Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool

Thomas Hoenen, Allison Groseth, Kyle Rosenke,
Robert J. Fischer, Andreas Hoenen,
Seth D. Judson, Cynthia Martellaro,
Darryl Falzarano, Andrea Marzi,
R. Burke Squires, Kurt R. Wollenberg,
Emmie de Wit, Joseph Prescott, David Safronetz,
Neeltje van Doremalen, Trenton Bushmaker,
Friederike Feldmann, Kristin McNally,
Fatorma K. Bolay, Barry Fields, Tara Sealy,
Mark Rayfield, Stuart T. Nichol, Kathryn C. Zoon,
Moses Massaquoi, Vincent J. Munster,
Heinz Feldmann

Rapid sequencing of RNA/DNA from pathogen samples obtained during disease outbreaks provides critical scientific and public health information. However, challenges exist for exporting samples to laboratories or establishing conventional sequencers in remote outbreak regions. We successfully used a novel, pocket-sized nanopore sequencer at a field diagnostic laboratory in Liberia during the current Ebola virus outbreak.

Disease outbreaks in resource-limited or remote areas pose unique challenges to outbreak responses. These challenges are exemplified by the ongoing Ebola virus (EBOV) outbreak in West Africa that began in 2014 (*1*) and is unprecedented in its size and duration. Correspondingly, the magnitude of the international response, encompassing ≈50 Ebola treatment units (ETUs) and >2 dozen diagnostic laboratories, has been equally unprecedented.

Author affiliations: Friedrich-Loeffler-Institut, Greifswald–Insel Riems, Germany (T. Hoenen, A. Groseth); National Institutes of Health, Hamilton, Montana, USA (T. Hoenen, A. Groseth, K. Rosenke, R.J. Fischer, S.D. Judson, C. Martellaro, D. Falzarano, A. Marzi, E. de Wit, J. Prescott, D. Safronetz, N. van Doremalen, T. Bushmaker, F. Feldmann, K. McNally, V.J. Munster, H. Feldmann); Independent scholar, Aachen, Germany (A. Hoenen); University of Saskatchewan, Saskatoon, Saskatchewan, Canada (D. Falzarano); National Institutes of Health, Bethesda, Maryland, USA (R.B. Squires, K.R. Wollenberg, K.C. Zoon); The Liberian Institute for Biomedical Research, Charles Ville, Republic of Liberia (F.K. Bolay); Centers for Disease Control and Prevention, Atlanta, Georgia, USA (B. Fields, T. Sealy, M. Rayfield, S.T. Nichol); Ministry of Health and Social Welfare, Monrovia, Republic of Liberia (M. Massaquoi)

These laboratories often are operated under improvised field conditions to keep them close to active, sometimes remote transmission sites (*2,3*).

Rapidly obtaining genome sequences during disease outbreaks is crucial for clarifying patterns of virus evolution, monitoring the validity of diagnostic assays, and investigating transmission chains (*4,5*). Further, rapid results may help determine the efficacy of sequence-dependent countermeasures, such as siRNAs or antibody treatments. In the past, obtaining timely genome sequences has been difficult because of political and logistical obstacles that limited the export of samples to laboratories capable of performing these analyses. As an example, during the first year of the outbreak in West Africa, only 2 reports of genome sequences from patients were published (*1,6*). Similarly, establishing conventional Sanger or next-generation sequencing technologies in affected countries is logistically challenging because of the size and weight (≈40 to ≈100 kg) of the necessary equipment, the high potential for transport damage related to the sensitive optics many of these machines incorporate, limitations on supportive infrastructure, and complex sample processing procedures. An additional challenge is the required installation or calibration of sequencing machines, which often has to be done by field engineers employed by the manufacturers, who may be reluctant to send their employees into outbreak areas. However, Kugelman et al. recently reported the successful deployment of an Illumina MiSeq, a well-established, conventional next-generation sequencing platform (Illumina, San Diego, CA, USA), to West Africa; the platform became operational in February 2015 (*5*).

