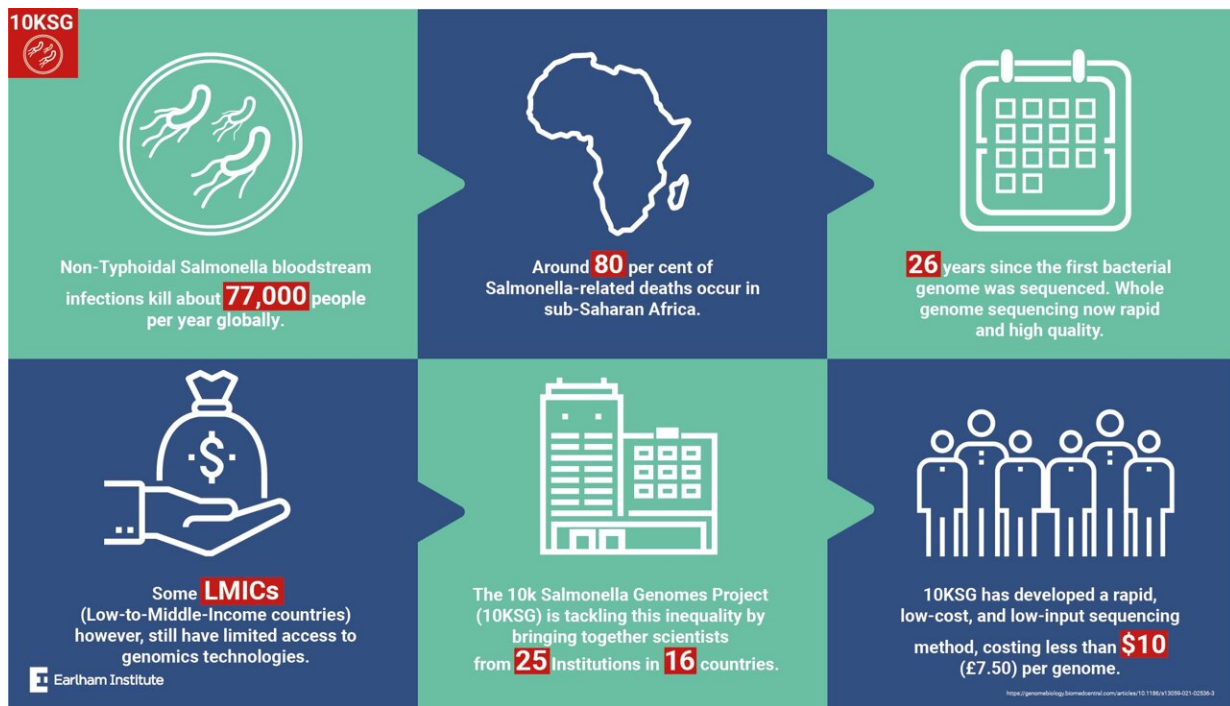


Affordable genome sequencing for pathogen analysis to help tackle global epidemics

December 20 2021



New bacterial pathogen pipeline makes whole genome sequencing accessible for public health scientists worldwide. Credit: Earlham Institute

A worldwide consortium of scientists, led by the Earlham Institute and the University of Liverpool in the UK, mark a significant milestone in equipping researchers in low- and middle-income countries (LMICs) with cheap and accessible methods for sequencing large collections of bacterial pathogens—at a cost of less than \$10USD per genome.

At a time when global genomic surveillance of coronavirus has been in the spotlight, the ability of countries to contribute through low-cost and rapid whole genome sequencing (WGS) has become increasingly important. The methods published in *Genome Biology* can be applied to large collections of bacterial pathogens and will strengthen global research collaborations to tackle future pandemics.

Over the past decade, WGS has revolutionized understanding of microbial epidemics. WGS data can be used for surveillance, functional genomics and the exploration of pathogen evolution, prompting both [public health](#) and research scientists to adopt genome-based approaches.

The genome sequencing of thousands of microorganisms has remained expensive—largely due to costs associated with sample transportation and the construction of DNA libraries—while the need to genome sequence collections of key pathogens has grown substantially in recent years.

Until now, large-scale bacterial genome projects could only be performed in a handful of sequencing centers around the world. With this study, the team of scientists have managed to make this technology accessible to laboratories worldwide.

"It has been 26 years since the first bacterial genome was sequenced, and it is now possible to sequence bacterial isolates at scale. However, access to this game-changing technology for scientists in low- and [middle-income countries](#) has remained restricted, said study author and Director of the Earlham Institute Prof Neil Hall.

"The need to 'democratise' the field of pathogen genomic analysis prompted us to develop a new strategy to sequence thousands of bacterial isolates with collaborators based in many economically-challenged countries."

10k Salmonella strains

Focusing on the organism *Salmonella enterica*, a pathogen with a global significance that causes infection and [deadly disease](#), this large-scale genomic sequencing initiative was led by the worldwide [10,000 Salmonella genomes research consortium \(10KSG\)](#) with scientists from 16 countries.

