SARS-CoV-2 genomic sequencing for public health goals

Interim guidance

8 January 2021

World Health Organization

Key messages:

- Global surveillance of SARS-CoV-2 genetic sequences and related metadata contributes to the COVID-19 outbreak response. This contribution includes tracking the spread of SARS-CoV-2 geographically over time and ensuring that mutations that could potentially influence pathogenicity, transmission or countermeasures (such as vaccines, therapeutics and diagnostics) are detected and assessed in a timely manner.
- While the cost and complexity of genetic sequencing have dropped significantly over time, effective sequencing programmes still require substantial investment in terms of staff, equipment, reagents and bioinformatic infrastructure. Additionally, effective collaboration is needed to ensure that generated data are of good quality and are used in a meaningful way.
- Countries are encouraged to rapidly deposit SARS-CoV-2 sequences in a public database in order to share them with the scientific community for public health purposes. Investments in a tiered global sequencing network for SARS-CoV-2 will contribute to the development of resilient, high-quality global sequencing programmes for the detection and management of other outbreak pathogens in the future.

Background

Over the last decade, genetic sequence data (GSD) of pathogens have come to play a pivotal role in the detection and management of infectious disease outbreaks, supporting the development of diagnostics, drugs and vaccines, and informing the outbreak response *(1–11)*. With the emergence of the novel coronavirus, later named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the importance of GSD has been further underlined. More than 280 000 full genome sequences have been shared via publicly accessible databases within a year of the initial identification of SARS-CoV-2 *(12)*. Near real-time analysis of data has directly impacted the public health response *(12–16)*. The public health objectives of SARS-CoV-2 genomic sequencing are listed in Table 1.

The growing understanding of how sequence information can contribute to improved public health is driving global investments in sequencing facilities and programmes. The falling cost and complexity of generating GSD provides opportunities for expanding sequencing capacity; however, challenges to widespread implementation remain, and sequencing capacity and data are not evenly distributed around the world, with an overrepresentation of SARS-CoV-2 GSD from high-income countries.

Table 1. Public health objectives of SARS-CoV-2 genomic sequencing

Activities that require a limited effort and once achieved might need either no sequencing or occasional sequencing for follow-up	**Activities that require sustained sequencing activities over a longer period of time**	
- Identify SARS-CoV-2 as the causative agent of disease. - Develop diagnostics for SARS-CoV-2. - Support the development of therapies and vaccines. - Investigate date of introduction into humans and investigate SARS-CoV-2 origin (ongoing). - Reinfection: • Evaluate and improve understanding of this phenomenon. • On the individual level, differentiate between prolonged infection and reinfection.	**SARS-CoV-2 evolution and its impact on:** - Change in viral behaviour (phenotypic change), e.g., transmissibility or pathogenicity; - Immunity (from vaccines or natural infection); - Diagnostics (i.e., molecular, serology, antigen assays); - Therapeutic interventions (e.g., monoclonal antibodies).	**Monitor viral movement and activity:** - Investigate geographic spread and reintroductions between populations. - Investigate outbreaks in specific settings and populations (e.g., in hospitals). - Track zoonotic reintroduction in both directions over the species barrier. - Monitor environmental and waste water. - Support classical surveillance by quantifying the period of transmission and evaluating drivers, and by estimating the transmission level in the population.

Purpose of this document

This document provides national-level policy-makers and stakeholders with guidance on how to maximize the public health benefit of SARS-CoV-2 genomic sequencing activities in the short and long term as the pandemic continues to unfold. Practical considerations for the implementation of a virus genomic sequencing programme and an overview of the public health objectives of genomic sequencing are covered. This guidance focuses on SARS-CoV-2 but is applicable to other pathogens of public health concern. It is recommended that countries wishing to build sequencing capacity for SARS-CoV-2 do so as part of a broader plan to build capacity to detect and monitor other pathogens of public health concern.

Additional WHO guidance

WHO has developed the implementation guide Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health in collaboration with sequencing experts around the world. This guide provides a more complete background on SARS-CoV-2 sequencing and is intended for those who are actively involved in implementing sequencing programmes *(17)*. It provides an in-depth review of the various uses of sequencing and gives technical advice on pathogen sequencing in the context of SARS-CoV-2. Beyond reading these and other published documents, laboratories with limited sequencing experience should actively look for opportunities to collaborate with experienced laboratories and/or join or form laboratory networks with sequencing expertise.

1. Introduction to SARS-CoV-2

SARS-CoV-2 is classified within the genus *Betacoronavirus* (subgenus *Sarbecovirus*) of the family *Coronaviridae (18)*. It is an enveloped, positive-sense, single-stranded ribonucleic acid (RNA) virus with an approximately 30kb genome *(19)*. Genetic sequencing enables the reading of viral genomes. As each pathogen has a unique genomic sequence, this method can be used to identify novel pathogens (as in the case of SARS-CoV-2) *(20)*. The SARS-CoV-2 genome encodes for non-structural proteins, four structural proteins (spike [S], envelope [E], membrane [M], nucleocapsid [N]) and several putative accessory proteins *(21–23)*. SARS-CoV-2 host cell entry requires binding of the viral S protein to the host cell angiotensin-converting enzyme 2 (ACE-2) receptor *(24–27)*. The SARS-CoV-2 spike protein, especially the receptor-binding domain (RBD), is a critical target for natural and vaccine-induced immunity *(28–32)*. Therefore, diversification of the gene encoding the spike protein could potentially impact vaccine efficacy, natural immunity and (monoclonal) antibody therapies *(33)*.

When viruses replicate, especially RNA viruses such as SARS-CoV-2, changes (mutations) occur in the genome. If an acquired mutation does not have an evolutionary disadvantage, it may become fixed in SARS-CoV-2 populations. The rate of evolutionary change in SARS-CoV-2 is currently estimated to be 1×10^{-3} substitutions per site per year at the nucleotide level *(34)*. This translates into approximately one substitution in the genome every two weeks *(35)*. This relatively low rate of evolution limits the time resolution of individual transmission events *(35)*. Studying SARS-CoV-2 evolution and rapidly identifying substitutions, insertions or deletions that could impact viral properties (phenotypic change) is an important tool for epidemic monitoring. Among the most obvious yields of such work is the detection of mutations that are associated with changes in the transmissibility and/or pathogenicity of the virus, or that could reduce the utility of medical countermeasures (diagnostics, vaccines and therapeutics). Following virus mutations over time and space can also help to track the spread of the pathogen and support an enhanced understanding of potential transmission routes and dynamics. Reconstructing the evolutionary history of pathogens can be achieved through phylogenetic analysis. Phylogenetic and phylodynamic (i.e., how virus phylogenies are shaped by epidemiological and evolutionary processes) analyses can provide extensive information to support outbreak response.

2. Establishing an optimal SARS-CoV-2 sequencing approach in the local context

2.1 Context-specific prioritization of sequencing objectives and approach

Although the cost of gene sequencing has fallen significantly over the past decades, sequencing still requires substantial investment in resources (financial, infrastructure and human). Before initiating a sequencing project, the critical first step is to determine whether sequencing is truly valuable for reaching a specific objective or whether there are other more time-effective or cost-effective approaches available. This decision may involve considering whether virus sequencing alone is sufficient to reach the defined objective or whether it should be included as a smaller component within a multidisciplinary approach. Epidemiologically focused activities that integrate genomic data analysts directly into public health investigation and response teams are likely to have a greater immediate impact than those in which virus genomic analysis exists as a separate or secondary activity.

Where resources to support sequencing are limited, it may be necessary to limit objectives of a sequencing programme to those activities with high clinical and/or public health potential, which can be sustained. Such a programme may prioritize the sequencing of SARS-CoV-2 i) from individuals vaccinated for SARS-CoV-2 but who later become infected with SARS-CoV-2 despite exhibiting an appropriate immune response to the vaccine; ii) in risk settings, such as where there is close human–animal interaction with a large number of animals that are susceptible to SARS-CoV-2 infection, or where there are immunocompromised patients with prolonged shedding, especially when receiving antibody therapy against SARS-CoV-2; iii) when there is an unexpected increase or change in SARS-CoV-2 transmissibility and/or virulence; iv) when there is suspicion of a change in the performance of diagnostic (antibody, antigen, molecular assays) methods or therapies; and v) during cluster investigations when sequencing can support understanding of transmission events and/or evaluate the efficacy of infection control procedures.

Fig. 1 provides an overview of the basic pillars required for sequencing. If there is no or limited capacity available in all three pillars, it will likely be necessary to build partnerships with other groups in order to achieve sequencing goals. Conversely, if there is adequate capacity and resources for one or more pillars, the laboratory could consider supporting other partners with nascent sequencing programmes. Variable demand on capacity will occur throughout different phases of an outbreak and might require laboratories to shift from one strategy to another.

