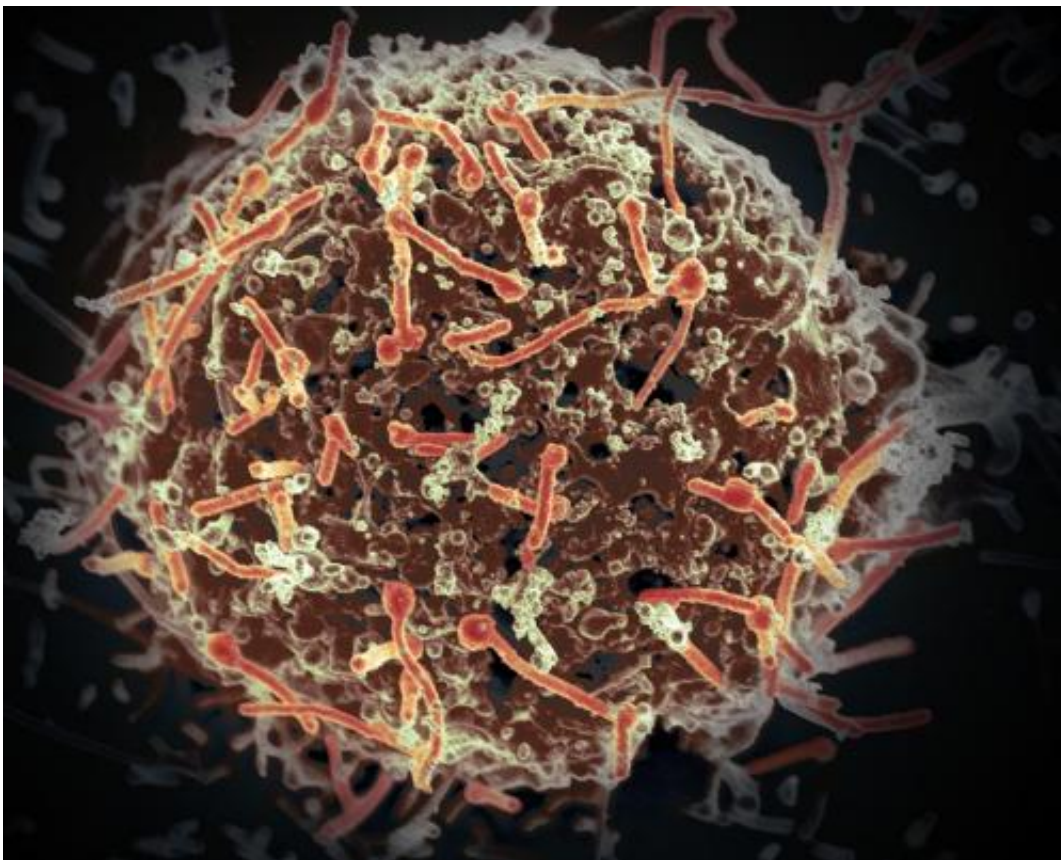


# **Ebola: Scientists show how real-time sequencing and data-sharing can help stop the next outbreak**

April 12 2017

---



The Ebola virus, isolated in November 2014 from patient blood samples obtained in Mali. The virus was isolated on Vero cells in a BSL-4 suite at Rocky Mountain Laboratories. Credit: NIAID

An international effort to analyze the entire database of Ebola virus genomes from the 2013-2016 West African epidemic reveals insights into factors that sped or slowed the rampage and calls for using real-time sequencing and data-sharing to contain future viral disease outbreaks.

Published today in the journal *Nature*, the analysis found that the epidemic unfolded in small, overlapping outbreaks with surprisingly few infected travelers sparking new outbreaks elsewhere, each case representing a missed opportunity to break the transmission chain and end the epidemic sooner.

"We calculated that 3.6 percent of cases traveled, basically meaning that if you were able to focus on those mobile cases and reduce their mobility, you might have had a disproportionate effect on the epidemic," said computational biologist Dr. Gytis Dudas, a Mahan Postdoctoral Fellow at Fred Hutchinson Cancer Research Center and the paper's lead author.

