

The Ancestry of an Outbreak: HIV Past to Present

HIV (human immunodeficiency virus) is a deadly virus that has infected 75 million people in the last one hundred years. Epidemiologists have known since the 1980s that the HIV **pandemic** (a widespread outbreak) traces back to somewhere in Central Africa in the early 1900s, but they have struggled to further pinpoint its origin and the details of how it has spread over time. In 2014, an international team of researchers filled in a lot of the missing information. The team compared HIV DNA found in patients across Central Africa and were able to precisely determine where and when the common ancestor of HIV infections in Africa first infected humans. Knowing this history helps medical professionals understand just how varied HIV viruses are, which in turn helps researchers to develop effective treatments and predict how the virus will evolve and spread in the future.

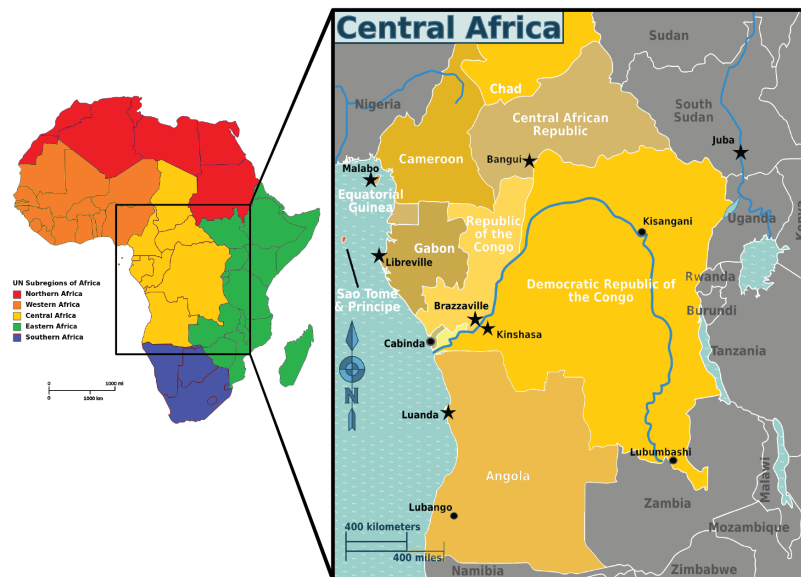


Figure 1. Map of Central Africa. Epidemiologists have known since the 1980s that HIV first infected humans living somewhere in Central Africa. Note that current names and boundaries are shown here, but that some have changed since the early 1900s.
Source: Modified from [Wikimedia Commons](https://commons.wikimedia.org/wiki/File:Map_of_Central_Africa.jpg)/Peter Fitzgerald.

The earliest sample of HIV that the research team could find was from 1959, but they suspected the virus began infecting humans much earlier. To figure out when and where HIV first arose and then how it spread within Central Africa, the researchers sequenced and compared the DNA of 792 HIV samples from the Central African countries of Cameroon, Central African Republic, Republic of Congo, Gabon and the Democratic Republic of the Congo (**Figure 1**).

To make comparisons, the scientists focused on a short segment (491 bases long) of the overall viral DNA sequence and identified mutations that made one sample's sequence different from another. Using computer modeling, the researchers constructed a phylogenetic tree of the virus samples based on how similar and different the sequences were. In general, the more similar the sequences, the more recently they shared a common viral ancestor. The researchers were then able to predict what the genetic code of the common ancestor of all the sequences would have been, as well as when and where that common ancestor would have been found. Look at **Figure 2** and read the caption to understand the data the researchers collected and the conclusions they were able to make.