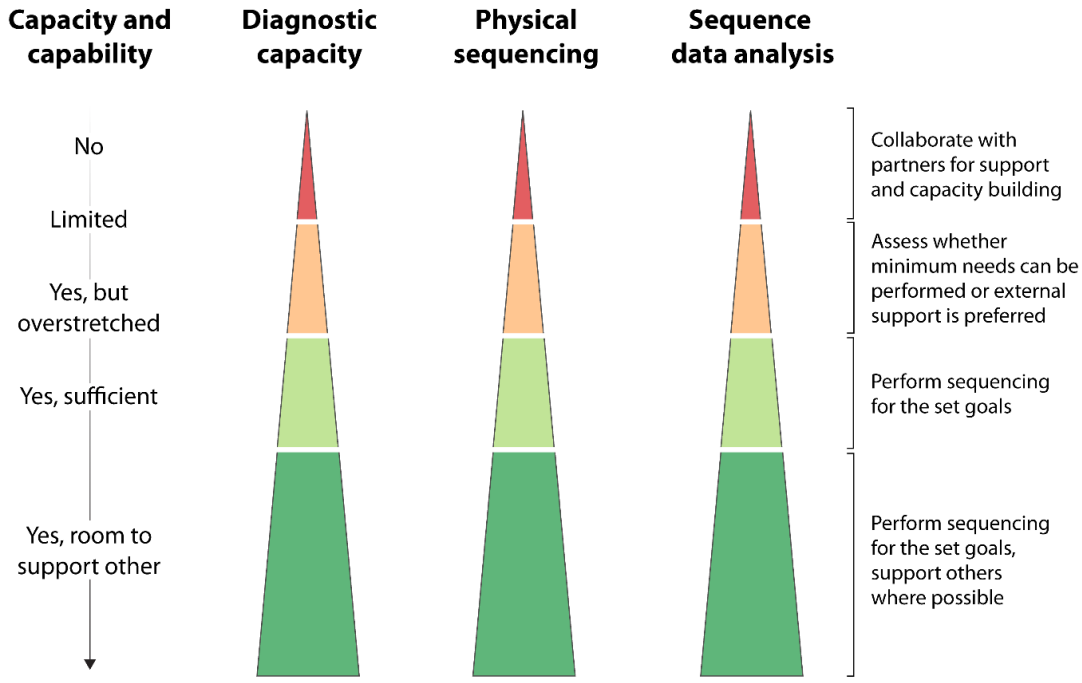

Fig. 1. Establishing capacity in the pillars required for sequencing in order to select the best approach for the setup of a sequencing programme

2.2 Investing in global sustainable sequencing capacity for SARS-CoV-2 and other (emerging/re-emerging) pathogens of public health concern

Tiered WHO laboratory networks have proven functionality and enable global collaboration, with regional adaptation of networks to specific national and regional needs *(36–39)*. Building such a strong and resilient global sequencing network can maximize the public health impact of sequencing for SARS-CoV-2 and emerging/re-emerging pathogens. Currently, the WHO reference laboratories providing confirmatory testing for COVID-19 are supporting some of these sequencing and analysis needs *(40)*. Several regions have, or are in the process of developing, sequencing capacities that will be able to join the global network of laboratories/sequencing groups. To determine the contribution laboratories in the network can make, a global estimation of capacity for each pillar listed in Fig. 1 can be undertaken. Various pathogen-specific laboratory networks (such as those working on antimicrobial resistance, MERS-CoV, influenza, measles, rubella, poliovirus and tuberculosis) have invested in sequencing capacity as part of their surveillance activities *(8, 9, 41–43)*. As the costs of sequencing are still substantial and many parts of the sequencing workstream can be used for various pathogens or sequencing objectives, national collaboration to ensure the optimal use of existing capacity is encouraged. Capacity-building programmes should focus on a stepwise approach to build competencies. The priorities of capacity building should be context-dependent. For some countries, building wet-laboratory capacity would make sense, whereas in other settings, outsourcing the actual sequencing and focusing on bioinformatics, data management and interpretation would have a bigger impact. For effective collaboration, data sharing, standardized protocols for sequencing, joint meetings and training, audits, proficiency testing (sequencing and analysis) and the development of reference standards for the evaluation of different procedures will support the further development of high-quality sequencing programmes for SARS-CoV-2 and for the detection of and response to future emerging pathogens. Where samples are shared in a network, appropriate mechanisms to ship samples under adequate requirements should also be in place.

3. Practical considerations for the implementation of a virus genomic sequencing programme

Here, we provide a general overview of the technical requirements to establish a sequencing programme. For detailed information regarding SARS-CoV-2 sequencing, refer to the full SARS-CoV-2 sequencing implementation guide *(17)*.

3.1 Practical considerations when developing a SARS-CoV-2 sequencing programme

The sequencing objectives will determine the design of the sequencing workstream (Table 1). Relevant key questions to aid this process can be found in Annex I. Annex II contains a checklist with considerations for planning a SARS-CoV-2 sequencing programme. Fig. 2 depicts the workstream for SARS-CoV-2 whole genome sequencing. All staff involved in a sequencing programme should receive the appropriate training and instruction to comply with the mandated task. Key stakeholders should be identified, consulted and involved at an early stage. Stakeholders to be engaged when developing sequencing programmes include public health bodies, diagnostic laboratories, sequencing facilities, analytical groups, and, depending on the setting, infection prevention and control teams or occupational health services, patient advocacy groups, and other institutions involved in human–animal interface research where appropriate. Communication channels and pathways that are aligned with the objectives of the programme should be developed and maintained throughout the project in order to ensure that the sequencing data are used to the greatest effect. Regular evaluations of the project's progress and end evaluation are key to ensure that lessons are learned and improvements are made where needed. Successfully achieving the objectives of a sequencing programme focused on emerging pathogens requires the involvement of experts in different fields: (i) wet-laboratory sequencing and safe handling of virus samples; (ii) generation of accurate genomes from raw data; (iii) analysis of genomes to generate meaningful results that are useful for the outbreak response; and (iv) pathogens. Most experts will be skilled in only one or two of these areas. For (ii) and (iii), powerful computer resources are required to achieve fast results. Collaboration between experts with different skillsets and pooling of resources is therefore often key to generate timely, accurate and effective results that can truly impact public health.

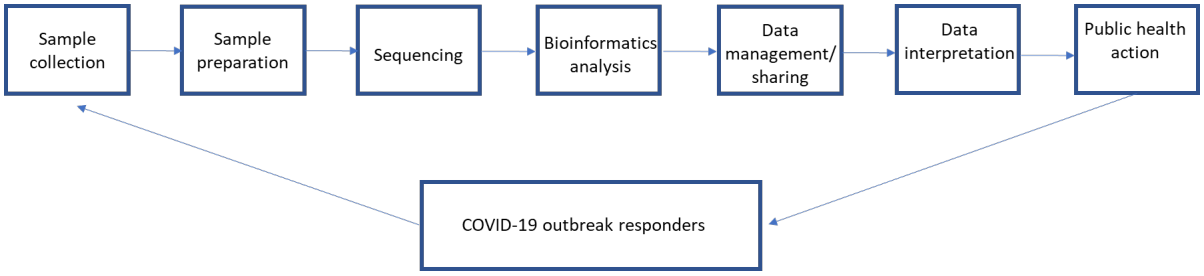

Fig. 2. Workstream for SARS-CoV-2 whole genome sequencing. Note that successfully implementing this workstream will involve bilateral communication between experts involved at different stages; for example, those conducting data interpretation would ideally directly discuss which samples will be chosen from sequencing with those involved in sample selection and preparation.

3.2 Ethical considerations

It is important to review ethical implications when designing a sequencing programme. Possible risks of social harm to research participants should be identified, and mitigation strategies should be defined. Any proposed investigations should be evaluated and approved by an ethical review committee, which takes into account the social value, scientific validity, participant selection, risk–benefit ratio, informed consent, and ongoing respect for participants *(44–46)*. Where researchers are not experienced in identifying possible ethical issues surrounding the sequencing of outbreak pathogens such as SARS-CoV-2, international collaboration and engagement of such expertise is strongly encouraged *(44)*. Collaboration between researchers across the world should ensure equitable and mutually beneficial collaborative research partnerships. Local researchers should be encouraged to take leading and active roles throughout the research process, as they are more likely to understand their health care and research systems and be able to translate results into policy *(44, 45)*. Ethical considerations for genomic sequence and metadata sharing are discussed in section 3.7.

3.3 Considerations for sampling strategy and sample preparation

Once goals have been identified, an appropriate sampling strategy needs to be developed with relevant stakeholders. Details on sampling can be found in the SARS-CoV-2 sequencing implementation guide *(17)*. Ideally, the reasons for the choice of specimens for sequencing should be recorded in the metadata, as the inclusion of non-random subsets of samples can affect the reliability of certain genetic analyses such as phylogenetic and phylodynamic analyses. Practical advice on how to collect clinical samples is covered in the SARS-CoV-2 diagnostic guidance *(47)*. Before sequencing, it is recommended to enrich the sample for SARS-CoV-2 genetic material relative to other genetic material. In this step, take care not to contaminate the sample *(17, 48, 49)*. PCR-based approaches are an inexpensive, rapid and convenient way of increasing the amount of virus genetic material available in a sample prior to sequencing, for example, the approach designed by the ARTIC Network *(51–53)*. For more technical details and how to select the optimal method for different settings, we refer to the SARS-CoV-2 sequencing implementation guide *(17)*. After initial sample preparation to enrich the SARS-CoV-2 genetic material, libraries can typically be prepared using standard sequencing protocols that are appropriate for any virus.

3.4 Laboratory considerations

Sequencing strategies for SARS-CoV-2 include targeted approaches that rely on knowledge of the genome, and metagenomic approaches that do not require prior knowledge of the genomic sequence *(54, 55)*. Annex III summarizes the key advantages and limitations of each commonly used sequencing technology. Before investing in sequencing capacity, consideration should be given to the requirements of the various technologies in terms of human resources, staff competencies, laboratory infrastructure, run-time, costs, ease of use, subsequent data processing, throughput (rate of data production) and sequencing accuracy. The number of samples that need to be analysed will depend on the sequencing objective. When calculating costs, consider not only the procurement of sequencing equipment, but also the recurrent costs for reagents, maintenance and service contracts. This guidance does not cover costs, but an extensive recent overview can be found in *(8)*. Basic infrastructure should be in place to support reliable sequencing, including reliable Internet connection and electricity supply, appropriate environment (e.g., vibration and dust free, temperature and humidity logged and regulated required for some platforms), and logged storage of

samples. Appropriate biosafety and biosecurity measures should be implemented. Assessing the costs and basic infrastructure requirements can help to decide whether the actual sequencing should be done in-house or would be better outsourced. Technology changes rapidly; consequently, certain techniques will become obsolete or manufacturers will shift to different machines and/or reagents. Before making large investments, it is recommended to establish how long the manufacturer will commit to supplying reagents and supporting the maintenance and troubleshooting of the selected platforms of interest. When planning a programme, the availability of ancillary reagents and additional equipment to support the sequencing work should also be taken into account (e.g., extraction methods [either automated or manual], instruments to quantify genetic material, amplification and incubation instruments, sample purification, and sample and reagent storage). Laboratories that conduct genomic sequencing should have high-quality SARS-CoV-2 PCR capacity confirmed by internal and external quality assurance. In addition, for each step in the process, quality indicators should be established and monitored.

