Next-Generation Sequencing Aids Researchers in the Fight Against the Ebola Virus

Genomic studies with the HiSeq® and MiSeq® Systems are enabling researchers to track Ebola outbreaks and understand the impact of the virus's rapid mutation rate.

Introduction

The 2014 Ebola outbreak sickened and killed thousands of people in West Africa, struck major urban centers and sent countries into chaos, and caused an international and humanitarian crisis. But first, the virus infected a child. Through careful history taking, researchers identified a two-year old boy as the first case in what would become the largest Ebola outbreak to date. He lived in a small village in Guinea, not far from the border of Sierra Leone and Liberia. In late December 2013, he became sick with fever, vomiting, and signs of hemorrhage. The boy died after several days, followed closely by several other family members. This strain of the virus is believed to have jumped to humans from an infected fruit bat, its natural wildlife reservoir.[1] Since the outbreak began nearly 18 months ago, the World Health Organization has documented 26,312 Ebola cases with 10,899 fatalities, as of April 29, 2015.[2]

The work to halt virus transmission has required input from a wide range of experts, including clinicians, public health engineers, and epidemiologists. Yet, stopping the Ebola virus in its tracks will require an understanding of the virus at its most basic level. Researchers are now using next-generation sequencing (NGS) platforms like the HiSeq and MiSeq Systems to understand the genetic structure of the virus. Deep sequencing of samples from people infected with the virus has revealed some of its many secrets, including a rapid mutation rate. Pardis Sabeti, PhD, a computational geneticist at Harvard University, and Christian Happi, PhD, a visiting professor of infectious disease at Harvard School of Public Health and Director of the African Center of Excellence for Genomics of infectious Diseases (ACEGID), Redeemer's University, Ede, Osun State, Nigeria, have spent much of the last year using the HiSeq System to understand how this virus spreads and changes over time.

Researchers Christian Happi, PhD, and Pardis Sabeti, PhD, in a large, open air church in Nigeria.

In the Center of the Storm

The pair first met at a conference in Senegal on computational genomics, where they began looking for signals in the human genome that might confer malaria resistance. Their work, however, revealed a signal that Lassa fever, another hemorrhagic virus endemic to West Africa, was placing a strong evolutionary pressure on humans. Within a month, they put together an Institutional Review Board (IRB) application and had a field site set up in Nigeria. They were deep into this work when they first heard of the Ebola outbreak in Sierra Leone in early 2014.

"We were in the middle of 1 of our weekly phone calls to discuss ongoing Lassa work when we received the news," Dr. Sabeti said. "We immediately moved to conduct surveillance at our sites in Nigeria and Sierra Leone, which were the best sites in this area to deal with a high-security virus. We knew that if the virus came to either of these countries that we would likely be dealing with it and we wanted to be prepared. When the first cases came in from Sierra Leone, we wanted to sequence the virus and understand if this was the same outbreak that had been declared in Guinea."

Team members from Nigeria, Senegal, and Sierra Leone who established virology laboratories with the MiSeq System to monitor and track the presence of Ebola and other pathogenic agents.

Their 7 years of work together on Lassa fever had placed them in a unique position to help. “These 2 diseases share several common features,” Dr. Happi said. “When the Ebola outbreak began, we were prepared.”

Like Lassa, Ebola is a single-stranded RNA virus. Working with the virus requires a maximum containment laboratory, also known as a BSL-4 lab. Ebola’s genetic material is RNA and therefore accumulates mutations very quickly, because there’s no way for the viral machinery to proofread sequences after replication. This fast mutation rate lets the virus evolve rapidly and adapt to new situations. It also enables researchers to track the virus over time by using the mutations as a guide.

When Ebola first struck, their lab was equipped with PCR technology, and other equipment that they used for their research. Yet, the scale of the outbreak and the work they wanted to do required genetic sequencing—lots of it. Drs. Sabeti and Happi and their colleagues evaluated several systems and ultimately chose the HiSeq 2500 System and the Nextera® XT Library Prep Kit.

“The HiSeq 2500 was by far the best performing sequencing system, based on its depth of coverage and data quality,” Dr. Sabeti said. “We tried about 5 or 6 different ways of preparing libraries and sequencing the virus. We found that the Nextera XT Kit worked exceptionally well and gave us the best reproducibility between technical replicates.”

