Whole genome sequencing as a tool to strengthen foodborne disease surveillance and response

Module 1. Introductory module

World Health Organization

Whole genome sequencing as a tool to strengthen foodborne disease surveillance and response

Module 1. Introductory module

World Health Organization

Whole genome sequencing as a tool to strengthen foodborne disease surveillance and response. Module 1. Introductory module

ISBN 978-92-4-002122-8 (electronic version)
ISBN 978-92-4-002123-5 (print version)

Suggested citation. Whole genome sequencing as a tool to strengthen foodborne disease surveillance and response. Module 1. Introductory module. Geneva: World Health Organization; 2023. 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.

Design and layout. Laura Ruiz & Velvet Pineda (original design and layout); Blossom Srl (design editing)

Printed in Switzerland

Contents

List of tables, figures, boxes

List of tables

List of figures

List of boxes

Contributors

Editorial guidance

Editorial Committee

Amy Cawthorne, Kirsty Hope, Kaouri Koume, Charlee Law, Megge Miller, Enrique Pérez Gutiérrez, Eric Stevens

Executive Editor

Amy Cawthorne (2017–2018), Jorge Matheu (2018–2021), Yuki Minato (2022), Charlee Law (2023)

Other members of the WHO secretariat

Jean Marc Gabastou, Marcelo Fabian Galas, Dhamari Naidoo, Christopher Oxenford

Authors and reviewers

Module 1. Introductory module

Authors

Kirsty Hope, Megge Miller, Eric Stevens

Peer reviewers

Marc W. Allard, Timothy Dallman, Xiangyu Deng, Rita Finley, Rene S. Hendriksen, Daneeta Henessey, William Hsiao, Kari Irvin, Claire Jenkins, Hilde Kruse, Ernesto Liebana, Patrick Munk, Stefano Pongolini, Valentina Rizzi, Nick Thomson, Moon Tay Yue Feng

Module 2. Outbreak module

Authors

Kirsty Hope, Megge Miller, Eric Stevens

Other authors

Claire Jenkins, Celine A. Nadon, Eric Device Ng'eno, Stefano Pongolini, Nick Thomson, Caterina Guzman Verri

Peer reviewers

Marc W. Allard, Jennifer Beal, Timothy Dallman, Peter Gerner-Smidt, Rene S. Hendriksen, William Hsiao, Kari Irvin, Hilde Kruse, Pimlapas Leekitcharoenphon, Valentina Rizzi

Module 3. Surveillance module

Authors

Kirsty Hope, Megge Miller, Eric Stevens

Other authors

Philippe Belanger, Peter Gerner-Smidt, Caterina Guzman Verri, Claire Jenkins, Celine A. Nadon, Eric Device Ng'eno, Stefano Pongolini, Nick Thomson

Peer reviewers

Marc W. Allard, Timothy Dallman, Ria Finley, Daneeta Hennessey, William Hsiao, Hilde Kruse, Patrick Munk, Valentina Rizzi

Acknowledgements

The World Health Organization would like to express our sincere gratitude to the many authors, peer reviewers and advisers whose support and expertise made these manuals possible. Special thanks go to the principal authors, Kirsty Hope, Megge Miller and Eric Stevens for their dedication and commitment to finalize the guidance. This guidance was developed as an outcome of a workshop entitled, "WHO Meeting on the application of whole genome sequencing as a tool to strengthen foodborne disease surveillance and response in developing countries" held in Washington DC, United States of America (USA), in January 2017, attended by experts and relevant stakeholders from various countries with a range of experience in public health, bioinformatics and microbiology. The World Health Organization would like to thank the contributions made by the meeting participants (see Annex 4 for a full list). The World Health Organization also wishes to thank the United States Food and Drug Administration for their generous financial support for the development and publication of this report.

Abbreviations and acronyms

AMR	antimicrobial resistance
cgMLST	core genome multi-locus sequence type
FAO	Food and Agriculture Organization of the United Nations
ISO	International Organization for Standardization
MLST	multi-locus sequence type
SNP	single nucleotide polymorphism
wgMLST	whole genome multi-locus sequence type
WGS	whole genome sequencing
WHO	World Health Organization

Executive summary

Whole genome sequencing (WGS) is a laboratory technique which is revolutionizing public health and more specifically the control and prevention of foodborne diseases. WGS is useful for understanding foodborne diseases through enhancing routine surveillance, outbreak detection, outbreak response and for source identification using a One Health approach. This document provides guidance on the capacities that need to be in place before WGS can be useful for foodborne disease surveillance and response, options for implementing WGS and how to integrate WGS within existing systems.

There are three modules in this guidance which cover the minimum requirements for using WGS, some options for introducing WGS for foodborne disease surveillance and response (introductory module (module 1)), using WGS for outbreak investigations (outbreak investigation module (module 2)) and for routine surveillance (routine surveillance module (module 3)).

Introductory module

The introductory module explains the minimum capacity requirements in the foodborne diseases surveillance and response system prior to considering the implementation of WGS. These requirements cover epidemiological capacity for detecting outbreaks and conducting outbreak investigations, laboratory capacities to test clinical specimens, and capacity within the food safety system in order to respond to events and undertake control measures.

For epidemiological capacity, there needs to be a functional event-based surveillance system, rapid risk assessment and the ability to conduct epidemiological investigations during outbreaks.

For a country's laboratory capacity, it is important that specimens from the field can be tested and foodborne pathogens identified. There needs to be a designated laboratory with the experience and resources to routinely culture and identify priority foodborne pathogens such as *Salmonella, Listeria,* and Shiga-toxin producing *Escherichia coli.* A laboratory quality management system also needs to be in place before a country considers using WGS.

Capabilities in the food safety sector depend on whether appropriate food and/or environmental samples can be collected during foodborne disease outbreaks, samples can be tested at a laboratory, there are food safety laws and regulations to support control measures, and food safety personnel can control the distribution of implicated food items. There is an assessment tool to audit existing capacities which can then be used to determine which option is chosen for implementation of WGS. There are two main options for implementing WGS: outbreak investigations and for enhanced surveillance of foodborne pathogens.

Outbreak and surveillance modules

The outbreak and surveillance modules examine how WGS can be used in each context. Often, the initial use of WGS in foodborne diseases is in the context of an outbreak investigation. It is important to define the goals and objectives of WGS within the current surveillance or outbreak response system. Once these have been articulated, it is then possible to review the options for sequencing which include outsourcing all or parts of WGS, or using a public health laboratory. The availability of human and financial resources is key to using WGS for surveillance and outbreak response.

Both the outbreak and surveillance modules contain step-by-step guidance on implementing WGS. The approach is the same, even though the context is different. The first step is to form a working group of the relevant stakeholders. The second step involves describing what already exists in the surveillance or outbreak response system. The third step contains advice on building a business case for WGS, and how to communicate effectively to engage decision-makers to see the value in WGS.

Once there is a commitment to using WGS, there is guidance on how to run and evaluate a pilot study. This will be useful in determining how WGS will work for future surveillance and outbreak investigations of foodborne diseases.

1. Background

Globally, there is growing recognition of foodborne diseases as a public health priority. The World Health Organization (WHO) estimates of the global burden of foodborne diseases indicate that infectious agents causing diarrhoeal diseases accounted for most foodborne illnesses, and estimate an annual incidence of 600 million cases, with 240 000 deaths *(1)*. From a public health perspective, foodborne diseases are largely preventable, and can be controlled through effective food safety systems that evaluate hazards along the food chain, from production to consumption. An integrated food chain surveillance system can detect and monitor foodborne bacteria, including antimicrobial resistant (AMR) bacteria, throughout the food chain.

Whole genome sequencing (WGS) has the potential to change how we detect and monitor microbial hazards in the food chain, as well as how we assess, investigate and manage food safety risks. It is anticipated that this new technology will help reduce the burden of foodborne diseases, given its advantages over previous low-resolution typing and detection methods.

1.1 What is WGS?

WGS is a revolutionary technology that can be used to determine the complete nucleic acid sequence of a given organism's genome , the order and number of nucleotides, adenine (A), guanine (G), cytosine (C) and thymine (T), which make up the microorganism's genetic code.

The technology also allows for detailed molecular subtyping of microbial pathogens. Public health professionals can use WGS data to quantify differences (and similarities) more accurately between individual strains, when compared with other molecular tools available to date.

Steps involved in foodborne pathogen WGS

- For foodborne bacteria (*Salmonella*, *Listeria*, and other genera), WGS begins with the extraction of a microbe's large, circular genome (deoxyribonucleic acid or DNA), followed by the application of an appropriate library preparation. For example, short-read sequencing requires breaking the genome into many smaller pieces. Note: sequencing viral genomes may involve a different process.
- Each piece of DNA is read by a sequencing machine, and the outputs are ordered as a string of A, T, C and G nucleotides, representing DNA components.
- A quality check on sequencing reads is performed and sequencing reads of poor quality are removed.
- Sequencing reads of good quality are then run through bioinformatic analyses to organize the information associated with these fragments. This process, called assembly, is used to align and merge fragments from the shorter DNA sequences into longer segments, in order to reconstruct the original genomic sequence.
- The genomic sequence is then analysed to see if it was re-assembled correctly.
- If the sequence is of good quality, it can be used to provide a molecular subtype and information about the presence of AMR and virulence genes.
- The genome sequence can also be compared with a standard reference genome for that microbe. Genetic differences among isolates can be displayed as phylogenetic trees, depicting genetic changes among strains so that the more similar strains are grouped (clustered) together.
- Comparisons between two microbial genomes can be performed by examining:
  - single nucleotide polymorphisms (SNPs) that is, differences in the nucleotides making up the DNA of one genome and another; or
  - alleles at the gene-by-gene level, using whole genome multi-locus sequence type (wgMLST).

A detailed description of WGS and the steps involved in WGS are provided in Annex 1, which also contains a discussion of data comparability.

1.2 Why is WGS of foodborne pathogens important?

WGS technology has revolutionized how public health and regulatory officials respond to and prevent and control foodborne contamination events, because WGS:

- has greater sensitivity and specificity in foodborne pathogen subtyping, vis-a-vis traditional subtyping methods;
- can be used to monitor preventative control measures aimed at improving good agricultural practices, it allows the food industry an effective tool for supply chain management, and it helps to distinguish between transient and resident pathogens within food processing facilities;
- can be used by the food industry and the public health sector to evaluate the impact of food safety interventions introduced to reduce the presence of foodborne pathogens in the food chain and their impact on human health;
- can be used for food traceability in international trade to better understand the source of contamination;
- provides the entire genomic content of a pathogen, yielding valuable information about AMR or virulence factors that may be present;
- allows public health officials the collection of information about a pathogen in a single microbiological workflow, as opposed to previous methods that required several individual microbiological tests; and
- makes strain comparison on an international scale possible, as sequencing data can be shared online, which is faster and easier than shipping actual pathogens for traditional typing.

Clinical microbiology

Using WGS in clinical microbiology will be important for those foodborne diseases that require antimicrobial treatment. WGS implementation will allow the monitoring of AMR transmission trends of clinically significant bacteria, as it will provide real-time data on emerging resistance at the local level, and inform antimicrobial treatment policy in clinical settings.

Comparability over time

Once the isolate's sequence has been obtained, it can be stored for historical and future comparisons. Over time, WGS bioinformatic analyses and outputs may change, but the actual sequence will remain available for future use, thus offering the opportunity to understand foodborne disease epidemiology in greater detail. For example, a study in Denmark *(2)* re-analysed historical WGS data for an AMR gene that was discovered after sequencing data had been generated, which led to insights into the emergence of said gene in the country.

1.3 How can WGS be useful for foodborne diseases?

There are four key areas where WGS can be useful for understanding, controlling and preventing foodborne diseases.

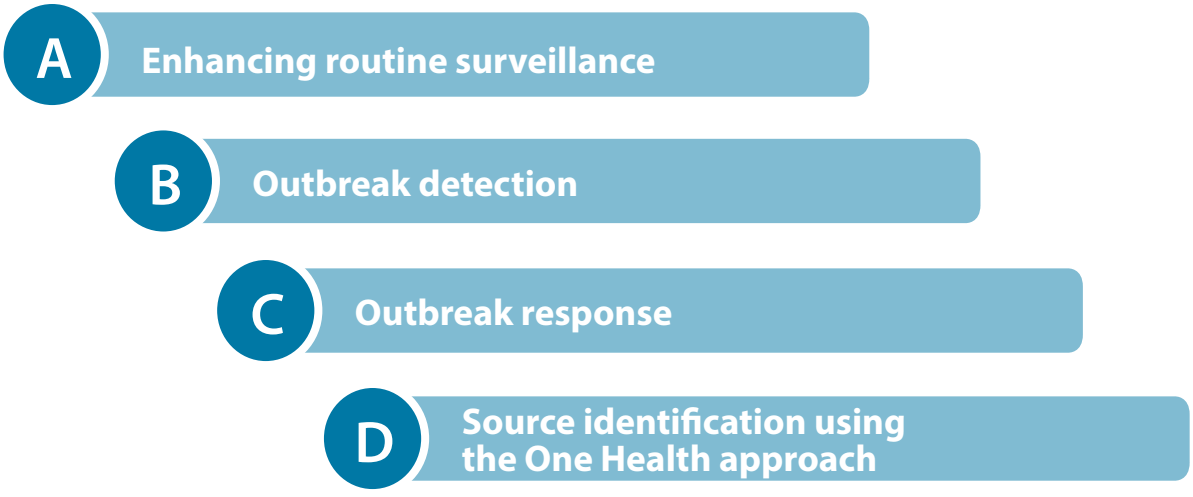
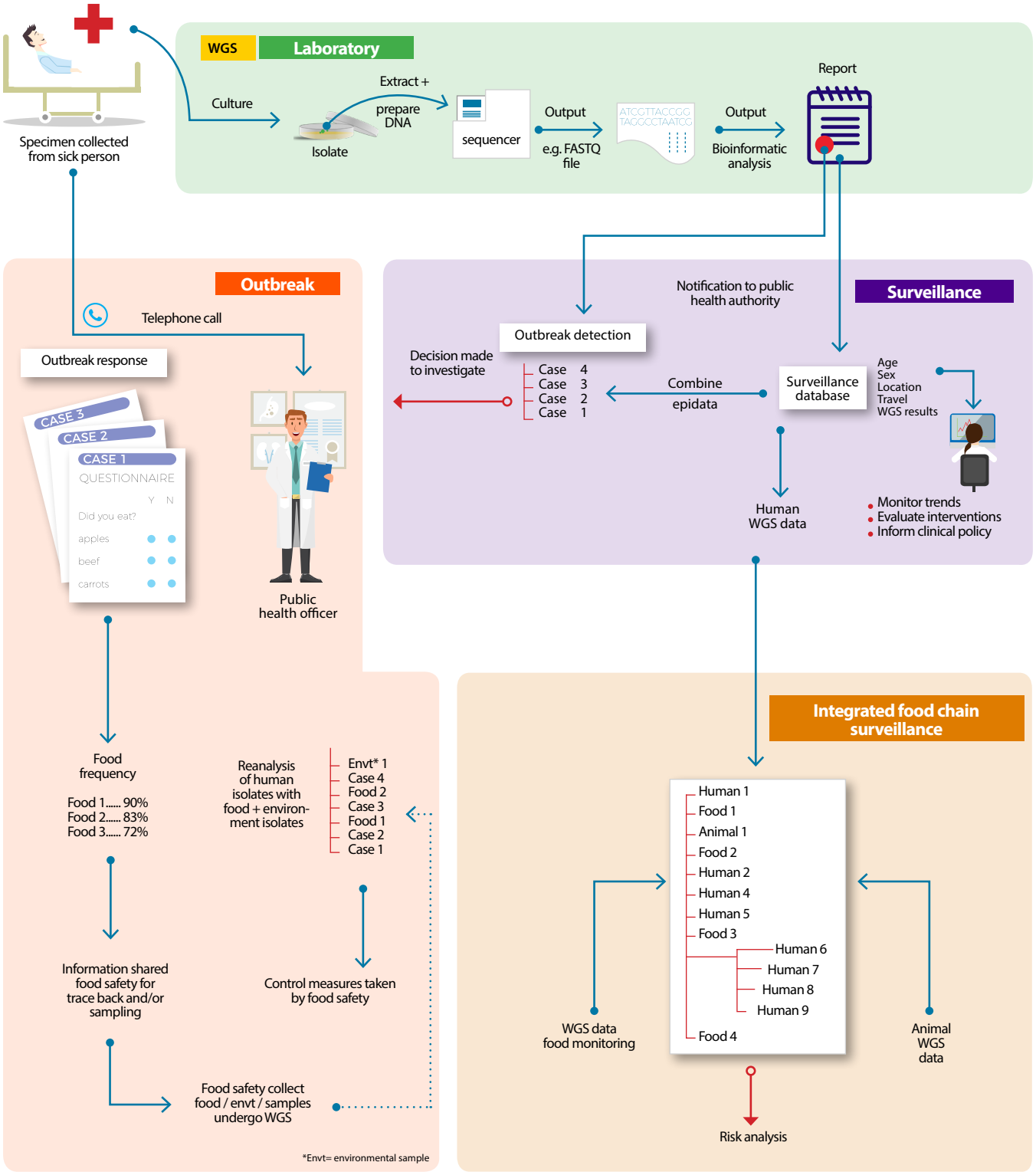


Fig. 1 illustrates how WGS results can be applied in each of the four key areas.