3.5 Bioinformatic and computational considerations

Hardware requirements differ depending of the approach taken (for details, we refer to the implementation guide *(17)*). The volume of raw data produced depends on the sequencing method (see Annex III) and the number of samples sequenced *(56)*. The computational power required for data analysis also differs according to the sequencing objective and method. For example, genome phylogenetics and alignment may require high-performance computational power, especially where datasets are large. The costs of the computational architecture required to store and handle these data should be considered when developing a sequencing pipeline. The bioinformatic pipeline will depend on the pre-sequencing laboratory stages, sequencing platform, and reagents used. For a detailed description of bioinformatic pipelines, we refer to the implementation guidance *(17)*.

3.6 Considerations for virus naming and nomenclature

A consistent nomenclature has not yet been established for SARS-CoV-2. In the absence of an agreed upon consistent nomenclature, three main nomenclature strategies are generally used. Lineages or clades can be defined based on viruses that share a phylogenetically determined common ancestor. Both GISAID and Nextstrain aim to provide a broad categorization of globally circulating diversity by naming different phylogenetic clades. Rambaut et al. proposed a dynamic nomenclature for SARS-CoV-2 lineages that focuses on actively circulating virus lineages and those that spread to new locations *(57)*. Software is available to automatically assign new sequences to a lineage and/or clade *(58–60)*. With the increasing diversity in SARS-CoV-2 genomes, the demand for a uniform nomenclature is growing *(57, 61, 62)*. While no consistent nomenclature exists, the best approach would be to list particular lineages and/or clades using all three of the commonly used systems, or at minimum state explicitly which nomenclature is being used.

3.7 Genomic sequence and metadata sharing

The rapid sharing of pathogen GSD, together with the relevant anonymized epidemiological and clinical metadata, will maximize the impact of genomic sequencing in the public health response *(63–65)*. The wide sharing of SARS-CoV-2 sequences, as well as diagnostic protocols, sequencing protocols and samples, has been globally beneficial to achieving worldwide molecular diagnostic capacity *(66–68)*. The scientific/medical community should continue to build on the global collaboration and timely data sharing during SARS-CoV-2 and future emerging outbreaks. There are two distinct choices available for SARS-CoV-2 genomic sequence data sharing: “public-domain” and “public-access” *(69)*. Public-domain databases provide access to data without requiring the identity of those accessing and using the data, for example, the INSDC, operated by DDBJ, EMBL-EBI and NCBI. In public-access databases, such as GISAID, users must identify themselves to ensure transparent use of the data, permit effective oversight, protect the rights of the data contributors, make best efforts to collaborate with data providers, and acknowledge their contribution in published results. The examples mentioned are free of charge and accessible to the public. When pathogen sequencing projects are developed, it is imperative to determine which, if either, of these choices is most appropriate and whether other methods to access and share GSD are necessary *(44)*. One of the critical factors to ensure continued sharing of genetic data is giving due acknowledgement to those who collect clinical samples and generate virus genomic sequences. Data sources should always be acknowledged where publicly available data are used, and related publications and pre-print articles should be cited where available.

Sequence data, including consensus sequences, partial consensus sequences and raw sequence data, can be valuably shared in multiple formats. The quality of the sequence data, including potential contamination with amplicons produced through PCR, should be carefully evaluated prior to sharing. Laboratories should contact sequence-sharing platforms to update previously submitted partial sequences if an error is identified and corrected. Sharing of raw

virus sequencing reads (i.e., all individual sequenced fragments of a virus genome before they are assembled into one consensus genome) is important because it enables the direct comparison of the effect of different bioinformatic approaches for consensus genome generation and facilitates the correction of errors if necessary. Given the large data size of sequenced libraries, the sharing of read-level data may be more challenging in settings that have limited Internet upload speeds or intermittent connections. Any shared data should protect patient anonymity. To ensure patient anonymity, raw data containing human reads must be filtered to retain only non-human (i.e., viral) GSD prior to sharing *(43)*. Sharing of linked metadata, such as date of sample collection or approximate sampling location, is necessary to enable sequence data to be used in many phylogenetic applications. However, which metadata can be reasonably shared without compromising patient anonymity should be carefully considered.

Preliminary analyses of GSD are frequently shared through forums, platforms and preprint servers *(70–72)*. Through their publications, as with all scientific reports, scientists should consider the strengths and weaknesses of their analyses and how analyses might be interpreted or presented by various audiences prior to peer-review. Scientists are encouraged to provide a clear interpretation of their findings so that misunderstandings or misuse of results are minimised.

4. Public health objectives of SARS-CoV-2 genomic sequencing

Summarized below are examples of the key public health objectives of SARS-CoV-2 sequencing; for detailed descriptions, please refer to the SARS-CoV-2 sequencing implementation guide *(17)*.

4.1 Identification and characterization of SARS-CoV-2 and development of countermeasures

The sharing of the complete genetic sequence of the novel virus in early January 2020 was fundamental to characterize SARS-CoV-2, enabling the rapid development of diagnostics and supporting the development of therapies and vaccines *(73–80)*. Genomic sequencing enhances our understanding of the origins and transmission of novel viruses. By studying the initial SARS-CoV-2 genomes available from Wuhan, People's Republic of China and surrounding areas, it was possible to determine the latest possible date of emergence in humans as November–December 2019 *(74, 75, 81, 82)*. Sampling of a wide range of animals is supporting research around the identification of the initial animal source and/or potential intermediate hosts *(81, 83, 84)*.

4.2 Monitoring transmission and geographic spread

Phylogenetics is a method for investigating evolutionary relationships between different organisms using their genetic sequences. It is used in almost every branch of biology and has many important applications in informing public health responses *(17, 85–87)*. The availability of epidemiological or clinical data related to the sampling of the virus genomic sequence (often referred to as metadata, e.g., date of sampling, location of patient, clinical parameters) enhances the interpretation of phylogenetic analyses. Which metadata are required differs according to the objective of the genomic sequencing. The technical aspects of phylogenetic and phylodynamic analyses, metadata and common risks of misinterpretation can be found in the SARS-CoV-2 sequencing implementation guide *(17)*.

4.2.1 Investigating geographical spread and reintroductions between populations

Phylogeographic analyses that use virus genomic sequences and information on sampling location are being used to track SARS-CoV-2 circulation globally *(13, 47, 88–90)*. Phylogeographic reconstructions are often computationally demanding, and subsampling strategies can help to reduce this computational burden. Inferring virus movement or country of origin of specific clades/lineages can be valuable, but should be done with caution because several factors can bias phylogeographic reconstruction. For example, the lack of available SARS-CoV-2 genomes from certain areas can make it less likely that those areas will be reconstructed as the geographic origin of a lineage/clade. Genomic sequences may be associated in some databases with the location of virus sampling, instead of with the suspected location of infection of a patient. Where these locations differ because a patient travelled between the times of infection and virus sampling, phylogeographic analysis can result in an inaccurate reconstruction of the origin of specific clades/lineages *(91)*. These results should be interpreted cautiously and not under the assumption that phylogeographic results represent the true patterns of spatio-temporal (time and space) viral spread.

Methods to infer spatio-temporal spread of an outbreak can also be used to investigate factors that have driven virus dispersal *(92)*. Identifying drivers of transmission may help to shape new strategies for preventing spread. This approach has been used, for example, in the Ebola virus disease outbreaks in West Africa *(93, 94)*. For SARS-CoV-2, several countries have used genomic sequencing to establish the contribution of local transmission compared to

imported cases, and used this information to help make policy decisions *(89, 90, 95–100)*. The phylodynamic identification of factors that are important for understanding transmission is often computationally demanding and requires the curation of extensive data on potential explanatory factors (e.g., human population density, human mobility). Analyses are therefore often completed weeks or months after virus genomic sequencing. However, even retrospective analyses are useful to guide interventions for SARS-CoV-2 or potential emerging pathogens.

4.2.2 Evaluating evidence on transmission routes or clusters

Phylogenetic clustering has been used to support cluster and outbreak investigations for SARS-CoV-2. Analyses of transmission clusters can guide local decisions on whether control measures are needed to prevent future transmission in identified outbreak settings *(101)*. Given the relatively slow evolutionary rate of SARS-CoV-2 (i.e., one nucleotide substitution every two weeks), it is expected that many individual transmission events will not be traceable based on genomic sequence data *(35)*. Phylogenetic clustering of sequences from patients with the same hypothesised source of exposure would be consistent with (although not strong evidence of) that exposure. By contrast, phylogenetic separation of virus sequences from patients with the same hypothesised source of exposure would strongly indicate that the common source of infection has been incorrectly identified.

4.2.3 Quantifying periods of transmission and following the reproduction number over time

Molecular clock phylogenetic approaches can help to estimate the upper and lower limits of time of circulation of the sampled virus's genetic lineages in a given population *(74, 90, 102–106)*. This approach can provide more accurate information on the period of viral transmission than the clinical identification of cases, particularly in the early or late phases of an outbreak when surveillance is limited. Studying the variation in the genomic sequences detected can determine whether there is clinically undetected local transmission. In these settings, improved diagnostic surveillance programmes would need to be implemented where undetected circulation is suspected.

Genomic sequence analysis can also estimate how many individuals are infected by one individual in a given population (the reproduction number [R_0]) and support the assessment of relative changes in outbreak size over time. This information could be used to evaluate the impact of specific control measures.

4.2.4 Environmental surveillance in wastewater and sludge

For pathogens such as poliovirus, wastewater monitoring is an important tool for tracing the silent circulation of viruses in a community. This approach provides opportunities to detect circulation (before the initial patients have been clinically detected), estimate prevalence, and understand the genetic linkage and diversity *(107, 108)*. Several countries have demonstrated molecular detection of SARS-CoV-2 RNA in wastewater *(109–115)*. Consequently, environmental surveillance is a promising approach, especially in low prevalence settings, to identify unrecognized carriers and serve as an "early warning" system for SARS-CoV-2 introduction or changes in prevalence *(109, 116, 117)*.

4.2.5 Investigating potential reinfections

Seasonal coronaviruses can reinfect humans *(118)*. For SARS-CoV-2, cases of reinfection have been documented *(119–124)*. In this context, the SARS-CoV-2 genomic sequences sampled from the first and subsequent episodes can be compared to determine whether the renewed detection of SARS-CoV-2 in an individual is a reinfection or the result of prolonged viral shedding *(125, 126)*. If sequences from each episode have strong genetic distinctions, such as occurring in different, well-supported lineages/clades, subsequent episodes can be considered to be reinfections. Concomitant serological investigations are necessary to understand whether reinfection is associated with an antigenically distinct strain or the lack of a protective immune response from the initial infection. Sequencing can therefore support the improved understanding of the frequency and potential risk factors for reinfection *(125, 126)*.

