperable, and reusable) framework to explicitly include equity.<sup>74,75</sup> These guidelines represent the start of the formalization of the data sharing discussion, and this must continue to find the best balance of protection vs usability, how to enforce usage agreements, and who should host the database itself.

## CASE STUDIES

In the following sections, we highlight how enhanced genomic surveillance could help the control efforts for different types of endemic viruses: RSV (respiratory, global distribution), dengue virus (mosquito-borne, tropic/subtropical distribution), and Lassa virus (zoonotic, regional distribution in West Africa).

### Respiratory syncytial virus

RSV is a leading cause of lower respiratory illness among infants and children under the age of 5 years,<sup>76</sup> especially in LMICs.<sup>77</sup> Globally, RSV transmission follows seasonal patterns and exhibits a short doubling time.<sup>78</sup> It typically causes over 34 million episodes of acute lower respiratory tract illness per year,<sup>77</sup> but

the public health practices implemented during the COVID-19 pandemic have significantly disrupted the global RSV dynamics. For example, in the US, there were almost no cases in the winter of 2020–21, an out-of-season epidemic in 2021,<sup>79</sup> and record highs in the fall of 2022. Many pediatric infectious disease wards are at or exceeding capacity with RSV and other respiratory diseases, including influenza.

RSV circulates as RSV-A and RSV-B serotypes which are defined by antigenic and genetic heterogeneity.<sup>80</sup> Both RSV-A and RSV-B comprise multiple genotypes, 10 for RSV-A and 13 for RSV-B, each of which is characterized by sequential variant replacement, and which often co-circulate during seasonal epidemics<sup>80</sup> and leads to frequent reinfections.<sup>81</sup> Effective specific treatment for RSV remains a challenge, and several vaccine candidates and monoclonal antibodies are at different stages of development.<sup>82</sup> Notably, the bivalent RSV perfusion vaccine by Pfizer was reported to be 82% effective in protecting infants in the first 3 months of life and reportedly effective in preventing RSV disease during the first six months of life.<sup>83</sup> For the vaccines to have a global impact, however, they need to be fairly priced, be available in LMICs, and be broadly effective at reducing transmission of the 23 genotypes.

Despite the high disease burden, there are only ~2,500 RSV genomes available in GenBank, and none from the case resurgence in 2022. The importance of sustained and continuous surveillance cannot be overstated given RSV's antigenic diversity and capacity for rapid genotype turnover. For example, genotype BA in RSV-B emerged in 1999 and replaced all other RSV-B genotypes in circulation, potentially through enhanced immune escape due to a 60-nucleotide duplication in the glycoprotein.<sup>84</sup> With the potential for vaccine roll-out in the near future, genotype-level dynamics must be monitored to detect any further immune escape; and with rapid variant turnover, vaccines may need to be matched to circulating strains, as with influenza. Further, connecting viral genotypes to phenotypes to better predict years with high RSV circulation or severity would enable pediatric healthcare to prepare, as well as raise the index of suspicion for children with respiratory illness. Routine genomic surveillance is critical to not only understand the drivers of epidemics, but also to help ensure that the new vaccines will be broadly effective.

### Dengue virus

Roughly half of the global population is at risk of dengue virus, causing ~100 million symptomatic infections per year.<sup>85</sup> In the Americas, the two highest years on record for reported dengue cases were 2019 and 2022.<sup>86</sup> Dengue virus is endemic to most of the tropical and subtropical regions where its primary mosquito vector, *Aedes aegypti*, resides. With rising global temperatures, the habitat for *Ae. aegypti* is likely to expand.<sup>85</sup> New dengue control methods, however, provide optimism. Other than Dengvaxia, a vaccine only recommended for those who have previously had dengue, there are several promising late-stage vaccines, as well as many other preclinical and clinical candidates.<sup>87</sup> Programs are also deploying *Ae. aegypti* mosquitoes infected with *Wolbachia* bacteria that, amongst other properties, limits dengue virus replication.<sup>88</sup> These new control tools could significantly reduce the future burden of dengue. As dengue virus is genetically diverse both within and between its four serotypes,<sup>89</sup> it is not certain that vaccines or modified

mosquitoes will have ubiquitous effectiveness. Thus, to monitor and help refine these approaches, and provide epidemiological data on emerging lineages, routine dengue virus genomic surveillance is desperately needed.

