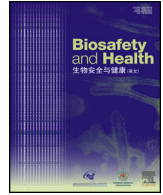




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Intra-host Ebola viral adaption during human infection

William J. Liu<sup>a,1</sup>, Weifeng Shi<sup>b,1</sup>, Wuyang Zhu<sup>a,1</sup>, Cong Jin<sup>a,1</sup>, Shumei Zou<sup>a</sup>, Ji Wang<sup>a</sup>, Yuehua Ke<sup>c</sup>, Xiaofeng Li<sup>d</sup>, Mi Liu<sup>a</sup>, Tao Hu<sup>b</sup>, Hang Fan<sup>d</sup>, Yigang Tong<sup>d</sup>, Xiang Zhao<sup>a</sup>, Wenbin Chen<sup>a</sup>, Yuhui Zhao<sup>e</sup>, Di Liu<sup>e</sup>, Gary Wong<sup>e</sup>, Chengchao Chen<sup>f</sup>, Chunyu Geng<sup>f</sup>, Weiwei Xie<sup>f</sup>, Hui Jiang<sup>f</sup>, Idrissa Laybor Kamara<sup>g</sup>, Abdul Kamara<sup>g</sup>, Matt Leiby<sup>g</sup>, Brima Kargbo<sup>g</sup>, Xiangguo Qiu<sup>h</sup>, Yu Wang<sup>i</sup>, Xiaofeng Liang<sup>i</sup>, Mifang Liang<sup>a</sup>, Xiaoping Dong<sup>a</sup>, Guizhen Wu<sup>a</sup>, George F. Gao<sup>a,e,i,\*</sup>, Yuelong Shu<sup>a,j,\*</sup>

<sup>a</sup> NHC Key Laboratory of Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

<sup>b</sup> Shandong Universities Key Laboratory of Etiology and Epidemiology of Emerging Infectious Diseases, Taishan Medical University, Taian 271000, China

<sup>c</sup> Institute of Disease Control and Prevention, Academy of Military Medical Sciences, Beijing 100071, China

<sup>d</sup> State Key Laboratory of Pathogen and Biosecurity, Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing 100071, China

<sup>e</sup> CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

<sup>f</sup> BGI-Shenzhen, Shenzhen 518083, China

<sup>g</sup> Sierra Leone Ministry of Health and Sanitation, Freetown, Sierra Leone

<sup>h</sup> Special Pathogens Program, National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Canada

<sup>i</sup> Chinese Center for Disease Control and Prevention, Beijing 102206, China

<sup>j</sup> Public Health School (Shenzhen), Sun Yat-sen University, Shenzhen 510006, China

### ARTICLE INFO

#### Article history:

Received 3 October 2018

Received in revised form 14 November 2018

Accepted 8 December 2018

Available online 15 February 2019

#### Keywords:

Ebola virus

Intra-host Nucleotide Variation

Adaptation

Genome sequencing

Clinical manifestations

Blood biochemistry

### ABSTRACT

The onsite next generation sequencing (NGS) of Ebola virus (EBOV) genomes during the 2013–2016 Ebola epidemic in Western Africa provides an opportunity to trace the origin, transmission, and evolution of this virus. Herein, we have diagnosed a cohort of EBOV patients in Sierra Leone in 2015, during the late phase of the outbreak. The surviving EBOV patients had a recovery process characterized by decreasing viremia, fever, and biochemical parameters. EBOV genomes sequenced through the longitudinal blood samples of these patients showed dynamic intra-host substitutions of the virus during acute infection, including the previously described short stretches of 13 serial T>C mutations. Remarkably, within individual patients, samples collected during the early phase of infection possessed Ts at these nucleotide sites, whereas they were replaced by Cs in samples collected in the later phase, suggesting that these short stretches of T>C mutations could emerge independently. In addition, up to a total of 35 nucleotide sites spanning the EBOV genome were mutated coincidentally. Our study showed the dynamic intra-host adaptation of EBOV during patient recovery and gave more insight into the complex EBOV-host interactions.

© 2019 Chinese Medical Association Publishing House. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

### 1. Introduction

The 2014–2016 epidemic of Ebola virus disease (EVD) in Western Africa is the largest EVD outbreak to date, with 28,616 confirmed, probable, and suspected cases and 11,310 deaths as of June 10, 2016 (<http://who.int/csr/disease/ebola/en/>). During this epidemic, the ability to perform next-generation sequencing of Ebola virus (EBOV) from patient specimens in the field [1,2] has generated EBOV genomes, which helped to trace the origin, evolution, and transmission of EBOV in West Africa

[3–9]. Based on the phylogenetic analysis, all of the EBOV in this outbreak can be traced back to a few cases in Guéckédou, Guinea. Subsequently, the virus spread to neighboring countries, such as Sierra Leone, Liberia, Nigeria, Senegal, and Mali [4,10,11].

Along with their geographical spread in West Africa, EBOV formed different lineages [12]. These lineages are characterized with different single nucleotide polymorphisms (SNPs) emerging from different time points during the outbreak [1,4–6,8,13]. These SNPs occur in both the non-coding and encoding regions, containing indispensable phylogenetic and evolutionary information [3,7,14]. Previous research demonstrated that the hotspots for non-synonymous substitutions are likely located in regions with a lower level of functional constraint of the encoded viral proteins [3]. Moreover, the intra-host selection for EBOV to escape from a developing humoral immune response may drive the diversifying selection of glycoprotein (GP) mucin-like domain, as

\* Corresponding authors: National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China.

E-mail addresses: [gaofu@chinacdc.cn](mailto:gaofu@chinacdc.cn) (George F. Gao), [shuyulong@mail.sysu.edu.cn](mailto:shuyulong@mail.sysu.edu.cn) (Yuelong Shu).

<sup>1</sup> William J. Liu, Weifeng Shi, Wuyang Zhu, and Cong Jin contributed equally.

## HIGHLIGHTS

### Scientific question

During 2013–2016 epidemic of Ebola virus disease (EVD) in Western Africa, the Ebola virus (EBOV) genomes generated from the patient specimens based on next-generation sequencing platforms in the field helped to quickly trace the origin, evolution, and transmission of EBOV. In particular, intra-host single nucleotide variations (iSNVs) appeared during the course of the epidemic across the EBOV genome. Whether and how these iSNVs could emerge from a single patient during the disease progression, are still largely unknown.