4.3 Monitoring the evolution of SARS-CoV-2

4.3.1 Structured evaluation of possibly relevant mutations

Genomic sequencing can be used to identify genetic substitutions that may change viral infection characteristics (phenotypic change), such as transmissibility or virulence. All viruses acquire genetic changes as they circulate, but the vast majority of acquired changes do not substantially affect virus behaviour. Nevertheless, rare genetic changes

in SARS-CoV-2 may cause relevant phenotypic changes of public health importance. Identifying and demonstrating the impact of such changes is challenging. In general, it is difficult to confidently establish whether the increase in relative prevalence of given mutation(s) over time is due to a phenotypic difference. The predominance of a specific viral clade/lineage in a population, for example, may be due more to the behaviour of the human population infected than to the behaviour of the virus itself. Most of the time, such patterns are likely to be stochastic. However, if phylogenetic analysis hints at the potential epidemiological or clinical impact of specific mutations/variants, properly conducted clinical genomic studies are required in order to evaluate candidate variants that might confer clinically observed phenotypic changes to the virus. Genetic changes proposed to cause phenotypic changes should be evaluated using standardized approaches, including protein modelling studies to assess potential impact and in vitro or in vivo experiments with a mutant virus (clones) with the specific mutations of concern to confirm or reject the specific properties of the candidate variants. A dedicated WHO Working Group has been established, derived from the WHO SARS-CoV-2 Reference Laboratory Network. This SARS-CoV-2 Evolution Working Group (SEWG) focuses on SARS-CoV-2 evolution to provide WHO with the timely identification and evaluation of potentially relevant mutations, as well as advice for risk mitigation *(16, 40, 127)*.

4.3.2 Monitoring the impact of SARS-CoV-2 evolution on countermeasures

At minimum, global surveillance of SARS-CoV-2 genomes should ideally detect the emergence of SARS-CoV-2 lineages with genetic variants that impact the effectiveness of countermeasures. Monitoring for SARS-CoV-2 genomic changes that might reduce vaccine efficacy should accompany the roll-out of SARS-CoV-2 vaccination campaigns. Monitoring and investigation into the possible causes of vaccine failure should include genomic evaluations to assess potential viral escape mutants. Additionally, sequencing can aid in the identification of escape mutants for monoclonal antibodies *(128)* and future therapeutics. Genomic monitoring to identify drug resistance has been used for other pathogens, including influenza, HIV and *Mycobacterium tuberculosis (9, 129)*.

Genomic sequencing can also be used to monitor viral genetic changes that affect molecular diagnostics. Using multiple targets for SARS-CoV-2 detection, such as a multiplex PCR targeted at two or more regions of the virus genome, is a cost-effective approach to reduce the chance of false negatives in assays as a result of the virus's evolution (*47,127*). Sequencing of the virus genome or target gene could be undertaken when there is consistent failure to detect one target or newly observed differences in the sensitivity of assays targeting different regions in order to identify the possible cause. For additional information, see the interim guidance on the diagnosis of SARS-CoV-2 infections *(47)*. Viral mutations can also impact the antigen or serological assay, and genomic sequencing can help to detect the potential failure of such assays at an early stage *(130–132)*.

4.3.3 SARS-CoV-2 evolution at the human–animal interface

When a virus transmits from one species to another, the virus may adapt to its new host. The ACE-2 receptor that is targeted by SARS-CoV-2 is similar across humans and a wide variety of animals (mostly mammals) *(133, 134)*. There is therefore potential for human to animal transmission (anthroponosis). While ACE-2 homology suggests that other animals may be susceptible to SARS-CoV-2, other proteins critical for viral replication might differ and prevent infections in these candidate animals. Therefore, appropriate real-world or experimental infection data are needed to determine the susceptibility of specific animals. Various animals have been shown to be susceptible to SARS-CoV-2 *(15, 127, 135–151)*, and SARS-CoV-2 is known to be transmissible in certain animal species (e.g., minks and hamsters). The resistance of some animal species to infection with SARS-CoV-2 has also been demonstrated. Genetic changes in the sequence that encodes the viral spike protein that binds to the ACE-2 receptors may emerge and facilitate a jump to new host species. The SARS-CoV-2 spike protein, especially the RBD, is a critical target for natural and vaccine-induced immunity *(28–32)*. Diversification of genomic regions encoding the spike protein has already been observed in cases where humans infected with SARS-CoV-2 have infected minks and there has been secondary zoonotic transmission back to humans *(149)*. Therefore, diversification of the spike gene following the human–animal exchange of SARS-CoV-2 likely increases the risk of strains emerging that can easily reinfect humans and may be associated with reduced vaccine efficacy or amenability to monoclonal antibody therapies *(33)*. To prevent these events from occurring, countries are encouraged to conduct risk assessments on the potential spread to other species living with or near humans in domestic, rural, agricultural or other zoological settings *(127, 152–154)*. Risk mitigation strategies need to be established and adequate monitoring is needed to ensure the timely detection of these events. Monitoring requires resources, and targeted strategies should be adopted where possible. This requires a One Health strategy in which different disciplines work together, including public, clinical and occupational health, veterinary and wildlife authorities, and forestry and natural resources management *(127, 152, 155–157)*. This collaboration should also focus on the development of joint outbreak investigation and infection prevention and

control protocols, testing of potentially infected humans and animals, and the sharing of sequence data. Where anthroponotic infection or secondary zoonotic infection is observed, sequencing of virus genomes can help to evaluate the possible new risks associated with these events.

Methods

This interim guidance was developed in conjunction with the implementation guide *Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health*. The implementation guide was developed in consultation with experts with experience in the various fields of genomic sequencing from the Global Laboratory Alliance of High Threat Pathogens (GLAD-HP), the reference network for confirmatory testing for COVID-19, and the Global Outbreak Alert and Response Network (GOARN). After initial discussions by a technical writing group led by a temporary advisor and members of the WHO COVID-19 Laboratory Team, contributions were sought from other experts within and outside WHO, and two online meetings were held to resolve outstanding questions. Subsequently, this interim guidance was drafted for national stakeholders, containing a summary of relevant information from this interim guidance and additional relevant information for this target audience. This interim guidance was subsequently circulated for input to the experts that supported the drafting of the implementation guide, the reference network for confirmatory testing for COVID-19, the regional laboratory focal points and other stakeholders, as listed in the acknowledgements.

Plans for updating

WHO continues to monitor the situation closely for any changes that may affect this interim guidance. Should any factors change, WHO will issue a further update. Otherwise, this interim guidance will expire one year after the date of publication.

Contributors

WHO steering group: Celine Barnadas; Sebastian Cognat; Roger Evans; Bruce Allan Gordon; Varja Grabovac; Rebecca Grant; Francis Inbanathan; Frank Konings; Karen Nahapetyan; Marco Marklewitz; Marie-jo Medina; Kate Olive Medlicott; Mick Mulders; Mark D Perkins; Magdi Samaan; Oliver Schmoll, Maria Van Kerkhove; Karin von Eije; Joanna Zwetyenga.

External contributors: Kim Benschop, Netherlands National Institute for Public Health and the Environment (RIVM), Bilthoven, Netherlands; Antonino di Caro, Instituto Nazionale per le Malattie Infettive Lazzaro Spallanzani, Italy; Nuno Rodrigues Faria, Imperial College, London and University of Oxford, Oxford, United Kingdom; Tanya Golubchik, University of Oxford, Oxford, United Kingdom; Keith Hamilton, World Organization for Animal Health (OIE); Edward Holmes, University of Sydney, Sydney, Australia; Sarah C Hill, Royal Veterinary College, London and University of Oxford, Oxford, United Kingdom; Erik Karlsson, Institut Pasteur de Cambodge, Cambodia; Meng Ling Moi, Nagasaki University, Nagasaki, Japan; Leo Poon, Hong Kong University, Hong Kong Special Administrative Region (SAR), China; James Shepherd, University of Glasgow, Glasgow, United Kingdom; Etienne Simon-Loriere, Pasteur Institute, Paris, France. And the other experts that contributed to the implementation guide for SARS-CoV-2 sequencing that served as the basis for this document: Kristian Andersen, Scripps Research, La Jolla, CA, USA; Julio Croda, Ministry of Health, Rio de Janeiro, Brazil; Túlio de Oliveira, University of KwaZulu-Natal, Durban, South Africa; Simon Dellicour, Free University of Brussels, Brussels, Belgium; Nathan Grubaugh, Yale University, New Haven, CT, USA; Liana Kafetzopoulou, KU Leuven – University of Leuven, Belgium; Marion Koopmans, Erasmus MC, Rotterdam, Netherlands; Tommy Lam, University of Hong Kong , Hong Kong SAR, China; Philippe Lemey, KU Leuven – University of Leuven, Belgium; Tze Minn Mak, National Centre for Infectious Diseases, Singapore; Marcio Roberto Nunes, Evandro Chagas Institute, Ananindeua, Pará, Brazil; Bas Oude Munnink, Erasmus MC, Rotterdam, Netherlands; Gustavo Palacios, United States Agency for International Development, Washington, DC, USA; Steven Pullan, Public Health England, London, United Kingdom; Timothy Vaughan, Eidgenössische Technische Hochschule Zurich (ETH Zurich), Zurich, Switzerland; Josh Quick, University of Birmingham, Birmingham, United Kingdom; Andrew Rambaut, University of Edinburgh, Edinburgh, United Kingdom; Chantal Reusken, RIVM, Bilthoven, Netherlands; Tanja Stadler, Eidgenössische Technische Hochschule Zurich (ETH Zurich), Switzerland; Marc Suchard, University of California at Los Angeles, Los Angeles, CA, USA;

Huaiyu Tian, Beijing Normal University, Beijing, China; Lia van der Hoek, Amsterdam Medical Centre, Amsterdam, Netherlands; Erik Volz, Imperial College, London, United Kingdom.

