

Genomic sequencing of SARS-CoV-2

A guide to implementation for maximum impact on public health

8 January 2021



World Health
Organization

Genomic sequencing of SARS-CoV-2

**A guide to implementation for maximum
impact on public health**

8 January 2021



**World Health
Organization**

Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health

ISBN 978-92-4-001844-0 (electronic version)

ISBN 978-92-4-001845-7 (print version)

© World Health Organization 2021

Some rights reserved. This work is available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo>).

Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited, as indicated below. In any use of this work, there should be no suggestion that WHO endorses any specific organization, products or services. The use of the WHO logo is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: “This translation was not created by the World Health Organization (WHO). WHO is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition”.

Any mediation relating to disputes arising under the licence shall be conducted in accordance with the mediation rules of the World Intellectual Property Organization (<http://www.wipo.int/amc/en/mediation/rules/>).

Suggested citation. Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health. Geneva: World Health Organization; 2021. Licence: [CC BY-NC-SA 3.0 IGO](#).

Cataloguing-in-Publication (CIP) data. CIP data are available at <http://apps.who.int/iris>.

Sales, rights and licensing. To purchase WHO publications, see <http://apps.who.int/bookorders>. To submit requests for commercial use and queries on rights and licensing, see <http://www.who.int/about/licensing>.

Third-party materials. If you wish to reuse material from this work that is attributed to a third party, such as tables, figures or images, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

General disclaimers. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by WHO in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by WHO to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall WHO be liable for damages arising from its use.

Contents

Preface	vi
Acknowledgements	vii
Abbreviations	ix
Executive summary	x
1. Introduction	1
2. Background	2
2.1 Growth in virus genomic sequencing	2
2.2 Growth in virus genomic applications.....	2
2.3 Phylogenetic and phylodynamic analyses	5
2.4 SARS-CoV-2 genomic and evolutionary features important for genomic applications.....	7
3. Practical considerations when implementing a virus genomic sequencing programme	8
3.1 Planning a sequencing programme	8
3.2 Ethical considerations	8
3.3 Identifying expected outputs and necessary data.....	9
3.4 Identifying and liaising with stakeholders	9
3.5 Project execution: acquisition of data, logistics and human resources	11
3.6 Project evaluation	11
4. Data-sharing	12
4.1 WHO recommendations on data-sharing.....	12
4.2 Sharing of appropriate metadata.....	12
4.3 Sharing of consensus sequences, partial consensus sequences and raw sequence data.....	12
4.4 Platforms for sharing	13
5. Applications of genomics to SARS-CoV-2	15
5.1 Understanding the emergence of SARS-CoV-2	15
5.1.1 Identifying the causative agent of COVID-19.....	15
5.1.2 Determining times of origin and early diversification	15
5.1.3 Identifying the zoonotic origin.....	16
5.2 Understanding the biology of SARS-CoV-2	17
5.2.1 Host receptor usage	17
5.2.2 SARS-CoV-2 evolution: identifying candidate genomic sites that may confer phenotypic changes	17
5.3 Improving diagnostics and therapeutics	18
5.3.1 Improving molecular diagnostics	18
5.3.2 Supporting the design and sensitivity monitoring of serological assays.....	19
5.3.3 Supporting vaccine design	19

5.3.4	Supporting design of antiviral therapy	20
5.3.5	Identifying antiviral resistance or vaccine escape mutations	20
5.4	Investigating virus transmission and spread	20
5.4.1	Supporting or rejecting evidence for transmission routes or clusters.....	20
5.4.2	Identifying and quantifying periods of transmission.....	21
5.4.3	Identifying importation events and local circulation.....	22
5.4.4	Evaluation of transmission drivers.....	25
5.4.5	Discerning involvement of other species	26
5.4.6	Discerning transmission chains between patients using intra-host viral diversity	26
5.5	Inferring epidemiological parameters.....	27
5.5.1	Reproduction number.....	27
5.5.2	Scale of outbreak over time and infection-to-case reporting ratio	28
6.	Practical guidance on technical aspects of genomic sequencing and analysis of SARS-CoV-2	30
6.1	Genome sampling strategies and study design	30
6.2	Appropriate metadata	32
6.3	Logistic considerations	36
6.3.1	Location.....	36
6.3.2	Biosafety and biosecurity	36
6.3.3	Ethical considerations	36
6.3.4	Human resources	37
6.4.	Choosing appropriate material for sequencing.....	39
6.4.1	Material for sequencing.....	39
6.4.2	Control samples.....	41
6.5	Enriching SARS-CoV-2 genetic material prior to library preparation	42
6.5.1	Metagenomic analyses of uncultured clinical samples	42
6.5.2	Metagenomic approaches following cell culture.....	43
6.5.3	Targeted capture-based approaches.....	43
6.5.4.	Targeted amplicon-based approaches	43
6.6	Selecting sequencing technology.....	44
6.7.	Bioinformatic protocols.....	47
6.7.1	Overview of typical bioinformatic steps	47
6.7.2	Dealing with multiplexed data	50
6.8	Analysis tools	51
6.8.1	Subsampling data prior to analysis.....	51
6.8.2	Sequence alignments	51
6.8.3	Quality control.....	52

6.8.4 Removing recombinant sequences	53
6.8.5 Phylogenetic tools	54
6.8.6 Visualization.....	55
6.8.7 Lineage classification.....	56
6.8.8 Phylogenetic rooting	56
7. Conclusions and future needs	57
References.....	59
Annex 1. Examples of sequencing studies for molecular epidemiology	73
Annex 2. Checklist for setting up a sequencing programme.....	78

Preface

The year 2020 was a turning point in history and in global health. The COVID-19 pandemic has highlighted the potential for deadly epidemic-prone diseases to overwhelm our globalized world. We have learned a hard lesson about the intrinsic vulnerability of our societies to a single pathogen.

Although COVID-19 has brought untold tragedy, it has also shown how science can respond when challenged by a massive global emergency. In short, the pandemic has opened great scientific opportunities and capitalized on them. A technological revolution, building over the past decade, provided several new capacities for a pandemic response. Development of vaccines at lightning speed is one of them. Genomic sequencing is another.

Sequencing enabled the world to rapidly identify SARS-CoV-2; and knowing the genome sequence allowed rapid development of diagnostic tests and other tools for the response. Continued genome sequencing supports the monitoring of the disease's spread and activity and evolution of the virus.

The COVID-19 pandemic is still ongoing, and new viral variants are emerging. The global response will have to continue for the foreseeable future. The progress made since the start of the pandemic with the use of genome sequencing can be consolidated and further expanded to new settings and new uses.

As more countries move to implement sequencing programmes, there will be further opportunities to better understand the world of emerging pathogens and their interactions with humans and animals in a variety of climates, ecosystems, cultures, lifestyles and biomes. This knowledge will shape a new vision of the world and open new paradigms in epidemic and pandemic prevention and control.

Increased urbanization and human mobility are providing the conditions for future epidemics and pandemics. The accelerated integration of genome sequencing into the practices of the global health community is a must if we want to be better prepared for the future threats. We hope this guidance will help pave the way for that preparedness.

预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=5_24148

