Whole Genome Sequencing (WGS) for Measles

WHAT IS WHOLE GENOME SEQUENCING?

Whole genome sequencing (WGS) is a powerful laboratory tool that allows us to analyze genetic material from an organism to identify its unique fingerprint.

WGS is used in public health to compare the DNA or RNA "fingerprints" of pathogens, like viruses and bacteria, from people who are sick, to identify clusters of infections. Once a cluster is detected, the next step involves investigating whether the infections originated from the same source. Epidemiologists (disease detectives) ask about recent travel history of infected individuals, where they've been, and with whom they have come into contact. Combining information from genomic sequencing with epidemiological data allows us to track disease transmission more quickly and accurately.

The WGS process reveals the complete genetic makeup of a pathogen, as opposed to partial sequencing, which allows us to analyze and compare only part of a pathogen's genome. WGS enables us to better understand variations between and within species using precision that other technologies lack.

For measles outbreak investigations, WGS has become especially important. Measles was declared eliminated in the United States in 2000. In the years following elimination, U.S. cases primarily resulted from importations and distinct transmission chains, rather than sustained endemic spread. The measles genome is relatively stable, and viral evolution is driven by global transmission dynamics. In 1990, 18 wild-type genotypes circulated worldwide; since 2021, only two (D8 and B3) have been detected globally. This sharp decline in genetic diversity reflects the long-term success of global vaccination efforts.

Historically, public health laboratories relied on sequencing a small portion of the measles virus genome known as N450. For many years, this targeted approach was sufficient to distinguish between different strains and track transmission chains. However, with only two genotypes and limited variation in the measles genome, distinguishing ongoing domestic transmission from repeated introductions of the same strain is challenging. When nearly every case looks identical using partial sequencing, it becomes extremely difficult to determine whether outbreaks are connected through domestic spread or represent separate international importations. The traditional method of only looking at N450 no longer provides enough detail to answer these critical questions.

That is where the utility of WGS becomes apparent. WGS offers more detail than partial sequencing that can help rule out links between cases in this context. Even when two virus strains appear genetically uniform based on a small segment, tiny differences may exist elsewhere in the genome. WGS can detect these subtle variations, offering the additional detail needed to distinguish between cases that seem identical. Note: WGS alone can help us know with certainty that two cases are **not** related, but it cannot prove that two cases are linked.

Understanding if measles cases are driven by international importation or by local transmission can help inform strategies for measles prevention and control.

When combined with epidemiologic data, information gleaned from whole genome sequencing may help us to understand how measles outbreaks are related and inform measles prevention and control efforts.

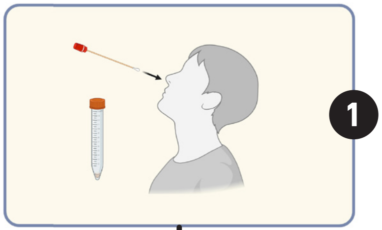
Today, advances in technology allow laboratories to sequence the entire measles genome quickly and at relatively low cost—capabilities that were not widely available when measles elimination was first achieved.

When coupled with epidemiological data, WGS offers insights into connections between cases for infectious diseases, allowing us to better understand the source and track the spread of specific pathogens, especially when there is little genetic variation among cases, as with measles.

2025 U.S. MEASLES OUTBREAK & WGS

- **Of the measles outbreaks reported in 2025,** there were four large outbreaks, each with more than 50 linked cases.
- **These four outbreaks accounted for 73% of all 2025 cases.**
  - Despite field investigations, the first cases in these outbreaks could not be connected to a known domestic or international source of infection and therefore the outbreaks cannot be epidemiologically linked.
  - Because there are now fewer distinct measles N450 genotypes circulating, we need broader genetic sequencing coupled with epidemiologic data to determine if there is sufficient information to tell outbreaks apart and to link with previous and current outbreaks elsewhere.
- **While we continue to work with state, tribal, local, and territorial public health agencies** to gather additional epidemiologic data in responding to measles cases, we are also collecting specimens for WGS.
- **The power of using WGS to address measles elimination status** lies in our ability to analyze how the virus changes genetically over time. This information can be used for comparison of measles strains found circulating in a specific region to others found in different geographic regions. If the strains are nearly identical genetically, then this could support transmission linkages (i.e., that importation may have introduced the strain). Conversely, identification of unique mutations in isolates from one region, distinct from what is circulating in other regions, could indicate that measles has been circulating, regardless of whether it has been previously clinically detected. However, support for transmission linkage using WGS alone is insufficient without epidemiological data to support this finding.
  - WGS sequencing data could be useful in determining if there are any genomic features, beyond the short region used for genotyping, that can rule out linkages among cases.
  - WGS is being applied to specimens collected since early 2025 to address this genetic linkage question.
  - This WGS data will be combined with epidemiological data to complete a comprehensive analysis of national data from 2025 onward.
  - Moving forward, use of WGS may provide valuable information to track overall genetic diversity among circulating Measles strains in addition to permitting focal regions for molecular diagnostics.
- **Once this comprehensive analysis is complete,** we can begin to support the annual process to verify elimination.

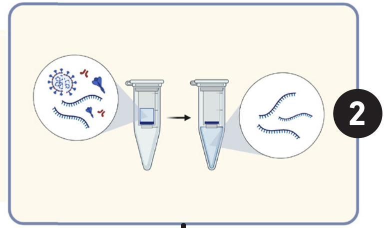
The Whole Genome Sequencing (WGS) Process



1

Specimen Collection

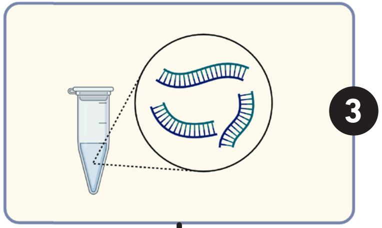
Specimen from suspected measles case is obtained.



2

Nucleic Acid Extraction

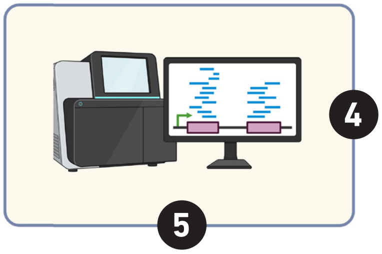
Scientists purify viral nucleic acids from the patient specimen.



3

Library Preparation

Scientists fragment the specimen RNA and further process the RNA in a mixed pool. The pool of fragments generated is called a “library”, which is then sequenced.



4

NGS Sequencing

Scientists load the library onto a sequencer, collecting millions of short “reads”.

5

Sequencing Analysis

Once the reads are collected, specialized computer programs are used to put them together in the correct order like pieces of a jigsaw puzzle. When completed, assembled genome sequences containing thousands of nucleotides are ready for further analyses.