2018 Ebola virus disease outbreak in Équateur Province, Democratic Republic of the Congo: a retrospective genomic characterisation

Placide Mbala-Kingebeni, Catherine B Pratt*, Michael R Wiley*, Moussa M Diagne*, Sheila Makiala-Mandanda, Amuri Aziza, Nicholas Di Paola, Joseph A Chitty, Mamadou Diop, Ahidjo Ayouba, Nicole Vidal, Ousmane Faye, Oumar Faye, Stormy Karhemere, Aaron Aruna, Justus Nsio, Felix Mulangu, Daniel Mukadi, Patrick Mukadi, John Kombe, Anastasie Mulumba, Sophie Duraffour, Jacques Likofata, Elisabeth Pukuta, Katie Caviness, Maggie L Bartlett, Jeanette Gonzalez, Timothy Minogue, Shanmuga Sozhamannan, Stephen M Gross, Gary P Schroth, Jens H Kuhn, Eric F Donaldson, Eric Delaporte, Mariano Sanchez-Lockhart, Martine Peeters, Jean-Jacques Muyembe-Tamfum, Amadou Alpha Sall†, Gustavo Palacios†, Steve Ahuka-Mundeke†*

Summary

Background The 2018 Ebola virus disease (EVD) outbreak in Équateur Province, Democratic Republic of the Congo, began on May 8, and was declared over on July 24; it resulted in 54 documented cases and 33 deaths. We did a retrospective genomic characterisation of the outbreak and assessed potential therapeutic agents and vaccine (medical countermeasures).

Methods We used target-enrichment sequencing to produce Ebola virus genomes from samples obtained in the 2018 Équateur Province outbreak. Combining these genomes with genomes associated with known outbreaks from GenBank, we constructed a maximum-likelihood phylogenetic tree. In-silico analyses were used to assess potential mismatches between the outbreak strain and the probes and primers of diagnostic assays and the antigenic sites of the experimental rVSVΔG-ZEBOV-GP vaccine and therapeutics. An in-vitro flow cytometry assay was used to assess the binding capability of the individual components of the monoclonal antibody cocktail ZMapp.

Findings A targeted sequencing approach produced 16 near-complete genomes. Phylogenetic analysis of these genomes and 1011 genomes from GenBank revealed a distinct cluster, confirming a new Ebola virus variant, for which we propose the name "Tumba". This new variant appears to have evolved at a slower rate than other Ebola virus variants ($0 \cdot 69 \times 10^{-3}$ substitutions per site per year with "Tumba" *vs* $1 \cdot 06 \times 10^{-3}$ substitutions per site per year without "Tumba"). We found few sequence mismatches in the assessed assay target regions and antigenic sites. We identified nine amino acid changes in the Ebola virus surface glycoprotein, of which one resulted in reduced binding of the 13C6 antibody within the ZMapp cocktail.

Interpretation Retrospectively, we show the feasibility of using genomics to rapidly characterise a new Ebola virus variant within the timeframe of an outbreak. Phylogenetic analysis provides further indications that these variants are evolving at differing rates. Rapid in-silico analyses can direct in-vitro experiments to quickly assess medical countermeasures.

Funding Defense Biological Product Assurance Office.

Introduction

On May 8, 2018, the Democratic Republic of the Congo declared an outbreak of Ebola virus disease (EVD) caused by Ebola virus (EBOV).[1] The outbreak, which was officially declared over on July 24, 2018, included 54 documented EVD cases (38 laboratory-confirmed, 16 probable) and 33 deaths (case fatality rate 61·1%). All of the cases originated from three health zones (Bikoro, Iboko, and Wangata) located in the northwestern region of the Democratic Republic of the Congo's Équateur Province near Lake Tumba (Ntomba).[2] This epidemic was the ninth recorded EVD outbreak in the country. Previous EVD outbreaks occurred in 1976, 1977, 1995, 2007, 2008–09, 2014, and 2017, and an additional EVD outbreak, caused by Bundibugyo virus, in 2012.[3] On Aug 1, 2018, the tenth EVD outbreak in the Democratic Republic of the Congo was declared in the country's northeastern North Kivu Province.[4] Multiple EVD outbreaks have also occurred in other parts of Middle Africa (defined according to UN Statistics Division geographical locations), including Republic of the Congo and Gabon, and a single outbreak occurred in Western Africa from 2013–16 (appendix).

As part of the public health response for the 2018 Équateur Province EVD outbreak, Institut National de Recherche Biomédicale (INRB, Kinshasa, the Democratic Republic of the Congo), in collaboration with international partners, deployed three mobile laboratories in July, 2018, to each of the affected areas: Bikoro (Bikoro health zone), Itipo (Iboko health zone), and Mbandaka (Wangata health zone). Clinical samples (n=432; serum, urine, and saliva) were collected from people with suspected EVD and tested in the mobile laboratories. Of these samples,

Lancet Infect Dis 2019; 19: 641–47

Published **Online** April 15, 2019 http://dx.doi.org/10.1016/S1473-3099(19)30124-0

See **Comment** page 567

*Contributed equally

†Joint senior authors

Institut National de Recherche Biomédicale, Kinshasa, Democratic Republic of the Congo (P Mbala-Kingebeni MD, S Makiala-Mandanda PhD, A Aziza MSc, S Karhemere DVM, P Mukadi MD, E Pukuta MSc, J-J Muyembe-Tamfum PhD, S Ahuka-Mundeke PhD); **TransVIHMI, Institut de Recherche pour le Développement, Institut National de la Santé et de la Recherche Médicale, Université de Montpellier, Montpellier, France** (P Mbala-Kingebeni, A Ayouba PhD, N Vidal PhD, E Delaporte PhD, M Peeters PhD); **Service de Microbiologie, Cliniques Universitaires de Kinshasa, Kinshasa, Democratic Republic of the Congo** (P Mbala-Kingebeni, S Makiala-Mandanda, D Mukadi MD, J-J Muyembe-Tamfum, S Ahuka-Mundeke); **Center for Genome Sciences** (C B Pratt MSc, M R Wiley PhD, N Di Paola ScD, J A Chitty BS, K Caviness PhD, M L Bartlett BS, J Gonzalez MS, M Sanchez-Lockhart PhD, G Palacios PhD), **and Diagnostics Services Division** (T Minogue PhD), **US Army Medical Research Institute of Infectious Diseases, Frederick, MD, USA; College of Public Health** (C B Pratt, M R Wiley), **and Department of Pathology and Microbiology** (M L Bartlett,