Fig. 1

How the results from WGS can be used for surveillance and outbreaks of foodborne diseases



A Enhancing routine surveillance

With WGS, foodborne pathogen genomes can be compared at the nucleotide level, providing greater insight into whether the isolates in question are genetically closely related, when compared with traditional subtyping methods. WGS can provide extensive information about an isolate, which can be added to the surveillance system. Such data include subtyping, the presence or absence of AMR genes, as well as virulence factors. With traditional methods, an isolate would need to undergo multiple laboratory tests, while WGS can provide most of the information in a single test. There are benefits in adding data to foodborne diseases routine surveillance systems, such as:

- garnering a better understanding of the epidemiology of the pathogens under surveillance;
- informing antimicrobial use policy in clinical settings for improved patient care; and
- informing policy-making to address antimicrobial use in animals used for human consumption *(3)*, if data from across the food chain are added to surveillance systems.

B Outbreak detection

Using WGS for surveillance purposes makes it possible to identify people who have been infected with closely related pathogens, which may point to a common source of infection. With greater discrimination of genetic relatedness, outbreaks can be detected earlier, when a small number of cases caused by pathogens sharing the same or nearly identical sequences can be observed. Thus, epidemiological information can be analysed, and hypotheses drawn on the source of exposure sooner than with traditional typing methods. The experience of some countries using WGS for surveillance purposes is that more foodborne outbreaks are detected, with fewer cases reported in each outbreak *(4)*.

C Outbreak response

Once an outbreak has been detected and assessed and an investigation is launched, WGS is useful in identifying the strain(s) involved. The outbreak strain is then incorporated into the outbreak's case definition. WGS offers greater specificity and sensitivity in defining cases, thus:

- minimizing misclassification bias in analytical epidemiological studies (e.g. case control or cohort studies), as case definitions are more precise; and
- strengthening case-related food history data to better inform traceback investigations *(5)*.

If a hypothesis about a potential food source is developed, samples of suspected food items or from the environment where the food item was processed or packaged can be collected. Once the isolate is available, it can undergo WGS and be analysed vis-à-vis human isolates.

WGS can also make case ascertainment more robust during outbreak investigations *(6)*. Sequencing data can show whether isolates are genetically close when epidemiological information may be lacking, or where there are issues of poor recall of exposure among patients. WGS can also provide clues for public health staff to re-interview cases or determine exposures in other ways.

D Source identification using the One Health approach

WHO defines One Health as an approach to designing and implementing programmes, policies, legislation and research in which multiple sectors communicate and work together to achieve better public health outcomes *(7)*. In the context of foodborne diseases, the One Health approach involves integrating data collected from points across the food chain, including the animal health, food safety and human health sectors. For example, when WGS of bacteria isolated from human clinical specimens are compared with isolates from food and environmental samples, a better indication can be obtained as to the likely source of the illness when compared with traditional typing methods. When combined with epidemiological investigation data, WGS can yield strong evidence of the link between a specific food or environmental source to foodborne diseases in humans. The information gathered from WGS surveillance and from outbreak investigations can be a driver for the allocation of resources to strengthen food control systems *(5)*.

WGS makes integrated food chain surveillance potentially more efficient and effective, as it allows for a more sensitive comparison of isolates from samples collected at different points in the food chain. When coupled with food monitoring in food processing environments and finished food products, WGS helps with early identification of a contaminated food source by determining which part of the food chain to target for corrective action, and consequently, with prevention of further human illness. This can include improved attribution of human infections to various reservoirs, leading to more targeted and effective allocation of resources for pathogen control in animals and across the food chain to prevent illness in humans.

1.4 Uncertainties of WGS for surveillance and response purposes

WGS will change the type of information reported to surveillance systems, as well as how data are used to meet the surveillance systems objectives. As the technology is introduced globally, WGS is becoming the typing method of choice for foodborne pathogens. However, there are still some uncertainties around the use of WGS for surveillance and outbreak response. This manual will focus on generic principles of sequencing and surveillance, and will highlight areas of uncertainty to ensure countries are fully aware of the limitations of using WGS at present. Currently, such uncertainty refers to the following facts.

- There is no internationally agreed standardized approach for the analysis of WGS for microbial subtyping, while traditional subtyping methods follow standardized approaches to foodborne pathogens typing. For example, there is the Kauffman-White-Le Minor scheme used for subtyping *Salmonella (8).* There are also standardized approaches for pulsed field gel electrophoresis (PFGE) testing for a range of foodborne pathogens. At the time of publication, there was no standardized approach for WGS typing, although PulseNet International recently agreed to use wgMLST *(9)*.

- Staff using WGS outputs have not traditionally been trained in the analysis and interpretation of genetic data. Epidemiologists and surveillance officers in public health institutions, such as the Ministry of Health, will need to acquire basic skills in molecular epidemiology. They will also need to have regular, ongoing communication with the laboratory to interpret results and make decisions about when public health measures are required.

- To harness WGS's full potential, it will be essential to share sequencing data across sectors and countries. There are issues around patient confidentiality, sharing commercial-in-confidence information, and ramifications beyond the health sector (e.g. food and food animal trade). These are important issues which are currently being addressed by ongoing work to develop data-sharing policies.

2. Introduction

2.1 Purpose of this manual

The purpose of this manual is to provide guidance on:

- capacities that need to be in place before WGS can be useful for foodborne diseases surveillance and response;
- options for implementing WGS; and
- how to implement WGS within existing surveillance and response systems.

2.2 Scope of the manual

This guidance on WGS is part of Stage 2 *(10)* and Stage 3 *(11)* of the publication, "Strengthening surveillance of and response to foodborne diseases (Fig. 2) *(12)*. If a country wishes to initiate laboratory-based surveillance of foodborne pathogens to strengthen their surveillance and response system, it would be ideal to begin with that manual.

Fig. 2

Strengthening surveillance of and response to foodborne diseases: stages 2 and 3

Source: World Health Organization *(12)*

The focus of the present manual is on the role of WGS in strengthening foodborne disease surveillance and responses. However, much of the information in this document will be relevant to sequencing for other diseases of public health importance. In addition, the current manual emphasizes competency development for using WGS in foodborne diseases surveillance and response within the human health sector. Developments within the human health sector should encourage a broader One Health approach addressing contamination along the food chain and involve colleagues in food safety, animal health and environmental health.

The use of WGS in public health surveillance is new and still evolving. Decisions are still being made about new internationally standardized subtyping schemes for foodborne pathogens based on sequencing outputs. This manual will provide general guidance based on the principles of implementing sequencing, rather than make recommendations about equipment, bioinformatic pipelines, subtyping schemes and analyses approaches.

The following topics are beyond the scope of this manual.

- Specific comprehensive advice for strengthening foodborne diseases surveillance response systems. There is an existing WHO publication on the topic of strengthening surveillance of and response to foodborne diseases *(13)*.
- Specific advice on using WGS in the One Health approach across all sectors to establish integrated food chain surveillance.
- Specific advice on the surveillance of AMR. There is an existing WHO publication on this topic, "Integrated surveillance of antimicrobial resistance in foodborne bacteria" *(14)*.
- Using WGS in food safety management. Information on this topic is available from the Food and Agriculture Organization (FAO), including:
  - the technical meeting: "One Health approach" *(15)*; and
  - the technical background paper: "Applications of whole genome sequencing (WGS) in food safety management" *(16)*.
- Using WGS in the animal health sector.
- Strengthening public health laboratory capacities to perform WGS. Some of the technical requirements for WGS are discussed in the publication, "Whole genome sequencing for foodborne disease surveillance: landscape paper" *(17)*.
- Strengthening laboratories to conduct WGS for further characterization of food and/or environmental samples, which is usually done by the food safety system.

2.3 Target audience

This manual is mainly for public health professionals, such as epidemiologists and laboratory staff who will utilize WGS as part of their foodborne diseases surveillance and response system. Usually, these professionals work in the Ministry of Health or other institutions of the human health sector. The manual should also be useful in the food safety sector, the animal health sector dealing with animal food sources for human consumption, and development partners, donors and international organizations.

2.4 Guiding principles

These are some guiding principles to keep in mind when considering the use of this manual and deciding if sequencing is appropriate for a given country.

- Every country has a national surveillance and response system for various diseases, including foodborne diseases. This manual focuses on building on existing systems that are part of the core capacity requirements of the International Health Regulation (IHR) (2005) *(18)*.
- Surveillance is needed to detect and respond effectively to acute public health events in a timely manner, in order to minimize the adverse impact on public health and the economy.
- Surveillance and response systems in countries are at different levels of development and complexity, and have different requirements and priorities for future development and capacity-building.
- Developments outside the health sector (e.g. food safety standards, international trade agreements and importation requirements) can also motivate the strengthening of foodborne disease surveillance and response systems.
- It is vital that senior policy- and decision-makers are committed to using WGS to improve foodborne disease surveillance and response systems.
- Sufficient resources, both human and financial, are required to support the introduction and ongoing use of WGS within the surveillance and response system.
- Countries should draw on existing international networks resources to strengthen national surveillance and response systems. Box 1 shows examples of networks and institutions that can provide technical assistance to countries.

Some terms used in this manual may have different meanings in different settings. Annex 2 provides a glossary for the terms in this document, as well as some key technical definitions related to WGS.

Also, throughout this manual, the terms isolate and DNA apply exclusively to foodborne bacteria. However, the general principles of using WGS for surveillance and response are applicable to foodborne viruses and parasites as well. For example, there is always a wet lab step, a dry lab step and the process of incorporating the outputs from WGS into the surveillance and response system.

BOX 1

International networks and institutions providing technical assistance

Examples of networks

COMPARE
http://www.compare-europe.eu/about

GenomeTrakr Network
https://www.fda.gov/food/whole-genome-sequencing-wgs-program/genometrakr-network

Global Microbial Identifier (GMI)
http://www.globalmicrobialidentifier.org/about-gmi

INFOSAN
https://www.who.int/groups/international-food-safety-authorities-network-infosan

PulseNet International
https://pulsenetinternational.org/

2.5 How to use this manual

There are three modules in this manual including: introductory module, outbreak module *(19)* and surveillance module *(20)*.

1 Introductory module

The present module introduces WGS and its relevance to foodborne diseases, defines the minimum capacities needed before a country can implement WGS for outbreak investigations and routine surveillance, and includes a section to assist countries in selecting an option for implementing WGS within their existing surveillance and response system.

2 Outbreak module

This module discusses how WGS can be used to support foodborne disease outbreak investigations. It is meant for countries in the initial stages of laboratory-based surveillance for selected foodborne pathogens. The module describes how WGS can be used in the investigation of outbreaks detected by existing surveillance systems.

3 Surveillance module

This module is about using WGS in routine surveillance of foodborne diseases. It is meant for countries experienced in laboratory-based surveillance of foodborne pathogens. WGS can be implemented where subtyping foodborne pathogens or replacing traditional typing methods is being considered. Routine surveillance includes outbreak detection, monitoring trends over time, and using WGS for AMR and virulence factor monitoring.

Readers are advised to review this introductory module. It is not necessary to read all three modules. Those wishing to use WGS for outbreak investigation will only need to refer to the outbreak module. For those wishing to use WGS for surveillance, the surveillance module will contain the relevant guidance. In the latter case, it will not be necessary to refer to the outbreak module, as any pathogens from outbreak cases detected will automatically be sequenced as part of routine surveillance.

2.6 Steps involved in using this manual

There are three steps involved in using this manual.

Step 1

Determine whether WGS will be appropriate in your country

There are some circumstances where the implementation of WGS may not be appropriate. This section describes minimum requirements of the foodborne diseases surveillance and response system that should be in place prior to considering the establishment of WGS. There is a self-assessment tool that may be used to decide whether WGS will be appropriate.

Step 2

Choose an option for implementation

There are two options for implementation:

- using WGS to support foodborne disease outbreak investigations
- using WGS to enhance routine surveillance of foodborne diseases.

Both options are analysed based on the minimum foodborne diseases surveillance and response system capacities required. Annex 3 is provided to help determine the most appropriate option for each country, based on the capacities that are already in place and available resources. The tool will help decide which module to use in planning and implementation of WGS for foodborne diseases.

Step 3

Manage implementation

Each module shows the steps needed for the implementation of WGS. Steps can be taken in any order and in parallel. However, before implementing WGS, it is important that its outbreaks or surveillance objectives are articulated, as well as any thoughts about how the system will work. The key steps for implementation are as follows:

- a description of how the system will operate, including its technical requirements
- development of a business case for senior management approval and funding
- a pilot study
- full implementation.

3. Minimum requirements for WGS

Certain capacities need to be in place in the foodborne diseases surveillance and response system prior to considering the implementation of WGS. Usually, the starting point for a country wishing to use WGS will be during foodborne disease outbreak investigations. Therefore, minimum requirements are focused on ensuring that capacities are in place to detect foodborne disease outbreaks and to respond with an appropriate investigation to try to determine its source, to control the outbreak and to prevent similar events in the future. Once the source is identified, the food safety system should be able to implement the necessary control measures.

Those minimum requirements focus mainly on:

- **epidemiological capacities for outbreak detection and response, including appropriate outbreak investigation capacity;**
- **laboratory capacities to test clinical specimens to identify the causal agent; and**
- **the food safety system to respond to events and undertake control measures, including testing food and environmental samples to determine potential source of the agent.**

Surveillance systems must be sensitive enough to detect foodborne disease outbreaks, either through indicator-based surveillance (e.g. notifiable diseases or syndromes system) or event-based surveillance. Once an outbreak has been detected and verified, an assessment should be made of its potential to spread and add new cases, and action taken to address it. Outbreak investigations should be thorough, and use standard epidemiological techniques to identify the source of the outbreak. Currently, for WGS to be useful, clinical specimens will need to be collected as part of the investigation, and the etiological agent isolated at a public health laboratory. In the future, sequencing may occur directly on clinical specimens at the point of care, however, such technology is not yet in place to make this option affordable.

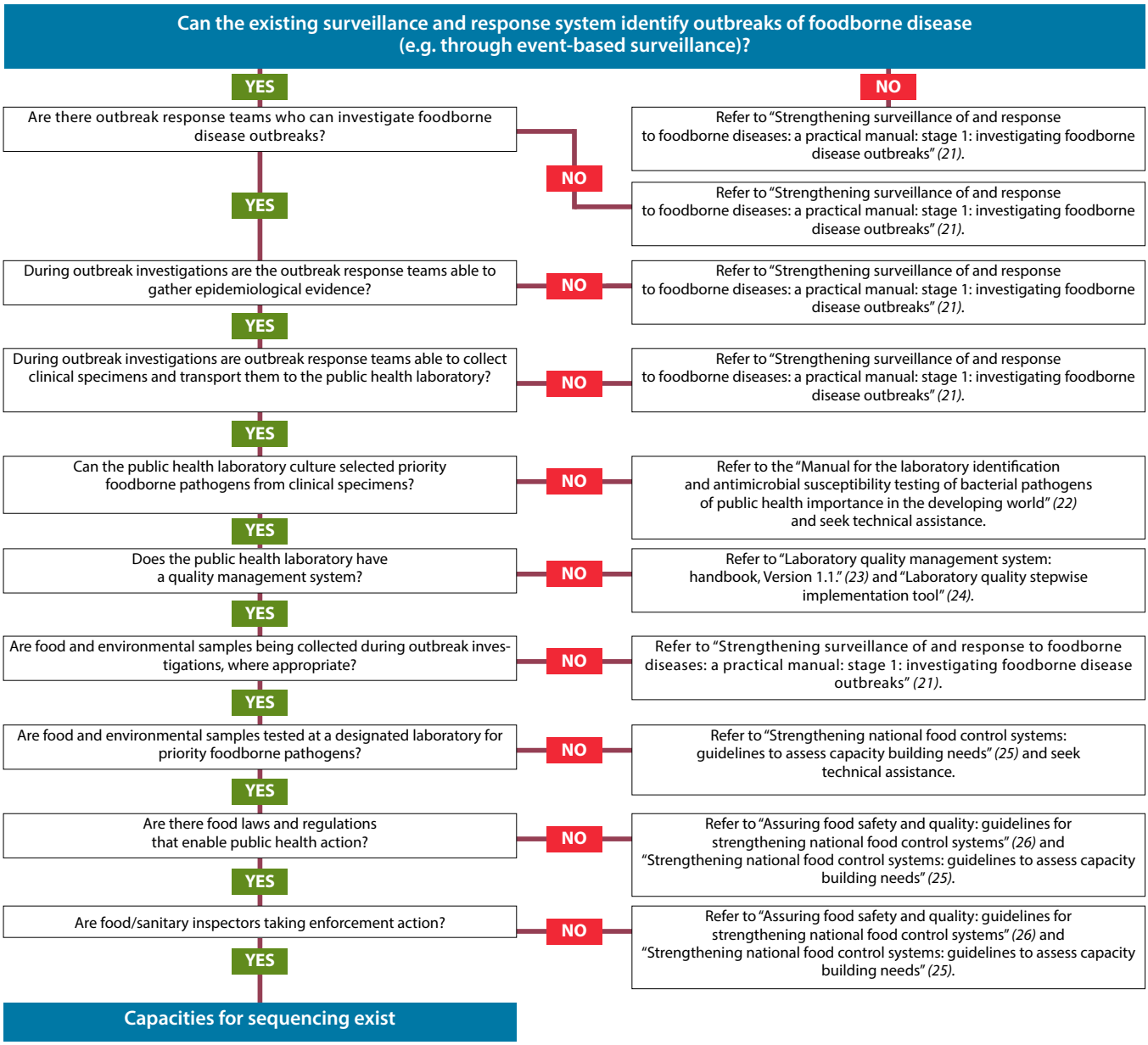
Some countries may have a laboratory-based surveillance system with the capacity to use WGS. In such instances, WGS can provide evidence of the close genetical relation of the microbial pathogens involved in human infections, data which must be used in conjunction with epidemiological information. Interviewing cases might be useful in ascertaining the nature of the link between agents. For example, there may have been a common food item consumed or an event attended by several cases.

