Genomic sequencing reveals mutations, insights into 2014 Ebola outbreak

August 28 2014

Mr. Augustine Goba, laboratory director at Kenema Government Hospital Lassa fever laboratory, diagnosed the first case of Ebola in Sierra Leone. Credit: Stephen Gire

In response to an ongoing, unprecedented outbreak of Ebola virus disease (EVD) in West Africa, a team of researchers from the Broad Institute and Harvard University, in collaboration with the Sierra Leone Ministry of Health and Sanitation and researchers across institutions and

continents, has rapidly sequenced and analyzed more than 99 Ebola virus genomes. Their findings could have important implications for rapid field diagnostic tests. The team reports its results online in the journal *Science*.

For the current study, researchers sequenced 99 Ebola virus genomes collected from 78 patients diagnosed with Ebola in Sierra Leone during the first 24 days of the outbreak (a portion of the patients contributed samples more than once, allowing researchers a clearer view into how the virus can change in a single individual over the course of infection). The team found more than 300 genetic changes that make the 2014 Ebola virus genomes distinct from the viral genomes tied to previous Ebola outbreaks. They also found sequence variations indicating that, from the samples sequenced, the EVD outbreak started from a single introduction into humans, subsequently spreading from person to person over many months.

The variations they identified were frequently in regions of the genome encoding proteins. Some of the genetic variation detected in these studies may affect the primers (starting points for DNA synthesis) used in PCR-based diagnostic tests, emphasizing the importance of genomic surveillance and the need for vigilance.

To accelerate response efforts, the research team released the full-length sequences on National Center for Biotechnology Information's (NCBI's) DNA sequence database in advance of publication, making these data available to the global scientific community.

"By making the data immediately available to the community, we hope to accelerate response efforts," said co-senior author Pardis Sabeti, a senior associate member at the Broad Institute and an associate professor at Harvard University. "Upon releasing our first batch of Ebola sequences in June, some of the world's leading epidemic specialists

contacted us, and many of them are now also actively working on the data. We were honored and encouraged. A spirit of international and multidisciplinary collaboration is needed to quickly shed light on the ongoing outbreak."

The 2014 Zaire ebolavirus (EBOV) outbreak is unprecedented both in its size and in its emergence in multiple populated areas. Previous outbreaks had been localized mostly to sparsely populated regions of Middle Africa, with the largest outbreak in 1976 reporting 318 cases. The 2014 outbreak has manifested in the more densely-populated West Africa, and since it was first reported in Guinea in March 2014, 2,240 cases have been reported with 1,229 deaths (as of August 19).

Augustine Goba, Director of the Lassa Laboratory at the Kenema Government Hospital and a co-first author of the paper, identified the first Ebola virus disease case in Sierra Leone using PCR-based diagnostics.

"We established surveillance for Ebola well ahead of the disease's spread into Sierra Leone and began retrospective screening for the disease on samples as far back as January of this year," said Goba. "This was possible because of our long-standing work to diagnose and study another deadly disease, Lassa fever. We could thus identify cases and trace the Ebola virus spread as soon as it entered our country."

The research team increased the amount of genomic data available on the Ebola virus by four fold and used the technique of "deep sequencing" on all available samples. Deep sequencing is sequencing done enough times to generate high confidence in the results. In this study, researchers sequenced at a depth of 2,000 times on average for each Ebola genome to get an extremely close-up view of the virus genomes from 78 patients. This high-resolution view allowed the team to detect multiple mutations that alter protein sequences—potential targets for

future diagnostics, vaccines, and therapies.

The Ebola strains responsible for the current outbreak likely have a common ancestor, dating back to the very first recorded outbreak in 1976. The researchers also traced the transmission path and evolutionary relationships of the samples, revealing that the lineage responsible for the current outbreak diverged from the Middle African version of the virus within the last ten years and spread from Guinea to Sierra Leone by 12 people who had attended the same funeral.

The team's catalog of 395 mutations (over 340 that distinguish the current outbreak from previous ones, and over 50 within the West African outbreak) may serve as a starting point for other research groups.

"We've uncovered more than 300 genetic clues about what sets this outbreak apart from previous outbreaks," said Stephen Gire, a research scientist in the Sabeti lab at the Broad Institute and Harvard. "Although we don't know whether these differences are related to the severity of the current outbreak, by sharing these data with the research community, we hope to speed up our understanding of this epidemic and support global efforts to contain it."

"There is an extraordinary battle still ahead, and we have lost many friends and colleagues already like our good friend and colleague Dr. Humarr Khan, a co-senior author here," said Sabeti. "By providing this data to the research community immediately and demonstrating that transparency and partnership is one way we hope to honor Humarr's legacy. We are all in this fight together."

More information: Gire, SK, Goba, A et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*, 2014. www.sciencemag.org/lookup/doi/ …

1126/science.1259657