M Sanchez-Lockhart), **University of Nebraska Medical Center, Omaha, NE, USA; Institut Pasteur de Dakar, Dakar, Senegal** (M M Diagne PhD, M Diop MSc, Ous Faye PhD, Oum Faye PhD, A Alpha Sall PhD)**; Direction Générale de Lutte contre la Maladie, Kinshasa, Democratic Republic of the Congo** (A Aruna MPH, J Nsio MPH, F Mulangu MPH, J Kombe MPH)**; Monsieur le Représentant de l'Organisation Mondiale de la Santé, Democratic Republic of the Congo** (A Mulumba MD, S Duraffour PhD)**; Bernhard-Nocht-Institut für Tropenmedizin, Hamburg, Germany** (S Duraffour)**; Laboratoire Provinciale, Mbandaka, Democratic Republic of the Congo** (J Likofata)**; Defense Biological Product Assurance Office, Joint Program Executive Office for Chemical, Biological, Radiological and Nuclear Defense—Joint Project Management Office for Guardian, Frederick, MD, USA** (S Sozhamannan PhD)**; Logistics Management Institute, Tysons, VA, USA** (S Sozhamannan)**; Illumina, San Diego, CA, USA** (S M Gross PhD, G P Schroth PhD)**; Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Frederick, MD, USA** (J H Kuhn MD)**; and Division of Antiviral Products, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA** (E F Donaldson PhD)

Correspondence to: Dr Gustavo Palacios, Center for Genome Sciences, US Army Medical Research Institute of Infectious Diseases, Frederick, MD 21702, USA
gustavo.f.palacios.civ@mail.mil

For more on **UN standard country and area codes** see https://unstats.un.org/

See **Online** for appendix

For more on the **MinION device** see https://nanoporetech.com/products/minion

Research in context

Evidence before this study

Whole-genome sequencing of epidemics has become a powerful tool for outbreak control. Viral whole-genome sequencing data enable the identification of possible target erosion in diagnostic assays and characterisation of responses to medical countermeasures within the context of the causative agent. We searched PubMed from database inception to Aug 15, 2018, without language restrictions for articles using the terms "ebola", "EVD", or "ebola virus", with the terms "genomic epidemiology", "diagnostic", "treatment", or "Equateur." There were more than 50 studies describing the genomic characterisation of the Ebola virus (EBOV) during the Western African Ebola virus disease (EVD) outbreak in 2013–16. However, only three articles were related to the Équateur EVD outbreak: genome sequencing of a fragment of the *VP35* gene, allowing identification of the outbreak as EVD; a traditional epidemiological assessment of the Équateur outbreak during April–May, 2018; and a general review of EVD outbreaks in the Democratic Republic of the Congo.

Added value of this study

We provide 16 near-complete EBOV genomes, sequenced from 16 EBOV-positive samples (serum and saliva) during the 2018 outbreak of EVD in Équateur Province, Democratic Republic of the Congo. We show that the outbreak was caused by a single spillover event from the host reservoir. The diagnostic assays, preventive vaccine, and immunotherapeutics were assessed against the EBOV variant circulating during the Équateur Province outbreak. No major erosion effects (ie, nucleotide changes in EBOV/"Tum" genomes that could disrupt primer or probe binding sites of EBOV-specific diagnostic PCR assays) were observed in either the diagnostic or the therapeutic options. The in-country sequencing capacity established through these efforts will aid future research and outbreak response. We provide a roadmap for the assessment and selection of appropriate tools to treat or prevent future EVD outbreaks.

Implications of all the available evidence

Advances in sequencing technology will start to bring the generation of sequencing data closer to the front lines of outbreak response. As in-country genomic sequencing capacities are established, sequencing data will need to be rapidly analysed and disseminated in a manner that can inform the public health response. Such capabilities could help public health responders, clinicians and epidemiologists monitor early effects of deployed medical countermeasures and help evaluate their appropriateness over time.

51 (collected from 38 confirmed cases) tested positive for EBOV RNA by quantitative reverse transcription PCR (RT-qPCR). Previous efforts at INRB using Oxford Nanopore Technologies' MinION device confirmed EBOV as the cause of the outbreak by sequencing a small fragment of the EBOV *VP35* gene (polymerase cofactor).[5]

To further characterise the EBOV variant causing the 2018 Équateur Province outbreak, we used the Illumina iSeq100 to obtain whole-genome sequences of EBOV from 38 positive cases, and compared the sequencing data with EBOV sequences from previous outbreaks in Africa. Additionally, we assessed the diagnostic assays, preventive vaccine, and immunotherapeutics against the EBOV variant circulating during the Équateur Province outbreak.

Methods

Sample processing and sequencing

Whole-blood or saliva samples from patients with confirmed EBOV during the 2018 Équateur Province EVD outbreak, which had been subsequently stored at INRB, were used. The samples had been processed in mobile laboratories using the Qiagen QIAamp viral RNA extraction kit (Qiagen, Hilden, Germany) and tested using RT-qPCR assays (appendix) by the outbreak response teams. Samples were re-tested at INRB with the Cepheid Xpert Ebola assay,[6] and RNA quality was assessed with an Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Because of the possibility of RNA degradation, samples were processed for sequencing with viral targeted-enrichment methods, as previously described,[7] and sequenced with an Illumina iSeq100 system (Illumina Technologies, San Diego, CA, USA). The consensus genome sequences were deposited in GenBank (MH733477–MH733491; appendix). Further details of sequencing methods can be found in the appendix.