### Evidence before this study

Some of the iSNVs, such as short stretches of intra-host T to C (T>C) mutations were shared by two or more patients, which represented a combination of human-to-human transmission and recurrent mutations. Furthermore, iSNVs could appear in the key sites, such as the B cell epitopes of GP and non-coding regions across the EBOV genome, which may influence the transcription level of an adjacent gene.

### New findings

In a cohort of EBOV-infected patients in the Ebola Treatment Units (ETUs) in Freetown, Sierra Leone, the recovery processes of the patients were represented by the dynamic viremia and biochemical features. By utilizing longitudinally collected samples during the recovery process, we successfully generated series of virus genomes from the patients. We observed coincident emergence of serial nucleotide variations, including the previously defined short stretch of T>C mutations during the recovery process. Phylogenetic analyses showed that after the T>C mutations occurred, all strains possessing such mutations were grouped together, but were not clustered together with their earlier sequences without such mutations from the same patient. Our results suggested such T>C mutations could arise independently within single patients, reflecting the host-adaptation of EBOV during infection.

### Significance of the study

Our data indicate that short stretches of T>C substitutions are part of the convergent evolution during the infection process of EVD patients, shedding light on the dynamic intra-host genomic variation of EBOV during the 2013–2016 epidemic.

shown by the enrichment of mutations within the B-cell epitopes of GP [4], although this was not observed in Ni's study [15]. On the other hand, short stretches of intra-host T to C (T>C) mutations were also observed [1,4]. This was speculated as a result of adenosine deaminases acting on RNA (ADARs), which is yet unclear [16]. It has been reported that viruses with the 13 T>C mutations (genome positions 5512–5631) continued to circulate in the Magazine Wharf area, Freetown, Sierra Leone, causing several infections [17].

In particular, intra-host single nucleotide variations (iSNVs) appeared during the course of the epidemic, within the B cell epitopes of GP and non-coding regions across the EBOV genome [4,5]. Interestingly, several iSNVs were shared by two or more patients, which represented a combination of human-to-human transmission and recurrent mutations [4]. These iSNVs were used to estimate the effective viral population size within a single patient, during a transmission bottleneck [18], and to identify human-to-human transmission chains [4,5,15]. Furthermore, two iSNVs were able to influence the transcription level of an adjacent gene of nucleocapsid protein (NP) by two-fold [15]. Recent

studies also showed that, during the epidemic, the EBOV isolates from early in the outbreak with amino acid substitutions in the GP protein possessed increased tropism for human cells, indicating human adaptation of Ebola virus during human-to-human transmission [19–21]. However, only a few of the previous studies described longitudinal sequence data from a single patient [5,22], and the intra-host dynamic evolution of EBOV during disease progression is still largely unknown.

At the Sierra Leone-China Friendship Biological Safety Laboratory (SLE-CHN Bio-safety Lab) [23], we cared for a cohort of EBOV-infected patients ( $n = 29$ ) in the Ebola Treatment Units (ETUs) in Freetown, Sierra Leone, from mid-March to late June 2015. The dynamic viremia [24] and biochemical features during disease progression were characterized. By utilizing a deep sequencing platform in the field [25], we successfully generated virus genomes from longitudinally collected blood samples from some of the patients. Surprisingly, we observed coincident emergence of serial nucleotide variations, including the previously defined short stretch of 13 T>C mutations during the recovery process. In a single patient, genome sequences obtained from samples during earlier stages of the acute infection phase possessed Ts at the 13 T>C positions, whereas Cs were found from samples collected during the recovery process. Phylogenetic analyses showed that after the T>C mutations occurred, all strains possessing such mutations were grouped together, but were not clustered together with their earlier sequences without such T>C mutations from the same patient. Our results suggested that such T>C mutations could arise independently within single patients, reflecting the host-adaptation of EBOV during infection.

## 2. Materials and methods

### 2.1. Patients

We undertook a cohort study of patients admitted to Jui Ebola treatment centre (Sierra Leone-China Friendship Hospital) between March 2015, and June 2015 [24]. We used a standard case definition consistent with WHO guidelines. All patients were included in the study except for those who died upon arrival, or those who had no blood results within 24 h of admission. Primary survival outcome measure was collected from the EBOV treatment centre.

### 2.2. Laboratory diagnosis and blood analysis

Samples were tested at the on-site laboratory SLE-CHN Bio-safety Lab for the presence of EBOV RNA by real-time RT-PCR against the glycoprotein (GP) and nucleoprotein (NP) gene targets [24], after inactivation and manual RNA extraction. Positive results were reported as cycle threshold values. All patients received a RT-PCR test upon admission. The Piccolo Express system (Abaxis, CA, USA) was used to generate metabolic and liver function profiles. Amylyte8 was used to assay for blood biochemistry and liver function profile in the same day as the sample collection.

### 2.3. Data collection and ethics statements

All data were collected as part of routine patient care, and recorded on standardized forms, which were kept securely. The clinical data extracted for research purposes were anonymized and stored on a password-protected database. The Sierra Leone ethics and scientific review committee provided approval for the study.

This work was conducted as part of the surveillance and public health response to contain the EVD outbreak in Sierra Leone. Blood samples from suspected individuals and oropharyngeal swab samples from corpses were collected for EVD testing and outbreak surveillance, with a waiver to provide written informed consent during the EVD outbreak under the agreement between the Sierra Leone and Chinese governments. The activities were coordinated by the Emergency Operations Centre in the charge of Sierra Leone Ministry of Health and Sanitation

**Table 1**  
Baseline information of the patients at the date of hospitalization.

Characteristics	Values (N = 29)
Male (%)	12 (41.4)
Age in years, (mean ± SD)	32.1 ± 17.6
Temperature	38.4 ± 1.3
CT value	28.4 ± 6.7
Comorbidities (%)	
Fever	23 (79.3)
Vomiting/nausea	16 (55.2)
Diarrhea	15 (51.7)
Intense fatigue/general weakness	16 (55.2)
Anorexia/loss of appetite	18 (62.1)
Abdominal pain	20 (69.0)
Chest pain	5 (17.2)
Muscle pain	12 (41.4)
Joint pain	16 (55.2)
Headache	13 (44.8)
Cough	9 (31.0)
Difficulty breathing	9 (31.0)
Difficulty swallowing	17 (58.6)
Sore throat	9 (31.0)
Jaundice	4 (13.8)
Conjunctivitis	13 (44.8)
Skin rash	9 (31.0)
Hiccups	8 (27.6)
Pain behind eyes	9 (31.0)
Coma	4 (13.8)
Confusion	2 (6.9)

and WHO. All the information regarding individual persons has been anonymized in the report.