Using a sample size calculator,<sup>53</sup> ~10,000 dengue virus infections would need to be sequenced per year to detect an emerging lineage in the Americas at 1% frequency with 95% confidence (assuming equal sampling distribution). From 2019–2022, that is ~40,000 of the 9.1 million (~0.4%) dengue cases reported in the Americas. From that period, only 367 sequenced dengue virus genomes (>70% coverage) were deposited in the GenBank data repository. This shortcoming is not entirely due to the COVID-19 pandemic as less than 3,000 dengue virus genomes have been deposited in GenBank from the region since 2000. On the other hand, ~400,000 COVID-19 cases have been sequenced in the Americas, demonstrating that the infrastructure is currently in place to implement genomic surveillance for dengue and other important endemic viruses. Now is the time to make the coordinated push before the widespread roll-out of dengue vaccines.

### Lassa virus

Lassa virus is transmitted by multimammate mice in West Africa where it can cause acute hemorrhagic fever. While there is no evidence of sustained human-to-human transmission, there was a large uptick in cases in Nigeria in 2018, raising concerns that it may be spreading between people. Two studies sequenced 50 cases in real-time, which was sufficient to confirm that the uptick was due to increased spillover from the reservoir and not a change in transmission route.<sup>67,90</sup> While successful in this case, relying on emergency research instead of routine surveillance is unsustainable and unreliable. Further, the response is slower: the authors of the former study were approached at the peak of cases, rather than being able to continuously monitor the situation.

Lassa virus is currently only known to be endemic to West Africa, but beyond the obvious moral imperative to prevent disease everywhere in the world, there is also a global benefit for establishing sustainable local surveillance programs. A comparison can be drawn with human monkeypox virus, which appeared to have stuttering transmission chains in rural West and Central Africa but found a new niche and is spreading in sexual networks across the world.<sup>91</sup> Local surveillance can detect changes in epidemiology, and therefore a rapid response if transmission between humans becomes more efficient: it is thought that human monkeypox virus was spreading locally between humans for five years before it began to spread globally.<sup>92</sup> By implementing local genomic surveillance of Lassa virus, significant changes to its epidemiology could be detected and contained prior to a global outbreak. This could be built on the existing, strong genomic capacity in Nigeria,<sup>90</sup> and Senegal, as well as strengthening the capacity in Guinea, initially developed during the 2013–2016 EVD epidemic.<sup>93</sup>

### Conclusion

Genomic surveillance is a crucial tool for public health response to epidemic and endemic viruses. Twenty years of infrastructure building and technical advancements in response to the HIV pandemic and acute emergencies has culminated in a staggering sequencing effort during the COVID-19 pandemic. The

key now is to transform this effort to build programs for endemic and historically neglected viruses. Moreover, the political will generated by public pressure to track SARS-CoV-2 virus evolution and diversity must now be converted into policy changes and financial commitments to maintain support for surveillance systems.

By using a modular technical framework so that sequencing and bioinformatic pipelines can easily swap between viruses, sustainable systems can be built for a relatively low cost. We have learned from COVID-19 how useful lineage classification systems are, as they mean that most people do not have to generate and interpret phylogenies to track viral diversity in their region. These lineages, and all analyses, require fair and open data sharing to contextualize local results, as well as a reduction in discriminatory economic punishments in the form of travel bans in response to that data sharing and corresponding positive incentives. We also saw that many of the most successful sequencing networks were built on strong collaborations, including sequencing centers, as well as networks of public health, clinical, and academic partners. These collaborations must be maintained in the future to keep sequencing engaged with public health.

A global virus genomic surveillance network has the potential to provide actionable insights for our public health systems. The monitoring of endemic viruses—such as RSV, dengue, and Lassa—through genomics can: (1) provide baseline genetic diversity to help design intervention strategies (e.g., multivalent vaccines); (2) track the effectiveness of vaccines or other control measures at the genotype level; (3) detect the emergence of new variants as an early warning for increased cases; (4) elucidate global dynamics; and (5) monitor changes in transmission routes. By keeping pipelines active between emergencies, the response when new outbreaks begin will also be much faster, providing the opportunity for more effective control. Global genomic surveillance provides a technological solution to a modern problem: diseases do not respect borders. With increasing globalization, urbanization, deforestation, and climate change, novel viruses will spillover more frequently<sup>94</sup> and spread faster once they do. We must not waste the impetus we acquired during the most recent pandemic and leverage it into improving health for everyone.

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.chom.2023.03.003>.

## ACKNOWLEDGMENTS

We would like to thank Emma Hodcroft for helping us generate Figure 2, Stefan Rooke at Public Health Scotland for his thoughts on centralized sequencing, Gage Moreno for sharing a Twitter thread outlining the state of RSV genomic surveillance, P. Jack and S. Taylor for technical discussions, and all of the data producers and those who work behind the scenes at GISAID for the work during the pandemic. Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number DP2AI176740 (NDG). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## DECLARATION OF INTERESTS

N.D.G. is a paid consultant for Tempus Labs and the National Basketball Association and has received speaking fees from Moderna.