The study was approved by the Ministry of Health of the Democratic Republic of the Congo and oral consent was obtained from all patients before sample collection.

Molecular evolution and rate analysis

All EBOV near-complete genomes (sequences >17 kb in length) associated with known outbreaks were downloaded from GenBank, along with sampling location and date, and sequences with long stretches of ambiguous bases and duplicates were removed. Alignments were done in Geneious, version 11.1.4, using MAFFT, version 7.388.[8]

To distinguish between a single spillover and multiple spillovers of EBOV into the human population, maximum-likelihood phylogenies were generated by use of RaxML,[9] with the GTRCAT model with 1000 bootstrap replicates. The maximum-likelihood phylogenies were used for estimates of root-to-tip distances, regression slopes, and correlations using TempEst[10] with the best-fitting root option. Tip divergence data were exported and mapped with linear regression (95% prediction interval) in Rstudio, version 3.2.3, with the R Stats and ggplot2 packages.

Bayesian Evolutionary Analysis by Sampling Trees, version 1.8.4,[11] was used to generate Bayesian phylogenies

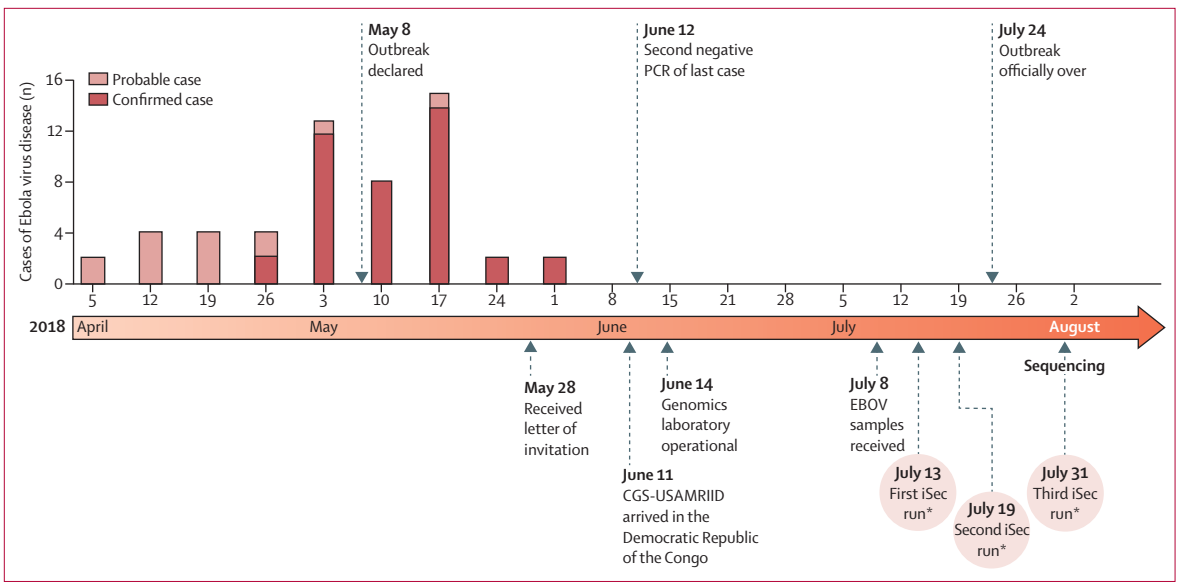

Figure 1: **A timeline of the Ebola virus (EBOV) disease outbreak and establishment of genomics capacity at the Institut National de Recherche Biomédicale, Kinshasa, Democratic Republic of the Congo**
The week of onset of illness for each probable and confirmed case is shown. *Sequences were generated on the following day. CGS-USAMRIID=Center for Genome Sciences-US Army Medical Research Institute of Infectious Diseases.

for each alignment. The initial dataset was reduced to representative sequences from each outbreak that were comparable with previous datasets (appendix).[12,13] These phylogenies incorporated tip date information (ie, sample collection dates) to fit the genetic data to a molecular clock. Epidemiological data from outbreak records were used to predict the month of collection when specific dates were not available for a historical sample; otherwise, dates were parsed with appropriate variable precision. Each analysis consisted of 10×10^8 to 50×10^8 Markov chain Monte Carlo steps (25% of which were discarded as burn-in), and parameters and trees were sampled every 100 000 generations. Extended information about partitions, substitution models, priors, and model testing is presented in the appendix.

We also assessed the time to the most recent common ancestor of the Équateur Province outbreak sequences. If the time to the most recent common ancestor is substantially older than any associated epidemiological data, it could indicate two or more spillover events because the amount of genetic diversity cannot be accounted for within the short timeframe.[14]

Diagnostic assay analysis

During the outbreak, readily available assays targeting the RNA-directed RNA polymerase (*L*) gene,[15] a modified version of a nucleoprotein (*NP*) gene assay,[16] and the multiplexed Cepheid Xpert Ebola assay (targeting glycoprotein [*GP*] and *NP* genes)[6] were used in the mobile laboratories to identify EBOV-positive samples. An in-silico analysis of 25 diagnostic assays was done by mapping primer and probe sequences for each assay to the available EBOV genome sequences from the 2018 Équateur Province outbreak (appendix). The number and position of nucleotide mismatches within each primer and probe were identified and confirmed by a pairwise alignment (appendix).