#### 2.4. Genome sequencing and assembly

The genome sequencing was performed as described previously [1]. Briefly, RNA samples extracted from whole blood (two positive oropharyngeal swab samples with high Ct values were not involved) from EVD patients were reverse transcribed to cDNA. PCR amplifications were

performed with 19 EBOV-specific primer pairs with overlaps. Amplicons from one patient were pooled for library preparation. NGS was performed using the BGISEQ-100 (Ion Proton) platform. All sequenced reads were filtered to remove the low quality and short reads. The genome sequences of the viruses were assembled by mapping the filtered reads to the 2014 EBOV consensus sequence (GenBank: KT013255) using Roche 454 Newbler version 2.9 (Roche), and the mutation site was manually checked with original sequencing data.

Clean reads were mapped to the EBOV genome (KJ660346) by TMAP3.4.1. We then scanned the EBOV genome site-by-site and determined the nucleotides for each genomic site according to mapping results. Finally, the ratios of four nucleotides at each site were obtained. Among the samples sequenced, 18 full-length or nearly full-length EBOV genomes were successfully generated in this study (Table S1). The other samples with only short partial of the EBOV genomes were not analyzed in this study.

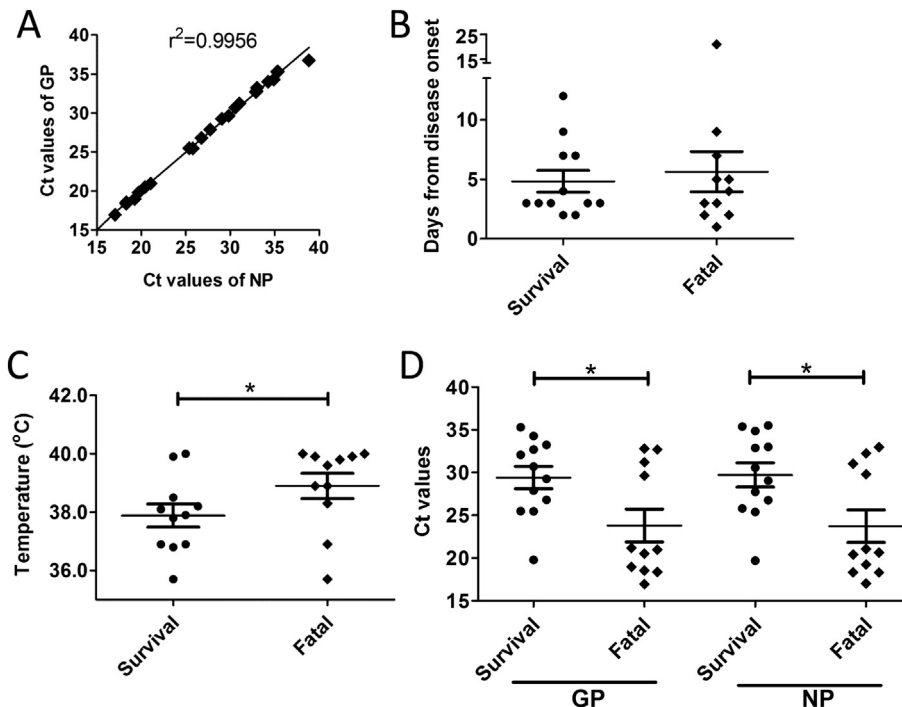
#### 2.5. Sequence analyses

All of the mutation sites were counted for each released EBOV full-length genome, using the Zaire ebolavirus isolate H.sapiens-wt/GIN/2014/Makona-Kissidougou-C15 (GenBank accession No. KJ660346) as the reference sequence. The existence of substitutions within a genome was screened within the 1078 EBOV genomes publicly available from GenBank. A “C strain” was defined as the existence of 3 or more T-C substitutions within a genome window width of 400 nt.