Declarations of interest

All contributors submitted declaration of interest documents for review. Contributors who were determined to have a potential conflict of interest or bias toward specific products were excluded from advising on platform selection.

Funder

Funded by WHO

References

1. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. Science. 2009;324:1557–61. doi: 10.1126/science.1176062.
2. Rambaut A, Holmes E. The early molecular epidemiology of the swine-origin A/H1N1 human influenza pandemic. PLoS Curr. 2009;1:RRN1003. doi: 10.1371/currents.rrn1003.
3. Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature. 2009;459:1122–5. doi: 10.1038/nature08182.
4. Mena I, Nelson MI, Quezada-Monroy F, Dutta J, Cortes-Fernández R, Lara-Puente JH, et al. Origins of the 2009 H1N1 influenza pandemic in swine in Mexico. eLife. 2016;5:e16777. doi: 10.7554/eLife.16777.
5. Ladner JT, Wiley MR, Mate S, Dudas G, Prieto K, Lovett S, et al. Evolution and spread of Ebola virus in Liberia, 2014–2015. Cell Host Microbe. 2015;18:659–69. doi: 10.1016/j.chom.2015.11.008.
6. Stadler T, Kühnert D, Rasmussen DA, Plessis dL. Insights into the early epidemic spread of Ebola in Sierra Leone provided by viral sequence data. PLoS Curr. 2014;6. doi: 10.1371/currents.outbreaks.02bc6d927ecee7bbd33532ec8ba6a25f.
7. Smits SL, Pas SD, Reusken CB, Haagmans BL, Pertile P, Cancedda C, et al. Genotypic anomaly in Ebola virus strains circulating in Magazine Wharf area, Freetown, Sierra Leone, 2015. Euro Surveill. 2015;20. doi: 10.2807/1560-7917.ES.2015.20.40.30035.
8. GLASS whole-genome sequencing for surveillance of antimicrobioal resistance. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/handle/10665/334354, accessed 20 November 2020).
9. The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guidance. Geneva: World Health Organization; 2018 (https://apps.who.int/iris/handle/10665/274443, accessed 15 November 2020).
10. Whole genome sequencing for foodborne disease surveillance: landscape paper. Geneva: World Health Organization; 2018 (https://apps.who.int/iris/handle/10665/272430, accessed 25 November 2020).
11. Next-generation sequencing of influenza viruses: general information for national influenza centres. Geneva: World Health Organization; 2020 (https://www.who.int/influenza/gisrs_laboratory/national_influenza_centres/NGS_guidance_for_NICs.pdf?ua=1, accessed 20 November 2020).
12. GISAID (https://www.gisaid.org/, accessed 5 January 2021).
13. Genomic epidemiology of novel coronavirus: global subsampling [website]. Nextstrain; 2020 (https://nextstrain.org/ncov/global, accessed 4 December 2020).
14. Volz E, Baguelin M, Bhatia S, Boonyasiri A, Cori A, Cucunuba Z, et al. Report 5 - phylogenetic analysis of SARS-CoV-2. London: Imperial College London; 2020 (http://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-5-phylogenetics-of-sars-cov-2/, accessed 26 June 2020).
15. Oude Munnink BB, Sikkema RS, Nieuwenhuijse DF, Molenaar RJ, Munger E, Molenkamp R, et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. Science. 2020. doi: 10.1126/science.abe5901.
16. Coronavirus disease (COVID-19): situation report – 185. Geneva: World Health Organization; 2020 (https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200723-covid-19-sitrep-185.pdf?sfvrsn=9395b7bf_2, accessed 15 November 2020).
17. Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health. Geneva: World Health Organization; 2020. (https://apps.who.int/iris/bitstream/handle/10665/338480/9789240018440-eng.pdf?sequence=1&isAllowed=y, accessed 8 January 2021)
18. International Committee on Taxonomy of Viruses (ICTV); 2020 (https://talk.ictvonline.org/, accessed 27 July 2020).
19. Gorbalenya ABS, Baric R, de Groot R, Drosten C, Gulyaeva A, Haagmans B, et al. The species *Severe acute respiratory syndrome-related coronavirus*: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020;5:536–44. doi: 10.1038/s41564-020-0695-z.
20. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382:727–33. doi: 10.1056/NEJMoa2001017.
21. Naqvi AAT, Fatima K, Mohammad T, Fatima U, Singh IK, Singh A, et al. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: structural genomics approach. Biochim Biophys Acta Mol Basis Dis. 2020;1866:165878. doi: 10.1016/j.bbadis.2020.165878.
22. Yoshimoto FK. The proteins of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2 or n-CoV19), the cause of COVID-19. Protein J. 2020;39:198–216. doi: 10.1007/s10930-020-09901-4.

23. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. The architecture of SARS-CoV-2 transcriptome. Cell. 2020;181:914–21 e10. doi: 10.1016/j.cell.2020.04.011.
24. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020;395:565–74. doi: 10.1016/S0140-6736(20)30251-8.
25. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science. 2020;367:1444–8. doi: 10.1126/science.abb2762.
26. Ni W, Yang X, Yang D, Bao J, Li R, Xiao Y, et al. Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. Crit Care. 2020;24:422. doi: 10.1186/s13054-020-03120-0.
27. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020;181:271–80 e8. doi: 10.1016/j.cell.2020.02.052.
28. Amanat F, Krammer F. SARS-CoV-2 vaccines: status report. Immunity. 2020;52:583–9. doi: 10.1016/j.immuni.2020.03.007.
29. Brouwer PJM, Caniels TG, van der Straten K, Snitselaar JL, Aldon Y, Bangaru S, et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. Science. 2020;369:643–50. doi: 10.1126/science.abc5902.
30. Tai W, He L, Zhang X, Pu J, Voronin D, Jiang S, et al. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. Cell Mol Immunol. 2020;17:613–20. doi: 10.1038/s41423-020-0400-4.
31. Shi R, Shan C, Duan X, Chen Z, Liu P, Song J, et al. A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. Nature. 2020;584:120–4. doi: 10.1038/s41586-020-2381-y.
32. Rogers TF, Zhao F, Huang D, Beutler N, Burns A, He WT, et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. Science. 2020;369:956–63. doi: 10.1126/science.abc7520.
33. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell. 2020;182:1284–94 e9. doi: 10.1016/j.cell.2020.07.012.
34. Candido DS, Claro IM, de Jesus JG, Souza WM, Moreira FRR, Dellicour S, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. Science. 2020;369:1255–60. doi: 10.1126/science.abd2161.
35. van Dorp L, Acman M, Richard D, Shaw LP, Ford CE, Ormond L, et al. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. Infect Genet Evol. 2020;83:104351. doi: 10.1016/j.meegid.2020.104351.
36. Xu W, Zhang Y, Wang H, Zhu Z, Mao N, Mulders MN, et al. Global and national laboratory networks support high quality surveillance for measles and rubella. Int Health. 2017;9:184–9. doi: 10.1093/inthealth/ihx017.
37. Mulders MN, Serhan F, Goodson JL, Icenogle J, Johnson BW, Rota PA. Expansion of surveillance for vaccine-preventable diseases: building on the global polio laboratory network and the global measles and rubella laboratory network platforms. J Infect Dis. 2017;216:S324–S30. doi: 10.1093/infdis/jix077.
38. Diop OM, Kew OM, de Gourville EM, Pallansch MA. The global polio laboratory network as a platform for the viral vaccine-preventable and emerging diseases laboratory networks. J Infect Dis. 2017;216:S299–S307. doi: 10.1093/infdis/jix092.
39. Hay AJ, McCauley JW. The WHO Global Influenza Surveillance and Response System (GISRS): a future perspective. Influenza Other Respir Viruses. 2018;12:551–7. doi: 10.1111/irv.12565.
40. Terms of reference for WHO reference laboratories providing confirmatory testing for COVID-19. Geneva: World Health Organization; 2020 (https://www.who.int/publications/m/item/terms-of-reference-for-who-reference-laboratories-providing-confirmatory-testing-for-covid-19, accessed 26 June 2020).
41. RubeNS database for rubella sequences (http://www.who-rubella.org/, accessed 26 June 2020).
42. MeaNS: Measles nucleotide surveillance (http://www.who-measles.org, accessed 26 June 2020).
43. Roy S, LaFramboise WA, Nikiforov YE, Nikiforova MN, Routbort MJ, Pfeifer J, et al. Next-generation sequencing informatics: challenges and strategies for implementation in a clinical environment. Arch Pathol Lab Med. 2016;140:958–75. doi: 10.5858/arpa.2015-0507-RA.
44. Mutenherwa F, Wassenaar DR, de Oliveira T. Experts' perspectives on key ethical issues associated with HIV phylogenetics as applied in HIV transmission dynamics research. J Empir Res Hum Res Ethics. 2019;14:61–77. doi: 10.1177/1556264618809608.
45. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? JAMA. 2000;283:2701–11. doi: 10.1001/jama.283.20.2701.
46. WHO guidelines on ethical issues in public health surveillance. Geneva: World Health Organization; 2017 (https://www.who.int/ethics/publications/public-health-surveillance/en/, accessed 15 November 2020).
47. Diagnostic testing for SARS-CoV-2: interim guidance. 11 September 2020. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2, accessed 6 December 2020).
48. MacCannell D. SARS-CoV-2 sequencing (https://github.com/CDCgov/SARS-CoV-2_Sequencing, accessed 1 November 2020).
49. Cesare MD. Probe-based target enrichment of SARS-CoV-2 [Protocol]. Univeristy of Oxford; 2020. doi: 10.17504/protocols.io.bd5di826.
50. Vogels CBF, Brito AF, Wyllie AL, Fauver JR, Ott IM, Kalinich CC, et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. Nat Microbiol. 2020:5:1299–1305. doi: 10.1038/s41564-020-0761-6.
51. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for minION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc. 2017;12:1261–76. doi: 10.1038/nprot.2017.066.
52. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. Genome Biol. 2019;20:8. doi: 10.1186/s13059-018-1618-7.
53. Matteson N, Grubaugh N, Gangavarapu K, Quick J, Loman N, Andersen K. PrimalSeq: Generation of tiled virus amplicons for MiSeq sequencing [Protocol]. 2020. doi: 10.17504/protocols.io.bez7jf9n.
54. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol. 2017;35:833–44. doi: 10.1038/nbt.3935.
55. Bragg L, Tyson GW. Metagenomics using next-generation sequencing. Methods Mol Biol. 2014;1096:183–201. doi: 10.1007/978-1-62703-712-9_15.