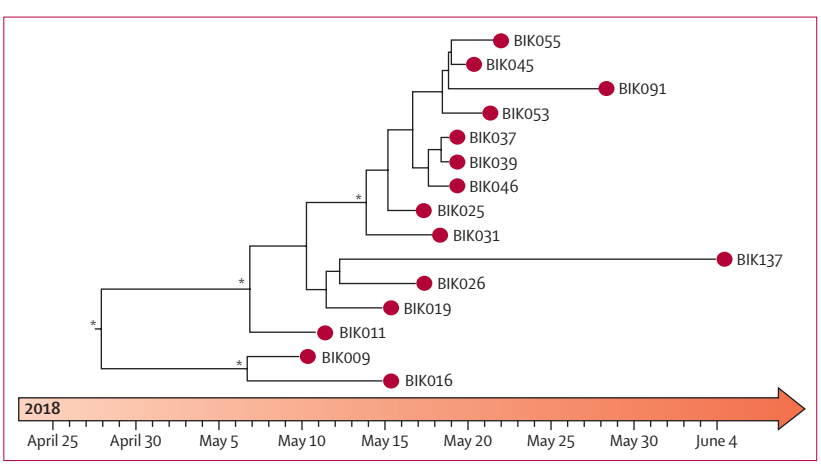

Figure 2: **Phylogenetic analysis of the 2018 Équateur Province Ebola virus disease (EVD) outbreak**
A temporally scaled maximum clade credibility tree from BEAST of the 2018 Équateur Province EVD outbreak with individual variants labelled. *Posterior probabilities of more than 0·7.

Antibody binding confirmation

Multiple experimental monoclonal antibody therapeutics, directed to the EBOV glycoprotein ($GP_{1,2}$), have been developed to treat EVD. Rapid in-silico analyses with the published binding regions for ZMapp and monoclonal antibody 114 were done to predict whether amino acid

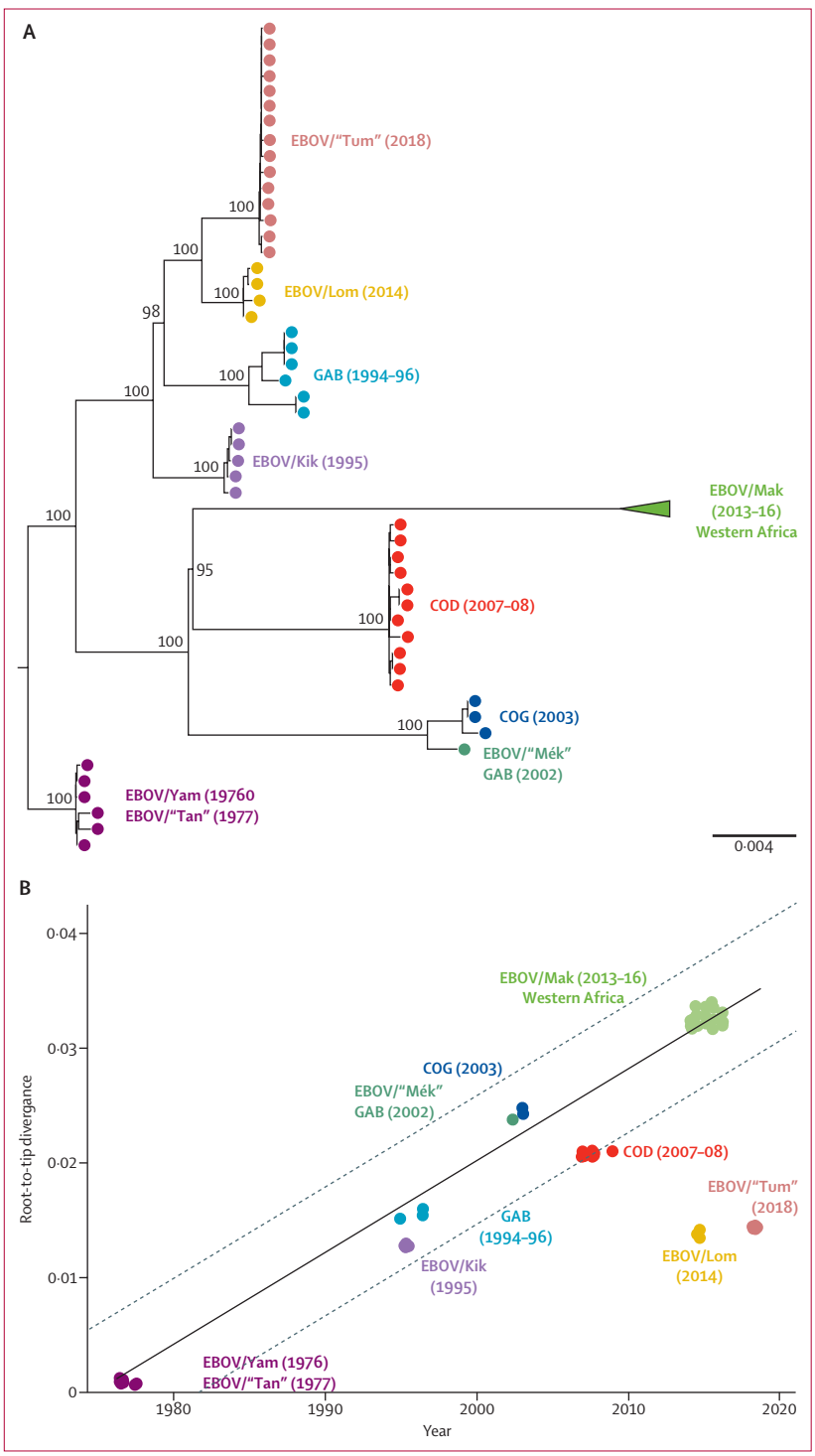

Figure 3: **Phylogenetic and root-to-tip distance analysis of near-complete Ebola virus (EBOV) sequences**
(A) A maximum-likelihood tree outgroup rooted with the EBOV/Yam and EBOV/"Tan" variants using 1026 sequences. Numbers near nodes indicate percentage bootstrap values after 1000 replicates. Trees are scaled by substitutions per site. (B) Root-to-tip divergence linear regression analysis using TempEst and R Studio software of 1026 sequences, with outgroup rooted to the 1976 EBOV Yambuku clade. A 95% prediction interval is shown using dashed lines. The R^2 is 0·52 using Pearson's product-moment correlation, p<0·001. COD=Democratic Republic of the Congo. COG=Republic of the Congo. GAB=Gabon.