When a food product or a food venue has been identified as the likely source of a foodborne disease outbreak, the food safety system should be able to intervene to control further distribution of the implicated food item, backed by adequate legislation and/or regulations.

Minimum capacities for WGS implementation are discussed below. Links to relevant guidance documents for countries wishing to explore these capacities in further detail are also provided. Annex 3 contains a tool for countries to use in assessing whether the capacities necessary to use WGS successfully and efficiently are available. Fig. 3 illustrates the broad capacities required prior to considering WGS implementation. When those capacities are not available, Fig. 3 provides links to relevant advice regarding building capacity in that area.

Fig. 3

Decision tree regarding the implementation of WGS



3.1 Capacities for epidemiological investigations

The existence of a foodborne disease surveillance and response system is essential to the appropriate and cost-effective implementation of WGS. The system should be able to detect and investigate outbreaks.

At a minimum, there needs to be a functional event-based surveillance system that can detect foodborne events such as outbreaks (Box 2). It is also important that, upon receiving a report of an event, it can be assessed to determine its risk and decisions can be made about whether an outbreak investigation is required. If so, a standard epidemiological investigation must take place.

BOX 2

Minimum requirements for applying WGS to the detection of and response to foodborne disease outbreaks

A functional event-based surveillance system should have:

- national focal points to receive reports about events;
- an event report form that systematically captures information about events;
- an event database to store information on reported events; and
- health care workers and food/sanitary inspectors trained to recognize and report events.

Functional rapid risk assessment:

- at a minimum, has a national-level team that can rapidly assess events; and
- ensures events are assessed within 24 hours of the first report.

The ability to conduct appropriate epidemiological investigation during outbreaks requires:

- trained personnel assigned to outbreak response teams;
- outbreak response teams who can develop and apply a case definition, interview people who meet the case definition using a standardized questionnaire, describe cases using a line list and conduct descriptive analyses by time, place and person;
- at least one epidemiologist in the country who can conduct analytical epidemiological studies when required; and
- outbreak investigations summarized in reports.

If a country does not have these capacities, the manual "Strengthening surveillance of and response to foodborne diseases" *(13)* provide step-by-step guidance on how the country can establish them.

Given that WGS is a new technology, epidemiologists using it will need additional training to understand its basic principles and to interpret its results.

3.2 Public health laboratory capacities

During an outbreak investigation, it is important to make sure that clinical specimens are collected from suspected cases, as those samples are crucial to confirm the causal pathogen. At the laboratory, specimens should be timely processed, following standard methods. The results must be reliable and replicable. To strengthen laboratories, the implementation of a quality management system is recommended, based on International Organization of Standardization (ISO) 15189, or another similar norm. Although many laboratories will not be prepared to meet all ISO 15189 requirements, it is important that laboratories implement as many quality management system components as possible (Box 3).

BOX 3

Minimum requirements for laboratory support to foodborne disease outbreak investigations

For clinical specimen collection in the field, ensure that:

- health care workers and outbreak response teams have been trained to obtain appropriate clinical specimens;
- specimens are regularly collected as part of outbreak investigations;
- a laboratory has been designated to test for priority foodborne pathogens;
- specimens can be transported to the designated laboratory in appropriate conditions; and
- the designated laboratory has experience and resources to culture and identify priority foodborne pathogens (e.g. *Salmonella* spp., *Shigella* spp., *Vibrio* spp., *Campylobacter* spp., Shiga-toxin producing *Escherichia coli* (STEC), *Listeria monocytogenes*).

Ensure there is a laboratory quality management system in place, which includes:

- regular internal quality control to monitor performance of specific assays for foodborne pathogen detection with standard laboratory methods documented; and
- participation in external quality assurance programs (proficiency testing) programmes covering the full range of assays offered by the laboratory.

If specimen collection and transport conditions in outbreak situations are not met, the publication "Strengthening surveillance of and response to foodborne diseases" *(13)* provides a step-by-step guide to build those capacities. The laboratory should also be able to conduct bioinformatic analyses or have access to such skills to support the implementation of WGS. The implementation of quality laboratory systems is a long-term goal, which should be planned and coordinated carefully.

WHO offers additional resources to assist national authorities committed to improving the quality of public health laboratories, such as the "Laboratory quality management system handbook" *(23)*, and the WHO website for strengthening laboratories that support surveillance and response systems, "Laboratory quality stepwise implementation tool" *(24)*. These guides show the steps required in the logical implementation of a quality management system, and activities are divided in four phases with a specific focus.

Phase 1 **Ensure that the primary process of the laboratory operates correctly and safely.**

Phase 2 **Control and assuring quality and creating traceability.**

Phase 3 **Ensure proper management, leadership and organization.**

Phase 4 **Create continuous improvement and prepare for accreditation.**

Even when a laboratory does not reach full implementation of the quality management system, it will have improved the quality of its service provision, from the moment it completes Phase 1.

The quality management template is an additional step in the implementation tool based on internationally accepted standards, providing guidance to public health and clinical laboratories on developing policies and procedures in support of a quality management system *(27)*. Countries can use the template and adapt it to the local context to develop their own quality management system manual.

3.3 Capacities in the food safety sector

During an outbreak investigation, food safety personnel responding to the outbreak need to be able to collect food and environmental samples. Those samples need to be tested in a designated laboratory that has the skills and expertise to detect priority foodborne pathogens. Confirmatory testing of agents identified in food and/or environmental samples provides important evidence on an outbreak's source. Microbiology laboratories that test food and environmental samples might be public health laboratories or a specialized laboratory. Strengthening food safety laboratory capacities are beyond the scope of this manual, but should, nonetheless, be considered. Furthermore, collaboration and information sharing among public health laboratories epidemiologists, food safety laboratories, and their staff is critical to respond efficiently to foodborne disease outbreaks. As WGS capacities improve, public health laboratory staff should communicate regularly with food/environmental laboratories to ensure data comparability and sharing.

The food safety system has responsibility over the implementation and enforcement of pertinent laws and regulations. During a foodborne disease outbreak, once there is evidence of a food item or food-related business being the source of a pathogen, the food safety system will need to take measures to protect the food chain, among them, issue public warnings regarding the consumption of suspect food items and halting further distribution to try to contain the outbreak and prevent the spread of the disease. Such measures might also help prevent future outbreak events. Box 4 lists the minimum capacities required of a food safety system.

BOX 4

Minimum requirements of a food safety system to undertake control measures during foodborne disease outbreaks

- Appropriate food and/or environmental samples are collected during foodborne disease outbreaks.
- Samples are tested at a laboratory capable of identifying priority foodborne pathogens.
- There are food safety laws and regulations that support the application of control measures.
- Food safety personnel can intervene to control the distribution of implicated food items.

For countries that do not meet the conditions in Box 4, guidance may be found in the following WHO publications: "Assuring food safety and quality: guidelines for strengthening national food control systems" *(26)* and "Strengthening national food control systems: guidelines to assess capacity building needs" *(25)*.

4. Options for implementing WGS

Once the decision has been made to use WGS to strengthen foodborne diseases surveillance and response systems, the two main options available (depending on in-country capacities and resources) are:

- supporting foodborne disease outbreak investigations (see more details in the outbreak module)
- enhancing routine surveillance of foodborne diseases (see more details in the surveillance module).

This section will address said options and provide guidance on choosing the appropriate one.

Detailed guidelines for implementing each option are provided in separate stand-alone modules. Fig. 4 summarizes the attributes of each option. A stepwise approach may be followed to implement WGS as part of surveillance and response systems. Countries with limited laboratory and epidemiological capacity may only use WGS for outbreak investigation. However, as the public health laboratory develops, epidemiological skills strengthen, the food safety system becomes stronger and a country may wish to use WGS for further characterization of foodborne pathogens included in their surveillance system. Usually, to assess the feasibility of either option, a pilot study will be conducted prior to full-scale implementation.

Both implementation options are briefly described below, as well as the vision of the role of WGS within surveillance and response systems, WGS objectives, minimum requirements, timeframes for public health action and some key considerations for implementing the chosen option. Each option will be discussed in more detail in the relevant module.

Fig. 4

Attributes of the implementation options

4.1 Foodborne disease outbreak investigations support

Vision for WGS to support foodborne disease outbreak investigations

- Outbreaks of foodborne diseases are detected through existing surveillance systems, whether indicator-based or event-based or both.
- An outbreak response team is set up, and a case definition developed.
- Discussions between epidemiologists, laboratory scientists and bioinformaticians take place to discuss how WGS may support outbreak investigations, desirable timeframes and outputs for WGS results for public health use.
- Initial isolates of the relevant foodborne pathogen from human samples are sequenced. The laboratory refines the reference genome, if required. An outbreak strain is identified and included in the case definition. All isolates collected as part of the outbreak investigation are then characterized and classified as either outbreak or non-outbreak strain.
- An identified microorganism that might be genetically related to the outbreak strain is jointly reviewed by the laboratory scientist, bioinformatician and epidemiologist, together with any epidemiological information available.
- Epidemiologists use the strain's characterization together with any epidemiological information to determine if the case meets the case definition.
- Epidemiologists and/or the outbreak response team interview cases to try to identify the suspect food item, which, depending on its nature, might need to be further investigated (traceback) to make a definite identification.
- Samples of the suspect food item, as well as relevant environmental samples are collected. Isolates obtained from food and/or environmental samples are sequenced and compared with those from individuals meeting the case definition of the outbreak under investigation.
- The food vehicle causing the illness is identified and control measures are implemented in the food safety sector.

Objectives of sequencing during an outbreak investigation

The objectives of sequencing priority foodborne pathogens during outbreak investigations will depend on the national capacity to conduct outbreak investigations and the routine subtyping methods already in place. One of more of the following may apply to:

- confirm the presence of an outbreak by determining whether strains are, genetically, closely related, if no other subtyping already exists;
- describe the magnitude of the outbreak, by including the suspect strain in the outbreak case definition;
- improve specificity in epidemiology data analysis, by removing any non-outbreak cases;
- determine if any cases detected prior to the outbreak's onset might be linked to it;
- assist in identifying the potential source of the outbreak, by providing food history data from cases carrying the outbreak strain, as this will help the food safety sector conduct trace-back activities and identify the microorganism from food or environmental samples, when they become available;
- increase specificity in analytical studies, by eliminating non-outbreak cases;
- determine any national or international linkages to the outbreak; and
- describe any changes in the pathogen's evolution.

As this approach will often be used where the laboratory and sequencing capacities are limited, its secondary objectives are to:

- determine whether WGS is an effective tool for use during outbreak investigations;
- determine whether the use of WGS as a tool for routine surveillance is appropriate in the country;
- identify the barriers to sequencing; and
- assist laboratory and surveillance staff to start building sequencing capacity, with adequate technical support.

Minimum requirements for WGS in outbreak investigations

Minimum requirements for a country wishing to conduct WGS for priority foodborne pathogens during outbreak investigations are listed below. The epidemiological, laboratory and food safety requirements need to be in place before considering WGS. However, the outbreak module, which is part of this manual, will provide options for meeting the bioinformatics, governance and resource requirements during the planning process for WGS.

Epidemiology

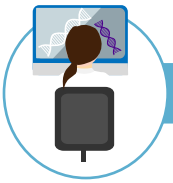
Minimum epidemiological requirements include:

- capacity to identify foodborne diseases outbreaks within the existing surveillance system
- capacity to investigate any outbreaks detected and complete standardized food history interviews
- capacity to conduct food frequency analyses
- availability of staff able of conducting analytical epidemiology studies to help identify a possible food source.

Laboratory

Minimum laboratory requirements include:

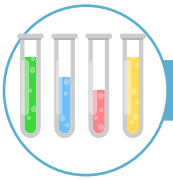
- capacity to culture and isolate selected priority foodborne pathogens from human specimens
- capacity to extract genomic DNA/RNA from the above pathogens
- availability of staff able to perform molecular techniques
- access to a sequencer and technical support.



Bioinformatics

Minimum bioinformatic requirements include:

- access to an analytics pipeline for processing sequencing data
- capacity for storing, sharing and interpreting sequencing data
- access to bioinformatics and microbial genomics experts to assist in data interpretation.



Governance

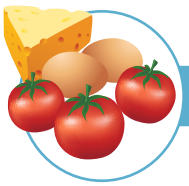
Minimum governance requirements include:

- capacity to develop contractual or collaborative agreements for:
  - access to sequencing technology and technical support (laboratory and bioinformatics)
  - materials transfer
  - sharing metadata and digital sequences.

Resources

Minimum resource requirements include:

- political commitment to sequencing foodborne pathogens during outbreak investigations;
- funding for sequencing during foodborne disease outbreaks.



Food safety

Ensure regulatory requirements are in place to enable food safety authorities to apply control measures in the food chain, informed by the investigation's findings.

Timeframe for public health action

Human specimens will be collected and analysed as part of the existing outbreak response. To meet the objectives specified above, human isolates will need to be sequenced as soon as possible to determine if the isolates are genetically related, suggesting a common source. If a food isolate is obtained as part of the outbreak investigation, it should be sequenced as soon as possible, as well, to determine whether it is closely related to the human isolate. This will help confirm the source of the illness and enable swift implementation of public health measures.

Key considerations

The main question that will need to be answered when exploring this option is who will perform WGS. It might be that the national public health laboratory has the capacity to do sequencing or that funding is available to establish said capacity. If not, a country may outsource WGS or parts of it. In the latter case, national authorities will need to:

- define the needs;
- determine the existing sequencing capacity within the country, for instance, in university settings or in private commercial institutions;
- consider sending nucleic acid to a facility abroad, if unable to conduct WGS nationally;
- enter into a service contract, if the sequencing will be done outside the public health laboratory; and
- ensure timely and accurate results to inform foodborne outbreak investigations.

4.2 Enhancing routine surveillance of foodborne diseases

Vision for using WGS to enhance the routine surveillance of foodborne diseases

- As part of the surveillance system, sick people who seek health care submit specimens which are cultured to identify the pathogen responsible for the illness.
- The isolates are sequenced, the outputs are analysed and the results reported to public health authorities.
- Decisions are made by public health authorities, in consultation with laboratory staff, about clusters that require an epidemiological investigation. Epidemiological responses can be launched quickly with a small number of cases in a cluster.
- Ideally, data from the surveillance system in the human health sector are combined with sequencing results from the animal health and food safety sectors in real time. Comparing human and non-human sequences of various pathogens can lead to the identification of potential sources of the pathogen and inform control measures.
- Sequence data can also be used to analyze AMR and virulence factors. The WGS data from the surveillance system can be used to gain a better understanding of the epidemiology of foodborne pathogens and inform clinical management policies.
- Sequence data can also be used to identify new or emerging strains of pathogens that are potentially more virulent.

Objectives of using WGS to enhance routine foodborne diseases surveillance

The objectives of using WGS in routine surveillance of foodborne diseases are to:

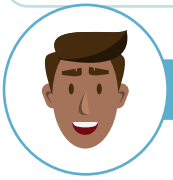
- monitor the emergence and spread of foodborne pathogen subtypes over time (including their resistance to antimicrobials or AMR);
- detect outbreaks of foodborne diseases, including low-intensity and geographically dispersed outbreaks;
- support the response to foodborne disease outbreaks;
- determine the magnitude of the problem of foodborne diseases;
- attribute food sources to specific foodborne diseases;
- inform clinical management policy, where appropriate (e.g. to reduce AMR in humans through the appropriate use of antimicrobials in a clinical setting);
- inform antimicrobial use policy in food-producing animals, where appropriate (e.g. to reduce AMR in food-producing animals and humans *(3)*;
- integrate human health data with data from other links of the food chain, to guide public health action to prevent and control foodborne diseases;
- inform risk-based food safety management; and
- monitor and evaluate interventions and measures to prevent and control foodborne diseases.

Secondary objectives of this approach are to:

- determine whether the ongoing use of WGS is appropriate for wider national routine surveillance; and
- build in-country capacity for sequencing in laboratories, bioinformatics support, and result interpretation by epidemiologists and public health staff.

Minimum requirements

Epidemiological, laboratory and food safety requirements need to be in place prior to considering WGS. However, the surveillance module will suggest options for meeting the bioinformatics, governance and resource requirements during the WGS planning process. Minimum requirements for a country wishing to enhance WGS surveillance are listed below.



Epidemiology

At a minimum, ensure:

- there is a fully functional notifiable disease surveillance system, that:
  - records case-based data that relies on laboratory confirmation of priority foodborne pathogens
  - can monitor trends in foodborne diseases
  - can detect foodborne disease outbreaks;
- there is a fully functional event-based surveillance system capable of detecting foodborne events; and
- outbreak response teams that can identify food sources with standardized tools for collecting case information and conducting analytical epidemiological studies.

Laboratory

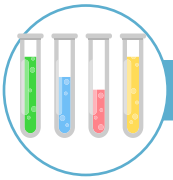
At a minimum, ensure:

- there are existing laboratories that routinely identify foodborne pathogens and report them to the notifiable disease surveillance system;
- the public health laboratory has equipment and infrastructure to handle molecular assays; and
- there is a public health laboratory or network of laboratories, which can lead the standardization of sequencing methods (e.g. platforms, analysis pipelines and outputs) to enable comparison of sequences nationally and internationally.