The objectives of 10KSG are to make genomic data more accessible to low and middle income countries, especially because mortality rates for Salmonella in sub-Saharan Africa are exceptionally high. Understanding the genetic makeup of significant collections of such bacteria strains was imperative, and the project sequenced and analyzed 10,000 Salmonella genomes from Africa and Latin America.

The researcher's innovative WGS approach aimed to streamline the large-scale acquisition and genome sequencing of bacteria, and amassed the genetic material of more than 10,400 clinical and environmental bacterial isolates from LMICs in under a year.

The sample logistics pipeline, developed by the University of Liverpool, was optimized by shipping the heat-inactivated bacterial isolates as 'thermolysates' in ambient conditions from across the world to the UK. Subsequently, isolates were sequenced at the Earlham Institute using the unique LITE protocol—a low cost, low input automated method for rapid genome sequencing. In total, the gene library construction and DNA sequencing bioinformatic analysis was done with a total reagent cost of less than USD\$10 (around £7.50GBP) per genome.

Prof of Microbial Pathogenesis and study author Jay Hinton from the University of Liverpool, said: "One of the most significant challenges facing public health researchers in LMI countries is access to state-of-the-art technology. For a combination of logistical and economic

reasons, the regions associated with the greatest burden of severe bacterial disease have not benefited from widespread availability of WGS. The 10,000 Salmonella genomes project was designed to begin to address this inequality."

Dr. Blanca Perez Sepulveda, Postdoctoral Research Associate and study author from the University of Liverpool, who led the global sample collection, optimisation and analysis, added: "The adoption of large-scale genome sequencing and analysis of bacterial pathogens will be an enormous asset to public health and surveillance in LMI countries. Here, we have established an efficient and relatively inexpensive pipeline for the worldwide collection and sequencing of bacterial genomes."

Deadly bacteria

Non-typhoidal Salmonella (NTS) have been widely associated with enterocolitis in humans, a zoonotic disease that is linked to the industrialisation of food production. Due to the scale of human cases of enterocolitis and concerns related to food safety, more genome sequences have been generated for Salmonella than any other genus.

In recent years, new lineages of NTS serovars Typhimurium and Enteritidis have been recognized as common causes of invasive bloodstream infections (iNTS disease), responsible for about 77,000 deaths per year worldwide.

Approximately 80 percent of deaths due to iNTS disease occur in sub-Saharan Africa. The new Salmonella lineages responsible for bloodstream infections can be identified by genomics, due to gene degradation, altered prophage repertoires and novel multidrug resistant plasmids.

Prof Neil Hall, added: "The number of publicly-available sequenced

Salmonella genomes reached 350,000 in 2021 and are available from several online repositories. However, limited genome-based surveillance of Salmonella infections has been done in LMI countries, and the existing dataset did not accurately represent the Salmonella pathogens that are currently causing disease across the world."

Global pipeline

Dr. Darren Heavens, Postdoctoral Scientist at the Earlham Institute, who developed the whole-genome sequencing pipeline, said: "We saw the need to simplify and expand [genome](#)-based surveillance of salmonellae from Africa and other parts of the world, involving isolates associated with invasive disease and gastroenteritis in humans, and extending to bacteria derived from animals and the environment.

"Our pipeline represents a cost-effective and robust tool for generating bacterial genomic data from LMI countries, to allow investigation of the epidemiology, drug resistance and virulence factors of isolates."

Development of the global [10KSG](#) consortium that involved collaborators from 25 institutions, research and reference laboratories across 16 countries. Members of the 10KSG provided access to 10,419 bacterial isolates sourced from 51 LMICs and regions—covering seven bacterial genera: Acinetobacter, Enterobacter, Klebsiella, Pseudomonas, Shigella, and Staphylococcus—coordinating the sample collection and transport of materials to be sequenced in the UK.

"Limited funding resources led us to design a genomic approach that ensured accurate sample tracking and captured comprehensive metadata for individual bacterial isolates while keeping costs to a minimum for the Consortium," said Prof Hall. "The pipeline streamlined the large-scale collection and sequencing of samples from LMICs. A key driver was to facilitate access to WGS and allow a worldwide collaborative

effort to generate a remarkably informative and robust set of genomic data."

The paper "An accessible, efficient and global approach for the large-scale sequencing of bacterial genomes" is published in *Genome Biology*.

More information: An accessible, efficient and global approach for the large-scale sequencing of bacterial genomes, *Genome Biology* (2021). [DOI: 10.1186/s13059-021-02536-3](https://doi.org/10.1186/s13059-021-02536-3)

Provided by Earlham Institute

Citation: Affordable genome sequencing for pathogen analysis to help tackle global epidemics (2021, December 20) retrieved 26 April 2024 from <https://phys.org/news/2021-12-genome-sequencing-pathogen-analysis-tackle.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.