56. Xiao M, Liu X, Ji J, Li M, Li J, Yang L, et al. Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. Genome Med. 2020;12:57. doi: 10.1186/s13073-020-00751-4.
57. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol. 2020;5:1403–7. doi: 10.1038/s41564-020-0770-5.
58. Singer J, Gifford R, Cotten M, Robertson D. CoV-GLUE: A web application for tracking SARS-CoV-2 genomic variation. Preprints. 2020:2020060225. doi: 10.20944/preprints202006.0225.v1.
59. CoVsurver: mutation analysis of hCoV-19. GISAID (https://www.gisaid.org/epiflu-applications/covsurver-mutations-app/, accessed 11 December 2020).
60. Pangolin COVID-19 lineage assigner (https://pangolin.cog-uk.io/, accessed 11 December 2020).
61. Fauquet CM, Fargette D. International Committee on Taxonomy of Viruses and the 3,142 unassigned species. Virol J. 2005;2:64. doi: 10.1186/1743-422X-2-64.
62. Alm E, Broberg EK, Connor T, Hodcroft EB, Komissarov AB, Maurer-Stroh S, et al. Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020. Euro Surveill. 2020;25. doi: 10.2807/1560-7917.ES.2020.25.32.2001410.
63. Policy statement on data sharing by WHO in the context of public health emergencies (as of 13 April 2016). Geneva: World Health Organization Geneva; 2016 (https://apps.who.int/iris/handle/10665/254440, accessed 25 November 2020).
64. Pandemic influenza preparedness framework, for the sharing of influenza viruses and access to vaccines and other benefits. Geneva: World Health Organization; 2011 (https://apps.who.int/gb/pip/pdf_files/pandemic-influenza-preparedness-en.pdf, accessed 20 November 2020).
65. Executive Board, 140th session, provisional agenda item 7.5 Review of the pandemic influenza preparedness framework, report by the Director-General. Geneva: World Health Organization; 2016 (https://apps.who.int/gb/ebwha/pdf_files/EB140/B140_16-en.pdf?ua=1, accessed 15 November 2020).
66. Laboratory testing strategy recommendations for COVID-19. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/handle/10665/331509, accessed 6 December 2020).
67. Guidance for laboratories shipping specimens to WHO reference laboratories that provide confirmatory testing for COVID-19 virus. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/handle/10665/331639, accessed 4 December 2020).
68. Molecular assays to diagnose COVID-19: summary table of available protocols. Geneva: World Health Organization; 2020 (https://www.who.int/who-documents-detail/molecular-assays-to-diagnose-covid-19-summary-table-of-available-protocols, accessed 4 December 2020).
69. Fact sheet: genetic sequence data and databases. Geneva: World Health Organization; 2018 (https://www.who.int/influenza/pip/GSD_EN_V2_10Sep2018.pdf?ua=1, accessed 11 December 2020).
70. medRxiv: The Preprint Server for Health Sciences (https://www.medrxiv.org/, accessed 1 November 2020).
71. bioRxiv: The Preprint Server for Biology (https://www.biorxiv.org/, accessed 1 November 2020).
72. Virological (https://virological.org/, accessed 1 November 2020).
73. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579:265–9. doi: 10.1038/s41586-020-2008-3.
74. Lu J, du Plessis L, Liu Z, Hill V, Kang M, Lin H, et al. Genomic epidemiology of SARS-CoV-2 in Guangdong Province, China. Cell. 2020;181:997–1003.e9. doi: 10.1016/j.cell.2020.04.023.
75. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579:270–3. doi: 10.1038/s41586-020-2012-7.
76. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 2020;25. doi: 10.2807/1560-7917.ES.2020.25.3.2000045.
77. Detection of 2019 novel coronavirus (2019-nCoV) in suspected human cases by RT-PCR. Hong Kong University Medical School; 2020 (https://www.who.int/docs/default-source/coronaviruse/peiris-protocol-16-1-20.pdf?sfvrsn=af1aac73_4, accessed 1 December 2020).
78. Melén K, Kakkola L, He F, Airenne K, Vapalahti O, Karlberg H, et al. Production, purification and immunogenicity of recombinant Ebola virus proteins: a comparison of Freund's adjuvant and adjuvant system 03. J Virol Methods. 2017;242:35–45. doi: 10.1016/j.jviromet.2016.12.014.
79. Draft landscape of COVID-19 candidate vaccines. Geneva: World Health Organization (https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines, accessed 26 June 2020).
80. Ren Y, Zhou Z, Liu J, Lin L, Li S, Wang H, et al. A strategy for searching antigenic regions in the SARS-CoV spike protein. Genomics Proteomics Bioinformatics. 2003;1:207–15. doi: 10.1016/s1672-0229(03)01026-x.
81. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med. 2020;382:1199–207. doi: 10.1056/NEJMoa2001316.
82. Andersen K. Clock and TMRCA based on 27 genomes. Scripps Research; 2020 (https://virological.org/t/clock-and-tmrca-based-on-27-genomes/347, accessed 26 June 2020).
83. Report of the WHO–China Joint Mission on coronavirus disease 2019 (COVID-19). Geneva: World Health Organization; 2020 (https://www.who.int/publications-detail-redirect/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19), accessed 15 July 2020).
84. Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol. 2019;17:181–92. doi: 10.1038/s41579-018-0118-9.
85. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. Science. 2004;303:327–32. doi: 10.1126/science.1090727.
86. Volz EM, Koelle K, Bedford T. Viral phylodynamics. PLoS Comput Biol. 2013;9. doi: 10.1371/journal.pcbi.1002947.
87. Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. Nat Rev Genet. 2009;10:540–50. doi: 10.1038/nrg2583.

88. Lai A, Bergna A, Caucci S, Clementi N, Vicenti I, Dragoni F, et al. Molecular tracing of SARS-CoV-2 in Italy in the first three months of the epidemic. Viruses. 2020;12. doi: 10.3390/v12080798.
89. Fauver JR, Petrone ME, Hodcroft EB, Shioda K, Ehrlich HY, Watts AG, et al. Coast-to-coast spread of SARS-CoV-2 during the early epidemic in the United States. Cell. 2020;181:990–6.e5. doi: 10.1016/j.cell.2020.04.021.
90. Candido DdS, Claro IM, Jesus dJG, Souza dWM, Moreira FRR, Dellicour S, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. Science. 2020;369:1255–60. doi: 10.1101/2020.06.11.20128249.
91. Lemey P, Hong S, Hill V, Baele G, Poletto C, Colizza V, et al. Accommodating individual travel history, global mobility, and unsampled diversity in phylogeography: a SARS-CoV-2 case study. bioRxiv. 2020. doi: 10.1101/2020.06.22.165464.
92. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. PLoS Pathog. 2014;10:e1003932. doi: 10.1371/journal.ppat.1003932.
93. Dudas G, Carvalho LM, Bedford T, Tatem AJ, Baele G, Faria NR, et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. Nature. 2017;544:309–15. doi: 10.1038/nature22040.
94. Dellicour S, Baele G, Dudas G, Faria NR, Pybus OG, Suchard MA, et al. Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak. Nat Commun. 2018;9:1–9. doi: 10.1038/s41467-018-03763-2.
95. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS Comput Biol. 2009;5:e1000520. doi: 10.1371/journal.pcbi.1000520.
96. Lemey P, Rambaut A, Welch JJ, Suchard MA. Phylogeography takes a relaxed random walk in continuous space and time. Mol Biol Evol. 2010;27:1877–85. doi: 10.1093/molbev/msq067.
97. Bloomquist EW, Lemey P, Suchard MA. Three roads diverged? Routes to phylogeographic inference. Trends Ecol Evol. 2010;25:626–32. doi: 10.1016/j.tree.2010.08.010.
98. Faria NR, Suchard MA, Rambaut A, Lemey P. Towards a quantitative understanding of viral phylogeography. Curr Opin Virol. 2011;1:423–9. doi: 10.1016/j.coviro.2011.10.003.
99. Lemey P, Hong S, Hill V, Baele G, Poletto C, Colizza V, et al. Accommodating individual travel history, global mobility, and unsampled diversity in phylogeography: A SARS-CoV-2 case study. bioRxiv. 2020:165464. doi: 10.1101/2020.06.22.165464.
100. Reusken CB, Buiting A, Bleeker-Rovers C, Diederen B, Hooiveld M, Friesema I, et al. Rapid assessment of regional SARS-CoV-2 community transmission through a convenience sample of healthcare workers, the Netherlands, March 2020. Euro Surveill. 2020;25. doi: 10.2807/1560-7917.ES.2020.25.12.2000334.
101. Worby CJ, Lipsitch M, Hanage WP. Shared genomic variants: identification of transmission routes using pathogen deep-sequence data. Am J Epidemiol. 2017;186:1209–16. doi: 10.1093/aje/kwx182.
102. Duchene S, Featherstone L, Haritopoulou-Sinanidou M, Rambaut A, Lemey P, Baele G. Temporal signal and the phylodynamic threshold of SARS-CoV-2. bioRxiv. 2020:077735. doi: 10.1101/2020.05.04.077735.
103. Volz E, Fu H, Wang H, Xi X, Chen W, Liu D, et al. Genomic epidemiology of a densely sampled COVID19 outbreak in China. medRxiv. 2020:20033365. doi: 10.1101/2020.03.09.20033365.
104. Bedford T, Greninger AL, Roychoudhury P, Starita LM, Famulare M, Huang M-L, et al. Cryptic transmission of SARS-CoV-2 in Washington State. Science. 2020;370:571–5. doi: 10.1101/2020.04.02.20051417.
105. Zehender G, Lai A, Bergna A, Meroni L, Riva A, Balotta C, et al. Genomic characterization and phylogenetic analysis of SARS-CoV-2 in Italy. J Med Virol. 2020;92:1637–40. doi: 10.1002/jmv.25794.
106. Worobey MA-O, Pekar JA-O, Larsen BA-O, Nelson MA-O, Hill V, Joy JB, et al. The emergence of SARS-CoV-2 in Europe and North America. Science. 2020;370:564–70. doi: 10.1126/science.abc8169.
107. Asghar H, Diop OM, Weldegebriel G, Malik F, Shetty S, El Bassioni L, et al. Environmental surveillance for polioviruses in the Global Polio Eradication Initiative. J Infect Dis. 2014;210 Suppl 1:S294–303. doi: 10.1093/infdis/jiu384.
108. Paul JR, Trask JD, Gard S. Ii. Poliomyelitic virus in urban sewage. J Exp Med. 1940;71:765–77. doi: 10.1084/jem.71.6.765.
109. Nemudryi A, Nemudraia A, Wiegand T, Surya K, Buyukyoruk M, Cicha C, et al. Temporal detection and phylogenetic assessment of SARS-CoV-2 in municipal wastewater. Cell Rep Med. 2020;1:100098. doi: 10.1016/j.xcrm.2020.100098.
110. Wu F, Zhang J, Xiao A, Gu X, Lee WL, Armas F, et al. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. mSystems. 2020;5. doi: 10.1128/mSystems.00614-20.
111. Ahmed W, Angel N, Edson J, Bibby K, Bivins A, O'Brien JW, et al. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. Sci Total Environ. 2020;728:138764. doi: 10.1016/j.scitotenv.2020.138764.
112. Wurtzer SMV, Mouchel JM, Maday Y, Teyssou R, Richard E, Almayrac JL, Moulin L. Evaluation of lockdown impact on SARS-CoV-2 dynamics through viral genome quantification in Paris wastewaters. medRxiv. 2020. doi: 10.1101/2020.04.12.20062679.
113. La Rosa G, Iaconelli M, Mancini P, Bonanno Ferraro G, Veneri C, Bonadonna L, et al. First detection of SARS-CoV-2 in untreated wastewaters in Italy. Sci Total Environ. 2020;736:139652. doi: 10.1016/j.scitotenv.2020.139652.
114. Gertjan Medema LH, Goffe Elsinga, Ronald Italiaander, Anke Brouwer. Presence of SARS-coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in the Netherlands. Environ Sci Technol Lett. 2020. doi: 10.1021/acs.estlett.0c00357.
115. Lodder W, de Roda Husman AM. SARS-CoV-2 in wastewater: potential health risk, but also data source. Lancet Gastroenterol Hepatol. 2020;5:533–4. doi: 10.1016/S2468-1253(20)30087-X.
116. Scientific brief: status of environmental surveillance for SARS-CoV-2 virus. Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/WHO-2019-nCoV-sci-brief-environmentalSampling-2020-1, accessed 12 December 2020).
117. Rapid expert consultation on environmental surveillance of SARS-CoV-2 in wastewater: summary report. Geneva: World Health Organization; 2020 (https://www.euro.who.int/en/health-topics/environment-and-health/water-and-sanitation/publications/2020/rapid-expert-consultation-on-environmental-surveillance-of-sars-cov-2-in-wastewater-summary-report-2020, accessed 12 December 2020).
118. Edridge AWD, Kaczorowska JM, Hoste ACR, Bakker M, Klein M, Jebbink MF, et al. Coronavirus protective immunity is short lasting. medRxiv. 2020. doi: 10.1101/2020.05.11.20086439.