changes within the 2018 Équateur Province EBOV variant could affect the binding of antibodies with defined epitopes (appendix).[17,18]

We did in-vitro analyses to test binding affinity between the individual ZMapp monoclonal antibodies (13C6, 2G4, and 4G7) and the $GP_{1,2}$ of the 2018 Équateur Province EBOV variant. Amino acid changes specific to the 2018 Equateur Province EBOV $GP_{1,2}$ were introduced into an EBOV $GP_{1,2}$ Kikwit expression plasmid by site-directed mutagenesis using Phusion polymerase (New England Biolabs, Ipswich, MA, USA) and the primers listed in the appendix (GP^{Kik}; GenBank accession number AAQ55048.1). We generated a construct encoding all the amino acid changes found in the $GP_{1,2}$ of the 2018 Équateur Province EBOV variant ($GP^{"Tum"}$) and an additional construct with the Arg266Gly mutation ($GP^{"Tum"+Arg266Gly}$). Constructs were confirmed with Sanger sequencing (Macrogen USA, Gaithersburg, MD) using in-house primers specific to the plasmid backbone.

Human embryonic kidney-derived 293T cells (CRL-3216; American Type Culture Collection, Manassas, VA, USA) were transfected with GP^{Kik} or one of the two constructs and processed by flow cytometry using the individual ZMapp monoclonal antibodies conjugated with phycoerythrin or allophycocyanin. Monoclonal antibody binding to $GP_{1,2}$ variants of the 2018 Équateur Province EBOV was compared with binding to GP^{Kik} (gating strategy shown in the appendix). Monoclonal antibody 6D8 served as a positive binding control because its epitope is conserved between the EBOV/Kik and the 2018 Équateur Province EBOV variants (appendix). Data were collected with the BD LSRFortessa (BD Biosciences, San Jose, CA, USA) and analysed with FlowJo software, version 10.4.2. Detailed methods are available in the appendix.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

On May 28, 2018, teams from the Institut Pasteur de Dakar (Dakar, Senegal) and the US Army Medical Research Institute of Infectious Diseases (Frederick, MD, USA) were invited to establish a genomics laboratory in Kinshasa, Democratic Republic of the Congo (figure 1). The genomics laboratory was fully operational by June 14, 2018. After training local scientists on the Illumina iSeq100 sequencer, excess diagnostic samples from serum and saliva obtained from 38 patients diagnosed with EVD arrived at INRB, and initial sequencing efforts yielded 16 near-complete viral genome sequences (15 coding-complete and one draft) with a mean genome coverage of 99·64% (SD 0·62; appendix).

Using the newly sequenced genomes, we estimated the most recent common ancestor to be April 26

(95% highest posterior density [HPD] April 10–May 8), 2018, which corresponds with the earliest probable EVD cases (figure 2). We estimated a maximum-likelihood phylogenetic tree incorporating the 15 coding-complete genomes from the 2018 Équateur Province EVD outbreak and 1011 other near-complete EBOV genomes from GenBank (figure 3; appendix). This phylogenetic tree showed that all the sequences from the 2018 Équateur Province EVD outbreak clustered together, which together with the most recent common ancestor, is evidence of a single spillover event. The Équateur Province outbreak sequences are distinct from previous outbreaks, justifying the establishment of a new EBOV variant. Because of the proximity of the outbreak to Lake Tumba, we propose the variant be named "Tumba" (abbreviated to "Tum"), following current conventions.[19]

The topology of the maximum-likelihood phylogenetic tree (figure 3) shows that the closest relative to EBOV/"Tum" is the EBOV variant Lomela (EBOV/Lom) from the 2014 Democratic Republic of the Congo outbreak that occurred in Tshuapa Province, adjacent to Équateur Province (appendix). Similar to EBOV/Lom,[15] EBOV/"Tum" had a shorter branch length than expected (the number of substitutions per site resembled variants sampled from over 20 years ago) demonstrating a lower substitution rate than other EBOV variants (figure 3). We did regression analysis of the root-to-tip distances over time (figure 3). The EBOV variants from outbreaks that started before 2014 lie close to the line of best fit, and within the 95% prediction interval, indicating a strong correlation between time and distance. However, both EBOV/"Tum" and EBOV/Lom fall below the linear regression line and outside the prediction interval. This position suggests a deviation from the clock-like evolution typically associated with EBOV and a slower than expected evolutionary rate.

A Bayesian phylogenetic inference was used to examine the temporal relationships of EBOV/"Tum" within 50 EBOV genomes. The topologies of the maximum-likelihood (figure 3) and Bayesian phylogenetic (appendix) trees were in agreement, and once again, the closest relative of EBOV/"Tum" was EBOV/Lom, sharing a common ancestor around mid-2011 (95% HPD 2007–14; appendix). Estimates of the inter-outbreak (substitutions per site per year estimated using all EBOV variants) rate with inclusion of EBOV/"Tum" and EBOV/Lom converge on a mean rate of $0{\cdot}69\times10^{-3}$ substitutions per site per year (95% HPD $0{\cdot}55$–$0{\cdot}83\times10^{-3}$; appendix).

The use of single representative datasets eliminates the potential bias created by the unequal representation of sequence data from past EVD outbreaks. We calculated inter-outbreak rates from single representative datasets before EBOV/Mak,[20] including EBOV/Mak,[13] and with all outbreaks including EBOV/"Tum" and EBOV/Lom (appendix). A trend was observed where the overall rate decreased with the addition of EBOV/"Tum" and EBOV/Lom.

