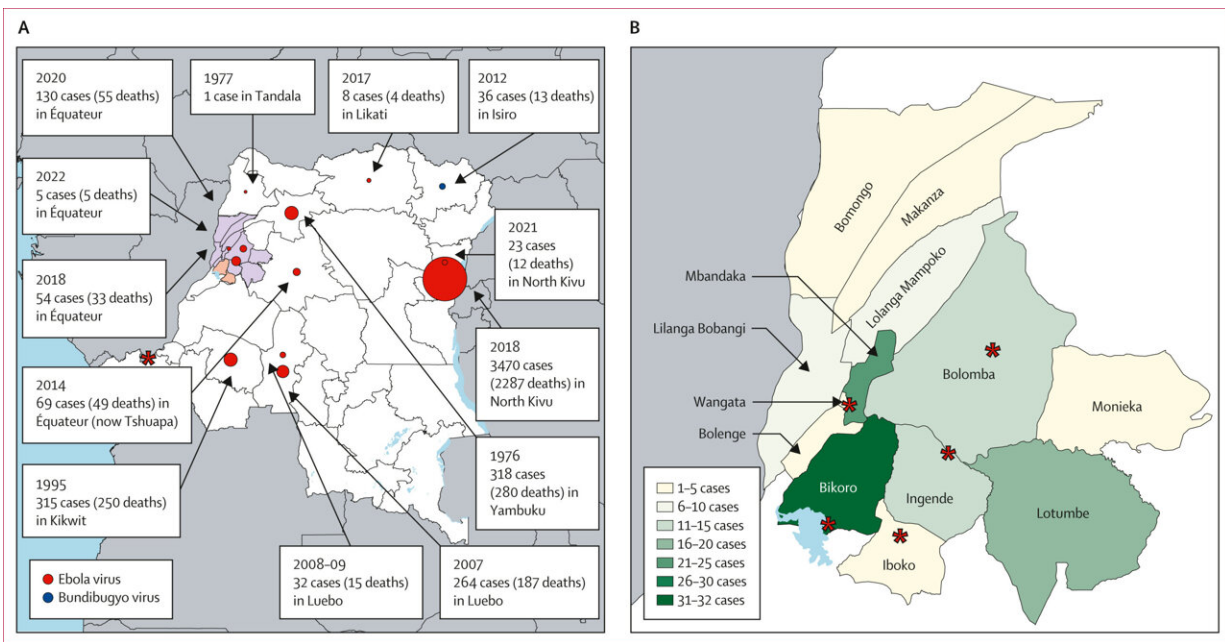# Retrospective genomic characterization of the 2020 Ebola outbreak

April 19 2024, by Paulette Campbell

Ebola virus outbreaks in the Democratic Republic of the Congo, 1976–2022. Credit: *The Lancet Microbe* (2024). DOI: 10.1016/S2666-5247(23)00259-8

Epidemiology researchers are harnessing tools developed by the Johns Hopkins Applied Physics Laboratory (APL) in Laurel, Maryland, to untangle the origins of the 2020 Ebola virus disease outbreak in the

Democratic Republic of Congo (DRC). An article in *The Lancet Microbe* explores how advanced tools developed by APL are revolutionizing infectious disease outbreak response and research.

Over the last four years, APL developed significant and enhanced capabilities for an application called ChainChecker, originally created by researchers at the Centers for Disease Control and Prevention (CDC) and Imperial College London, to help visualize and assess chains of transmission during disease outbreaks like Ebola.

ChainChecker supports outbreak investigations and response by combining epidemiological and genomic sequencing evidence to enable, as noted in the paper, investigative fact-checking and verify patient-reported sources of possible zoonotic (diseases that spread between humans and animals) spillover.

APL has incorporated genomic sequencing data into the tool and enabled it to rapidly upload and visualize data, identify hospital-acquired infections, and operate in internet-sparse environments. Because of these enhancements, ChainChecker has established itself as a novel tool in epidemiology.

Most recently, APL researchers expanded its functionality, improving its ability to assist responders in future outbreaks by applying the algorithms to the Ebola virus as well as other diseases, and enhancing the investigation workflow through data quality assessment and anomaly detection.

"The outputs of the tool help inform responders in closer to real time on aspects of the outbreak, such as novel transmission modes, differences between variants of the pathogen, and better infection prevention and

control measure selection," explained Miles Stewart, a project lead and [software engineer](#) in APL's Global Health Mission Area and one of three APL co-authors on the paper, along with Julia Eng and Abraham Rajan.

"In the middle of a response, these types of analysis can be labor intensive and complicated to do on top of other efforts to stop the outbreak, sometimes leading to critical insights being discovered after the response has concluded."

The DRC has faced 15 Ebola outbreaks since 1976. The 11th outbreak, declared in the country's western Équateur Province in June 2020, occurred at the same time as an outbreak in the DRC's eastern North Kivu province. Scientists wanted to understand how these and concurrent outbreaks were connected to the eighth Ebola outbreak—in 2018, also in Équateur Province—by combining epidemiological and genomic data.

"Gaining an understanding of the connections between the Ebola outbreaks in the Democratic Republic of Congo presents significant challenges due to the vast geographical spread across the country, the genetic diversity of the Ebola virus strains involved and the complex sociopolitical landscape," Stewart said.

"The simultaneous occurrence of outbreaks in distant provinces complicates tracing the chains of transmission and understanding any potential connection," he added. "And limited access to some affected areas, due to [security concerns](#) and logistical constraints, hinders comprehensive data collection, sampling and genetic sequencing efforts necessary to map the spread and evolution of the virus accurately."

Visualizing disease transmission

ChainChecker played a crucial role in visualizing and confirming the chains of Ebola transmission in the DRC and aided in the understanding

and analysis of the Ebola outbreak discussed in the research paper, said Stewart.

The researchers studied various features of the 11th outbreak, such as details about the patients, how the virus was spreading, and where and when it was spreading in the Équateur Province. They also collected information about people who might have the virus, those who probably had it, and those who definitely had it. This included information like patient age, possible ways they were exposed to the virus, the symptoms they experienced, and the date of symptom onset.

Between May and September 2020, there were 130 cases of Ebola in the Équateur Province, with the first reported case being an individual who frequently ate bat meat but had no contact with sick individuals, suggesting the outbreak was due to zoonotic spillover. Genetic analysis showed two subtypes of the virus, one present in the majority of cases and another that showed genetic similarities similar to a previous outbreak.

Analysis revealed that the second virus subtype mutated at a slower rate and may have originated in a survivor of the disease. Combining genetics and information on how the virus spreads helped confirm the transmissions, helped identify other potential, previously undiscovered connections, and showed how useful this approach is for understanding transmission and disease dynamics.

Since the conclusion of this outbreak, the APL team has continued to work with the CDC to apply ChainChecker to emerging Ebola outbreaks and adapt the tool to other viral hemorrhagic fevers such as Marburg virus disease, which has a very high case fatality rate. ChainChecker is also being used in coordination with international ministries of health and alongside existing contact-tracing tools.

"Our advanced tools are improving the ability of responders to make better decisions in the field during outbreak response through rapid, more accurate interpretation of data, and this paper showcases their transformative impact," said Stewart.

"From data analysis to interdisciplinary collaboration, APL's capabilities are revolutionizing the way we approach complex biological challenges."

More information: Eddy Kinganda-Lusamaki et al, 2020 Ebola virus disease outbreak in Équateur Province, Democratic Republic of the Congo: a retrospective genomic characterisation, *The Lancet Microbe* (2024). DOI: 10.1016/S2666-5247(23)00259-8

Provided by Johns Hopkins University