Tracing the Origins of the Ebola Outbreak

Dr. Sabeti’s team worked around the clock to sequence all the samples and their hard work paid off. By late August, the team had traced the origin and transmission of the 2014 West African Ebola virus outbreak in unprecedented detail.[3] They sequenced 99 Ebola virus genomes from 78 patients at approximately 2000× coverage. Their results revealed that Ebola first moved to West Africa from its historic home in Central and East Africa in 2004. There was substantial genetic variation as the virus moved from human to human, and even as it multiplied within the same host. “Rather than 1 consensus viral sequence, we found thousands of different genetic snapshots of the virus as it mutated in an individual,” Dr. Sabeti said.

The study also revealed that the virus likely only jumped to humans from its animal reservoir 1 time, but that 1 spillover event created a sustained chain of human-to-human transmission. In October 2014, a separate group of scientists traced the Ebola transmission chain back to the toddler in Guinea.[4]

“If we had seen that each of these outbreaks was independent, and that they weren’t genetically related and coming from the same evolutionary tree, then we might have thought these were different infections from a natural reservoir and the result of different entries from the environment,” Dr. Sabeti said. “However, these outbreaks were closely related in time and the contact tracing supported that. It suggested that this was a single event in a one long viral transmission chain.”

These results confirmed the importance of contact tracing to prevent the spread of disease. Because the virus wasn’t being continually reintroduced from the environment, epidemiologists knew that if they could stop people from passing the virus to each other, then they could stop the outbreak.

Thousand Days – One Truth

The insights into Ebola were not gained without significant cost to the physicians and scientists on the ground during the outbreak. More than 800 health care workers were infected with the virus and 492 have died, including some of the people Drs. Sabeti and Happi worked with. Out of that pain came a song that means a lot to Dr. Sabeti and her research team in West Africa.

Her team included a group of 11 scientists from Nigeria and Senegal, who traveled to Sabeti’s lab during the outbreak for training in genomics. Sabeti, a gifted musician and vocalist, enjoyed singing with them.

“The women all have beautiful voices,” Dr. Sabeti said. “When they came to visit, we’d gather and sing once a week. Even though we were working hard and I often wasn’t sleeping, I promised to keep this date. It was in the midst of 1 of our weekly sessions that the inspiration came. I looked at the faces of these women and the song just wrote itself.”

Watch the video.

The deep sequencing data provided Drs. Sabeti and Happi with other insights. “With NGS, you get high-resolution data and can see more than just the common variants,” Sabeti said. “If we’d had only the common variants, we would have missed a lot and it would have been hard to see who was infecting whom at a high scale. At 1% frequency or lower, you can see all the rare things that are percolating and can begin to understand the transmissions. NGS provided us with the depth of information we needed to understand the human-to-human viral transmission chain.”

Open Data Allows Deeper Exploration

The reams of sequencing data that Drs. Sabeti and Happi released allowed many other groups to advance our understanding of how the outbreak evolved and informed the development of therapeutics to combat outbreaks in the future. In January 2015, they worked with a team at the U.S. Army Medical Research Institute of Infectious Diseases (USAMRIID) investigating how Ebola’s rapid mutation rate could potentially impact the efficacy of currently available treatments.[5] These include sequence-based therapeutics such as small interfering RNAs (siRNAs), phosphorodiamidate morpholino oligomers (PMOs), and antibodies that target genes and proteins within the Ebola virus. For each outbreak, the specific targets of these drugs will need to be modified to make sure that candidate treatments and vaccines are strain-specific.

Sequencing with the HiSeq System also enabled examination of the patient samples for other microbes, and thus new discoveries about factors that might help people survive an Ebola infection. Led by the University of Wisconsin, the team found that individuals carrying the pegivirus GB virus-C, which is known for its ability to modulate the

immune system, were significantly less likely to die of Ebola than individuals who weren't coinfected with GBV-C.[6] "These are amazing insights you can only find with NGS," Dr. Sabeti said.

New Laboratories to Monitor Viral Outbreaks

With financial support from USAID and Illumina, Drs. Sabeti and Happi obtained a MiSeq System in January 2015 as the foundation of Nigeria's first next-generation virology lab. They also obtained a MiSeq system for a laboratory in Senegal, and collaborated with USAMRIID on their work to establish a laboratory with a MiSeq System in Liberia. Illumina will soon provide a MiSeq System for a virology laboratory in Sierra Leone. For them, the MiSeq System is an obvious choice.