Seeking a platform that would be more rapidly deployable and reliable under field conditions, we established protocols and evaluated the feasibility of nanopore sequencing technology under outbreak conditions using a pocket-sized (≈10 × 4 × 2 cm, 75 g) MinION sequencing device (Oxford Nanopore Technologies, [https://www.nanoporetech.com/]). Because of its small size, this device can easily be transported into remote locations; furthermore, it requires no special setup or calibration procedures and can be operational immediately after arrival in an outbreak area. Further, data turnaround is very rapid, and consequently, nanopore sequencing is being developed as a rapid diagnostic tool for management of outbreaks of various diseases (*7,8*). The MinION device senses individual DNA molecules based on modulation of ion currents across nanopores as the molecules are passing through. These modulations are

dependent on the physical properties of the nucleotides and allow determination of the nucleotide sequence (*9*).

The Study

To facilitate sequencing of the RNA genome of EBOV, we developed and tested an approach based on reverse transcription PCR, in which whole virus genomes were amplified in overlapping fragments (Figure 1, panels A, B; online Technical Appendix, http://wwwnc.cdc.gov/EID/article/22/2/15-1796-Techapp1.pdf). This approach was first validated in a regular laboratory setting in the Rocky Mountain Laboratories of the National Institutes of Health (NIH) by using blood samples from nonhuman primates experimentally inoculated with EBOV strain Makona-Gueckedou-C07 (*10,11*). This validation showed that sequencing information was obtainable for the complete genome with an average of 7,038 reads at every nucleotide position (read depth; online Technical Appendix Figure 1, panel A). We observed no sequence differences when comparing the consensus sequence derived from these data to those obtained by using Sanger sequencing (online Technical Appendix Figure 1, panel B). Furthermore, by analyzing linearized plasmid DNA of known sequence, we

established the accuracy of the MinION device as ≈84% for a single read (online Technical Appendix Figure 1, panels C, D). On the basis of this information, and the fact that read depth can compensate for miscalled nucleotides in individual reads by piling up reads covering the same region, we determined the theoretical probability for ≥1 miscalled base (TPMB) in a complete MinION-sequenced EBOV genome to be <5% when the read depth is ≥33 at all positions (online Technical Appendix Figure 1, panels E and F).

After having validated this approach, MinION devices were taken to the Centers for Disease Control and Prevention (CDC)/NIH field laboratory that provided diagnostic support for ETUs in Monrovia, Liberia, during August 2014–May 2015. All equipment and reagents necessary for sequencing could be easily transported as checked luggage by a single person on a commercial carrier. In Liberia, temperatures in the laboratory area used for sequencing ranged from 28 to 32°C, necessitating the use of an improvised heat sink for the devices, which consisted of a metal plate of ≈30 × 30 cm (online Technical Appendix Figure 2, panels A, B). Under field conditions, we initially failed to produce complete genomes with high confidence because of problems with
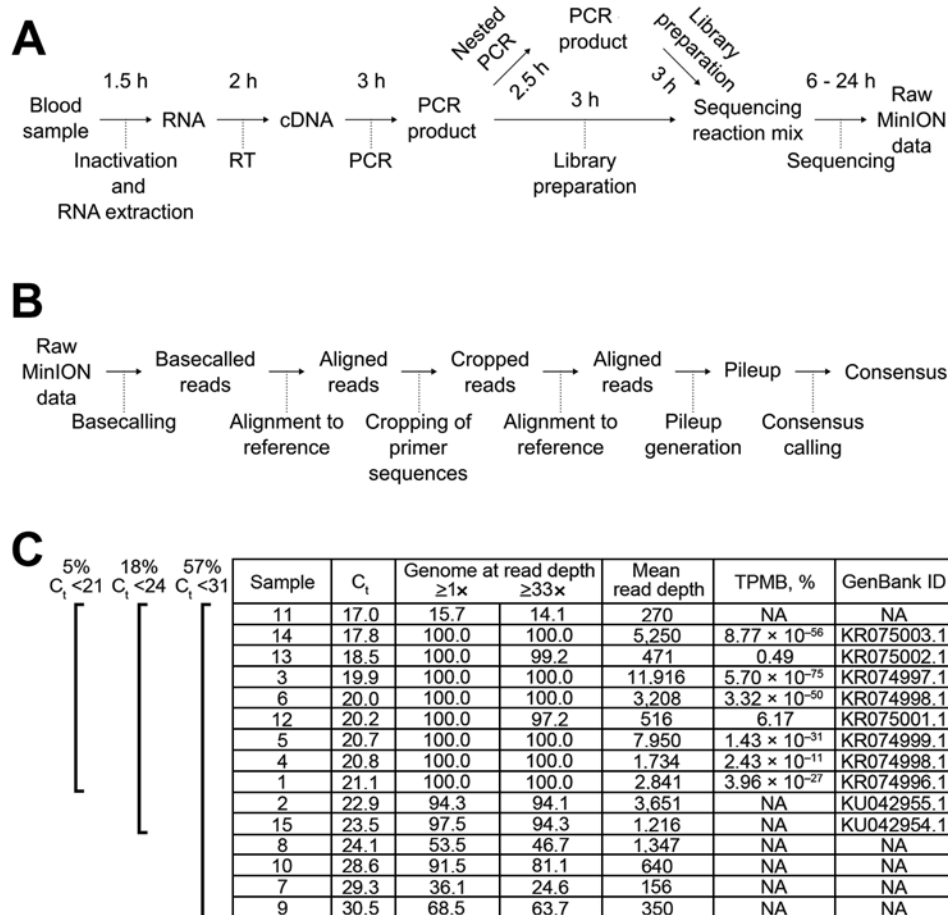