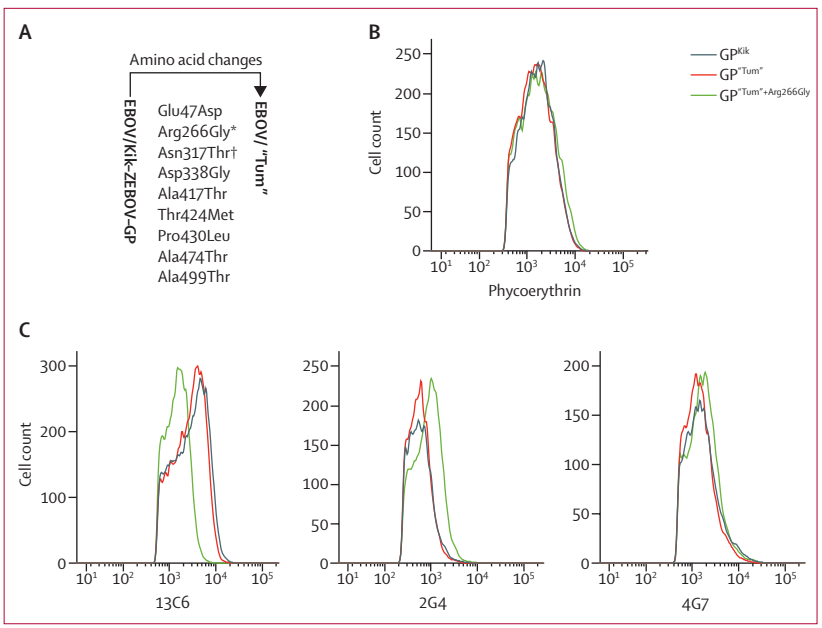

Figure 4: **Binding of ZMapp monoclonal antibodies to EBOV/"Tum" glycoprotein**
(A) Amino acid changes between the EBOV/Kik $GP_{1,2}$ and the 16 EBOV/"Tum" $GP_{1,2}$ sequences. Changes were noted in all sequences, except where marked. Histogram plots compare cell counts with log fluorescence intensity and overlay cell populations bound by the (B) phycoerythrin-conjugated binding control monoclonal antibody 6D8 or (C) the individually allophycocyanin-conjugated monoclonal antibodies 13C6, 2G4, and 4G7. EBOV=Ebola virus. GP=glycoprotein. *Only found in patient sample BIK012. †Only found in patient sample BIK091.

The intra-outbreak (substitutions per site per year estimated from an outbreak of a single EBOV variant) rate is an estimation of the rate of evolution within an outbreak. Before EBOV/"Tum", EBOV/Mak was the only variant with a large number of sequences across a suitable timeframe to calculate an intra-outbreak rate. With the 15 coding-complete EBOV/"Tum" genomes, we estimated an intra-outbreak rate of $3{\cdot}6\times10^{-3}$ (95% HPD $1{\cdot}22$–$5{\cdot}99\times10^{-3}$), a considerably higher mean rate than observed for EBOV/Mak (appendix).

We tested three diagnostic PCR assays in silico for potential target erosion (appendix). Only a single mismatch was found in the middle of the reverse primer in the Cepheid Xpert Ebola *NP* assay. An analysis of additional assays and the polymorphisms and amino acid changes can be found in the appendix.

The EBOV/"Tum" $GP_{1,2}$ sequences were analysed to determine the proximity of any mutations to the known ZMapp monoclonal antibody binding sites (figure 4). Only the Arg266Gly mutation found in patient sample BIK012 raised concerns because of its close proximity to residues Thr270 and Lys272, amino acids known to be part of the 13C6 monoclonal antibody binding site (appendix).

An in-vitro assay was used to assess binding affinity between each individual ZMapp monoclonal antibody (13C6, 2G4, and 4G7) and EBOV/"Tum" $GP_{1,2}$ ($GP^{"Tum"}$ and $GP^{"Tum"+Arg266Gly}$). The monoclonal antibodies readily bound to $GP^{"Tum"}$ and $GP^{"Tum"+Arg266Gly}$ (figure 4), except monoclonal

antibody 13C6, which exhibited reduced fluorescence when tested with $GP^{"Tum"+Arg266Gly}$, indicating decreased binding efficiency of 13C6.

Discussion

The generation of near-complete EBOV genomes was used extensively to characterise the 2013–16 Western African EVD outbreak. Approximately 5% of the confirmed cases were characterised genomically.[21] Initially, data were generated in laboratories outside of the affected countries, with the establishment of in-country sequencing later in the outbreak.[22–24] Within 10 days of the 2018 Équateur Province EVD outbreak declaration, the MinION platform was used to sequence a short amplicon from the first two samples that were transported back to INRB.[5] Less than 6 weeks after the declaration of the outbreak, a sequencing capacity incorporating the Illumina iSeq100 and protocols for generation of coding-complete EBOV genomes was established at INRB. As the outbreak was quickly contained, the near-complete viral genome data generated from the 2018 Équateur Province outbreak became part of a retrospective study. All sequencing data were generated at INRB within a productive international collaboration that included scientists from the Democratic Republic of the Congo, France, Senegal, and the USA, providing a functional in-country genomics capacity that was available for the start of the 2018 EVD outbreak in North Kivu.