## REFERENCES

- Ou, C.Y., Ciesielski, C.A., Myers, G., Bandea, C.I., Luo, C.C., Korber, B.T., Mullins, J.I., Schochetman, G., Berkelman, R.L., and Economou, A.N. (1992). Molecular epidemiology of HIV transmission in a dental practice. *Science* 256, 1165–1171.
- Faria, N.R., Azevedo, R.d.S.d.S., Kraemer, M.U.G., Souza, R., Cunha, M.S., Hill, S.C., Théze, J., Bonsall, M.B., Bowden, T.A., Rissanen, I., et al. (2016). Zika virus in the Americas: Early epidemiological and genetic findings. *Science* 352, 345–349.
- Stockdale, J.E., Liu, P., and Colijn, C. (2022). The potential of genomics for infectious disease forecasting. *Nat. Microbiol.* 7, 1736–1743.
- Hill, S., Perkins, M., and von Eije, K. (2021). Genomic Sequencing of SARS-CoV-2 (World Health Organization).
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- Vasilyeva, T.I., du Plessis, L., Pineda-Peña, A.C., Kühnert, D., Lemey, P., Vandamme, A.-M., Gomes, P., Camacho, R.J., Pybus, O.G., Abecasis, A.B., and Faria, N.R. (2019). Tracing the Impact of Public Health Interventions on HIV-1 Transmission in Portugal Using Molecular Epidemiology. *J. Infect. Dis.* 220, 233–243.
- Marra, M.A., Jones, S.J.M., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S.N., Khattri, J., Asano, J.K., Barber, S.A., Chan, S.Y., et al. (2003). The Genome sequence of the SARS-associated coronavirus. *Science* 300, 1399–1404.
- Ruan, Y.J., Wei, C.L., Ee, A.L., Vega, V.B., Thoreau, H., Su, S.T.Y., Chia, J.-M., Ng, P., Chiu, K.P., Lim, L., et al. (2003). Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. *Lancet* 361, 1779–1785.
- Chinese SARS Molecular Epidemiology Consortium; Molecular Epidemiology Consortium (2004). Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. *Science* 303, 1666–1669.
- Smith, G.J.D., Vijaykrishna, D., Bahl, J., Lycett, S.J., Worobey, M., Pybus, O.G., Ma, S.K., Cheung, C.L., Raghwani, J., Bhatt, S., et al. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459, 1122–1125.
- Morris, D.H., Gostic, K.M., Pompei, S., Bedford, T., Łuksza, M., Neher, R.A., Grenfell, B.T., Lässig, M., and McCauley, J.W. (2018). Predictive Modeling of Influenza Shows the Promise of Applied Evolutionary Biology. *Trends Microbiol.* 26, 102–118. <https://doi.org/10.1016/j.tim.2017.09.004>.
- Zost, S.J., Parkhouse, K., Gumina, M.E., Kim, K., Diaz Perez, S., Wilson, P.C., Treanor, J.J., Sant, A.J., Cobey, S., and Hensley, S.E. (2017). Contemporary H3N2 influenza viruses have a glycosylation site that alters binding of antibodies elicited by egg-adapted vaccine strains. *Proc. Natl. Acad. Sci. U. Proc. Natl. Acad. Sci. USA.* 114, 12578–12583.
- Lam, T.T., and Pybus, O.G. (2018). Genomic surveillance of avian-origin influenza A viruses causing human disease. *Genome Med.* 10, 50. <https://doi.org/10.1186/s13073-018-0560-3>.
- Macklin, G.R., O'Reilly, K.M., Grassly, N.C., Edmunds, W.J., Mach, O., Santhana Gopala Krishnan, R., Voorman, A., Vertefeuille, J.F., Abdelwahab, J., Gumedde, N., et al. (2020). Evolving epidemiology of poliovirus serotype 2 following withdrawal of the serotype 2 oral poliovirus vaccine. *Science* 368, 401–405.
- Cowger, T.L., Burns, C.C., Sharif, S., Gary, H.E., Jr., Iber, J., Henderson, E., Malik, F., Zahoor Zaidi, S.S., Shaikat, S., Rehman, L., et al. (2017). The role of supplementary environmental surveillance to complement acute flaccid paralysis surveillance for wild poliovirus in Pakistan – 2011–2013. *PLoS One* 12, e0180608. <https://doi.org/10.1371/journal.pone.0180608>.
- Yakovenko, M.L., Gmyl, A.P., Ivanova, O.E., Ereemeeva, T.P., Ivanov, A.P., Prostova, M.A., Baykova, O.Y., Isaeva, O.V., Lipskaya, G.Y., Shakaryan, A.K., et al. (2014). The 2010 outbreak of poliomyelitis in Tajikistan:

- epidemiology and lessons learnt. *Euro Surveill.* 19, 20706. <https://doi.org/10.2807/1560-7917.es2014.19.7.20706>.
17. Dudas, G., Carvalho, L.M., Bedford, T., Tatem, A.J., Baele, G., Faria, N.R., Park, D.J., Ladner, J.T., Arias, A., Asogun, D., et al. (2017). Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* 544, 309–315.
18. Quick, J., Loman, N.J., Duraffour, S., Simpson, J.T., Severi, E., Cowley, L., Bore, J.A., Koundouno, R., Dudas, G., Mikhail, A., et al. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530, 228–232.
19. Gardy, J.L., and Loman, N.J. (2018). Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat. Rev. Genet.* 19, 9–20.
20. Mate, S.E., Kugelman, J.R., Nyenswah, T.G., Ladner, J.T., Wiley, M.R., Cordier-Lassalle, T., Christie, A., Schroth, G.P., Gross, S.M., Davies-Wayne, G.J., et al. (2015). Molecular Evidence of Sexual Transmission of Ebola Virus. *N. Engl. J. Med.* 373, 2448–2454.
21. Diallo, B., Sissoko, D., Loman, N.J., Bah, H.A., Bah, H., Worrell, M.C., Conde, L.S., Sacko, R., Mesfin, S., Loua, A., et al. (2016). Resurgence of Ebola Virus Disease in Guinea Linked to a Survivor With Virus Persistence in Seminal Fluid for More Than 500 Days. *Clin. Infect. Dis.* 63, 1353–1356. <https://doi.org/10.1093/cid/ciw601>.
22. Faria, N.R., Quick, J., Claro, I.M., Théze, J., de Jesus, J.G., Giovanetti, M., Kraemer, M.U.G., Hill, S.C., Black, A., da Costa, A.C., et al. (2017). Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature* 546, 406–410. <https://doi.org/10.1038/nature22401>.
23. Quick, J., Grubaugh, N.D., Pullan, S.T., Claro, I.M., Smith, A.D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T.F., Beutler, N.A., et al. (2017). Multiplex PCR method for MiniON and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* 12, 1261–1276.
24. Faria, N.R., Sabino, E.C., Nunes, M.R.T., Alcantara, L.C.J., Loman, N.J., and Pybus, O.G. (2016). Mobile real-time surveillance of Zika virus in Brazil. *Genome Med.* 8, 97.
25. Kinganda-Lusamaki, E., Black, A., Mukadi, D.B., Hadfield, J., Mbala-Kingebeni, P., Pratt, C.B., Aziza, A., Diagne, M.M., White, B., Bisento, N., et al. (2021). Integration of genomic sequencing into the response to the Ebola virus outbreak in Nord Kivu, Democratic Republic of the Congo. *Nat. Med.* 27, 710–716.
26. O'Toole, Á., Scher, E., Underwood, A., Jackson, B., Hill, V., McCrone, J.T., Colquhoun, R., Ruis, C., Abu-Dahab, K., Taylor, B., et al. (2021). Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol.* 7, veab064.
27. Brito, A.F., Semenova, E., Dudas, G., Hassler, G.W., Kalinich, C.C., Kraemer, M.U.G., Ho, J., Tegally, H., Githinji, G., Agoti, C.N., et al. (2022). Global disparities in SARS-CoV-2 genomic surveillance. *Nat. Commun.* 13, 7003. <https://doi.org/10.1038/s41467-022-33713-y>.
28. Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., et al. (2020). A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269.
29. Tegally, H., San, J.E., Cotten, M., Moir, M., Tegomoh, B., Mboowa, G., Martin, D.P., Baxter, C., Lambisia, A.W., Diallo, A., et al. (2022). The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance. *Science* 378, eabq5358.
30. Wilkinson, E., Giovanetti, M., Tegally, H., San, J.E., Lessells, R., Cuadros, D., Martin, D.P., Rasmussen, D.A., Zekri, A.-R.N., Sangare, A.K., et al. (2021). A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *Science* 374, 423–431.
31. Sahadeo, N.S.D., Nicholls, S., Moreira, F.R.R., O'Toole, Á., Ramkissoon, V., Whittaker, C., Hill, V., McCrone, J.T., Mohammed, N., Ramjag, A., et al. (2022). Implementation of Genomic Surveillance of SARS-CoV-2 in the Caribbean: Lessons Learned for Sustainability in Resource-Limited Settings. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4184797>.
32. COVID-19 Genomics UK COG-UK consortiumcontact@cogconsortiumuk (2020). An integrated national scale SARS-CoV-2 genomic surveillance network. *Lancet. Microbe* 1, e99–e100.