119. To KK, Hung IF, Ip JD, Chu AW, Chan WM, Tam AR, et al. COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing. Clin Infect Dis. 2020. doi: 10.1093/cid/ciaa1275.
120. Goldman JD, Wang K, Roltgen K, Nielsen SCA, Roach JC, Naccache SN, et al. Reinfection with SARS-CoV-2 and failure of humoral immunity: a case report. medRxiv. 2020. doi: 10.1101/2020.09.22.20192443.
121. Gupta V, Bhoyar RC, Jain A, Srivastava S, Upadhayay R, Imran M, et al. Asymptomatic reinfection in two healthcare workers from India with genetically distinct SARS-CoV-2. Clin Infect Dis. 2020. doi: 10.1093/cid/ciaa1451.
122. Mulder M, van der Vegt D, Oude Munnink BB, GeurtsvanKessel CH, van de Bovenkamp J, Sikkema RS, et al. Reinfection of SARS-CoV-2 in an immunocompromised patient: a case report. Clin Infect Dis. 2020. doi: 10.1093/cid/ciaa1538.
123. Tillett RL, Sevinsky JR, Hartley PD, Kerwin H, Crawford N, Gorzalski A, et al. Genomic evidence for reinfection with SARS-CoV-2: a case study. Lancet Infect Dis. 2020. doi: 10.1016/S1473-3099(20)30764-7.
124. Van Elslande J, Vermeersch P, Vandervoort K, Wawina-Bokalanga T, Vanmechelen B, Wollants E, et al. Symptomatic SARS-CoV-2 reinfection by a phylogenetically distinct strain. Clin Infect Dis. 2020. doi: 10.1093/cid/ciaa1330.
125. Common investigation protocol for investigating suspected SARS-CoV-2 reinfection. Atlanta: United States Centers for Disease Control and Prevention; 2020 (https://www.cdc.gov/coronavirus/2019-ncov/php/reinfection.html, accessed 1 November 2020).
126. Reinfection with SARS-CoV-2: considerations for public health response. Stockholm: European Centre for Disease Prevention and Control; 2020 (https://www.ecdc.europa.eu/sites/default/files/documents/Re-infection-and-viral-shedding-threat-assessment-brief.pdf, accessed 1 November 2020).
127. Emergencies preparedness, response: SARS-CoV-2 variants. Disease outbreak news. Geneva: World Health Organization; 31 December 2020 (https://www.who.int/csr/don/31-december-2020-sars-cov2-variants/en/, accessed 31 December 2020)
128. Baum A, Fulton BO, Wloga E, Copin R, Pascal KE, Russo V, et al. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. Science. 2020;369:1014–8. doi: 10.1126/science.abd0831.
129. Inzaule SC, Hamers RL, Paredes R, Yang C, Schuurman R, Rinke de Wit TF. The evolving landscape of HIV drug resistance diagnostics for expanding testing in resource-limited settings. AIDS Rev. 2017;19:219–30 (https://www.ncbi.nlm.nih.gov/pubmed/28182618, accessed 15 November 2020).
130. Sepulveda N, Phelan J, Diez-Benavente E, Campino S, Clark TG, Hopkins H, et al. Global analysis of *Plasmodium falciparum* histidine-rich protein-2 (pfhrp2) and pfhrp3 gene deletions using whole-genome sequencing data and meta-analysis. Infect Genet Evol. 2018;62:211–9. doi: 10.1016/j.meegid.2018.04.039.
131. Cremer J, Hofstraat SHI, van Heiningen F, Veldhuijzen IK, van Benthem BHB, Benschop KSM. Genetic variation of Hepatitis B surface antigen among acute and chronic Hepatitis B virus infections in the Netherlands. J Med Virol. 2018;90:1576–85. doi: 10.1002/jmv.25232.
132. Hollinger FB. Hepatitis B virus genetic diversity and its impact on diagnostic assays. J Viral Hepat. 2007;14 Suppl 1:11–5. doi: 10.1111/j.1365-2893.2007.00910.x.
133. Lam SD, Bordin N, Waman VP, Scholes HM, Ashford P, Sen N, et al. SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals. Sci Rep. 2020;10:16471. doi: 10.1038/s41598-020-71936-5.
134. Damas J, Hughes GM, Keough KC, Painter CA, Persky NS, Corbo M, et al. Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. Proc Natl Acad Sci U S A. 2020;117:22311–22. doi: 10.1073/pnas.2010146117.
135. Freuling CM, Breithaupt A, Müller T, Sehl J, Balkema-Buschmann A, Rissmann M, et al. Susceptibility of raccoon dogs for experimental SARS-CoV-2. bioRxiv. 2020. doi: 10.1101/2020.08.19.256800v1.
136. Schlottau K, Rissmann M, Graaf A, Schon J, Sehl J, Wylezich C, et al. SARS-CoV-2 in fruit bats, ferrets, pigs, and chickens: an experimental transmission study. Lancet Microbe. 2020;1:e218–e25. doi: 10.1016/S2666-5247(20)30089-6.
137. Shi J, Wen Z, Zhong G, Yang H, Wang C, Huang B, et al. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. Science. 2020;368:1016–20. doi: 10.1126/science.abb7015.
138. Kim YI, Kim SG, Kim SM, Kim EH, Park SJ, Yu KM, et al. Infection and rapid transmission of SARS-CoV-2 in ferrets. Cell Host Microbe. 2020;27:704–9 e2. doi: 10.1016/j.chom.2020.03.023.
139. Halfmann PJ, Hatta M, Chiba S, Maemura T, Fan S, Takeda M, et al. Transmission of SARS-CoV-2 in domestic cats. N Engl J Med. 2020;383:592–4. doi: 10.1056/NEJMc2013400.
140. Ruiz-Arrondo I, Portillo A, Palomar AM, Santibanez S, Santibanez P, Cervera C, et al. Detection of SARS-CoV-2 in pets living with COVID-19 owners diagnosed during the COVID-19 lockdown in Spain: a case of an asymptomatic cat with SARS-CoV-2 in Europe. Transbound Emerg Dis. 2020. doi: 10.1111/tbed.13803.
141. Richard M, Kok A, de Meulder D, Bestebroer TM, Lamers MM, Okba NMA, et al. SARS-CoV-2 is transmitted via contact and via the air between ferrets. Nat Commun. 2020;11:3496. doi: 10.1038/s41467-020-17367-2.
142. Sia SF, Yan LM, Chin AWH, Fung K, Choy KT, Wong AYL, et al. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. Nature. 2020. doi: 10.1038/s41586-020-2342-5.
143. Munster VJ, Feldmann F, Williamson BN, van Doremalen N, Pérez-Pérez L, Schulz J, et al. Respiratory disease and virus shedding in rhesus macaques inoculated with SARS-CoV-2. medRxiv. 2020. doi: 10.1101/2020.03.21.001628v1.
144. Zhao Y, Wang J, Kuang D, Xu J, Yang M, Ma C, et al. Susceptibility of tree shrew to SARS-CoV-2 infection. Sci Rep. 2020;10:16007. doi: 10.1038/s41598-020-72563-w.
145. Woolsey C, Borisevich V, Prasad AN, Agans KN, Deer DJ, Dobias NS, et al. Establishment of an African green monkey model for COVID-19. bioRxiv. 2020. doi: 10.1101/2020.05.17.100289.
146. Lu S, Zhao Y, Yu W, Yang Y, Gao J, Wang J, et al. Comparison of nonhuman primates identified the suitable model for COVID-19. Signal Transduct Target Ther. 2020;5:157. doi: 10.1038/s41392-020-00269-6.
147. Sit THC, Brackman CJ, Ip SM, Tam KWS, Law PYT, To EMW, et al. Infection of dogs with SARS-CoV-2. Nature. 2020. doi: 10.1038/s41586-020-2334-5.
148. Newman A, Smith D, Ghai RR, Wallace RM, Torchetti MK, Loiacono C, et al. First reported cases of SARS-CoV-2 infection in companion animals - New York, March–April 2020. MMWR Morb Mortal Wkly Rep. 2020;69:710–3. doi: 10.15585/mmwr.mm6923e3.