Figure 1. MinION sequencing. A) Experimental and B) bioinformatics workflows. Times indicated are the approximate duration for each procedure. RT, reverse transcription. C) Sequencing results showing Ebola virus load (expressed as C_t value), percentage of the genome with a minimum read depth of ≥1 or ≥33, mean read depth, theoretical probability for a miscalled base (TPMB), and GenBank accession numbers of complete and nearly complete genomes. Brackets at left indicate percentage of Ebola virus–positive patient samples below each of the 3 cutoff cycle threshold (C_t) values used in this study (C_t <21, <24, <31). Sample 8 was from an oral swab; all others were from blood. NA, not available.

Sample	C_t	Genome at read depth ≥1×	≥33×	Mean read depth	TPMB, %	GenBank ID
11	17.0	15.7	14.1	270	NA	NA
14	17.8	100.0	100.0	5,250	8.77×10^{-56}	KR075003.1
13	18.5	100.0	99.2	471	0.49	KR075002.1
3	19.9	100.0	100.0	11.916	5.70×10^{-75}	KR074997.1
6	20.0	100.0	100.0	3,208	3.32×10^{-50}	KR074998.1
12	20.2	100.0	97.2	516	6.17	KR075001.1
5	20.7	100.0	100.0	7.950	1.43×10^{-31}	KR074999.1
4	20.8	100.0	100.0	1.734	2.43×10^{-11}	KR074998.1
1	21.1	100.0	100.0	2.841	3.96×10^{-27}	KR074996.1
2	22.9	94.3	94.1	3,651	NA	KU042955.1
15	23.5	97.5	94.3	1.216	NA	KU042954.1
8	24.1	53.5	46.7	1,347	NA	NA
10	28.6	91.5	81.1	640	NA	NA
7	29.3	36.1	24.6	156	NA	NA
9	30.5	68.5	63.7	350	NA	NA

Brackets at left: 5% C_t <21, 18% C_t <24, 57% C_t <31

PCR yields (online Technical Appendix Figure 3, panels A, B). However, by implementing a second PCR step, we circumvented this problem and obtained high quality complete genome sequences for 8 of 9 high-virus load samples (cycle threshold ≤21) (Figure 1, panel C; online Technical Appendix Figure 4, panel A). In lower virus load samples, we could obtain only incomplete genome sequences; however, even in those samples regions for which sequencing information was available generally showed high read depths (online Technical Appendix Figure 4, panel B), suggesting that further optimization of PCRs might also allow complete coverage for these samples. Furthermore, even incomplete genome sequences can provide valuable information during an outbreak, allowing analysis of individual genes and the tracing of transmission chains (*12*).

Using this updated protocol, we achieve a sustained capacity of 4 full-length genomes per day for a single person conducting the laboratory work using 2 MinION devices (Figure 1, panels A, B). However, with the exception of the first 2 sequencing runs, bioinformatics analysis during this mission was mainly completed after returning to the NIH, to maximize the time for raw data acquisition (online Technical Appendix).