33. Nicholls, S.M., Poplawski, R., Bull, M.J., Underwood, A., Chapman, M., Abu-Dahab, K., Taylor, B., Colquhoun, R.M., Rowe, W.P.M., Jackson, B., et al. (2021). CLIMB-COVID: continuous integration supporting decentralised sequencing for SARS-CoV-2 genomic surveillance. *Genome Biol.* 22, 196–221.
34. O'Toole, Á., Hill, V., Jackson, B., Dewar, R., Sahadeo, N., Colquhoun, R., Rooke, S., McCrone, J.T., McHugh, M.P., Nicholls, S., et al. (2021). Genomics-informed outbreak investigations of SARS-CoV-2 using civet. *medRxiv*. <https://doi.org/10.1101/2021.12.13.21267267>.
35. Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., and Pybus, O.G. (2020). A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* 5, 1403–1407.
36. Alm, E., Broberg, E.K., Connor, T., Hodcroft, E.B., Komissarov, A.B., Maurer-Stroh, S., Melidou, A., Neher, R.A., O'Toole, Á., and Pereyaslov, D.; The WHO European Region sequencing laboratories and GISAID Epi-CoV groupÁ., Pereyaslov, D. (2020). Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020. *Euro Surveill.* 25, 10. <https://doi.org/10.2807/1560-7917.ES.2020.25.32.2001410>.
37. Konings, F., Perkins, M.D., Kuhn, J.H., Pallen, M.J., Alm, E.J., Archer, B.N., Barakat, A., Bedford, T., Bhiman, J.N., Cally, L., et al. (2021). SARS-CoV-2 Variants of Interest and Concern naming scheme conducive for global discourse. *Nat. Microbiol.* 6, 821–823.
38. Githinji, G., de Laurent, Z.R., Mohammed, K.S., Omuoyo, D.O., Macharia, P.M., Morobe, J.M., Otieno, E., Kinyanjui, S.M., Agweyu, A., Maithe, E., et al. (2021). Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. *Nat. Commun.* Now. 12, 1–10.
39. Cahyani, I., Putro, E.W., Ridwanuloh, A.M., Wibowo, S., Hariyatun, H., Syahputra, G., Akbariani, G., Utomo, A.R., Ilyas, M., Loose, M., et al. (2022). Genome Profiling of SARS-CoV-2 in Indonesia, ASEAN and the Neighbouring East Asian Countries: Features, Challenges and Achievements. *Viruses* 14, 778. <https://doi.org/10.3390/v14040778>.
40. Worobey, M., Pekar, J., Larsen, B.B., Nelson, M.I., Hill, V., Joy, J.B., Rambaut, A., Suchard, M.A., Wertheim, J.O., and Lemey, P. (2020). The emergence of SARS-CoV-2 in Europe and North America. *Science* 370, 564–570.
41. Viana, R., Moyo, S., Amoako, D.G., Tegally, H., Scheepers, C., Althaus, C.L., Anyaneji, U.J., Bester, P.A., Boni, M.F., Chand, M., et al. (2022). Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* 603, 679–686.
42. Isidro, J., Borges, V., Pinto, M., Sobral, D., Santos, J.D., Nunes, A., Mixão, V., Ferreira, R., Santos, D., Duarte, S., et al. (2022). Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. *Nat. Med.* 28, 1569–1572.
43. Worobey, M., Watts, T.D., McKay, R.A., Suchard, M.A., Granade, T., Teuwen, D.E., Koblin, B.A., Heneine, W., Lemey, P., and Jaffe, H.W. (2016). 1970s and “Patient 0” HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature* 539, 98–101.
44. Grubaugh, N.D., Gangavarapu, K., Quick, J., Matteson, N.L., De Jesus, J.G., Main, B.J., Tan, A.L., Paul, L.M., Brackney, D.E., Grewal, S., et al. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol.* 20, 8–19.
45. Faria, N.R., Kraemer, M.U.G., Hill, S.C., Goes de Jesus, J., Aguiar, R.S., Iani, F.C.M., Xavier, J., Quick, J., du Plessis, L., Dellicour, S., et al. (2018). Genomic and epidemiological monitoring of yellow fever virus transmission potential. *Science* 361, 894–899.
46. Lambisia, A.W., Mohammed, K.S., Makori, T.O., Ndwigwa, L., Mburu, M.W., Morobe, J.M., Moraa, E.O., Musyoki, J., Murunga, N., Mwangi, J.N., et al. (2022). Optimization of the SARS-CoV-2 ARTIC Network V4 Primers and Whole Genome Sequencing Protocol. *Front. Med.* 0, 10. <https://doi.org/10.3389/fmed.2022.836728>.
47. Gohl, D.M., Garbe, J., Grady, P., Daniel, J., Watson, R.H.B., Auch, B., Nelson, A., Yohe, S., and Beckman, K.B. (2020). A rapid, cost-effective tailored amplicon method for sequencing SARS-CoV-2. *BMC Genom.* 21, 863–910.
48. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing



- Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
49. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760.
50. Salyer, S.J., Maeda, J., Sembuche, S., Kebede, Y., Tshangela, A., Mousif, M., Ihekweazu, C., Mayet, N., Abate, E., Ouma, A.O., and Nkengasong, J. (2021). The first and second waves of the COVID-19 pandemic in Africa: a cross-sectional study. *Lancet* 397, 1265–1275. [https://doi.org/10.1016/S0140-6736\(21\)00632-2](https://doi.org/10.1016/S0140-6736(21)00632-2).
51. Krieger, N. (2022). Fixing US public health infrastructure for 1% of annual military budget. *Lancet* 400, 1581–1582.
52. France. (2022). Pfizer lifts 2022 forecast for Covid-19 vaccine sales as profits rise. *France 24*. <https://www.france24.com/en/live-news/20221101-pfizer-lifts-2022-forecast-for-covid-19-vaccine-sales-as-profits-rise>.
53. Wohl, S., Lee, E.C., DiPrete, B.L., and Lessler, J. (2022). Sample Size Calculations for Variant Surveillance in the Presence of Biological and Systematic Biases. *medRxiv*. <https://doi.org/10.1101/2021.12.30.21268453>.
54. MacCannell, D. (2019). Platforms and Analytical Tools Used in Nucleic Acid Sequence-Based Microbial Genotyping Procedures. *Microbiol. Spectr.* 7, 10. <https://doi.org/10.1128/microbiolspec.AME-0005-2018>.
55. Hill, V., Ruis, C., Bajaj, S., Pybus, O.G., and Kraemer, M.U.G. (2021). Progress and challenges in virus genomic epidemiology. *Trends Parasitol.* 37, 1038–1049.
56. Rambaut, A., Dudas, G., de Carvalho, L.M., Park, D.J., Yozwiak, N.L., Holmes, E.C., and Andersen, K.G. (2016). Comment on "Mutation rate and genotype variation of Ebola virus from Mali case sequences". *Science* 353, 658.
57. Hadfield, J., Megill, C., Bell, S.M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., and Neher, R.A. (2018). Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34, 4121–4123.
58. Grubaugh, N.D., Saraf, S., Gangavarapu, K., Watts, A., Tan, A.L., Oidman, R.J., Ladner, J.T., Oliveira, G., Matteson, N.L., Kraemer, M.U.G., et al. (2019). Travel Surveillance and Genomics Uncover a Hidden Zika Outbreak during the Waning Epidemic. *Cell* 178, 1057–1071.e11.
59. Lemey, P., Hong, S.L., Hill, V., Baele, G., Poletto, C., Colizza, V., O'Toole, A., McCrone, J.T., Andersen, K.G., Worobey, M., et al. (2020). Accommodating individual travel history and unsampled diversity in Bayesian phylogeographic inference of SARS-CoV-2. *Nat. Commun.* 11, 5110–5114.
60. Wegrzyn, R.D., Appiah, G.D., Morfino, R., Milford, S.R., Walker, A.T., Ernst, E.T., Darrow, W.W., Li, S.L., Robison, K., MacCannell, D., et al. (2022). Early detection of SARS-CoV-2 variants using traveler-based genomic surveillance at four US airports, September 2021–January 2022. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciac461>.
61. Klapsa, D., Wilton, T., Zealand, A., Bujaki, E., Saxentoff, E., Troman, C., Shaw, A.G., Tedcastle, A., Majumdar, M., Mate, R., et al. (2022). Sustained detection of type 2 poliovirus in London sewage between February and July, 2022, by enhanced environmental surveillance. *Lancet* 400, 1531–1538.
62. Karthikeyan, S., Levy, J.I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H.M., Valles, T., Tribelhorn, C.E., et al. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature* 609, 101–108.
63. Peccia, J., Zulli, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-Massana, A., Ko, A.I., Malik, A.A., Wang, D., Wang, M., et al. (2020). Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat. Biotechnol.* 38, 1164–1167.
64. Smyth, D.S., Trujillo, M., Gregory, D.A., Cheung, K., Gao, A., Graham, M., Guan, Y., Guldenpfennig, C., Hoxie, I., Kannoly, S., et al. (2022). Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nat. Commun.* 13, 1836.
65. Hadei, M., Mohebbi, S.R., Hopke, P.K., Shahsavani, A., Bazzazpour, S., Alipour, M., Jafari, A.J., Bandpey, A.M., Zali, A., Yarahmadi, M., et al. (2021). Presence of SARS-CoV-2 in the air of public places and transportation. *Atmos. Pollut. Res.* 12, 302–306.
66. Bhardwaj, J., Hong, S., Jang, J., Han, C.-H., Lee, J., and Jang, J. (2021). Recent advancements in the measurement of pathogenic airborne viruses. *J. Hazard Mater.* 420, 126574.