149. Oreshkova N, Molenaar RJ, Vreman S, Harders F, Oude Munnink BB, Hakze-van der Honing RW, et al. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. Euro Surveill. 2020;25. doi: 10.2807/1560-7917.ES.2020.25.23.2001005.
150. Cahan E. COVID-19 hits U.S. mink farms after ripping through Europe. Science Magazine. 2020 (https://www.sciencemag.org/news/2020/08/covid-19-hits-us-mink-farms-after-ripping-through-europe, accessed 13 November 2020).
151. Abdel-Moneim AS, Abdelwhab EM. Evidence for SARS-CoV-2 infection of animal hosts. Pathogens. 2020;9. doi: 10.3390/pathogens9070529.
152. Exposure of humans or animals to SARS-CoV-2 from wild, livestock, companion and aquatic animals. Rome: Food and Agriculture Organization of the United Nations; 2020 (http://www.fao.org/3/ca9959en/ca9959en.pdf, accessed 1 December 2020).
153. Guidelines to mitigate the impact of the COVID-19 pandemic on livestock production and animal health. Rome: Food and Agriculture Organization of the United Nations; 2020 (http://www.fao.org/3/ca9177en/CA9177EN.pdf, accessed 1 December 2020).
154. Guidance on working with farmed animals of species susceptible to infection with SARS-CoV-2. Paris: World Organisation for Animal Health (OIE); 2020 (https://www.oie.int/fileadmin/Home/MM/Draft_OIE_Guidance_farmed_animals_cleanMS05.11.pdf, accessed 8 December 2020).
155. El Zowalaty ME, Jarhult JD. From SARS to COVID-19: a previously unknown SARS-related coronavirus (SARS-CoV-2) of pandemic potential infecting humans: call for a One Health approach. One Health. 2020;9:100124. doi: 10.1016/j.onehlt.2020.100124.
156. Leroy EM, Ar Gouilh M, Brugere-Picoux J. The risk of SARS-CoV-2 transmission to pets and other wild and domestic animals strongly mandates a One-Health strategy to control the COVID-19 pandemic. One Health. 2020;10:100133. doi: 10.1016/j.onehlt.2020.100133.
157. Guidelines for working with free-ranging wild mammals in the era of the COVID-19 pandemic. Paris: World Organisation for Animal Health (OIE); 2020 (https://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/COV-19/A_WHSG_and_OIE_COVID-19_Guidelines.pdf, accessed 10 December 2020).

Annex I: Key questions to consider before initiating a sequencing programme

(1) What are the expected outputs of the sequencing programme?
(2) Which samples should be sequenced to achieve the expected outputs identified in step 1? Which metadata or additional data sources are critical?
(3) Who are the key stakeholders and what are their responsibilities? How can they be effectively engaged?
(4) How can samples and information be transferred rapidly and appropriately between stakeholders, as required?
(5) Is the project designed in accordance with local, national and international laws and ethical guidelines?
(6) Are adequate funding, equipment and human resources available to deliver all stages of specimen retrieval, wet-laboratory sequencing, bioinformatic, phylodynamic and other analyses, data sharing, and communication of timely results to appropriate stakeholders?
(7) How can goals be achieved without disrupting other areas of laboratory work, such as clinical diagnostics, and avoiding duplication of effort?
(8) How will the programme be evaluated for cost-effectiveness and impact?

Annex II: Checklist for setting up a SARS-CoV-2 sequencing programme

Aims

- ☐ Define the expected aims of the sequencing programme; what information will sequencing be likely to provide that is additional to or more cost-effective than existing approaches?

Stakeholder identification and engagement

- ☐ Identify key stakeholders.
- ☐ Discuss the programme aims with senior representatives of stakeholder groups and define the responsibilities of each group.
- ☐ Consider sharing educational materials about the potential and requirements of SARS-CoV-2 sequencing with stakeholders.
- ☐ Identify the links needed between key stakeholders to enable rapid movement of samples, requests for information and use of results.
- ☐ Ensure that clear, appropriate links between stakeholders are established.

Technical considerations

- ☐ Determine the level of genomic sampling required to achieve the desired goals, in discussion with senior members of case-identification and analytical teams.
- ☐ Identify the metadata required to achieve the desired goals, in discussion with senior members of case-identification and analytical teams.
- ☐ Choose appropriate sample and library preparation protocols.
- ☐ Choose appropriate bioinformatic protocols.
- ☐ Choose appropriate analytical protocols.

Logistical considerations

- ☐ Consider where sequencing and analysis will be conducted (e.g., an existing diagnostic laboratory or external commercial or academic laboratory).
- ☐ Identify appropriate sources of funding that will be adequate to support laboratory sequencing, data storage and data analysis.
- ☐ Ensure that sufficient reagents and computational resources are available and can be sustainably obtained as required.
- ☐ Ensure that there are sufficient and appropriate human resources to deliver the programme at every stage.
- ☐ Ensure that sample integrity can be maintained at all steps throughout the pipeline via cold-chain or other measures.
- ☐ Ensure adequate collection and storage of metadata and correct association with biological samples.
- ☐ Consider the possible additional pressure that sequencing will place on existing arms of the public health response, and seek ways to alleviate this.
- ☐ For large-scale sequencing programmes, identify how to streamline the sharing of data and samples between participating groups (e.g., the feasibility of using a single-sample identification and identical metadata formats).

Ensuring a safe and ethical environment

- ☐ Conduct appropriate ethical reviews for the generation, use and storage of sequence data and associated metadata.
- ☐ Conduct risk assessments of sequencing activities to ensure appropriate biosafety at all stages.
- ☐ Conduct risk assessments of sequencing activities to ensure appropriate biosecurity, if relevant under national and regional law.

- Consider the impact on human resources, including the reallocation of staff or hiring of additional staff to maintain the individual workload at reasonable levels.
- Ensure that staff can commute to work and be in the workplace safely and in accordance with national guidelines on preventing transmission during the COVID-19 outbreak.
- Define strategies for maintaining the sequencing programme if key staff members become ill or must self-isolate.

Data sharing

- Ensure that all stakeholders are in agreement as to which sequences and metadata will be shared publicly, via which platforms and when.
- Ensure that all stakeholders are in agreement as to whether any metadata are to be restricted to a limited number of local users and devise strategies for securely sharing those data.
- Ensure data sharing complies with national and international regulatory frameworks.

Evaluation

- Ensure regular opportunities for evaluating the sequencing programme, including successes and continuing challenges.
- Ensure that a monitoring and evaluation framework is implemented to assess performance of the sequencing programme both technically (quality, etc.) and in terms of the programme's success in meeting its objectives.

Annex III: Commonly used platforms for SARS-CoV-2 sequencing analysis and their characteristics

Instrument[a]	Advantages	Limitations	Instrument run-time	Sequencing throughput	Relative cost comparison
Sanger sequencing	Widely accessible Easy to use Cost-effective sequencing if few targets required	Very low throughput Amplicons (often no more than 1000 bp) must be individually amplified and sequenced Expensive for full genomes Inappropriate for metagenomics	Typically a few hours	100 kB-2 Mb per single run	Relatively low cost for a few targets
Illumina (e.g., iSeq, MiniSeq, MiSeq, NextSeq, HiSeq, NovaSeq)	Very high sequencing yields possible Very high accuracy iSeq is portable Methods for handling data are well established	With the exception of Illumina iSeq, expensive to purchase and maintain compared to some other platforms Maximum read length 2 x 300 bp	10–55 h, depending on the instrument	1.2–6000 Gb, depending on instrument	High maintenance and start-up costs Moderate running costs
Oxford Nanopore Technologies (Flongle, MinION, GridION, PromethION)	Portable, direct sequencing Real-time data Low start-up and maintenance costs Can stop sequencing as soon as sufficient data are achieved Very long read lengths achievable (exceeding the full length of the SARS-CoV-2 genome)	Challenges with homopolymers Error rate per read is ~5% (R9.4 flowcells) so use of appropriate pipelines is critical to obtain high-accuracy consensus sequences Currently unsuitable for determining intra-host variation unless replicate sequencing is used *(52)*	Reads available immediately Can be monitored and run for up to several days as required	Ranging from < 2 Gb for Flongle flow cell to 220 Gb for PromethION flow cell Up to 48 flow cells can be used on PromethION	No maintenance and relatively low start-up costs Moderate running costs
Ion Torrent	Fast turnaround once sequencing starts	Challenges with homopolymers Expensive to purchase Maximum typical read lengths around 400 bp	2 h–1 day, depending on chip and device	30 Mb–50 Gb depending on device and chips	Moderate costs

[a] This listing of the various instruments is to provide an overview of the most commonly used tools for SARS-CoV-2 genomic sequencing and does not imply WHO endorsement of these products.

[b] Different cost estimations can be found in *(8)*.

WHO continues to monitor the situation closely for any changes that may affect this interim guidance. Should any factors change, WHO will issue a further update. Otherwise, this interim guidance document will expire 2 years after the date of publication.

WHO reference number: WHO/2019-nCoV/genomic_sequencing/2021.1