Bioinformatics

At a minimum, ensure:

- there is a pipeline or a set of protocols available to process sequencing data;
- there is enough capacity to store sequence information securely, where data are regularly backed up, and there are access controls to the data storage warehouse;
- there is enough computing and network capacity for sequence sharing to enable national comparison of sequences for all foodborne pathogens; and
- bioinformaticians are available to provide advice on and support sequencing data output interpretation.



Governance

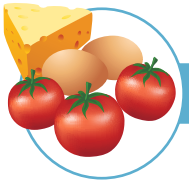
Enhancing governance can be achieved by building:

- referral pathways for sharing isolates, as well as agreed reporting mechanisms among laboratories;
- political commitment to WGS for surveillance of foodborne diseases; and
- a history of successful multi-sectoral collaboration among the human health, food safety and animal health sectors.

Resources

In terms of resources, ensure the system has, at a minimum:

- committed ongoing funds for sequencing foodborne pathogens as part of the notifiable disease surveillance system;
- adequate laboratory staffing for processing isolates, analysing the results, and reporting to the notifiable diseases surveillance system;
- adequate staffing in the human health sector to detect and respond to foodborne disease outbreaks; and
- adequate staffing in the food safety sector to undertake public health interventions.



Food safety

At a minimum, ensure the system has:

- the ability to perform product tracing investigations to determine the source of foods suspected as outbreak sources;
- adequate resources to investigate, inspect and collect samples from businesses that produce food suspected as outbreak sources; and
- regulatory requirements in place to enable food safety authorities to implement measures in the food chain, informed by surveillance or outbreak investigation data.

Timeframe for public health action

Because one objective of the surveillance system is to detect outbreaks early, turnaround time from pathogen isolation to sequencing results needs to be as short as possible. Data need to be reported to public health authorities to ensure that case clusters can be followed up and rapidly investigated. Swift action by the food safety sector is needed to correct any problems detected and to improve public health and safety.

Key considerations

The main questions that need to be answered regarding the role of WGS in the surveillance system will relate to pathogen selection, geographical coverage, and identifying components that need to be initially outsourced.

It is important to plan the laboratory's workflow to determine the outputs that will be shared with public health authorities. Matching sequencing results to epidemiological information will require careful planning as well as a policy framework to allow data-sharing.

4.3 Deciding on an option

While WGS is a powerful tool with many benefits when choosing an option, it is important to be aware of its limitations and those related to applying the technology in a public health setting. For WGS to be useful for foodborne diseases, turnaround time needs to be quick to ensure outbreaks are timely detected, verified and responded to as quickly as possible. Delays in identifying the source of a foodborne disease outbreak can lead to more exposures and, consequently, more cases.

The main factors to consider when choosing an option for implementing sequencing are the:

- objectives for using WGS within the surveillance and response system
- potential availability of financial resources
- availability of human resources
- current structure and objectives of the surveillance and response system
- role of the laboratory in the surveillance and response system
- food safety system and its ability to implement public health measures.

Fig. 5 illustrates an algorithm that can be useful for deciding which option to use, based on existing laboratory and surveillance infrastructure. There are other considerations which may affect decision-making (Table 1). For example, if funds and human resources are limited in the surveillance and response system, using WGS to support outbreak investigations might be the most appropriate option. However, if there is an established public health laboratory that routinely cultures priority foodborne pathogens, it may be possible to scale up its performance by using WGS for surveillance purposes.

Fig. 5

Decision tool on options for implementing WGS in the foodborne diseases surveillance and response system

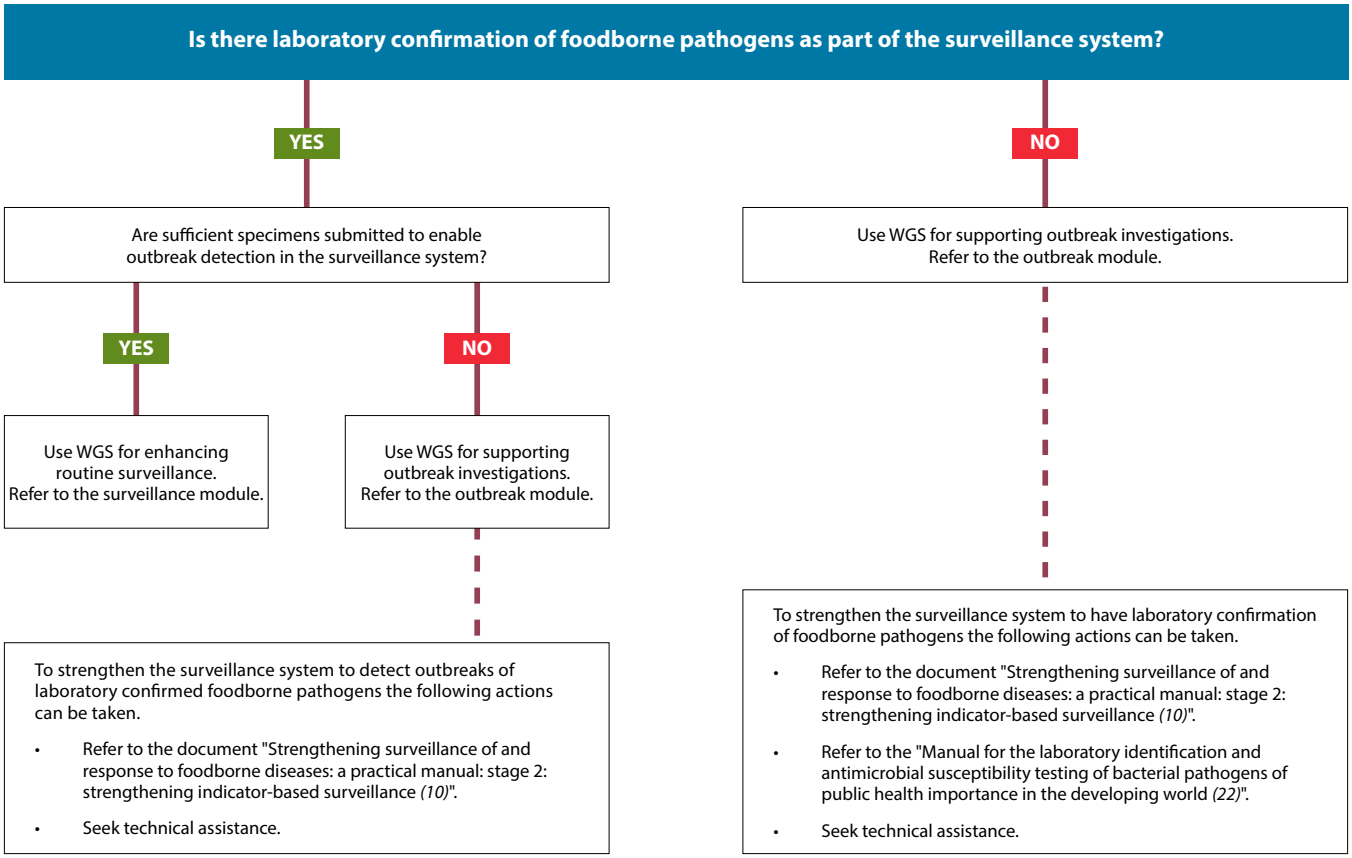


Table 1

Features of WGS implementation options in the foodborne diseases surveillance and response system

Features	Outbreak Investigations	Routine Surveillance
Objective of using WGS		
Outbreak detection		✔
Outbreak response	✔	✔
Surveillance		✔
Assess feasibility of WGS in the surveillance and response system	✔	
Financial resources available[1]		
Low	✔	
Medium	✔	✔
High		✔
Human resources available[2]		
Low	✔	
Medium	✔	✔
High		✔
Type of disease surveillance system		
Mostly syndromic surveillance (i.e. laboratory has minimal role in disease surveillance)	✔	
Routine laboratory confirmation of priority foodborne pathogens		✔
Computational resources		
Limited computer and internet resources	✔	
Dedicated computing and IT resources for WGS		✔
Bioinformatic resources		
Limited bioinformatics resources available	✔	
Dedicated bioinformatics support available		✔

[1]Low = small amount of funding available; high = large amount of funding available.

[2]Low = few trained staff available within the system; High = enough trained staff available in the system.

There are advantages and disadvantages to each implementation option as discussed in Table 2.

Table 2

Advantages and disadvantages of each implementation option

WGS option	Advantages	Disadvantages
 **Supporting outbreak investigations**	• Initially, WGS capacity does not need established in the public health laboratory. • Does not require large financial investments in infrastructure, if there is a sequencer that can be accessed in-country. • WGS results can answer specific questions related to outbreak investigations. • Provides guidance regarding the feasibility of implementing WGS within the existing surveillance and response system.	• If outsourcing, it does not help to build capacity for WGS within the public health laboratory. • WGS is not financially viable if only a small number of isolates are collected during outbreaks. • Results may not be obtained early enough to inform public health decision-making during outbreak investigations. • When used in an ad-hoc manner, it does not provide valuable information for surveillance purposes.
 **Enhancing routine surveillance of foodborne pathogens**	• Sequencing results are used to monitor trends in foodborne diseases, detect outbreaks and answer questions during outbreaks. • Sequencing becomes the preferred method for further characterization of foodborne pathogens within the surveillance and response system. • May be more cost effective overall, as there is only one laboratory method for further typing of multiple pathogens, rather than various typing tests for different pathogens. • Opportunities for integrated food chain surveillance, whereby data from the animal health, food safety and human health sectors are shared to identify microbiological food hazards along the food chain.	• Costly to implement. • Currently, there are few internationally agreed standards for WGS-based typing (only PulseNet International members have agreed to use wgMLST). • Early stages of the project may not be able to provide timely results to detect and respond to case clusters. As analytics and reporting processes are tested and revised, they may be limited by the small number of isolates included in the analyses. • If there is only one laboratory with sequencing capacity and specimens are being referred from peripheral sites, quick reporting might be affected, thus limiting timely detection of outbreaks. • Needs enough public health resources to investigate detected case clusters, and gather good quality epidemiological data. • Public health action must be ensured within the food safety system when food items or food businesses are identified as the source of an outbreak.

References

1. WHO estimates of the global burden of foodborne diseases. Geneva: World Health Organization; 2015 (https://apps.who.int/iris/handle/10665/199350, accessed 31 May 2022).
2. Hasman H, Hammerum AM, Hansen F, Hendriksen RS, Olesen B, Agersø Y et al. Detection of mcr-1 encoding plasmid-mediated colistin-resistant *Escherichia coli* isolates from human bloodstream infection and imported chicken meat, Denmark 2015. Euro Surveill. 2015;20(49):30085. doi:10.2807/1560-7917.ES.2015.20.49.30085.
3. Tang KL, Caffrey NP, Nóbrega DB, Cork SC, Ronksley PE, Barkema HW et al. Restricting the use of antibiotics in food-producing animals and its associations with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis. Lancet Planet Health. 2017;1(8):e316–e327. doi:10.1016/S2542-5196(17)30141-9.
4. Jackson BR, Tarr C, Strain E, Jackson KA, Conrad A, Carleton H et al. Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. Clin Infect Dis. 2016;63(3):380–86. doi:10.1093/cid/ciw242.
5. Inns T, Ashton PM, Herrera-Leon S, Lighthill J, Foulkes S, Jombart T et al. Prospective use of whole genome sequencing (WGS) detected a multi-country outbreak of *Salmonella Enteritidis*. Epidemiol Infect. 2017;145(2):289–98. doi:10.1017/S0950268816001941.
6. Butcher H, Elson R, Chattaway MA, Featherstone CA, Willis C, Jorgensen F et al. Whole genome sequencing improved case ascertainment in an outbreak of Shiga toxin-producing *Escherichia coli* O157 associated with raw drinking milk. Epidemiol Infect. 2016;144(13):2812–23. doi:10.1017/S0950268816000509.
7. One Health. Geneva: World Health Organization; 2017 (https://www.who.int/news-room/questions-and-answers/item/one-health, accessed 16 May 2023).
8. Grimont PA, Weill FX. Antigenic formulae of the *Salmonella* serovars, 9th edition. Paris: WHO Collaborating Centre for Reference and Research on Salmonella, Pasteur Institute; 2007 (https://www.pasteur.fr/sites/default/files/veng_0.pdf, accessed 4 April 2023).
9. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J et al. PulseNet International: vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Euro Surveill. 2017;22(23):30544. doi:10.2807/1560-7917.ES.2017.22.23.30544.
10. Strengthening surveillance of and response to foodborne diseases: a practical manual. Stage 2: strengthening indicator-based surveillance. Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/259472).
11. Strengthening surveillance of and response to foodborne diseases: a practical manual. Stage 3: integrating surveillance data to better understand risks across the food chain. Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/259476, accessed 31 May 2022).
12. Strengthening surveillance of and response to foodborne diseases: introductory module. Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/259469, accessed 31 May 2022).
13. Strengthening surveillance of and response to foodborne diseases: a practical manual. Stage 1: using indicator and event-based surveillance to detect foodborne events. Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/259471, accessed 31 May 2022).
14. Integrated surveillance of antimicrobial resistance in foodborne bacteria: application of a one health approach: guidance from the WHO Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR). Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/255747, accessed 11 April 2022).
15. Technical meeting on the impact of whole genome sequencing (WGS) on food safety management: within a One Health approach. The 9th meeting of the Global Microbial Identifier (GMI9): final meeting report. Rome: Food and Agriculture Organization of the United Nations; 2016 (https://www.fao.org/documents/card/en/c/ada03815-5184-461f-995e-5475ae2fd198/, accessed 11 April 2022).
16. Technical background paper: applications of whole genome sequencing (WGS) in food safety management. Rome: Food and Agriculture Organization of the United Nations; 2016 (https://www.fao.org/documents/card/en/c/61e44b34-%20b328-4239-b59c-a9e926e327b4/, accessed 11 April 2022).
17. Whole genome sequencing for foodborne disease surveillance: landscape paper. Geneva: World Health Organization; 2018 (https://apps.who.int/iris/handle/10665/272430, accessed 11 April 2022).

18. International Health Regulations (2005), 2nd edition. Geneva: World Health Organization; 2008 (https://apps.who.int/iris/handle/10665/43883, accessed 11 April 2022).

19. Whole genome sequencing as a tool to strengthen foodborne disease surveillance and response. Module 2. Whole genome sequencing in foodborne disease outbreak investigations. Geneva: World Health Organization; 2023.

20. Whole genome sequencing as a tool to strengthen foodborne disease surveillance and response. Module 3. Whole genome sequencing in foodborne disease routine surveillance. Geneva: World Health Organization; 2023.

21. Strengthening surveillance of and response to foodborne diseases: a practical manual. Stage 1: investigating foodborne disease outbreaks. Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/259475, accessed 12 April 2022).

22. Manual for the laboratory identification and antimicrobial susceptibility testing of bacterial pathogens of public health importance in the developing world: *Haemophilus influenzae, Neisseria meningitidis, Streptococcus pneumoniae, Neisseria gonorrhoea, Salmonella* serotype Typhi, *Shigella,* and *Vibrio cholerae.* Geneva: World Health Organization; 2003 (WHO/CDS/CSR/RMD/2003.6; https://apps.who.int/iris/handle/10665/68554, accessed 12 April 2022).

23. Laboratory quality management system: handbook, version 1.1. Geneva: World Health Organization; 2011 (https://apps.who.int/iris/handle/10665/44665, accessed 11 April 2022).

24. Laboratory quality stepwise implementation tool. Geneva: World Health Organization; 2015 (https://extranet.who.int/lqsi/content/homepage, accessed 31 May 2022).

25. Strengthening national food control systems: guidelines to assess capacity building needs. Rome: Food and Agriculture Organization of the United Nations; 2006 (https://www.fao.org/3/a0601e/a0601e.pdf, accessed 12 April 2022).

26. Assuring food safety and quality: guidelines for strengthening national food control systems. Rome: Food and Agriculture Organization of the United Nations; 2003 (Food and Nutrition Paper No. 76; https://www.fao.org/documents/card/en/c/92f82d38-5557-4ca1-b361-be14cd129db6/, accessed 11 April 2022).