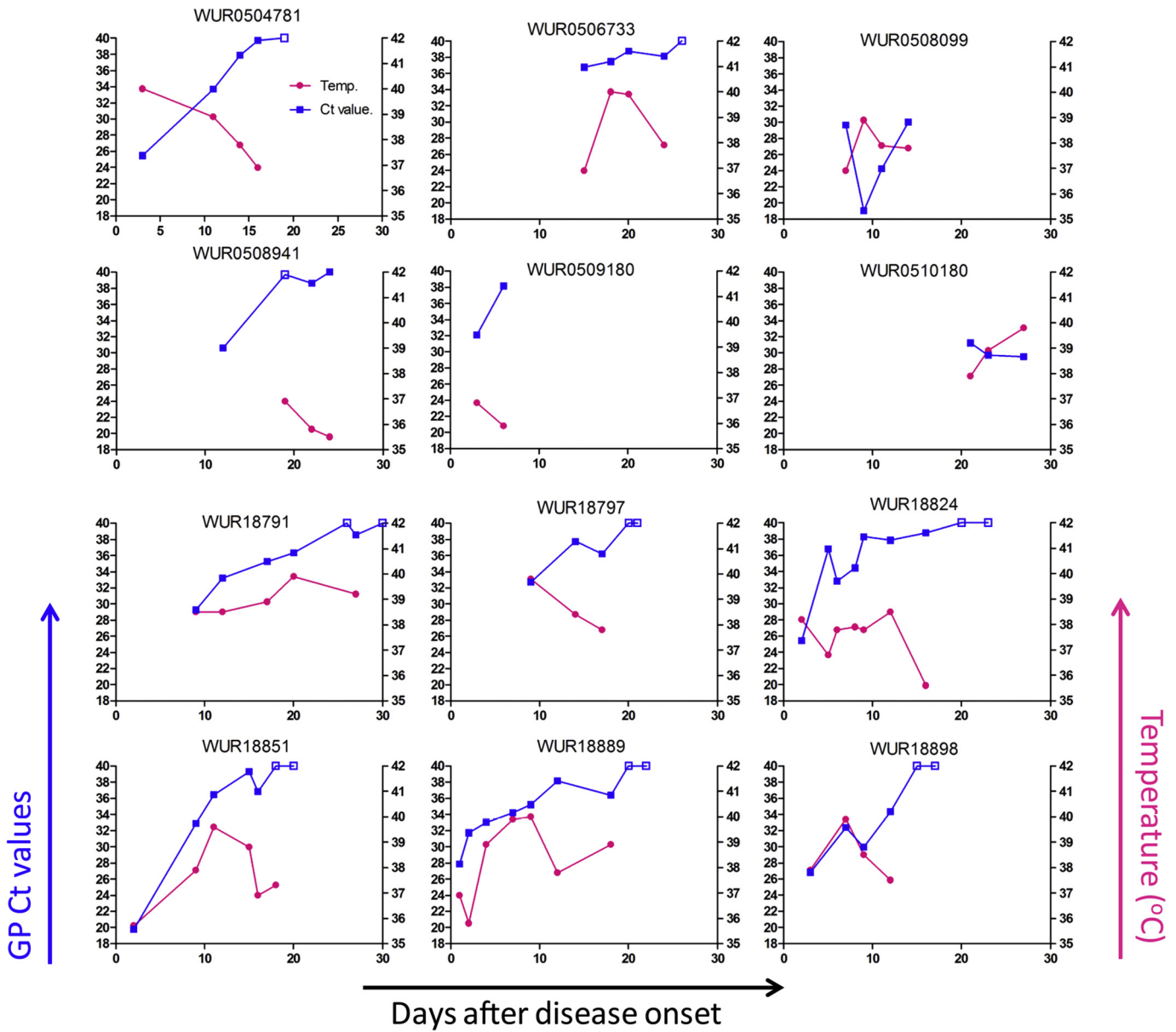
Our dataset included 18 full-length or nearly full-length EBOV genomes sequenced in this study and 1078 EBOV genomes publicly available from GenBank. A maximum-likelihood phylogenetic tree was inferred using the software RAxML, with the GTRGAMMA model and 1000 bootstrap replicates.

#### 2.6. Accession numbers

The genomes and NGS data of the 18 EBOV viruses were deposited in GenBank with the access numbers MF599504–MF599522 (Table S1).



**Figure 1.** The diagnosis and initial clinical features of EBOV patients at hospitalization. A. Correlation of the Ct values from RT-PCR testing in patient blood samples for EBOV GP and NP. B. Average days from symptom onset to initial sampling, as categorized by survival and fatal cases. C. Comparison of temperatures during initial sampling between survival and fatal cases. D. Ct values between survival and fatal cases.



**Figure 2.** The longitudinal clinic variation of the EBOV patients during the disease process. The EBOV shedding and body temperature. The EBOV shedding was represented by the Ct values (blue) tested by RT-PCR targeting EBOV GP in blood. The closed boxes indicate the Ct ≤ 38 (Ebola RNA Positive); while the open boxes denote the Ct > 38 or no Ct available (Ebola RNA Negative). When there are two longitudinally-collected blood samples with Ebola RNA Negative, the patient will be discharged. The corresponding temperatures (purple) during the blood sampling.

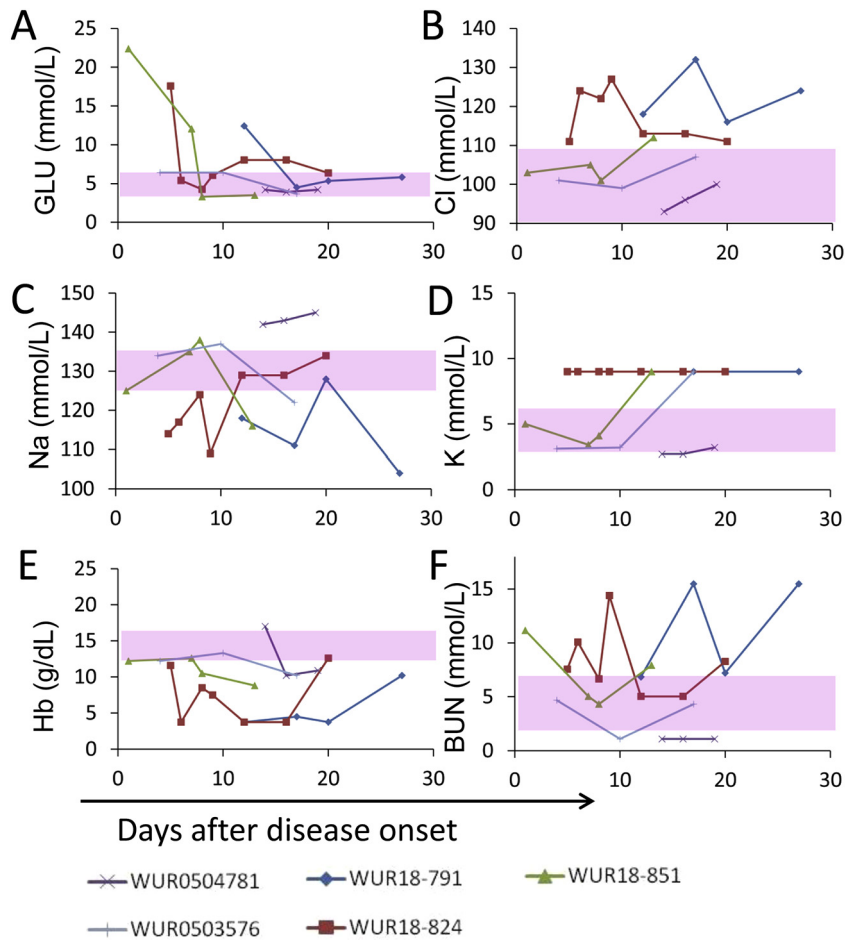
### 3. Results

#### 3.1. Clinical characteristics of the EBOV patients

From March 11th to June 28th, 2015, 473 patients who had symptoms meeting the definition of suspected EVD were admitted to the Sierra Leone-China Friendship Hospital in Freetown. The blood specimens were collected and delivered by the Sample Center of Ministry of Health and Sanitation (MOHS), Sierra Leone. The SLE-CHN Bio-safety Lab tested the blood samples for EBOV through a double-channel real-time RT-PCR detection kit targeting both GP and NP genes of EBOV. A total of 29 (6.1%) patients were confirmed to have EVD (Table 1). The first blood samples collected from these patients after hospitalization were tested for laboratory confirmation, and the Ct values from the GP and NP channels matched well with a high linear correlation (Figure 1A). Of these 29 patients, the most common clinical features at presentation included fever in 23 patients (79.3%; mean temperature, 38.4 °C) and