The sequencing analysis of the 2018 Équateur Province EVD outbreak revealed that the causative agent was a novel EBOV variant that we are calling "Tumba". EBOV/"Tum" has evolved at a slower rate than other EBOV variants on the basis of on maximum-likelihood phylogenetic estimation, root-to-tip distances, and Bayesian analysis (figures 3; appendix). Reduced rates of evolution have been documented before from EVD cases caused by persistently infected sources[25,26] and from the EBOV variant that caused the 2014 Democratic Republic of the Congo outbreak, EBOV/Lom.[27] EBOV/Lom is, phylogenetically, the closest relative of EBOV/"Tum", and the outbreaks occurred in adjacent provinces. When EBOV/"Tum" and EBOV/Lom sequences were used in the calculation of inter-outbreak rates, we saw a reduction in the mean rate compared with previously calculated rates. All three of the estimates presented in the appendix used datasets of different sizes, which has an effect on the value of the mean rate (appendix). The use of single representatives for each outbreak (appendix) produced the same trend (ie, addition of EBOV/Lom and EBOV/"Tum" reduced the mean inter-outbreak rate) and is a simple way to standardise inter-outbreak rate analysis.

We estimated a high intra-outbreak evolutionary rate compared with EBOV/Mak. However, the high mean value and wide distribution is probably the result of having a small sample size and incomplete purifying selection.[12,13] An early estimation with the first 81 EBOV/Mak genomes calculated a mean rate of $1{\cdot}5\times10^{-3}$ substitutions per site per year, but this decreased to $1{\cdot}2\times10^{-3}$ substitutions per site per year when it was re-estimated at the end of the outbreak with more than 1500 genomes[28] (appendix). Despite differences in rates of evolution between EBOV/Mak and EBOV/"Tum" (figure 3), the intra-outbreak rates of these variants are similar. This similarity suggests that EBOV/"Tum" did not intrinsically acquire mutations at a lower rate than EBOV/Mak because of functional changes within the viral genome—something that can be confirmed through viral growth kinetic studies. An alternative explanation is that EBOV/Lom and EBOV/"Tum" are both maintained in a reservoir characterised by a lower mutation rate, similar to the observation in the persistently infected human cases from the 2013–16 Western African outbreak. However, there is still much to be learned regarding the yet-to-be discovered reservoir hosts of EBOV, the complexity of its maintenance in nature, and the bottlenecks that might occur during spillover into the human population.

As part of the retrospective analysis of the 2018 Équateur Province EVD outbreak, we did rapid in-silico testing on available and deployed diagnostics and medical countermeasures. Of the three deployed RT-qPCR assays, only a single mismatch was found in the Cepheid Xpert Ebola *NP* target region. However, a single mismatch distant from the 3' end of a primer is unlikely to affect sensitivity or specificity.[29] Our data suggest that the amino acid differences in EBOV/"Tum" should have little effect on the rVSVΔG-ZEBOV-GP vaccine (appendix) and immunotherapies. We identified a novel amino acid change (Arg266Gly) in patient sample BIK012 that was in close proximity to the known binding epitope of monoclonal antibody 13C6 (appendix).[17] Within an 8-day timeframe, in-vitro testing revealed a reduced binding affinity, showing that this mutation affected binding of the 13C6 monoclonal antibody to $GP^{"Tum"+Arg266Gly}$. The Arg266Gly mutation was only found in one sample, and the two remaining monoclonal antibodies were not affected by any mutations; thus, if deployed, overall ZMapp effectiveness would likely have been maintained. Overall, these results emphasise the power of sequencing data to enable the quick assessment of immunotherapies targeting the EBOV $GP_{1,2}$.

Our heavy reliance on in-silico analyses is the main limitation in determining whether available medical countermeasures would be effective in treating a novel EBOV variant. However, in-silico analyses are a fast and affordable approach to generate data while an outbreak is occurring, and they help to guide the downstream functional in-vitro and in-vivo studies that are time consuming and costly. We showed the usefulness of in-silico observations directing in-vitro assays, with the rapid discovery of the reduced binding of 13C6 monoclonal antibody to $GP^{"Tum"+Arg266Gly}$. Although these data should not be used in isolation to determine the use of experimental medical countermeasures, they add to the very limited body of evidence. In such cases, we believe these data, coupled with a physician's post-treatment

observations, could be used to direct the choice of treatment when multiple experimental therapies are available.

With the ability to rapidly identify and characterise the causative agent of an outbreak, the most appropriate diagnostics, preventive measures, and treatment options can be deployed. As sequencing becomes more readily available, we anticipate these applications and analyses will become more streamlined and a part of the routine outbreak response procedures. INRB is now equipped to do these analyses in real-time for future disease outbreaks, as already shown in the response to the 2018 North Kivu and Ituri EVD outbreak.

Contributors

PM-K, CBP, MRW, SK, SMG, GPS, EFD, ED, MS-L, MP, J-JM-T, AAS, GP, and SA-M conceived and designed the analysis. PM-K, CBP, AAz, MMD, SM-M, AAy, NV, OusF, OumF, AAr, JN, FM, DM, PM, JK, AM, SD, JL, EP, KC, and JG collected the data. PM-K, CBP, MRW, MMD, SM-M, AAy, NDP, JAC, MD, KC, MLB, TM, and SS did the analysis. PM-K, CBP, MRW, MMD, NDP, KC, JHK, GP, and SA-M wrote the Article.

Declaration of interests

Illumina donated the iSeq100, reagent cartridges, and flow cells to INRB. Illumina paid the costs for attendance of PM-K at the American Society of Tropical Medicine & Hygiene 2018 conference. MRW reports non-financial support from Illumina during the conduct of the study. CBP, MMD, SM-M, AAz, NDP, JAC, MD, AAy, NV, OusF, OumF, SK, AAr, JN, FM, DM, PM, JK, AM, SD, JL, EP, KC, MLB, JG, TM, SS, SMG, GPS, JHK, EFD, ED, MS-L, MP, J-JM-T, AAS, GP, SA-M declare no competing interests.