The West African Ebola epidemic dwarfed all previous central African outbreaks of the virus, sickening more than 28,000 people and killing more than 11,000 of them.

The 1,610 Ebola virus genomes analyzed by the researchers represented more than 5 percent of the known cases, the largest sample analyzed for a single human epidemic. The analysis is the first to look at how Ebola spread, proliferated and declined across all three countries most affected: Guinea, Sierra Leone and Liberia. Previous analyses used fewer sequences or focused primarily on either a single country or a limited time frame.

The new paper also amounts to a manifesto for collaborative science, with 93 scientists from 53 institutions in 16 countries listed as authors. Many of them had worked on earlier papers as clinicians gathering blood

samples, researchers doing genome sequencing or analysts drawing on portions of the dataset. Dudas and senior author Dr. Andrew Rambaut of the Institute of Evolutionary Biology at Scotland's University of Edinburgh were involved in the analyses for many of these efforts.

The authors' intention, they wrote, was for this comprehensive analysis to "provide a framework for predicting the behavior of future outbreaks for Ebola virus" and other human pathogens and to guide targeted, life-saving responses.

## **Cities aided virus' spread, distance slowed it**

The new analysis assessed 25 factors that could have contributed to the spread and duration of the West African epidemic. It confirmed the common perception that cities played a major role in the magnitude of the epidemic compared to central African outbreaks that had occurred in remote, sparsely populated regions.

Distance between cities also played a role, with the shorter the distance, the more likely that infected travelers would arrive and seed an infection. Distance was key to sparing nearby Guinea-Bissau, Senegal, Mali, Cote d'Ivoire and northern Guinea from severe and protracted epidemics. Some of these regions had large cities in which Ebola would likely have exploded had the virus been introduced.

"Essentially, it was entirely down to chance that the outbreak didn't spread further and cause an even bigger crisis," Dudas said.

Other variables, such as shared languages, economic output and climate were not found to be significantly associated with speeding or slowing the epidemic.

The analysis did see correlations between border closure dates and virus

traffic reduction; once the borders were closed, virus movement occurred mostly within countries rather than among them.

But by the time Sierra Leone, Liberia and Guinea closed their borders, cross-border travel had already seeded outbreaks in each country. And although international traffic of viruses was reduced after the closures, it didn't stop completely.

"That was part of the problem in Sierra Leone and Guinea in the final stages of the epidemic, where a particularly mobile chain [of infected people] was moving back and forth between the countries," Dudas said.

## **What the genome knows**

In previous genome analyses, scientists traced the epidemic's origin to December 2013, when a two-year-old who had been playing near a bat-filled tree died in a small village in the southeastern part of Guinea. It took until March 2014 for hospital workers to detect and report the spread of a disease with an unusually high death rate. By later that month it was identified as Ebola.

Bats are the suspected—but not proven—reservoir for Ebola virus. (A reservoir refers to an animal that harbors a virus, allowing the virus to live and multiply between outbreaks in humans.) The virus spreads to humans and then from person to person through direct contact.

Sequencing virus genomes from even a fraction of people infected in an epidemic and comparing mutation patterns can give researchers valuable information about how big the epidemic is, how long it has been spreading and where transmissions chains start and end, said Dr. Trevor Bedford, a Fred Hutch evolutionary biologist and one of the paper's authors.

Some of this information can be and was obtained the old-fashioned way by public health workers going door-to-door, tracing contacts of those infected. But in West Africa and other areas where public health resources and even basic infrastructure are limited, real-time genome sequencing and analysis can speed the response by telling health officials where to initiate contact tracing, put beds on the ground, quarantine those infected and implement other infection controls.

It can also provide information unavailable by other methods. In earlier studies, for example, genome sequencing had confirmed that Ebola deaths in Sierra Leone and Liberia came from Guinea and were not a new introduction from the virus' natural reservoir.

"Genome sequencing can tell you epidemiologically relevant things that are unobtainable by traditional methods," Bedford said.

And synthesizing that information with data on population size, travel distances, geography, language and other factors can provide context for which factors influenced the epidemic's spread and duration—and where to target treatment and interventions.