gastrointestinal symptoms, e.g. abdominal pain, in 20 patients (69.0%) and anorexia/loss of appetite in 18 patients (62.1%) (Table 1). Ocular signs were also common in the patients (13 with conjunctivitis [44.8%] and 9 with pain behind the eyes [31.0%]).

The mean age of the patients was 32 years (range, 10 to 73) and 12 patients (41.4%) were male (Table 1). The case fatality rate of the 29 EVD patients was 41.4% (12/29), similar to the overall ratio of 39.5% (11,310/28,616) during this EBOV epidemic. There was no significant difference ( $P = 0.67$ ) in the average interval from symptom onset to presentation between survivors and non-survivors (Figure 1B). The mean body temperature on the day of hospitalization was 37.9 °C among survivors, significantly lower ( $P = 0.048$ ) than that of non-survivors, 38.9 °C (Figure 1C). The initial viremia of survivors and non-survivors was also significantly different, with non-survivors possessing lower mean Ct values for both GP (23.8 for non-survivors vs 29.4 for survivors,  $P = 0.023$ ) and NP (23.8 for non-survivors vs 29.7 for survivors,  $P = 0.017$ ) (Figure 1D).



**Figure 3.** Blood biochemistry of Ebola patients. Blood parameters were tested at the SLE-CHN Bio-safety Lab in Sierra Leone. The range of values in healthy individuals is denoted as a pink area.

### 3.2. The kinetics of viremia and biochemical features during EVD process

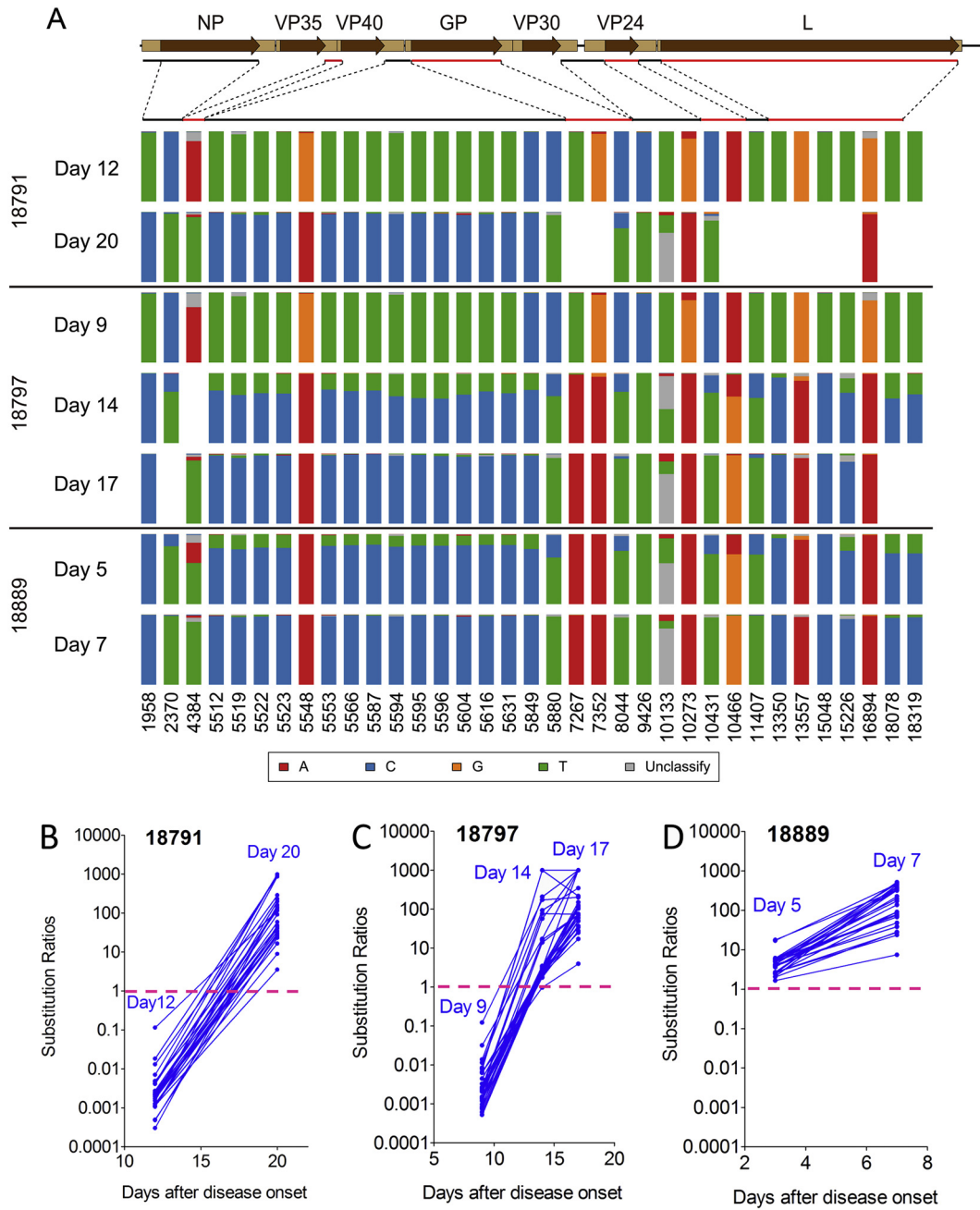
Blood samples from 12 patients (11 survivors and 1 fatality 10180) were available for longitudinal collection during the hospitalization. As in the previous study [26], the viremia of the survivors ameliorated after the presentation, as revealed by increased Ct values (Figure 2). Interestingly, in our study, two different trends in the variation of body temperatures of the patients after hospitalization were observed. Consistent with viremia, body temperatures of patients 4781, 8941, 9180, and 18797 decreased after hospitalization. On the other hand, in patients 6733, 8099, 18791, 18824, 18851, 18889, and 18898, a transient increase in body temperature after hospitalization was detected. Especially in patient 8099, the Ct value had a dramatic decrease from 29.6 on day 7 (hospitalization) to 19.0 on day 9 (2 days after hospitalization), and the body temperature of the patient elevated from 36.9 °C on day 7 to 38.9 °C on day 9. On day 11, the viremia and fever of this patient decreased. In contrast to the survivors, the moribund patient 10180 had a continuous viremia and the body temperature increased from 37.9 °C on day 21 to 39.8 °C on day 27. The patient died on day 30.