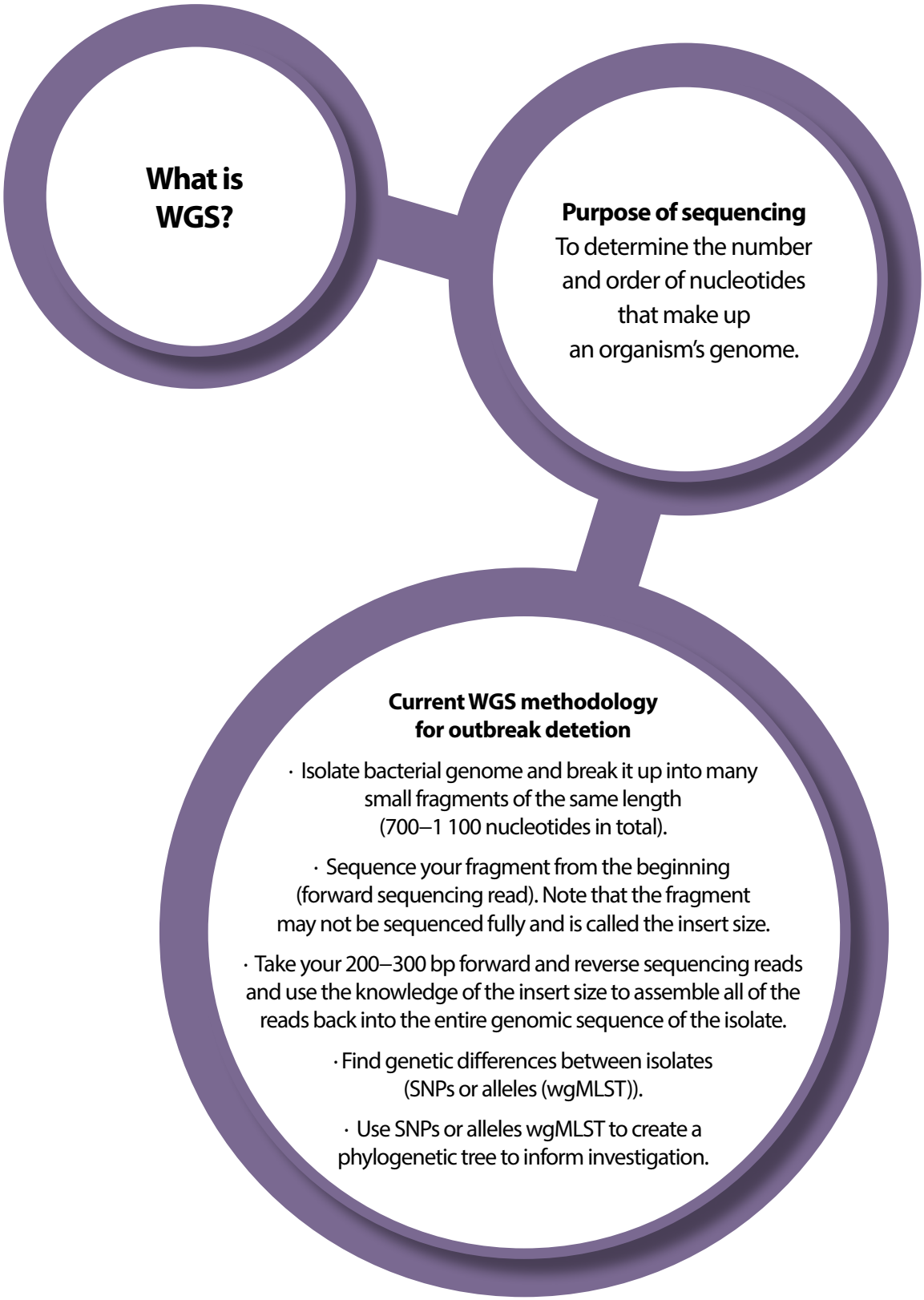
27. Quality manual template. Geneva: World Health Organization; 2014 (https://www.who.int/publications/m/item/laboratory-quality-manual, accessed 31 May 2022).

Annex 1. Overview of WGS

This annex contains an overview of WGS and how WGS data can be compared to make conclusions about genetic relatedness. The areas covered are broken into two parts. Annex 1.1 consists of a discussion, in a slide format, of the following: defining WGS (Slide 1); how WGS works (Slide 2–8); and some outputs from WGS used in foodborne disease surveillance and response, i.e. SNP, wgMLST and phylogenetic trees (Slide 9–26). Annex 1.2 is a more in-depth discussion on data comparability and genetic relatedness based on the slides in Annex 1.1 and how WGS is used in foodborne disease surveillance and response.

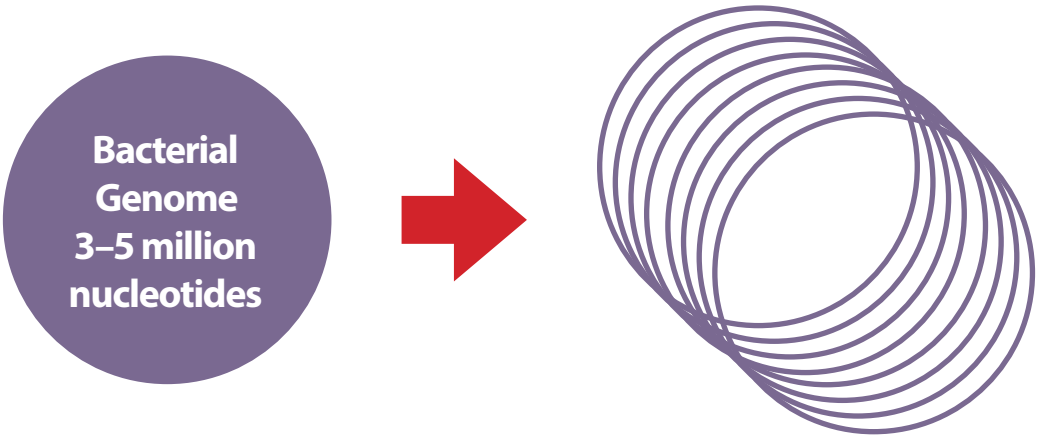
Annex 1.1 What is WGS and how does it work?

Slide 1 **Defining WGS**



Slide 2

Isolate and copy bacterial genome

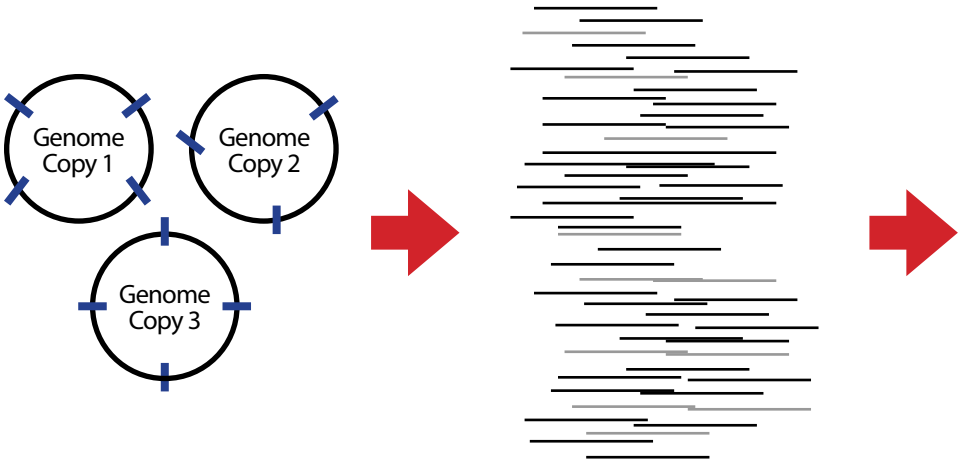


OVERVIEW:

Current WGS methodology involves growing up the bacterial genome to be sequenced into a pure culture (one bacterial colony) so that there are many copies of the same bacterial genome. Then the bacterial DNA is extracted.

Slide 3

Break up multiple copies of same genome into different pieces



- Depending on sequencing chemistry you will have thousands to millions of reads.
- Long-read: up to 40 kb in length (thousands of reads) and fragmentation is different.
- Short-read: up to ~500 bp (can be millions of reads) and involves paired-end-sequencing.

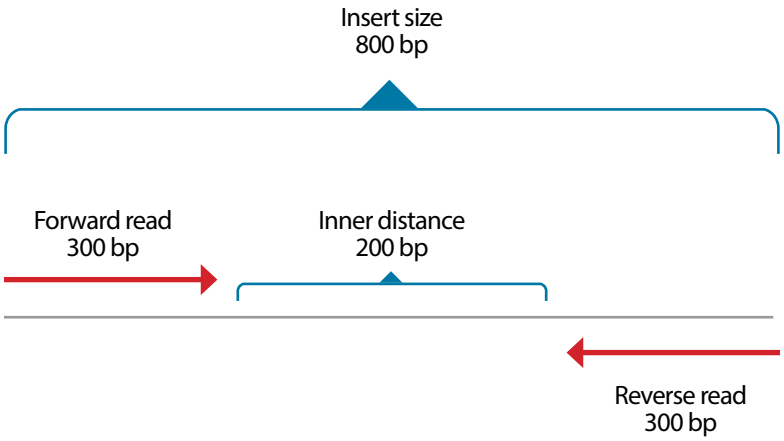
OVERVIEW:

Current WGS methodology involves cutting multiple copies of the same genome in different positions to generate shorter fragments (of the same length that you can size select for) that will overlap later on when re-assembling the genome.

Note: This is key to generating a WGS.

Slide 4

Forward and reverse reads (and insert size) or paired-end sequencing

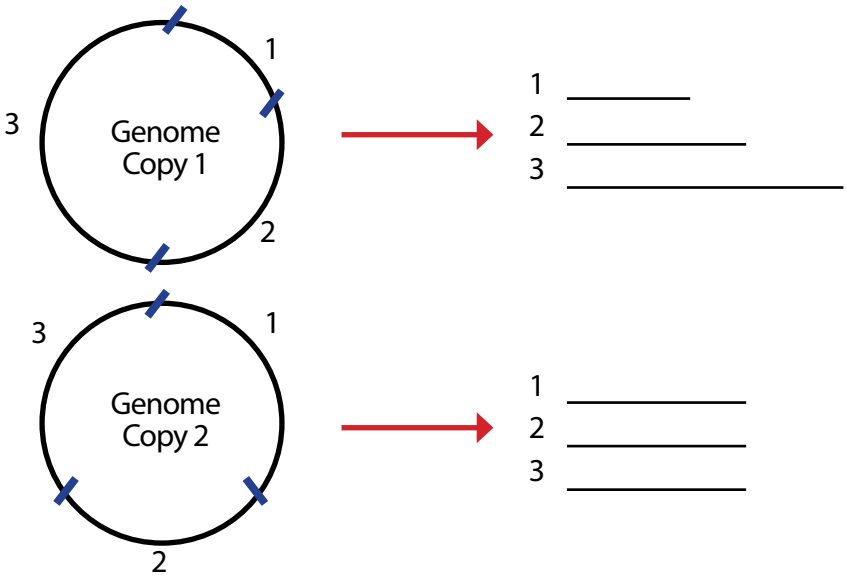


OVERVIEW:

Instead of sequencing the entire 800 bp (more time and more expensive) you can sequence the first 300 bp (forward read) and the last 300 bp (reverse read) and then use the knowledge of the inner distance (200 bp) to help ensure the forward read and reverse read are aligned/assembled together.

Slide 5

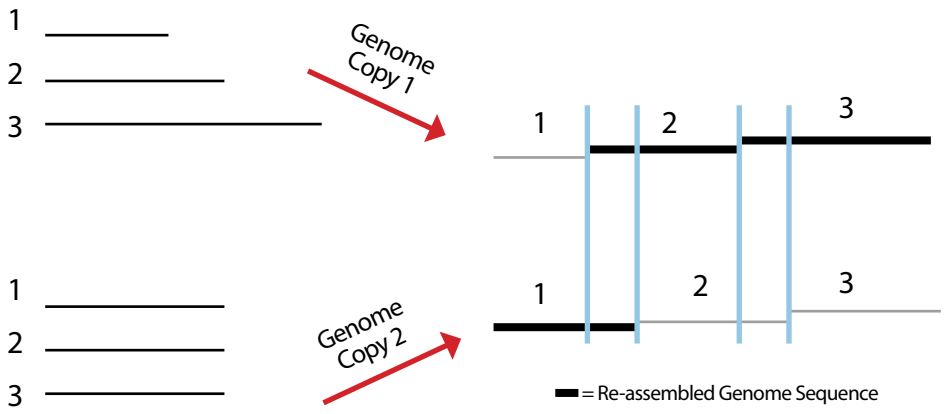
Break up same genome into different pieces



/ = breakpoint that cuts up the circular genome. In this example, there are three breakpoints for Genome Copy 1 and Genome Copy 2. The numbers on the outside of each circle correspond to the fragment sizes to the right of each respective circle.

Slide 6

How assembly works

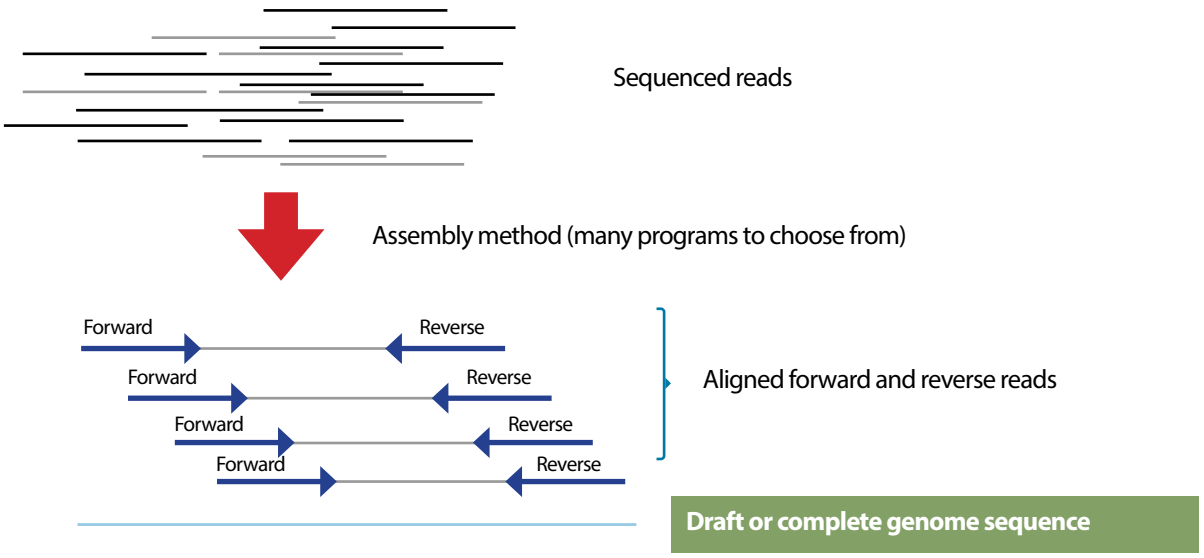


OVERVIEW:

Find overlaps (same string of nucleotides) between reads with different breakpoints. In this example, fragment 1 (Genome Copy 1) is aligned with fragment 1 (Genome Copy 2). The bold line represents the overlap (same string of nucleotides), and because fragment 1 (Genome Copy 2) is longer than fragment 1 (Genome Copy 1), part of its sequence will now overlap with fragment 2 (Genome Copy 1). The bold line is now able to be extended further when fragment 2 (Genome Copy 1) aligns with fragment 2 (Genome Copy 2). This continues until all the reads are connected.

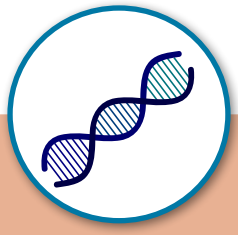
Slide 7

Assemble reads from all copies in order and use insert size between forward and reverse reads



Slide 8

Condensing multiple reads into a draft or complete genome



AGGATTGTTGGCAG
GGAATGTTGGCAGT
GAATGTTGGCAGTC
AATGTTGGCAGTCG

- 4 sequence reads that have been aligned together from the same isolate.
- Red letter indicates possible sequencing error that is resolved by looking at other nucleotides from other reads at that position.

AGGAATGTTGGCAGTCG

- Derived genomic sequence that takes into account the other sequencing reads.

Slide 9

Comparing a single isolate's reads against a reference genome

CCCGGCCCTTTGGCCCTTTGGGAAATCGCCCCAATGGAAATTT
CCCTTTGGCCCTATGGGAAATCGCCCCAATGGAAATTT
CGGCCCTTTGGCCCTATGGGAAATCGCCCCAATGGAAATTT
CCCGGCCCTTTGGCCCTATGGGAAATCGCCCC
CCTATGGGAAATCGCCCCAATGGAAATTT
GCCCTTTGGCCCTATGGGAAATCGCCCCAATGGAAA
CCCGGCCCTTTGGCCCTTTGGGAAATCGCCCCAATGGAAAT
CCCGGCCCTTTGGCCCTATGGGAAATCGCCCCAATGG
CCCGGCCCTTTGGCCCTATGGGAAATCGCCCCAATGGAAATT
CCCGGCCCTTTGGCCCTATGGGAAATCGCCCCAATGGA
CCCGGCCCTTTGGCCCTATGG

Reference genome

- Sequencing reads that reveal a variant (A) when compared with a reference genome.
- Note that the reference genome can change and a SNP would not be called if the reference genome chosen had an A instead of a T at this position.
- The selection of a reference genome (e.g. complete genome has been resolved with no gaps) is important to the downstream analysis and should be chosen on a case-by-case basis that has neither to many SNPs or too few SNPs between the outbreak cluster *(1-4)*.

Slide 10

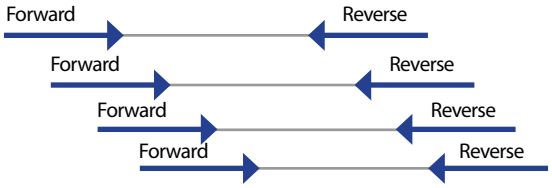
Comparing the sequences of different bacterial isolates and finding the genetic differences

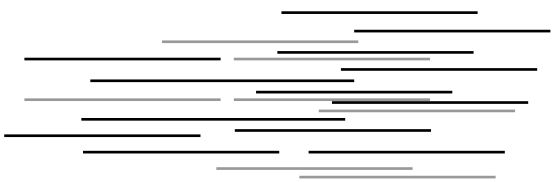
Isolate 1	TGGAATGTTGGCAGTCG
Isolate 2	AGGAATGTTGGCAGTCG
Isolate 3	AGGAATGTTGGCAGTCG
Isolate 4	AGGAATGTTGGCAGTCC

- Four genome sequences have been aligned together from different isolates.
- The red letter indicates a SNP or a different nucleotide in one sequence when compared with another at the same position.
- In this example, the first and the fourth isolates have sequences that are different than the sequences from the second and third isolate which are identical.
- Based on this data, you could say that Isolate 2 and Isolate 3 are more closely genetically related because they have no SNPs.