67. Kafetzopoulou, L.E., Pullan, S.T., Lemey, P., Suchard, M.A., Ehichioya, D.U., Pahlmann, M., Thielebein, A., Hinzmann, J., Oestereich, L., Wozniak, D.M., et al. (2019). Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. *Science* 363, 74–77.
68. Schmidt, J.P., Park, A.W., Kramer, A.M., Han, B.A., Alexander, L.W., and Drake, J.M. (2017). Spatiotemporal Fluctuations and Triggers of Ebola Virus Spillover. *Emerg. Infect. Dis.* 23, 415–422.
69. Petersen, E., Kantele, A., Koopmans, M., Asogun, D., Yinka-Ogunleye, A., Ihekweazu, C., and Zumla, A. (2019). Human Monkeypox: Epidemiologic and Clinical Characteristics, Diagnosis, and Prevention. *Infect. Dis. Clin. North Am.* 33, 1027–1043.
70. Vogels, C.B.F., Brackney, D.E., Dupuis, A.P., Robich, R.M., Fauver, J.R., Brito, A.F., Williams, S.C., Anderson, J.F., Lubelczyk, C.B., Lange, R.E., et al. (2022). Phylogeographic reconstruction of the emergence and spread of Powassan virus in the northeastern United States. *bioRxiv*. <https://doi.org/10.1101/2022.10.14.512245>.
71. Schermerhorn, J., Case, A., Graeden, E., Kerr, J., Moore, M., Robinson-Marshall, S., Wallace, T., Woodrow, E., and Katz, R. (2022). Fifteen days in December: capture and analysis of Omicron-related travel restrictions. *BMJ Glob. Health* 7, e008642. <https://doi.org/10.1136/bmjgh-2022-008642>.
72. Agencies. (2020). Europe bans travel from the UK over new Covid strain: what we know so far. *The Guardian*.
73. Khare, S., Gurry, C., Freitas, L., Schultz, M.B., Bach, G., Diallo, A., Akite, N., Ho, J., Lee, R.T., Yeo, W., et al. (2021). GISAI's Role in Pandemic Response. *China CDC Wkly.* 3, 1049–1051.
74. World Health Organization. (2022). WHO Guiding principles for pathogen genome data sharing.
75. Boeckhout, M., Zielhuis, G.A., and Bredenoord, A.L. (2018). The FAIR guiding principles for data stewardship: fair enough? *Eur. J. Hum. Genet.* 26, 931–936.
76. Berkley, J.A., Munywoki, P., Ngama, M., Kazungu, S., Abwao, J., Bett, A., Lassaunière, R., Kresfelder, T., Cane, P.A., Venter, M., et al. (2010). Viral Etiology of Severe Pneumonia Among Kenyan Infants and Children. *JAMA* 303, 2051–2057.
77. Shi, T., McAllister, D.A., O'Brien, K.L., Simoes, E.A.F., Madhi, S.A., Gessner, B.D., Polack, F.P., Balsells, E., Acacio, S., Aguayo, C., et al. (2017). Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *Lancet* 390, 946–958.
78. Munywoki, P.K., Koech, D.C., Agoti, C.N., Lewa, C., Cane, P.A., Medley, G.F., and Nokes, D.J. (2014). The source of respiratory syncytial virus infection in infants: a household cohort study in rural Kenya. The Source of Respiratory Syncytial Virus Infection in Infants: A Household Cohort Study In Rural Kenya. *J. Infect. Dis.* 209, 1685–1692.
79. Zheng, Z., Warren, J.L., Artin, I., Pitzer, V.E., and Weinberger, D.M. (2022). Relative timing of respiratory syncytial virus epidemics in summer 2021 across the United States was similar to a typical winter season. *Influenza Other Respi. Influenza Other Respir. Viruses* 16, 617–620.
80. Cantú-Flores, K., Rivera-Alfaro, G., Muñoz-Escalante, J.C., and Noyola, D.E. (2022). Global distribution of respiratory syncytial virus A and B infections: a systematic review. *Pathog. Glob. Health* 116, 398–409.
81. Hall, C.B., Walsh, E.E., Long, C.E., and Schnabel, K.C. (1991). Immunity to and frequency of reinfection with respiratory syncytial virus. *JID (J. Infect. Dis.)* 163, 693–698. <https://doi.org/10.1093/infdis/163.4.693>.
82. Rodríguez-Fernández, R., Mejías, A., and Ramilo, O. (2021). Monoclonal Antibodies for Prevention of Respiratory Syncytial Virus Infection. *Pediatr. Infect. Dis. J.* 40, S35–S39. <https://doi.org/10.1097/INF.0000000000003121>.
83. Pfizer. (2022). Pfizer announces positive top-line data of phase 3 global maternal immunization trial for its bivalent respiratory syncytial virus (RSV) vaccine candidate. <https://www.pfizer.com/news/press-release/press-release-detail/pfizer-announces-positive-top-line-data-phase-3-global>.