Blood biochemical parameters were also tested in the SLE-CHN Bio-safety Lab in the field (Figure 3). Hematological and biochemical abnormalities were observed in the patients upon hospitalization, which was also reported in the previous studies (Biochemical testing in a laboratory tent and semi-intensive care of Ebola patients on-site in a remote part of Guinea: a paradigm shift based on a bleach-sensitive point-of-care device. Clinical, virological, and biological parameters associated with outcomes of Ebola virus infection in Macenta, Guinea.). Hyperglycemia occurred in patients 18791, 18824, 18851 upon hospitalization

and ameliorated afterwards. Abnormal concentrations of blood urea nitrogen, potassium, sodium, and chlorine were observed in the patients, especially in patients 18791 and 18824. Hematological abnormalities detected during the disease progression included reduced concentrations of hemoglobin and hematocrit. These biochemical abnormalities alleviated gradually during the recovery process.

### 3.3. Coincident substitutions during the EVD recovery process

Finally, 14 full-length and four nearly full-length EBOV genomes from 12 patients were obtained based on the NGS platform in the field. Through analysis of the NGS data of longitudinally collected blood samples from the same patient, we found that the previously reported short stretch of T>C mutations within positions 5512–5631 appeared in the late blood samples of the patients 18791 (day 20) and 18797 (days 14 and 17), while these patients still possessed Ts in early blood samples (day 12 for patient 18791 and day 9 for 18797) (Figure 4A). In particular, in the day 9 sample of patient 18797, almost no Cs were observed at the 13 positions, whereas Cs accounted for approximately 25% on day 14 and ~100% on day 17 at all of the 13 T>C nucleotide sites (Figure 4A), indicating that day 14 may be an intermediate status between days 9 and 17. The similar ratios of T:C at the different T>C nucleotide sites on the same day may indicate that these T>C mutations occurred concurrently, though not all these mutations can be found on the same nucleic acid strands (reads). However, in patient 18889, dominant percentages of Cs (>86%) were already observed on day 5, and continued increasing in percentage on day 7 (~99%) (Figure 4A).



**Figure 4.** The coincident transition of the 35 coincident nucleotide variations of EBOV during the recovery process. A. Only nucleotide positions with >100 reads were estimated. The red and black lines below the genome structure were used to highlight the specific genomic regions of EBOV with iSNVs. B–D. The similar substitution trends of the EBOV iSNV in longitudinally collected blood samples of patients 18791, 18797 and 18889. The substitution ratio of each point = the reads of the dominant iSNV at the later time/the reads of the dominant iSNV at the early time of the disease. The 50% substitution (ratio = 1) was shown in purple dashed line. The list of the iSNV sites is present in Table 2. The exceptional sites which do not have the substitution trends are shown in Supplementary Figure S1.

Meanwhile, we found 22 additional coincident substitutions distributed across the entire EBOV genome (Table 2 and Figure 4), which has the similar ratios for the substitutions as in the short stretch of the 13 T>C mutations between positions 5512 and 5631. This may also indicate that all these substitutions were correlated to each other. These nucleotide variations not only included T>C substitutions, but also other types of substitutions, e.g. C>T, G>A, A>G, A>T, T>A, and even one T>Deletion mutation at site 10133 (Table S2). These nucleotide variations occurred coincidentally among different patients during the acute infection phase.

The intra-host adaptation of EBOV was also illustrated as the continuous change of iSNV substitution ratios (the ratios of latter dominant

iSNV/former dominant iSNV, for instance, ratios of C/T for the 13 T>C nucleotide sites) across longitudinal sampling points in patients 18791, 18797, and 18889 (Figures 4B–D). The iSNVs at most sites have a synchronized change in the three patients during infection. The substitution ratios for patients 18791 increased from 0.0021 (median, with lower quartile [Q1] = 0.0012 and upper quartile [Q3] = 0.0042) on day 12 to 59.17 (median, with Q1 = 30.33 and Q3 = 140.5) on day 20 (Figure 4B). For patient 18797, the substitution ratios increased from 0.0023 (median, with Q1 = 0.0013 and Q3 = 0.0062) on day 9 to 75.22 (median, with Q1 = 40.1 and Q3 = 189.375) on day 17, with a transition ratio of 2.87 (median, with Q1 = 2.22 and Q3 = 16.24)