Phylogenetic analysis of the complete genomes generated in Monrovia, Liberia, showed them being clearly distinct from Sierra Leone or early Guinea sequences of EBOV-Makona (online Technical Appendix Figure 5) but clustering well with all other sequences found in samples from Liberia. These results suggest that EBOV in Liberia resulted from a single introduction or a limited number of introductions with genetically similar viruses. When analyzing the obtained full-length sequences and comparing them to a consensus sequence from the outbreak (*13*), we observed few mutations, most in noncoding regions or synonymous mutations (Figure 2); none affected siRNA target sequences or the diagnostic targets used in the CDC/NIH laboratory.

Using Bayesian analysis including these sequences, we estimated the nucleotide substitution rate during the outbreak at 1.36×10^{-3}, consistent with recently published values (*5,13–15*). In a root-to-tip-analysis, the sequences we obtained showed substitution rates comparable to other sequences from the outbreak (online Technical Appendix Figure 6). Overall, these data suggest that EBOV has remained relatively stable genetically during the outbreak.

Conclusions

We found that, because of the device's small size and comparatively modest resource requirements, nanopore sequencing has tremendous potential for use in remote and resource-limited areas, and its implementation could revolutionize the capacity of public health professionals to perform sequencing during future disease outbreaks. Although we used a directed approach to sequencing, approaches not dependent on prior pathogen identification (i.e. for diagnostic use of the MinION device) are currently being developed (*7*) and will even further increase this technology's usefulness in future outbreaks.
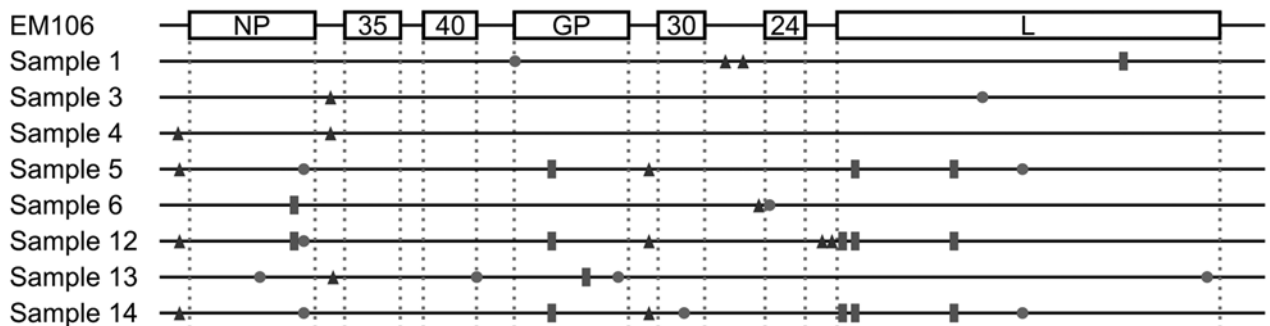
Figure 2. Observed mutations in the 8 fully nanopore-sequenced Ebola-positive blood samples compared to a reference sequence from June 2014 (SLI/Makona-EM106, GenBank accession number KM233036.1). Squares indicate nonsynonymous mutations, circles indicate synonymous changes, and triangles indicate changes in noncoding regions. NP, nucleoprotein; GP, glycoprotein.

T.H., A.G., K.R., and J.P. are participants in the ONT MinION Access Programme, and received some of the MinION devices and Flow Cells used for this study free of charge or at reduced cost. T.H. was invited by Oxford Nanopore Technologies to present part of this work at the London Calling meeting in London, UK. The deidentified patient specimens used in this manuscript were collected as part of public health surveillance and not as human subjects research. The NIH Office of Human Subjects Research has determined that federal regulations for the protection of human subjects do not apply.

Dr. Hoenen was a visiting postdoctoral fellow at the Laboratory of Virology of the National Institutes of Health in Hamilton, Montana, USA, for major phases of this study, and has recently joined the Friedrich-Loeffler-Institut on the island of Riems, near Greifswald, Germany, where he is head of the Laboratory for Molecular Biology of Filoviruses. His research focuses on the molecular biology of filoviruses and their interactions with host cells, with the ultimate goal of finding new antivirals against filoviruses and other hemorrhagic fever viruses.