## **Technology—and data-sharing—advances**

Virus genome analysis has played a bigger role in understanding the West African Ebola epidemic than for any other infectious disease outbreak for two reasons: modern advances in sequencing technologies and scientists who were unusually willing to share data.

So-called "next-generation" sequencing equipment have dramatically lowered both the costs and time to prepare samples and do sequencing, making it much easier to sequence an entire viral genome.

And scientists relatively early on in the epidemic decided to share viral

genome sequences they were collecting from patients rather than waiting until they published a research paper. Publications are considered the currency of science, but the admonition to "publish or perish" took on a new meaning in the midst of such a devastating outbreak.

"The old model where you hold onto the sequences until the publication comes out, which as any academic knows is going to be months, is morally wrong if those sequences can be used to affect a response on the ground," Dudas said.

Early posting of data on the public database GenBank led to a surge of collaboration from experts in diverse fields. It was when one of those early researchers, Harvard University's Dr. Pardis C. Sabeti (another author on the new paper) and her team sequenced 99 Ebola genomes from patients in Sierra Leone and uploaded their data that Rambaut became involved. Bedford had been a postdoc in Rambaut's lab before coming to the Hutch in 2013, six months before the Ebola epidemic started. Dudas, who is now a postdoc in Bedford's lab, did his doctorate research under Rambaut.

Analysts like Dudas, Rambaut and Bedford bring to data a good understanding of how evolution works, along with strong critical thinking skills and an ability to spot trends and anomalies.

"Andrew is wonderfully curious," Bedford said of Rambaut. "The community working on Ebola was really lucky that he was involved in all of this. And Gytis is faster than anyone I can think of. If Andrew or I think of a question and ask Gytis about it, he'll have some beautiful figure to show us an hour later."

## **Next steps: more speed, more data-sharing**

Because speed is critical in an outbreak, Bedford wants to make the



analysis process even faster, just as new technologies have sped up sequencing.

"We'd like basically to make some of the core analyses in this paper, and the animation [that Dudas made of the virus' spread] something that can just happen," he said.

To this end, he and a longtime collaborator, Dr. Richard Neher of the University of Basel in Switzerland, have designed a tool called nextstrain to analyze and track genetic mutations during outbreaks. Anyone can download the source-code from GitHub, run genetic-sequencing data for the outbreak they are following through the pipeline and build a web page showing a phylogenetic tree, or genetic history, of the outbreak. The innovation recently won the first-ever international Open Science Prize.

"If there is a next [epidemic], and there is faster data-sharing," Bedford said, "you can have analyses out the door very quickly."

But the challenge of getting scientists to share data remains. Despite the precedent set by the response to the Ebola epidemic, Bedford and Dudas point out that fewer researchers have shared Zika virus genomes from the more recent crisis in Brazil, Central America and the Caribbean. In part, they said, that may be because the Zika virus is more difficult to sequence than Ebola, making researchers more prone to guard their rare sequences for publication.

The Ebola epidemic began while Dudas was working on his Ph.D., and the data-sharing that ensued impressed the young researcher deeply. "My standards for what collaboration is supposed to look like have been set pretty high," he said.

"There are still some people who think that [genome sequencing](#) is

effectively stamp collecting," Dudas continued. "You might collect samples, and you might sequence them and look in retrospect at the outbreak. But all the sequencing that's been done leading up to this publication was essentially being done in real time. And each analysis was then used to go back to the field and make decisions. It's a way to understand what's driving an [epidemic](#)."

**More information:** Gytis Dudas et al, Virus genomes reveal factors that spread and sustained the Ebola epidemic, *Nature* (2017). [DOI: 10.1038/nature22040](#)

Provided by Fred Hutchinson Cancer Research Center

Citation: Ebola: Scientists show how real-time sequencing and data-sharing can help stop the next outbreak (2017, April 12) retrieved 10 April 2024 from <https://medicalxpress.com/news/2017-04-ebola-scientists-real-time-sequencing-data-sharing.html>

<p>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.</p>
--