**Table 2**  
The linked and convergent substitutions of Ebola viruses emerged during human infection.

Number	Site	Type	Former	Latter	18791			18797			18889			C stains (5512–5631) <sup>h</sup>	
					Day 12	Day 20	Figure	Day 9	Day 14	Day 17	Figure	Day 5	Day 7		Figure
1	1958	<b>NP(S497P)<sup>a</sup></b>	T	C	0.01 <sup>b</sup>	289.00 <sup>b</sup>	1 <sup>c</sup>	0.01	207.00	1000.00 <sup>e</sup>	1	1103.31	660.36	0 <sup>d</sup>	8/8
2	2370	<b>NP(A634V)</b>	C	T	0.00	48.50	1	0.00	2.82	N/A <sup>f</sup>	1	4.89	76.14	1	8/8
3	4384	VP35–VP40 <sup>g</sup>	A	T	0.00	26.08	1	0.00	2.11	17.27	1	2.04	28.00	1	7/8
4	5512	VP40–GP	T	C	0.00	865.00	1	0.00	3.06	76.20	1	4.21	346.40	1	7/8
5	5519	VP40–GP	T	C	0.00	43.50	1	0.00	2.22	26.00	1	3.90	48.13	1	8/8
6	5522	VP40–GP	T	C	0.00	23.24	1	0.00	2.47	35.00	1	4.19	38.47	1	8/8
7	5523	VP40–GP	T	C	0.00	135.33	1	0.00	2.50	69.80	1	4.25	516.36	1	8/8
8	5548	VP40–GP	G	A	0.02	873.00	1	0.01	92.50	347.00	1	431.71	631.98	0	8/8
9	5553	VP40–GP	T	C	0.00	137.83	1	0.00	3.38	108.33	1	5.47	388.47	1	8/8
10	5566	VP40–GP	T	C	0.00	199.50	1	0.00	2.94	150.00	1	5.43	195.73	1	8/8
11	5587	VP40–GP	T	C	0.00	43.67	1	0.00	3.07	1000.00	1	5.68	75.12	1	8/8
12	5594	VP40–GP	T	C	0.00	231.00	1	0.01	2.13	121.00	1	5.36	344.04	1	8/8
13	5595	VP40–GP	T	C	0.00	59.17	1	0.00	1.82	49.40	1	5.16	172.81	1	8/8
14	5596	VP40–GP	T	C	0.00	92.00	1	0.00	1.76	27.78	1	5.06	79.54	1	8/8
15	5604	VP40–GP	T	C	0.00	140.50	1	0.00	2.35	37.00	1	6.12	468.33	1	8/8
16	5616	VP40–GP	T	C	0.00	40.62	1	0.00	2.87	59.40	1	5.70	67.39	1	8/8
17	5631	VP40–GP	T	C	0.00	119.00	1	0.01	2.56	103.33	1	5.77	315.38	1	8/8
18	5849	VP40–GP	T	C	190.00	85.00	0 <sup>d</sup>	163.22	3.20	44.91	0	3.74	137.60	1	8/8
19	5880	VP40–GP	C	T	0.00	161.70	1	0.00	2.14	113.29	1	2.08	441.73	1	7/8
20	7267	<b>GP(R410S)</b>	T	A	0.00	N/A	N/A	0.00	57.00	1000.00	1	1281.65	2044.63	0	0/8
21	7352	<b>GP(E439K)</b>	G	A	0.03	N/A	N/A	0.03	17.33	1000.00	1	1102.30	495.67	0	8/8
22	8044	GP(F669F)	C	T	0.00	3.55	1	0.00	2.87	24.68	1	3.60	225.08	1	7/8
23	9426	VP30–VP24	C	T	0.01	1000.00	1	0.01	1000.00	1000.00	1	792.13	473.41	0	8/8
24	10133	VP30–VP24	T	Del	0.00	1000.00	1	0.01	0.97	3.95	1	1.65	7.45	1	6/8
25	10273	VP30–VP24	G	A	0.11	116.00	1	0.12	173.00	202.50	1	330.20	317.62	0	8/8
26	10431	VP24(F29F)	C	T	0.00	30.33	1	0.00	2.91	125.67	1	2.70	321.75	1	7/8
27	10466	<b>VP24(Y41C)</b>	A	G	0.00	28.33	1	0.00	2.11	50.57	1	2.53	382.21	1	7/8
28	11407	VP24–L	C	T	113.88	N/A	N/A	133.23	1.86	17.14	0	2.44	91.21	1	8/8
29	13350	<b>L(END591Q)</b>	T	C	0.00	N/A	N/A	0.00	16.24	72.00	1	17.68	321.12	1	7/8
30	13557	<b>L(E657K)</b>	G	A	0.00	N/A	N/A	0.00	13.98	71.33	1	16.96	466.17	1	7/8
31	15048	<b>L(L1157P)</b>	T	C	0.01	N/A	N/A	0.01	1000.00	219.33	1	1216.72	937.44	0	8/8
32	15226	L(P1216P)	T	C	0.00	N/A	N/A	0.00	3.49	34.14	1	3.91	90.47	1	7/8
33	16894	<b>L(A1773T)</b>	G	A	0.00	34.50	1	0.00	77.00	74.25	1	93.06	61.84	0	8/8
34	18078	L(H2187H)	T	C	0.00	16.50	1	0.00	1.75	N/A	1	2.68	23.47	1	8/8
35	18319	5'-UTR	T	C	0.00	9.00	1	0.00	2.35	N/A	1	2.64	27.41	1	8/8

<sup>a</sup> The nonsynonymous mutations are emphasized by bold fonts.

<sup>b</sup> The ratio of the reads of Latter/Former.

<sup>c</sup> “1” means that the corresponding data of the survivor was shown in Figure 4B–D.

<sup>d</sup> “0” means that the corresponding data of the survivor was shown in Supplementary Figure S1.

<sup>e</sup> The ratio is termed as “1000” when the read of the former nucleotide is 0.

<sup>f</sup> “N/A” means that the reading depths of the NGS was lower than 50.

<sup>g</sup> “VP35–VP40” means that the substitutions locate in the non-coding region between VP35 and VP40.

<sup>h</sup> “C stains (5512–5631)” indicate the 8 stains from publicly released full-length EBOV genome sequences from GenBank, including KT961624, KP759706, KU296370, KU296319, LN877955, KT357858, KT357859 and KT357860. These 8 stains possess the short stretch of T>C mutations within positions 5512–5631. Furthermore, these strains also have the linked mutations in other sites. The proportions of the strains which possess the mutations were shown in fractions.

on day 14 (Figure 4C). For patient 18889, substitution ratios remained >1, indicating a dominant role of the latter iSNV on both day 5 and day 7 (Figure 4D), however, the increasing trend of the iSNV substitution ratios could still be observed from day 5 to day 7.

However, there were still some exceptional sites which had consistent iSNVs during the disease process (Supplementary Figure S1). All exceptional sites possessed the latter form at the target sites (e.g. Site 5849 of viral genome had already mutated to C in patient 18791 on day 12), indicating early substitution events at these sites.