Slide 11

Understanding genomes and long-read vs short-read sequencing





Short-read sequencing (200–300 bp forward and reverse reads with insert size depending on sequencing machine and platform):

- can get draft or nearly complete genome;
- might have gaps where the reads were not able to be resolved; and
- is cost-effective for surveillance of clinical illness and able to be used with food or environmental samples to find a link.

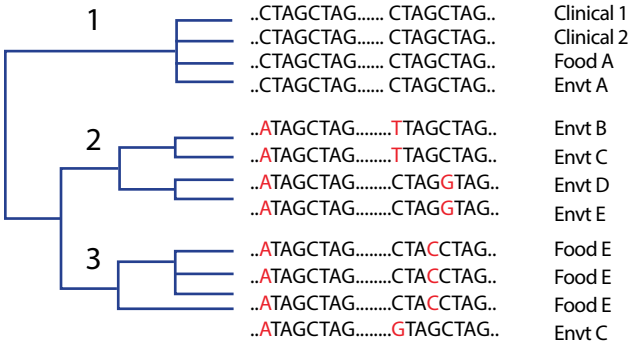
Long-read sequencing (5 kb + reads):

- can get complete genome with information on plasmids;
- is more expensive than short-read and not cost effective for routine use in surveillance or environmental monitoring; and
- reference genomes should be sequenced using this method is possible.

Draft or complete genome sequence derived from either short or long-read sequencing

Slide 12

Goal of wgs: constructing phylogenetic trees to infer the genetic relatedness of isolates



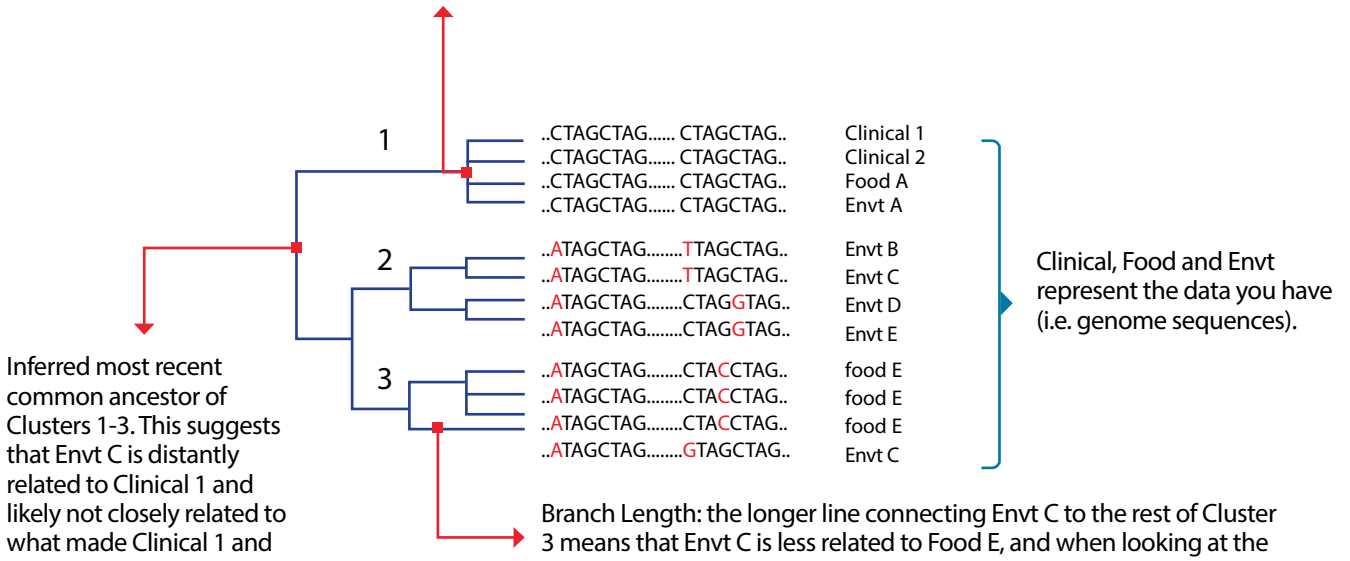
- Genetic relationship (evolutionary relationship) is inferred based on the number of SNP or allele (wgMLST) differences.
- In this example, Clinical isolates 1 and 2 have identical sequences to Food and Environmental A samples and thus form Cluster 1.
- Clusters 2 and 3 have SNP differences indicated by a red letter and are able to be separated from the sequences in Cluster 1.
- This clustering and partitioning could help support an outbreak and traceback investigation if we had epidemiological data to link Food A as something Clinical 1 or 2 ate.

Envt = environmental sample

Slide 13

Understanding the basics of a phylogenetic tree

Inferred most recent common ancestor of Clinical 1 and 2, Food A and Envt A. This suggests that Food A and Envt A are closely related to what made Clinical 1 and 2 ill.



Inferred most recent common ancestor of Clusters 1-3. This suggests that Envt C is distantly related to Clinical 1 and likely not closely related to what made Clinical 1 and 2 ill.

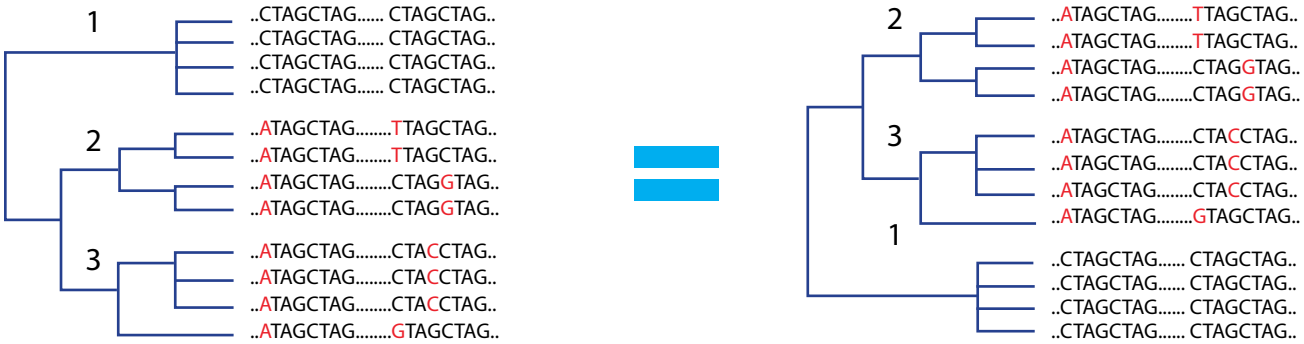
Branch Length: the longer line connecting Envt C to the rest of Cluster 3 means that Envt C is less related to Food E, and when looking at the sequence you see that it is 2 SNPs different.

Envt = environmental sample

Slide 14

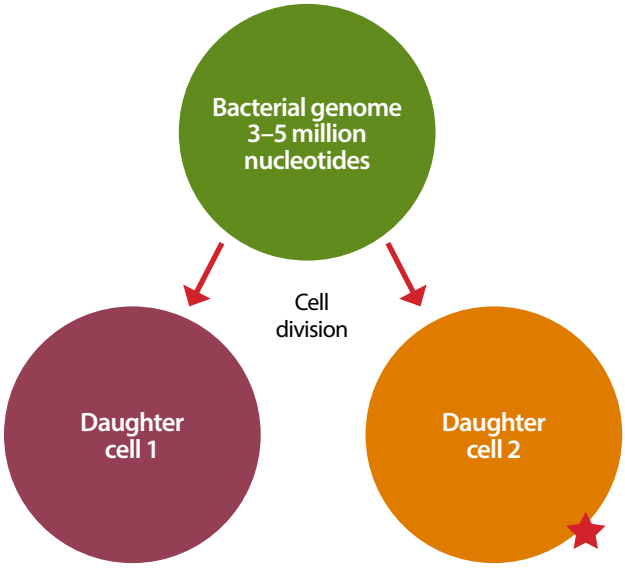
Understanding the basics of a phylogenetic tree

Same phylogenetic tree but the first branch has been flipped.



Slide 15

Understanding bacterial evolution and genetic relatedness

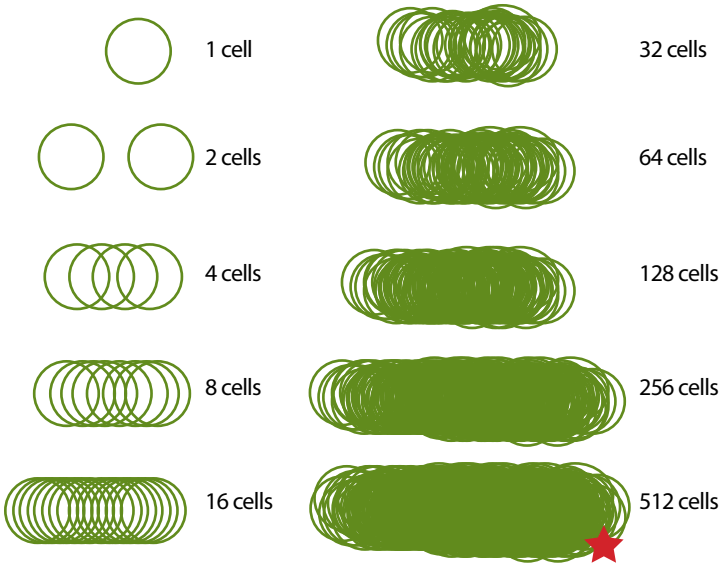


- A bacterial cell replicates its genetic material and then divides in half.
- Sometimes during replication a mutation occurs in the DNA and the genome of the daughter cell might be slightly different than the parent.
- As an example, we will use 0.003 as the estimated mutation rate. When the bacterial genome divides into Daughter cells 1 and 2 there is a 0.003 probability that a mutation will happen. In other words, there is a 0.003 probability that the genome of Daughter cell 2 will have a 1 SNP difference when compared with what the genome of the original bacterial genome was.

★ = 0.003 mutations per genome per cell generation

Slide 16

Detecting a mutation based on cell division



OVERVIEW

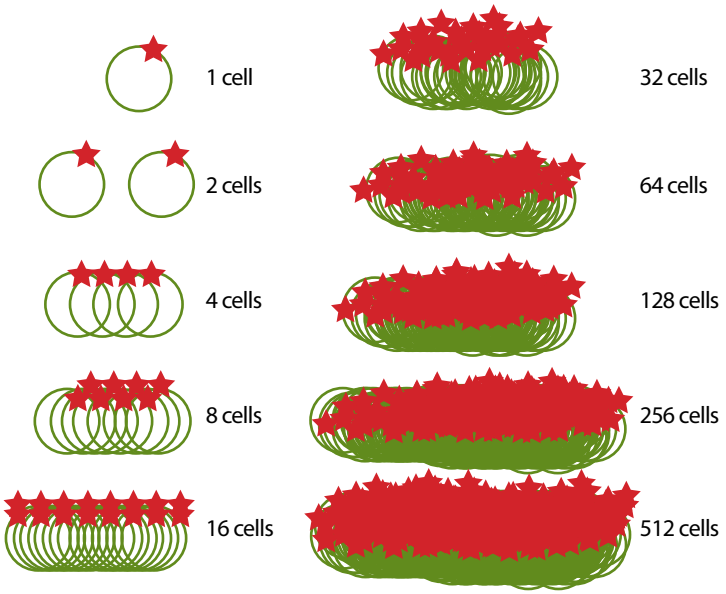
- Assuming a mutation rate of 0.003 mutations per genome per cell division, it would take nine cell divisions to see a mutation in a single cell (out of 512 cells).
- If the group of 512 cells were used for sequencing, then the fraction of reads with the mutation (or variant) would not be high enough to detect by current sequencing technologies.
- Remember, when condensing four reads into one genome sequence, if three of the reads show A and the other shows a T, you will select the A and the other shows a T, you will select the A as what the nucleotide is at that position.

AGGATTGTTGGCAG
GGAATGTTGGCAGT
GAATGTTGGCAGTC
AATGTTGGCAGTCG
Reads

AGGAATGTTGGCAGTCG
Draft genome

Slide 17

Detecting a mutation based on cell division



OVERVIEW

- To detect that mutation the bacterial cell having that variant needs to proliferate and be present in a majority of the sequencing reads.
- If the group of 512 cells were used for sequencing, then that variant could now be detected because it has proliferated.

AGGATTGTTGGCAG
GGAATGTTGGCAGT
GAATGTTGGCAGTC
AATGTTGGCAGTCG
Reads

AGGATTGTTGGCAGTCG
Draft genome

Slide 18 WGS: Two data analysis methods *(1)* wgMLST

- Analysis begins by generating a species library of known variants (categorized as alleles) using annotated reference genomes. For example, more than 200 complete genome sequences for *Listeria monocytogenes* (around 3000 genes in a single genome) generates a library of approximately 5000 genes. This reference library catalogues all the different variants seen within a particular gene and assigns them an allele number. Then, when a new *L. monocytogenes* genome is queried against that wgMLST library, that sequence can be assigned one of the allele types that were seen. With wgMLST, all of the genes present within a queried genome can be assigned an allele profile.

  NOTE: For a variant in a gene in a queried genome to be called as an allele, the reference library must contain that allele.

- WgMLST allows consistent nomenclature for all genomes queried using that reference library.

  NOTE: Nomenclature derived from one wgMLST library is not comparable to nomenclature derived from another wgMLST library.

- These alleles can also be used to generate phylogenetic trees and are especially useful by public health professionals in detecting outbreak clusters that should have nearly identical wgMLST profiles.
- Requires bioinformatics expertise and data analysis pipelines that are standardized and referenced.

Slide 19 WGS: Two data analysis methods (2) SNP

- Unlike wgMLST, a typical SNP-based approach does not usually incorporate insertions, deletions or copy number variants.
- Uses a single reference genome to describe all of the point substitutions between two or more genomes.

  NOTE: SNP assignments will change depending on the reference genome used.

- This is as specific as technology allows for comparing the similarity between two genomes (important for linking clinical to food or environmental isolates). Good for regulatory decision-making.
- Requires bioinformatics expertise and data analysis pipelines that are standardized and referenced.

  NOTE: Nomenclature can be assigned based on the phylogenetic results and when the same reference genome is used for all SNP assignments.

Slide 20 — Understanding the WGS data analysis methods

WgMLST gene-by-gene approach:

- WgMLST.
- Assess variations ('alleles') within each gene and assign allele number to different combinations seen.

  Build a library of SNFs, indels, re-arrangements from representative genomes. Combinations at the same loci are considered single evolutionary event.

NOTE: Each wgMLST scheme is specific to each library.

'Locus' (gene)	Strain 1	Strain 2	Strain 3
A	ACTAGAGGGAA Allele 1	ACTAGAGGCAA Allele 2	ACT GAGGGTAA Allele 3
B	TAGCCAGGGTC Allele 1	TAGCAAGGGTC Allele 2	TAGC GGTC Allele 3

SNP gene-by-gene approach:

- Represents a change from one nucleotide (e.g. ATGC) to another (e.g. TACG) at the same position when comparing two different genomic sequences.
- SNPs are called based on the reference genome that is used.
- Each SNP considered its own evolutionary event.
- SNPs retain information on location and nucleotide change within the genome since they are specific to the reference used.

..CTAGCTAG.....CTAGCTAG..
..CTAGCTAG.....CTAGCTTG..

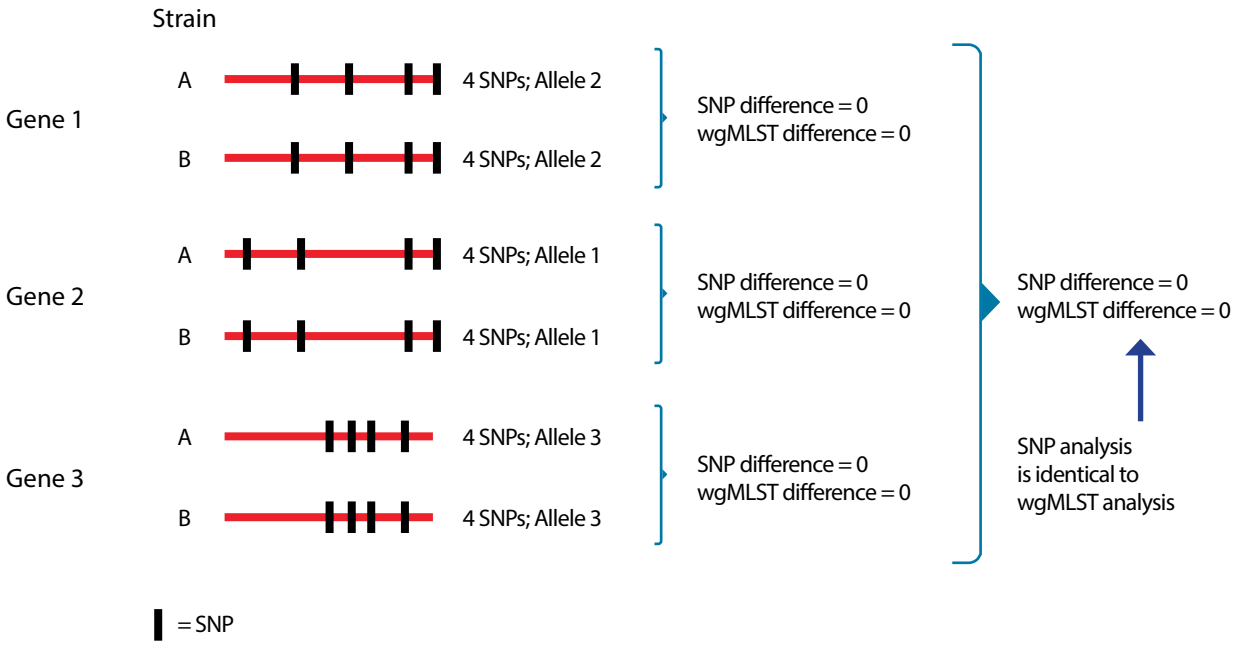
Slide 21 — Relationship between SNPs and wgMLST alleles

OVERVIEW:

Strain A and B being compared with each other at one locus. Strain A has one SNP at this locus (compared with a reference sequence) and Strain B has four SNPs. Using wgMLST you can assign the one SNP in Strain A as being Allele 1 and the four SNPs in Strain B as being Allele 2. Thus, when you compare the differences at this locus between Strain A and B using a SNP-based approach, there are three SNP differences (Strain A and B share one SNP in common) but using a wgMLST approach there is only one Allele difference (because Allele 1 (one SNP) is different than Allele 2 (that has four SNPs)).