Acknowledgments

This work was supported by the Defense Biological Product Assurance Office through a task order awarded to the National Strategic Research Institute (FA4600-12-D-9000). We thank Bojan Obradovic, Bryan Crane, Phill Schaecher, Daniel Brekken, Amanda Young, and Jonathan Hetzel from Illumina for reagents and support setting up the iSeq100. We thank Laura Bollinger (Integrated Research Facility-Frederick, Frederick, MD, USA) and Jason Ladner (Northern Arizona University) for critically editing this manuscript. JHK's participation was funded, in part, through Battelle Memorial Institute's prime contract with the US National Institute of Allergy and Infectious Diseases under contract number HHSN272200700016I. This Article represents the opinions of the authors and does not represent the official policy or views of the US Department of the Army, the US Department of Defense, the US Department of Health and Human Services, the US Government, nor the institutions or companies affiliated with the authors.

References

1 Ebola Outbreak Epidemiology Team. Outbreak of Ebola virus disease in the Democratic Republic of the Congo, April–May, 2018: an epidemiological study. *Lancet* 2018; **392:** 213–21.

2 WHO. Ebola virus disease. Democratic Republic of the Congo. External situation report 17: declaration of end of Ebola virus disease outbreak. 2018. https://reliefweb.int/report/democratic-republic-congo/democratic-republic-congo-ebola-virus-disease-external-situation-14 (accessed April 2, 2019).

3 Kuhn JH. Ebolavirus and marburgvirus infections. In: Jameson JL, Fauci AS, Kasper DL, Hauser SL, Longo DL, Loscalzo J, eds. Harrison's principles of internal medicine. 20th edn. Columbus: McGraw-Hill Education, 2018: 1509–15.

4 Mbala-Kingebeni P, Aziza A, Di Paola N. Medical countermeasures during the 2018 Ebola virus disease outbreak in the Nord Kivu and Ituri provinces of the Democratic Republic of the Congo: a rapid genomic assessment. *Lancet Infect Dis* 2019; published online April 15. http://dx.doi.org/10.1016/S1473-3099(19)30118-5.

5 Mbala Kingebeni P, Villabona-Arenas CJ, Vidal N, et al. Rapid confirmation of the Zaire Ebola Virus in the outbreak of the Equateur province in the Democratic Republic of Congo: implications for public health interventions. *Clin Infect Dis* 2019; **68:** 330–33.

6 Pinsky BA, Sahoo MK, Sandlund J, et al. Analytical performance characteristics of the Cepheid GeneXpert Ebola assay for the detection of Ebola virus. *PLoS One* 2015; **10:** e0142216.

7 Mate SE, Kugelman JR, Nyenswah TG, et al. Molecular evidence of sexual transmission of Ebola virus. *N Engl J Med* 2015; **373:** 2448–54.

8 Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013; **30:** 772–80.

9 Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014; **30:** 1312–13.

10 Rambaut A, Lam TT, Max Carvalho L, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2016; **2:** vew007.

11 Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 2012; **29:** 1969–73.

12 Gire SK, Goba A, Andersen KG, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science* 2014; **345:** 1369–72.

13 Park DJ, Dudas G, Wohl S, et al. Ebola virus epidemiology, transmission, and evolution during seven months in Sierra Leone. *Cell* 2015; **161:** 1516–26.

14 Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Genet* 2009; **10:** 540–50.

15 de Wit E, Rosenke K, Fischer RJ, et al. Ebola laboratory response at the eternal love winning Africa campus, Monrovia, Liberia, 2014–2015. *J Infect Dis* 2016; **214** (suppl 3)**:** S169–76.

16 Weidmann M, Muhlberger E, Hufert FT. Rapid detection protocol for filoviruses. *J Clin Virol* 2004; **30:** 94–99.

17 Davidson E, Bryan C, Fong RH, et al. Mechanism of binding to Ebola virus glycoprotein by the ZMapp, ZMAb, and MB-003 cocktail antibodies. *J Virol* 2015; **89:** 10982–92.

18 Murin CD, Fusco ML, Bornholdt ZA, et al. Structures of protective antibodies reveal sites of vulnerability on Ebola virus. *Proc Natl Acad Sci USA* 2014; **111:** 17182–87.

19 Kuhn JH. Guide to the correct use of filoviral nomenclature. *Curr Top Microbiol Immunol* 2017; **411:** 447–60.

20 Carroll SA, Towner JS, Sealy TK, et al. Molecular evolution of viruses of the family Filoviridae based on 97 whole-genome sequences. *J Virol* 2013; **87:** 2608–16.

21 Dudas G, Rambaut A. Phylogenetic analysis of Guinea 2014 EBOV Ebolavirus outbreak. *PLoS Curr* 2014; **6**.

22 Kugelman JR, Wiley MR, Mate S, et al. Monitoring of Ebola virus Makona evolution through establishment of advanced genomic capability in Liberia. *Emerg Infect Dis* 2015; **21:** 1135–43.

23 Quick J, Loman NJ, Duraffour S, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 2016; **530:** 228–32.

24 Smits SL, Pas SD, Reusken CB, et al. Genotypic anomaly in Ebola virus strains circulating in Magazine Wharf area, Freetown, Sierra Leone, 2015. *Euro Surveill* 2015; **20**.

25 Blackley DJ, Wiley MR, Ladner JT, et al. Reduced evolutionary rate in reemerged Ebola virus transmission chains. *Sci Adv* 2016; **2:** e1600378.

26 Whitmer SLM, Ladner JT, Wiley MR, et al. Active Ebola virus replication and heterogeneous evolutionary rates in EVD survivors. *Cell Rep* 2018; **22:** 1159–68.

27 Lam TT, Zhu H, Chong YL, Holmes EC, Guan Y. Puzzling origins of the Ebola outbreak in the Democratic Republic of the Congo, 2014. *J Virol* 2015; **89:** 10130–32.

28 Holmes EC, Dudas G, Rambaut A, Andersen KG. The evolution of Ebola virus: insights from the 2013–2016 epidemic. *Nature* 2016; **538:** 193–200.

29 Lefever S, Pattyn F, Hellemans J, Vandesompele J. Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clin Chem* 2013; **59:** 1470–80.