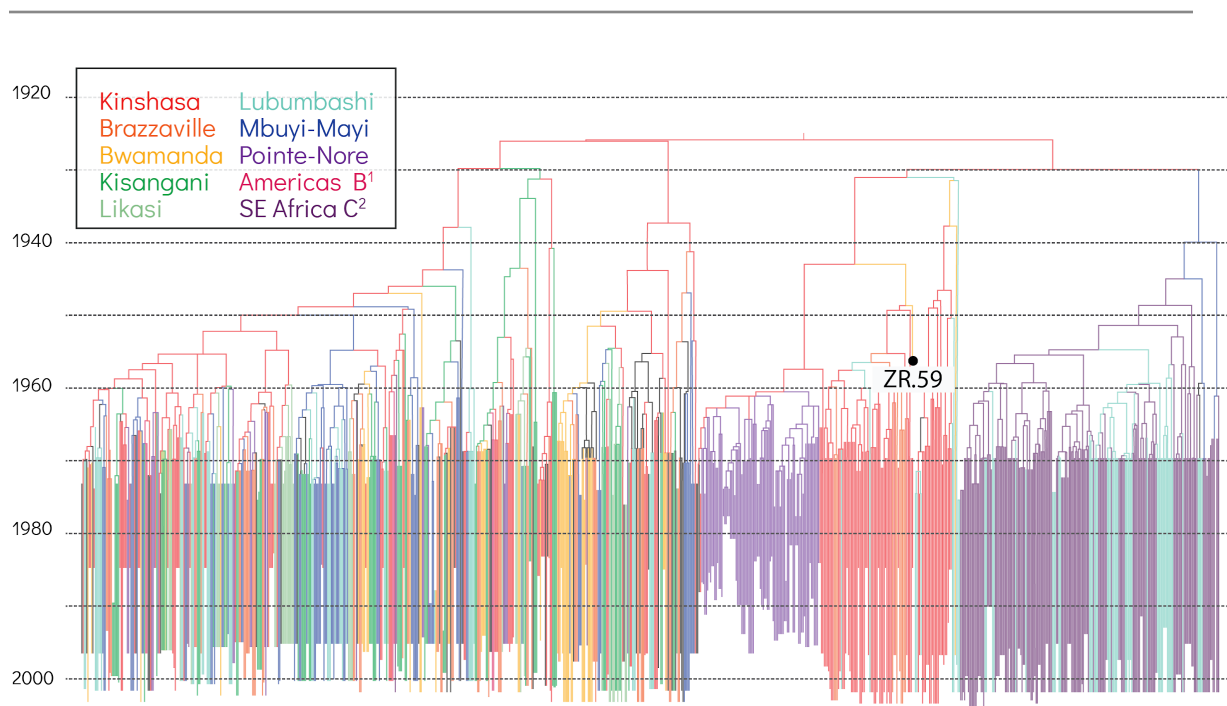


Figure 2. Phylogenetic Tree of HIV Diversity in Central Africa. This figure looks complicated, but you can get a lot of information from it. Each vertical line represents an individual patient sample, and the bottom point of the line ends at the time the sample was collected, with years on the Y-axis. Each line is also colored based on where the sample was collected. See the black circle labeled ZR.59? That represents the earliest actual sample of HIV. Can you tell when and where the common HIV ancestor was found? You can! Look for the “trunk” of the tree and use the left axis to find the date, and the color to identify the region. *Source: Modified from Faria et. al. 2014.*

The evidence collected supports the conclusion that the common ancestor of HIV first infected humans in Kinshasa, now in the Democratic Republic of the Congo, around 1920. Knowing that, the researchers were able to trace the virus forward in time. They showed that from the 1920s onward, HIV traveled with its human hosts along waterways and railways to other urban centers around Central Africa. Within these centers, the virus spread locally and led to the establishment of different virus populations in different cities.

These findings confirm that HIV was circulating within Central Africa starting in the 1920s, a date far before the 1980s infections in American and Europe that brought greater attention to the AIDS crisis. This research shows how fast HIV is able to evolve and diversify as it spreads through populations worldwide. HIV remains a global epidemic, and one of the leading causes of death and disease around the world. It is important to understand how the disease began its spread, so we can learn to predict how it will continue to evolve and diversify. Furthermore, the context and speed of how disease spreads is important for understanding how to combat any disease.

Though in this case the research looks backwards in evolutionary time, these kinds of genetic analyses of viral samples can be done in real time during an epidemic. Understanding how epidemics spread and which populations are particularly susceptible can help epidemiologists and healthcare workers design interventions and public health campaigns. Knowing how the virus is evolving can help researchers identify how the virus might be susceptible to different treatments or what vaccination strategies might work. All of this knowledge is critical to effectively stopping an ongoing epidemic.

Reference

Faria, Nuno R., Andrew Rambaut, Marc A. Suchard, Guy Baele, Trevor Bedford, Melissa J. Ward, Andrew J. Tatem, João D. Sousa, Nimalan Arinaminpathy, Jacques Pépin, David Posada, Martine Peeters, Oliver G. Prybus and Philippe Lemey. 2014. "The Early Spread and Epidemic Ignition of HIV-1 in Human Populations." *Science*, no. 346 (6205): 56-61. doi: [10.1126/science.1256739](https://doi.org/10.1126/science.1256739)

BiteScientist Profiles



Madeleine Jennewein is a graduate student in Harvard's virology program where she studies antibodies and the immune system in the context of acute HIV infection and in pregnancy. Outside of the lab she runs Harvard's LGBTQ student group and is passionate about diversity in the sciences.

Since 2016, **Tess Bernhard** has been teaching biology and statistics at Boston Prep in Hyde Park, Massachusetts. She formerly was a researcher at Princeton University and studied RNA's role in the origins of life and helped build a database on the traits of the last universal common ancestor.