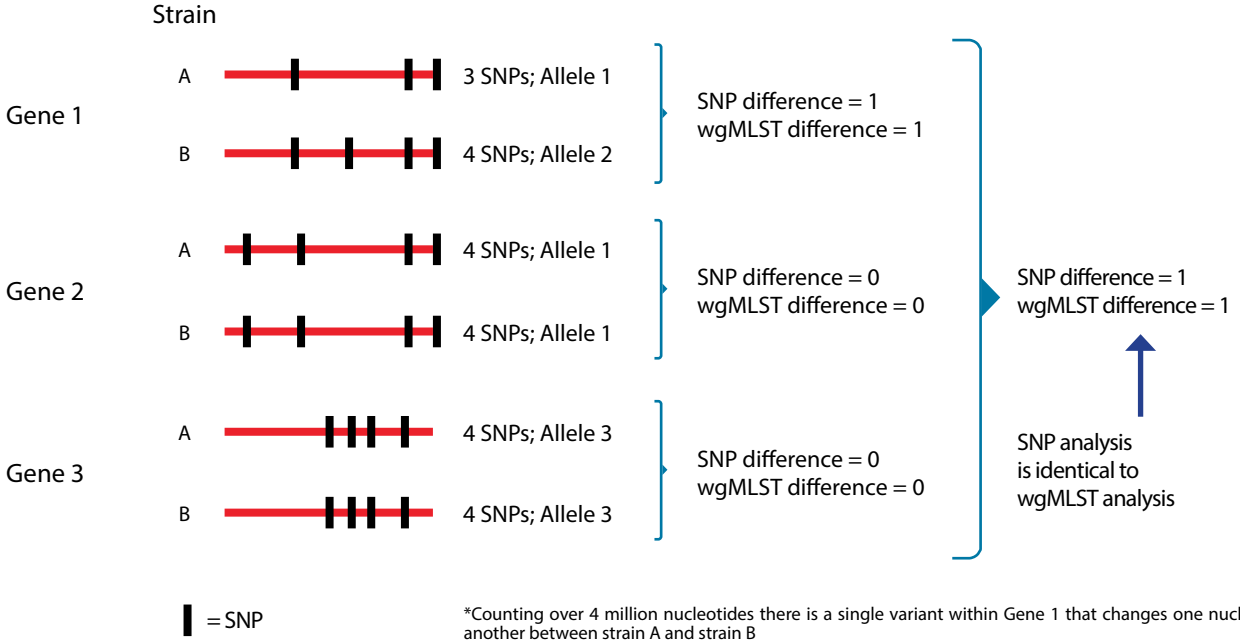
Slide 22

Comparison of SNPs and wgMLST alleles differences for identical strains A and B over three genes



Slide 23

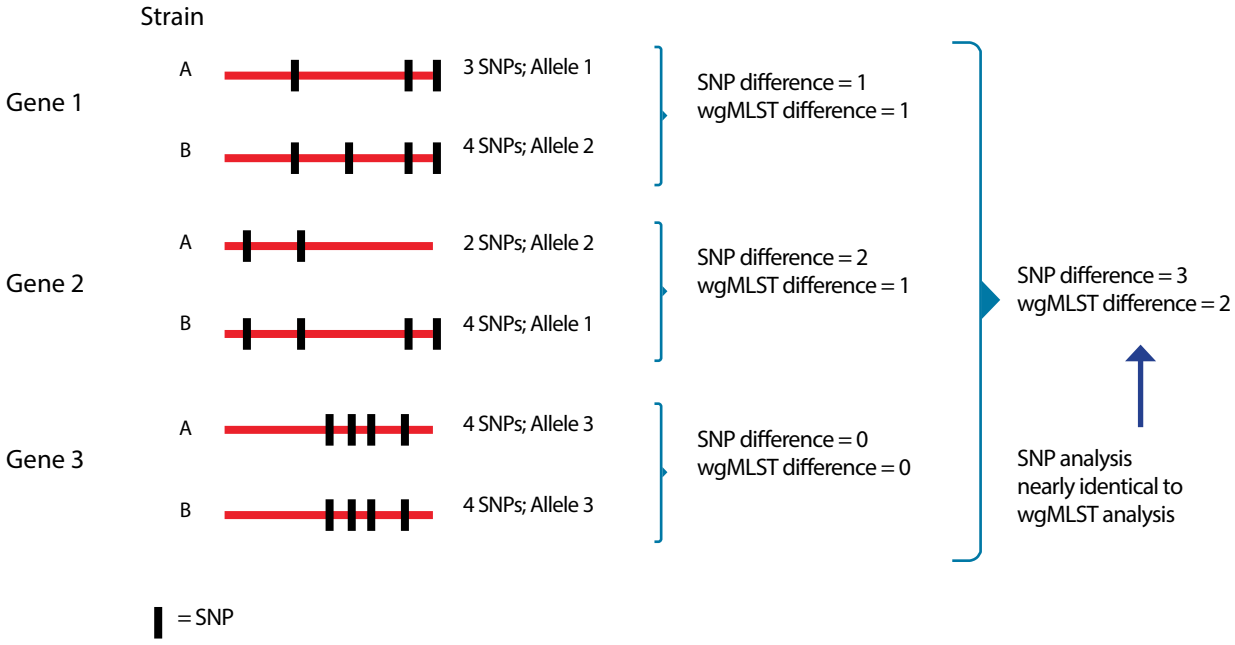
Comparison of SNPs and wgMLST alleles differences for very close* strains A and B over three genes



*Counting over 4 million nucleotides there is a single variant within Gene 1 that changes one nucleotide for another between strain A and strain B

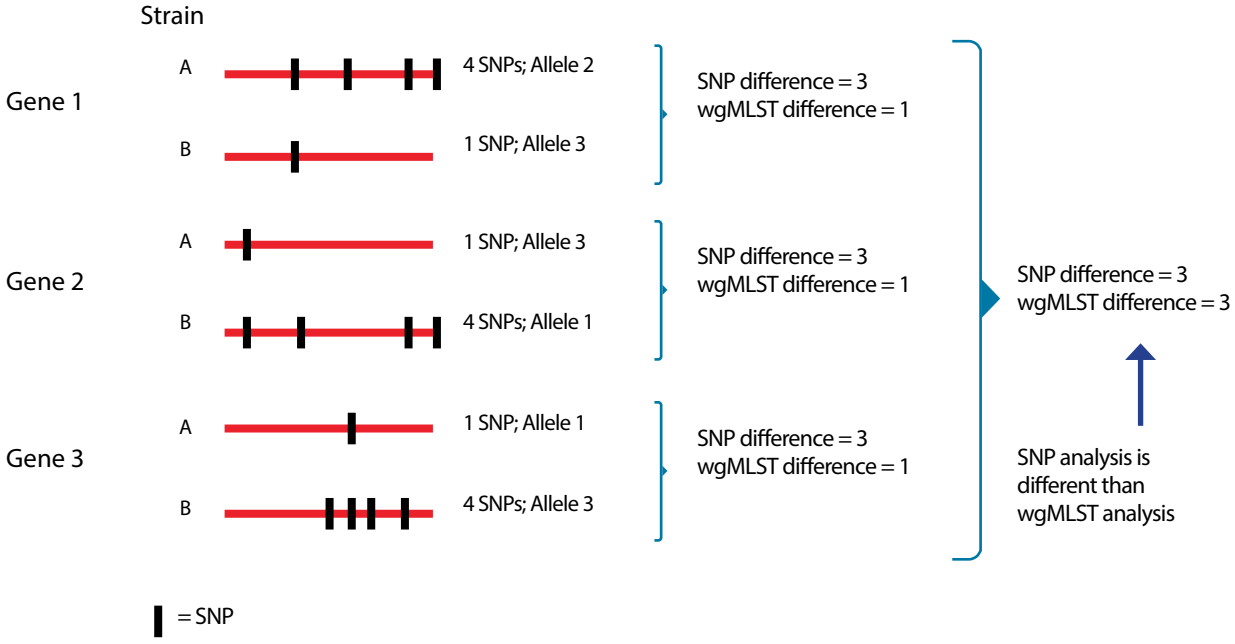
Slide 24

Comparison of SNPs and wgMLST alleles differences for very close* strains



Slide 25

Comparison of SNPs and wgMLST alleles differences for different strains



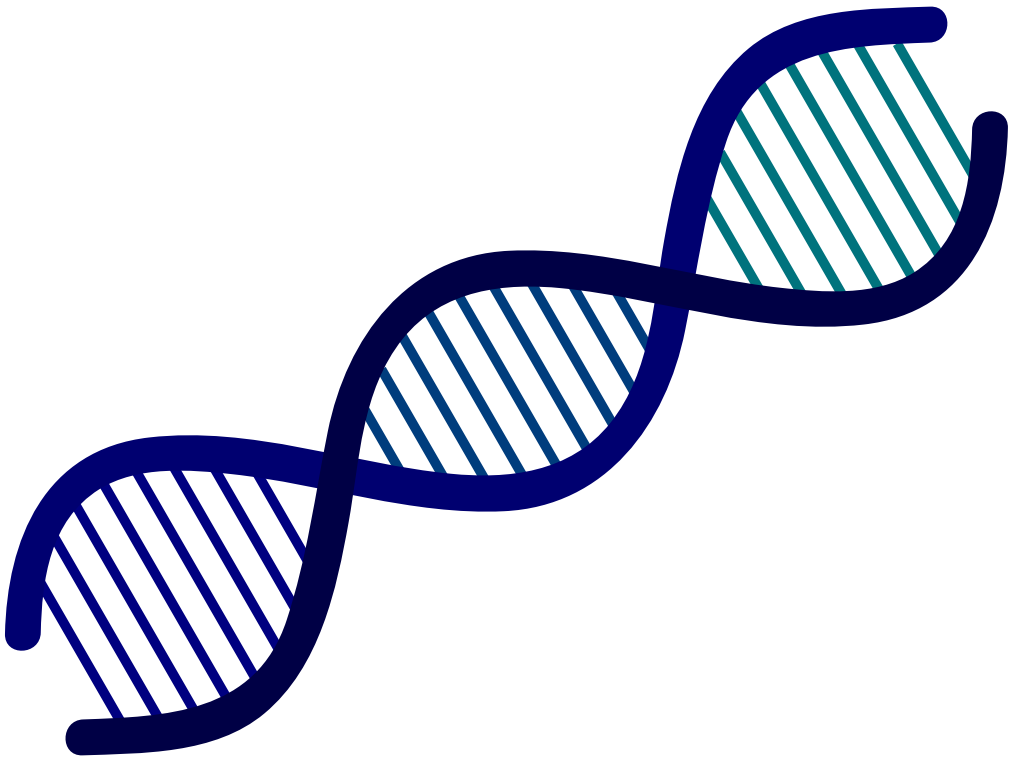
Slide 26 Comparability of SNPs and wgMLST

Closely related isolates

- Isolates that share a closely related ancestor (between days, weeks, or even years).
- In the context of foodborne disease surveillance, a single foodborne outbreak with a single cause will involve bacteria that are nearly identical to each other.
- This means that the number of mutations (labelled as variants or SNPs or different alleles) expected between comparing two or more clinical, food, or environmental isolates involved with this outbreak will be small.
- SNP analysis and wgMLST analysis will provide nearly identical results (though the number of allele or SNP differences might vary slightly).

Not closely related isolates

- Isolates do not share a closely related ancestor (many years could separate them).
- SNP analysis will provide many differences when comparing two strains, as will wgMLST.
- The number of differences might closely agree or might not but the number of total differences will be high enough (e.g. 100 SNP or allele differences) to indicate that the strains are unrelated within the context of surveillance or linking food to a clinical insolate.



Annex 1.2 Genetic relatedness and data comparability

Comparing two or more bacterial isolates (e.g. isolates cultured from different sources, as in clinical specimen A and clinical specimen B) in the context of foodborne disease surveillance requires an understanding of the term 'genetically closely related'.

In a foodborne outbreak, at some point a food item is contaminated by a foodborne pathogen. If the food item has not been properly treated to kill the pathogen, bacteria will multiply and grow. The food will be later ingested and make individuals ill. Some of these individuals will seek medical care, and a portion of them will provide faecal samples to detect and culture the bacteria that made them sick. Sometimes, it is even possible to collect some of the culprit food, and recover some bacteria which are, genetically, closely related to the bacteria isolated from those individuals who became ill.

Slide 16 (Part 1) illustrates an underlying mutation rate that may change the bacterial nucleotide sequence every time the cell divides. Any such mutation will then be passed down to cells with each cell division of that lineage. Because bacteria from the same outbreak are expected to be very closely related, that is, to have very few rounds of cell divisions between them, the genomes of bacteria obtained from patients and those from the food ingested will be very closely related. In consequence, very few genetic differences between those bacteria should be expected, such as mutations, variants, SNP, alleles, etc. Slide 12 describes a phylogenetic tree in which bacteria that are closely genetically related (i.e. they have very few or zero SNPs difference between them) are grouped together. This is how WGS links related clinical samples (or links bacteria from food samples to one or more clinical samples).

There are data analysis methods, also called workflows or pipelines, to assess the genetic relatedness of bacteria, and determine whether they are part of an outbreak (Slide 18). The first method is wgMLST. Using this approach, genomic sequences of multiple strains of a pathogen, such as *Listeria monocytogenes*, are compared, and the known variants for each gene are catalogued. Following, the different combinations of mutations are assigned as a different allele, or variant of a gene. For example, Allele A could represent a gene with no mutations, while Allele B, a gene with a SNP at position 2, and Allele C, could indicate multiple SNPs always found together within that gene. This approach allows similar bacteria to share similar alleles, and provides a mechanism for nomenclature (provided that the same wgMLST library is used for all comparisons). Just like using pulsed-field gel electrophoresis, nomenclature makes it easy to label similar isolates; it is also a tool for foodborne diseases surveillance, because similar clinical isolates will have alleles, and thus share a similar name. Also, when comparing two or more strains using wgMLST, isolates that are, genetically, more closely related will have fewer allele differences (Slide 22).

Slide 19 provides an overview of the second approach, using SNPs. Unlike wgMLST, in this case, there is no reference library to catalogue all mutations. Instead, the genomic sequences of two or more isolates are compared with find the SNPs. The results are described as the number of shared SNPs or the number of nucleotide differences between them. The premise is still that isolates that are part of the same outbreak will have nearly identical genomic sequences, and thus very few SNPs. Like wgMLST, an SNP-based approach allows for comparability between two or more genomes only if the same reference genome is used to make all SNP determinations. The approach is often used in regulatory settings when public health officials want to provide technologically supported evidence (along with epidemiological proof) linking an isolate from a specific food commodity to clinical isolates.

With regard to data comparability, often, public health officials involved in detecting growing clusters of clinical samples use wgMLST, while other colleagues might prefer the SNP-based approach, which makes it necessary to determine the comparability of both methods. Slides 18-26 above, refer to this issue, and the short answer is that the results of both methods consistent, that is, there is agreement between each other. For instance, the variants of a single gene across Strains A and B illustrated in Slide 21, show that A has Allele 1 at this gene (or one SNP), while strain B has Allele 2 (or four SNPs). In other words, there is one allele difference between Strains A and B using multi-locus sequence type (MLST) and three SNP differences. Nonetheless, usually, either the SNP-based approach is used along the entire genome or all the present genes. Slides 22-25 are meant to illustrate the fact that identical isolates (same genome) display no SNP or allele differences (Slide 22). Isolates that have a single nucleotide difference will also have identical SNP or allele differences (Slide 23), and when isolates are still closely related, the number of SNP differences will be very similar to the number of allele differences (Slide 24). However, when the strains are not closely related, there should be high numbers of both SNP and allele differences (Slide 25). The output, therefore, is comparable and consistent when closely related bacteria are compared with one another, and describing outbreak samples based on SNP and wgMLST results is basically interchangeable. The specific pipelines might differ, but the interpretation of the data is nearly identical for closely related isolates. However, different thresholds might need to be set, depending on whether SNP or wgMLST are being used (ten differences or five differences, for example), but is important to understand that, for public health purposes, the results are the same.