84. Trento, A., Viegas, M., Galiano, M., Videla, C., Carballal, G., Mistchenko, A.S., and Melero, J.A. (2006). Natural History of Human Respiratory Syncytial Virus Inferred from Phylogenetic Analysis of the Attachment (G) Glycoprotein with a 60-Nucleotide Duplication. *J. Virol.* **80**, 975–984.
85. Messina, J.P., Brady, O.J., Golding, N., Kraemer, M.U.G., Wint, G.R.W., Ray, S.E., Pigott, D.M., Shearer, F.M., Johnson, K., Earl, L., et al. (2019). The current and future global distribution and population at risk of dengue. *Nat. Microbiol.* **4**, 1508–1515.
86. Pan American Health Organization Dengue Fact Sheet. <https://www.paho.org/en/topics/dengue>
87. Pinheiro-Michelsen, J.R., Souza, R., da, S.O., Santana, I.V.R., da Silva, P. de S., Mendez, E.C., Luiz, W.B., and Amorim, J.H. (2020). Anti-dengue Vaccines: From Development to Clinical Trials. *Front. Immunol.* **11**, 1252.
88. Walker, T., Johnson, P.H., Moreira, L.A., Iturbe-Ormaetxe, I., Frentiu, F.D., McMeniman, C.J., Leong, Y.S., Dong, Y., Axford, J., Kriesner, P., et al. (2011). The wMel Wolbachia strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature* **476**, 450–453.
89. Katzelnick, L.C., Coello Escoto, A., Huang, A.T., Garcia-Carreras, B., Chowdhury, N., Maljkovic Berry, I., Chavez, C., Buchy, P., Duong, V., Dussart, P., et al. (2021). Antigenic evolution of dengue viruses over 20 years. *Science* **374**, 999–1004.
90. Siddle, K.J., Eromon, P., Barnes, K.G., Mehta, S., Oguzie, J.U., Odia, I., Schaffner, S.F., Winnicki, S.M., Shah, R.R., Qu, J., et al. (2018). Genomic analysis of Lassa virus from the 2018 surge in Nigeria. *N. Engl. J. Med.* **379**, 1745–1753.
91. Vivancos, R., Anderson, C., Blomquist, P., Balasegaram, S., Bell, A., Bishop, L., Brown, C.S., Chow, Y., Edeghere, O., Florence, I., et al. (2022). Community transmission of monkeypox in the United Kingdom, April to May 2022. *Euro Surveill.* **27**, 2200422. <https://doi.org/10.2807/1560-7917.ES.2022.27.22.2200422>.
92. O'Toole, Á., and Rambaut, A. (2022). Initial observations about putative APOBEC3 deaminase editing driving short-term evolution of MPXV since 2017. *Virological*. <https://virological.org/t/initial-observations-about-putative-apobec3-deaminase-editing-driving-short-term-evolution-of-mpxv-since-2017/830>.
93. Keita, A.K., Koundouno, F.R., Faye, M., Dux, A., Hinzmann, J., Diallo, H., Ayoub, A., Le Marcis, F., Soropogui, B., Ifono, K., et al. (2021). Resurgence of Ebola virus in 2021 in Guinea suggests a new paradigm for outbreaks. *Nature* **597**, 539–543.
94. Carlson, C.J., Albery, G.F., Merow, C., Trisos, C.H., Zipfel, C.M., Eskew, E.A., Olival, K.J., Ross, N., and Bansal, S. (2022). Climate change increases cross-species viral transmission risk. *Nature* **607**, 555–562.