References

1. Baize S, Pannetier D, Oestereich L, Rieger T, Koivogui L, Magassouba N, et al. Emergence of Zaire Ebola virus disease in Guinea. N Engl J Med. 2014;371:1418–25. http://dx.doi.org/10.1056/NEJMoa1404505
2. Grolla A, Jones S, Kobinger G, Sprecher A, Girard G, Yao M, et al. Flexibility of mobile laboratory unit in support of patient management during the 2007 Ebola-Zaire outbreak in the Democratic Republic of Congo. Zoonoses Public Health. 2012;59 (Suppl 2):151–7. http://dx.doi.org/10.1111/j.1863-2378.2012.01477.x
3. Allaranga Y, Kone ML, Formenty P, Libama F, Boumandouki P, Woodfill CJ, et al. Lessons learned during active epidemiological surveillance of Ebola and Marburg viral hemorrhagic fever epidemics in Africa. East Afr J Public Health. 2010;7:30–6.
4. Mate SE, Kugelman JR, Nyenswah TG, Ladner JT, Wiley MR, Cordier-Lassalle T, et al. Molecular Evidence of Sexual Transmission of Ebola Virus. N Engl J Med. 2015. http://dx.doi.org/10.1056/NEJMoa1509773
5. Kugelman JR, Wiley MR, Mate S, Ladner JT, Beitzel B, Fakoli L, et al. Monitoring of Ebola Virus Makona Evolution through Establishment of Advanced Genomic Capability in Liberia. Emerg Infect Dis. 2015;21:1135–43. http://dx.doi.org/10.3201/eid2107.150522
6. Gire SK, Goba A, Andersen KG, Sealfon RS, Park DJ, Kanneh L, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science. 2014;345:1369–72. http://dx.doi.org/10.1126/science.1259657
7. Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med. 2015;7(1):99. http://dx.doi.org/10.1186/s13073-015-0220-9
8. Quick J, Ashton P, Calus S, Chatt C, Gossain S, Hawker J, et al. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. Genome Biol. 2015;16:114. http://dx.doi.org/10.1186/s13059-015-0677-2
9. Ashton PM, Nair S, Dallman T, Rubino S, Rabsch W, Mwaigwisya S, et al. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. Nat Biotechnol. 2015;33:296–300. http://dx.doi.org/10.1038/nbt.3103
10. Hoenen T, Groseth A, Feldmann F, Marzi A, Ebihara H, Kobinger G, et al. Complete genome sequences of three Ebola virus isolates from the 2014 outbreak in west Africa. Genome Announc. 2014;2:e01331–14. http://dx.doi.org/10.1128/genomeA.01331-14
11. Marzi A, Feldmann F, Hanley PW, Scott DP, Gunther S, Feldmann H. Delayed disease progression in cynomolgus macaques infected with Ebola virus Makona strain. Emerg Infect Dis. 2015;21:1777–83. http://dx.doi.org/10.3201/eid2110.150259
12. Christie A, Davies-Wayne GJ, Cordier-Lasalle T, Blackley DJ, Laney AS, Williams DE, et al. Possible sexual transmission of Ebola virus – Liberia, 2015. MMWR Morb Mortal Wkly Rep. 2015;64:479–81.
13. Hoenen T, Safronetz D, Groseth A, Wollenberg KR, Koita OA, Diarra B, et al. Virology. Mutation rate and genotype variation of Ebola virus from Mali case sequences. Science. 2015;348:117–9. http://dx.doi.org/10.1126/science.aaa5646
14. Tong YG, Shi WF, Liu D, Qian J, Liang L, Bo XC, et al. Genetic diversity and evolutionary dynamics of Ebola virus in Sierra Leone. Nature. 2015;524:93–6. http://dx.doi.org/10.1038/nature14490
15. Carroll MW, Matthews DA, Hiscox JA, Elmore MJ, Pollakis G, Rambaut A, et al. Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa. Nature. 2015;524:97–101. http://dx.doi.org/10.1038/nature14594

Address for correspondence: Thomas Hoenen, Friedrich-Loeffler-Institut, Südufer 10, 17493 Greifswald–Insel Riems, Germany; email: thomas.hoenen@fli.bund.de