References

1. Pightling AW, Petronella N, Pagotto F. Choice of reference-guided sequence assembler and SNP caller for analysis of *Listeria monocytogenes* short-read sequence data greatly influences rates of error. BMC Res Notes. 2015;8:748. doi:10.1186/s13104-015-1689-4.
2. Olson ND, Lund SP, Colman RE, Foster JT, Sahl JW, Schupp JM et al. Best practices for evaluating single nucleotide variant calling methods for microbial genomics. Front Genet. 2015;6:235. doi:10.3389/fgene.2015.00235.
3. Davis S, Pettengill JB, Luo Y, Payne J, Shpuntoff A, Rand H et al. CFSAN SNP pipeline: an automated method for constructing SNP matrices from next-generation sequence data. PeerJ Comput Sci. 2015;1:e20. doi:10.7717/peerj-cs.20
4. Lynch T, Petkau A, Knox N, Graham M, Van Domselaar G. A primer on infectious disease bacterial genomics. Clin Microbiol Rev. 2016;29(4):881–913. doi:10.1128/CMR.00001-16.

Annex 2. Glossary

The following definitions to the terms as used in this manual, may be different from those of other documents.

Bioinformatics	The study of biological data, including computational and statistical methods.
Bioinformatics pipeline	The integration of several bioinformatics applications in an automated or semiautomated computational workflow to process biological data to produce the desired output.
Case	Any person who meets a case definition, either for surveillance purposes or during an outbreak investigation. The term case is used because: • case is a standard way to describe a person counted as part of an outbreak investigation; • not all suspect and confirmed cases are patients (patient refers to someone who receives medical care or is registered with a medical service); • people or persons does not cover all cases, however, in certain situations asymptomatic cases may be included, as well as people who are well at the time of the investigation but had symptoms in the past; and • case-finding and case definition are standard epidemiological terms.
Case definition	A set of criteria (not necessarily diagnostic) that must be fulfilled in order to identify a person as having a particular disease or condition. Case definitions can be based on time, geographical, clinical and/or laboratory criteria *(1)*.
Capacity	The ability of individuals, institutions and societies to perform functions, solve problems and set and achieve objectives in a sustainable manner *(2)*.
Core genome MLST (cgMLST)	A set of loci (e.g. genes) present in all strains undergoing analysis. In other words, for a locus to be included in a cgMLST, it must be present in more than 95% of the strains analysed. cgMLST loci usually represent well-conserved and necessary genes for a specific species.
Dry lab	Bioinformatic activities relating to data processing, analysis and management (e.g. the bioinformatics pipeline used to re-assemble the genome and the outputs of the bioinformatics analyses).
Event-based surveillance	The organized collection, monitoring, assessment and interpretation of unstructured information about health events that may represent a risk to public health *(3)*.
Food	Any substance, whether processed, semi-processed or raw, that is intended for human consumption; it includes drink, chewing gum and any substance used in the manufacture, preparation or treatment of food, but does not include cosmetics, tobacco or substances used only as drugs *(4)*.
Food chain	The series of processes that food goes through; it includes primary production (including feeds, agricultural practices and environmental conditions), product design and processing, transport, storage, distribution, marketing, preparation and consumption *(5)*.
Foodborne disease	Any disease of an infectious or toxic nature caused by the consumption of food *(6)*.
Foodborne disease outbreak	For common diseases (such as salmonellosis), the occurrence of two or more cases resulting from ingestion of the same food. For rare diseases, e.g. botulism, one case may be considered an outbreak *(6)*.
Food safety	Actions taken to ensure that food will not cause harm to the consumer when it is prepared and eaten according to its intended use *(7)*.

Food safety system	All the systems that are in place in the food safety sector for the control and management of food safety hazards.
Integrated food chain surveillance	The routine sharing of data and information between the public health, food safety and animal health sectors, in order to direct control measures to minimize the burden of foodborne diseases.
Laboratory-based surveillance	A form of indicator-based surveillance of cases that have been confirmed by a laboratory test. The laboratories that perform the testing report the results to the surveillance system, as well as informing the clinicians who requested the tests.
Notifiable disease	A disease that, because of its public health importance, must be reported to the public health authority under legislation or decree, in the pertinent jurisdiction when a diagnosis is made.
One Health	An approach to designing and implementing programmes, policies, legislation and research in which multiple sectors communicate and work together to achieve better public health outcomes *(8)*.
Public health action	Any action aimed at protecting public health (either through control or prevention measures) anywhere along the food chain. The term is also used in surveillance and outbreak response to describe any activities taken to control or prevent foodborne disease in humans. The actions can be taken anywhere along the food chain, either in response to an outbreak, or to target interventions to reduce contamination during production or distribution of food. These actions are typically undertaken by food safety authorities and/or animal and plant health sectors.
Public health authority	Refers to work unit(s) responsible for communicable disease surveillance, outbreak detection and outbreak investigation within the Ministry of Health at the national or sub national level. This includes surveillance officers, public health officers, outbreak response teams and epidemiologists.
Public health surveillance	The systematic continuous collection, collation and analysis of data for public health purposes and the timely dissemination of public health information for assessment and public health response as necessary *(9)*.
Response	Any public health action (e.g. event monitoring, providing information to the public, field investigations and control or mitigation measures) triggered by the detection of a public health risk *(10)*.
Sensitivity	The proportion of actual cases in a population detected and reported through the surveillance system *(11)*. A surveillance system with low sensitivity would not be able to detect all outbreaks when they occur in the community.
Sequence	The order of nucleotides, adenosine (A), thymine (T) (or uracil (U) in RNA), cytosine (C) and guanine (G) in a nucleic acid molecule (i.e. DNA or RNA). Sequencing is the method to determine the linear order of these nucleotide bases.
Sequencing	Is a general terms used to describe the process of detering the nucleotide order of a given DNA or RNA fragment. Throughout this manual, the term 'sequencing' refers to whole genome sequencing and the term 'sequence' refers to the 'whole genome sequence'.
SNP	Represents a change from one nucleotide (e.g. ATGC) to another (e.g. TACG) at the same position when comparing two different genomic sequences (e.g. bacteria A versus bacteria B) at the same position. For example, at position 1 in the genome (out of an average of 5 million nucleotides for a bacteria) a SNP would be present if Bacteria A had a G nucleotide, while Bacteria B had a C nucleotide.
Specificity	The proportion of persons without the disease in a population that are considered by the surveillance system as not having the disease. A surveillance system with low specificity would pick up false outbreaks, resulting in a waste of resources for their investigation *(11)*.
Subtyping	Determines the similarity between separate isolates of bacteria of the same species. If two bacteria have the same subtype, they are more likely to be related to each other than if they had different subtypes *(12)*. Subtyping methods are diverse, but a growing number are genetically based, such as MLST and WGS.
Surveillance and response system	The existing infrastructure, staff and processes used for surveillance of, and response to, communicable diseases.

Surveillance and response for foodborne diseases	Use of existing surveillance and response systems for foodborne diseases.
Syndrome	A group of clinical signs and symptoms that consistently occur together, or a condition characterized by a set of associated clinical signs and symptoms *(13)*.
Syndromic surveillance	The use of health-related data based on clinical observations rather than laboratory confirmation of diagnosis (e.g. influenza-like illness or acute watery diarrhoea). Such data can be used to signal sufficient probability of a case or outbreak to warrant further public health investigation and response *(14)*.
Timeliness	Refers to an acceptable length of time between steps in a public health surveillance system *(11)*. For example, the time between onset of illness and specimen collection, or between laboratory confirmation of disease and notification, needs to be within certain limits.
Wet lab	Laboratory processes up to, and including DNA sequencing. This includes isolating a pathogen, culturing it, extracting the DNA, library preparation and actual DNA sequencing (as opposed to dry lab techniques which are computer-performed (in silico).
WgMLST	In contrast to cgMLST schemes, wgMLST includes: 1) all closely conserved loci (e.g. all cgMLST loci) and 2) accessory loci that might define certain strains within a species and thus a much higher typing resolution can be obtained with wgMLST.
Whole genome sequencing	A laboratory technique that can be used to determine the complete nucleic acid sequence of an organism's genome (i.e. the order and number of the nucleotides, adenine (A), guanine (G), cytosine (C) and thymine (T), which make up an organism's genetic code).
Workflow	Step-by-step tasks involved in a specific process.

References

1. Porta M. A dictionary of epidemiology, 6th edition. New York: Oxford University Press; 2014.
2. Capacity development: a UNDP primer. New York: United Nations Development Programme; 2009 (https://www.adaptation-undp.org/resources/relevant-reports-and-publications/capacity-development-undp-primer, accessed 12 April 2022).
3. Early detection, assessment and response to acute public health events: implementation of early warning and response with a focus on event-based surveillance, interim version. Geneva: World Health Organization; 2014 (https://apps.who.int/iris/handle/10665/112667, accessed 12 April 2022).
4. Codex Alimentarius Commission: procedural manual, twenty-seventh edition. Rome: Food and Agriculture Organization of the United Nations and World Health Organization; 2019 (https://www.fao.org/3/ca2329en/CA2329EN.pdf, accessed 16 May 2022).
5. Principles and guidelines for the conduct of microbiological risk management (MRM): CAC/GL 63-2007. Rome: Food and Agriculture Organization of the United Nations and World Health Organization; 2007.
6. Foodborne disease outbreaks: guidelines for investigation and control. Geneva: World Health Organization; 2018 (https://apps.who.int/iris/handle/10665/43771, accessed 12 April 2022).
7. Assuring food safety and quality: guidelines for strengthening national food control systems. Rome: Food and Agriculture Organization of the United Nations; 2003 (Food and Nutrition Paper No. 76; https://www.fao.org/documents/card/en/c/92f82d38-5557-4ca1-b361-be14cd129db6/, accessed 11 April 2022).
8. One Health basics. Atlanta (GA): United States Centers of Disease Control and Prevention; 2022 (https://www.cdc.gov/onehealth/basics/index.html, accessed 16 May 2022).
9. Whole genome sequencing for foodborne disease surveillance: landscape paper. Geneva: World Health Organization; 2018 (https://apps.who.int/iris/handle/10665/272430, accessed 11 April 2022).
10. Strengthening national food control systems: guidelines to assess capacity building needs. Rome: Food and Agriculture Organization of the United Nations; 2006 (https://www.fao.org/3/a0601e/a0601e.pdf, accessed 12 April 2022).
11. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN et al. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep. 2001;50(RR-13)1–35.
12. Sandora TJ, Gerner-Smidt P, McAdam AJ. What's your subtype? The epidemiologic utility of bacterial whole-genome sequencing. Clin Chem. 2014;60(4):586–88. https://doi.org./10.1373/clinchem.2013.217141.
13. Rapid risk assessment of acute public health events. Geneva: World Health Organization; 2012 (https://apps.who.int/iris/handle/10665/70810, accessed 10 May 2022).
14. Public health for mass gatherings: key considerations. Geneva: World Health Organization; 2015 (https://apps.who.int/iris/handle/10665/162109, accessed 10 May 2022).

Annex 3. Assessment of existing capacity and capability

When a country is considering implementing WGS, it is necessary to assess the existing capacities and capabilities of the foodborne diseases surveillance and response system. If such capacities are insufficient, it is unlikely that sequencing will be useful, and it may hamper longer term inclusion of sequencing into the surveillance and response system.

If an assessment has already been undertaken based on the WHO publication "Strengthening surveillance and response for foodborne diseases" *(12)*, those results may be used. Otherwise, in order to conduct a capacity assessment, a team of relevant staff who are part of the surveillance and response system should be convened. These participants should include those who:

- conduct surveillance and outbreak investigation of foodborne diseases
- manage laboratories and perform laboratory tests for foodborne pathogens
- take control measures in the food safety system.

The team should go through the assessment tool that follows to determine whether each capacity exists (Table 3). The notes column can be used to indicate:

- if the capacity exists, and document what is in place, e.g. the names of focal points, laboratories and any pertinent food laws or regulations; and
- if the capacity does not exist or is only partially present, document what the gaps are (e.g. no event report form developed).

All gaps identified should be prioritized when moving towards implementing WGS for foodborne disease surveillance and response.

Table 3 Assessment of capacities needed for strengthening surveillance and response for foodborne diseases

Capacity	Does the capacity exist?			Notes
	Yes	Partially	No	
Epidemiological capacities				
A national focal point to receive reports about events				
An event report form which systematically captures information about events				
An event database to store information about reported events				
Health care workers and sanitary or food inspectors have been trained to recognize and report events				
The event-based surveillance system has successfully received reports of foodborne outbreaks over the past 12 months				
The event-based surveillance system has successfully received reports of foodborne outbreaks over the past 12 months				
At a minimum, there is a team at the national level who can rapidly assess events				
Events are assessed within 24 hours of the first report				
Rapid risk assessments of foodborne outbreaks have been conducted over the past 12 months				
Trained personnel have been designated in outbreak response teams				
The outbreak response teams can: • develop and apply a case definition • interview people who meet the case definition using a standardized questionnaire • describe cases using a line list • conduct a descriptive analysis by time, place and person				
At least one epidemiologist in the country who can conduct analytical epidemiology studies when required				
Outbreak investigations are summarized in reports				
Laboratory capacities				
Outbreak response teams have been trained to collect the appropriate clinical specimens during an outbreak investigation				
Clinical specimens are collected in foodborne disease outbreak investigations				
There is at least one designated laboratory to conduct the testing for priority foodborne pathogens				
Specimens can be transported to the designated laboratory under appropriate conditions				
The designated laboratory has experience in, and the resources to, culture and identify priority foodborne pathogens (e.g. *Salmonella* spp., *Shigella* spp., *Vibrio* spp., *Campylobacter* spp., Shiga-toxin producing *Escherichia coli*, *Listeria monocytogenes*, etc.)				

Laboratory capacities
There is a laboratory quality management system in place which includes: • regular performance of internal quality control • documentation of standard laboratory methods • participation in external quality assessment schemes
Food safety system capacities
Appropriate food and/or environmental samples are collected during foodborne disease outbreaks, where possible
The samples are tested at a laboratory that can culture and identify priority foodborne pathogens
There are food laws and regulations that enable control measured to be taken
Food/sanitary inspectors can take the necessary enforcement action to stop the sale and distribution of contaminated food

References

1. Strengthening surveillance of and response to foodborne diseases. Geneva: World Health Organization; 2017 (https://www.who.int/publications/i/item/strengthening-surveillance-of-and-response-to-foodborne-diseases, accessed 31 May 2022).

Annex 4.

List of participants at the WHO Meeting, Washington DC, USA, January 2017

Expert participants

Marc W. Allard
United States Food and Drug Administration
Washington DC, USA

Jennifer Beal
United States Food and Drug Administration
Washington DC, USA

Eric Brown
United States Food and Drug Administration
Washington DC, USA

Josefina Campos
Servicio Enterobacterias
Departamento de Bacteriología
Instituto Nacional de Enfermedades Infecciosas
Buenos Aires, Argentina

Hugo Castelan Sánchez
Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria
Mexico City, Mexico

Tom Chiller
United States Centers for Disease Control and Prevention
Atlanta, USA

Uday Dessai
United States Department of Agriculture
Washington DC, USA

Pilar Donado
Corporación Colombiana de Investigación Agropecuaria
Bogotá, Colombia

Peter Evans
United States Department of Agriculture
Washington DC, USA

Eelco Franz
National Institute for Public Health and the Environment
Bilthoven, Kingdom of the Netherlands

Jim Gallarda
Bill & Melinda Gates Foundation
Seattle, USA

Alejandra Garcia Molina
Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria
Mexico City, Mexico

Peter Gerner-Smidt
United States Centers for Disease Control and Prevention,
Atlanta, USA

Caterina Guzman Verri
Universidad Nacional
Heredia, Costa Rica

Rene S. Hendriksen
Technical University of Denmark
Lyngby, Denmark

Juan Carlos Hormazabal
Instituto de Salud Pública de Chile
Santiago, Chile

William Hsiao
British Columbia Centre for Disease Control
University of British Columbia
Vancouver, British Colombia, Canada

Claire Jenkins
Public Health England
London, United Kingdom of Great Britain and Northern Ireland

Ole Lund
Technical University of Denmark
Lyngby, Denmark
Rahman Mahmadur
Bangladesh

Celine A. Nadon
Public Health Agency of Canada
Ottawa, Canada

Eric Device Ng'eno
Kenya Medical Research Institute
Nairobi, Kenya

Stefano Pongolini
Instituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia-Romagna
Brescia, Italy

Valentina Rizzi
European Food Safety Authority
Parma, Italy

Jørgen Schlundt
Nanyang Technological University
Singapore, Singapore

Sharon Shea
Association of Public Health Laboratories
Silver Spring, USA

Nick Thomson
Wellcome Trust Sanger Institute
Cambridgeshire, United Kingdom of Great Britain and Northern Ireland

Ian Williams
United States Centers for Disease Control and Prevention
Atlanta, USA

FAO participants

Hilde Kruse
Food and Agriculture Organization of the United Nations
Rome, Italy

WHO participants

Amy Cawthorne
World Health Organization
Geneva, Switzerland

Jean Marc Gabastou
Pan American Health Organization
(WHO Regional Office for the Americas)
Washington DC, USA

Marcelo Galas
Pan American Health Organization
(WHO Regional Office for the Americas)
Washington DC, USA

Danilo Lo Fo Wong
WHO Regional Office for Europe
Copenhagen, Denmark

Jorge Matheu
World Health Organization
Geneva, Switzerland

Enriques Perez Gutierrez
Pan American Health Organization
(WHO Regional Office for the Americas)
Washington DC, USA

Nutrition and Food Safety Department
World Health Organization
20 avenue Appia
1211 Geneva
Switzerland

Email: nfs@who.int

Website: https://www.who.int/teams/nutrition-and-food-safety/about

9789240021228