"Of all the NGS systems, the MiSeq System is a standout," Dr. Sabeti said. "There is no other NGS system right now that has this kind of capacity to move into the field. It has very high efficiency and you get great reads and rapid turnaround. In some places, we have to consider power efficiency. The MiSeq is a just a beautiful system that works very well."

"Based on our own interaction, the MiSeq System is also very user-friendly," Dr. Happi added.

They are also continuing to use the Nextera XT Library Prep Kits. "We use the Nextera XT kit because it has good efficiency and very good accuracy," Dr. Sabeti said. "We have made a few modifications to the kit, enabling us to remove more of the human material so that we can enrich as much of the viral material as possible."

They hope to use the MiSeq System to gain a better understanding of the whole range of viruses and pathogens that affect people living in West Africa. Currently, the lab is focusing on Ebola sequencing, although they plan to perform sequencing on Lassa fever samples shortly.

"Pardis and I are always moving forward with technology, making sure that we're using the latest core technology to understand disease and health problems in Africa," Dr. Happi said. "By bringing together knowledge, know-how, and technology, we're giving people the ability to start looking inward to solve problems in this region. Now that we have the MiSeq System, I think it's important that we enable African scientists to start using this technology and start answering questions that are important to people in Africa."

"The Nigeria laboratory is a landmark facility," Dr. Sabeti said. "It is developing into a training center, with a graduate program that Christian is in charge of for West African students. There's a strong interest forming in West Africa and beyond in understanding the genomics in infectious disease. Christian is leading that effort here."

A Constantly Mutating Foe

The Ebola virus strain responsible for the 2014 outbreak is still mutating. In addition to the 99 strains sequenced early during the outbreak, the team has since cataloged another 150 Ebola genomes and made them available in a public database.[7]

Even as the outbreak eventually winds down, it won't necessarily be the end for Ebola sequencing. Dr. Sabeti says that scientists need to keep on top of what viral strains are in play and monitor any changes that are occurring. Both researchers believe that NGS will open a new chapter in the never-ending fight against viruses.

References

1. Mari Saéz A, Weiss S, Nowak K, et al. Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol Med.* 2014; 7(1): 17-23.
2. WHO Ebola Situation Report, April 29, 2015: http://apps.who.int/ebola/current-situation/ebola-situation-report-29-april-2015.
3. Gire SK, Goba A, Andersen KG, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science.* 2014; 345 (6202): 1369–1372.
4. Baize S, Pannetier D, Oestereich L, et al. Emergence of Zaire Ebola Virus Disease in Guinea. *N Engl J Med.* 2014; 371(15): 1418–1425.
5. Kugelman JR, Sanchez-Lockhart M, Andersen KG, et al. Evaluation of the Potential Impact of Ebola Virus Genomic Drift on the Efficacy of Sequence-Based Candidate Therapeutics. *MBio.* 2015; 6(1): pii: e02227-14.
6. Lauck M, Bailey AL, Andersen KG, et al. GB Virus C coinfections in west African Ebola Patients. *J Virol.* 2015; 89(4): 2425–2429.
7. NCBI Ebolavirus database: http://www.ncbi.nlm.nih.gov/genomes/VirusVariation/Database/nph-select2.cgi?cmd=database&taxid=186536

Learn more about the Illumina systems and products mentioned in this article, and our partnership with the Broad Institute of MIT and Harvard, and USAID to combat the Ebola epidemic in West Africa:

- HiSeq System, www.illumina.com/systems/hiseq_2500_1500.html
- MiSeq System, www.illumina.com/systems/miseq.html
- Nextera XT DNA Library Prep Kit, www.illumina.com/products/nextera_xt_dna_library_prep_kit.html
- USAID, Broad Institute, and Illumina Form a Public-Private Partnership Combating the Ebola Epidemic in West Africa, investor.illumina.com/phoenix.zhtml?c=121127&p=irol-newsArticle&ID=1990982

Illumina • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.

 Pub. No. 1270-2015-001 Current as of 14 May 2015

illumina®