### 3.4. Phylogenetic association of the T>C strains

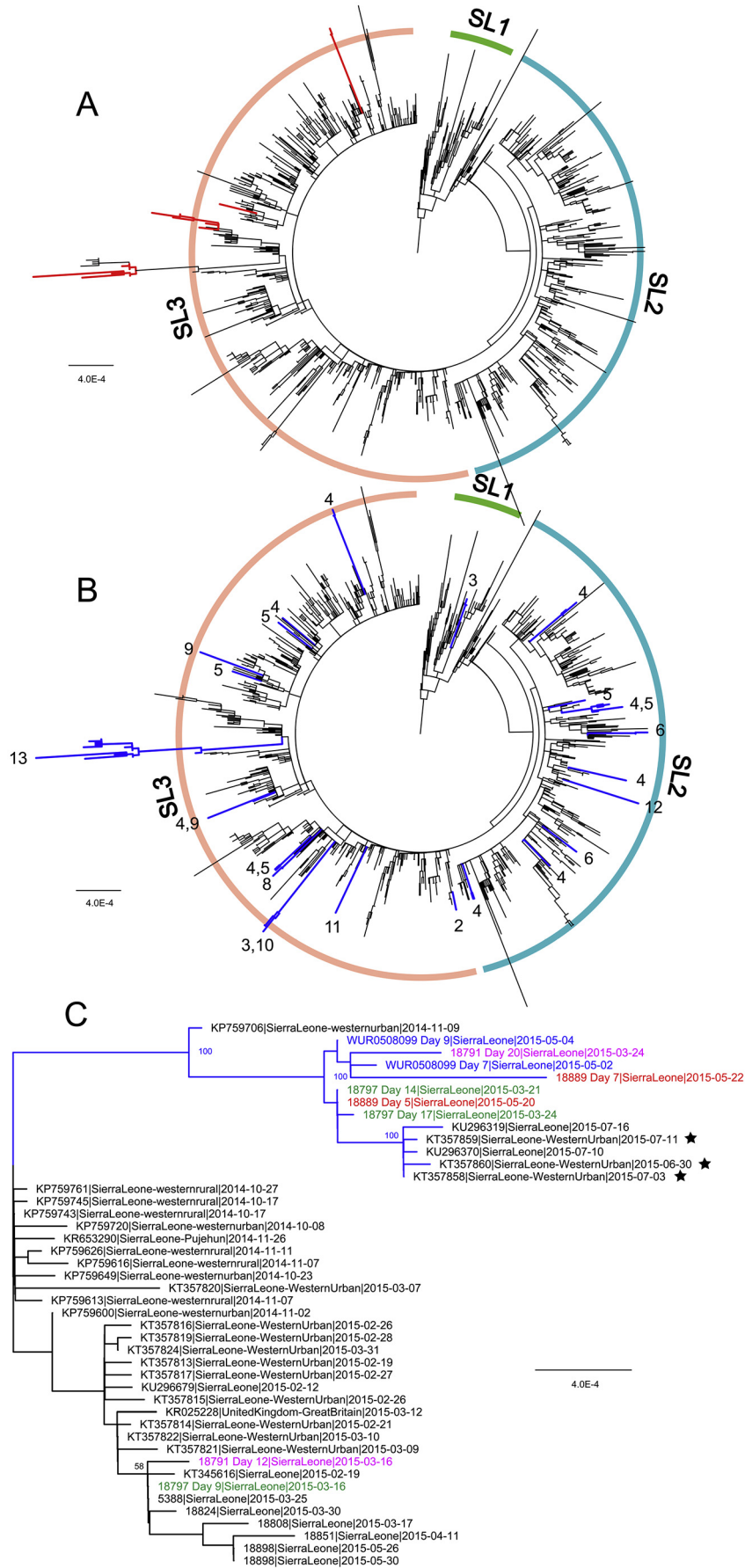
Phylogenetic analysis of the 18 novel and 1078 publicly released full-length EBOV genome sequences from GenBank showed that the novel EBOV sequences did not cluster together (Figure 5A). Alternatively, they formed four small independent clusters scattered across lineage SL3, which was circulating during the late stage of the outbreak in 2015.

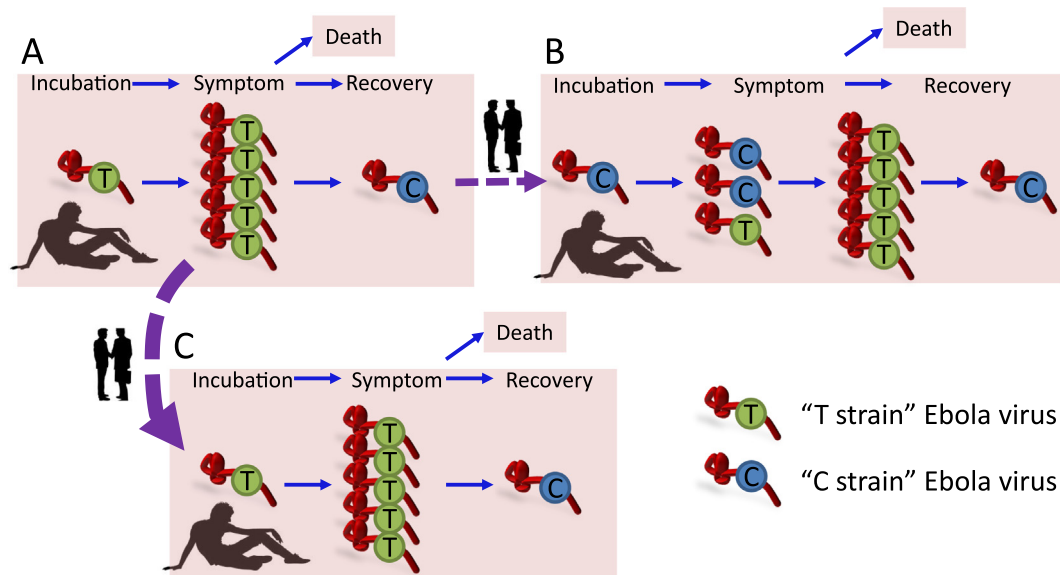
We further analyzed all EBOV genomes available online and identified the strains with multiple T>C substitutions within a short genomic region in different positions of the genomes (termed as “C strains”) (Figure 5B and Table S3). A total of 49 strains, including six sequenced in this study, possessed multiple T>C nucleotide substitutions (Table

S3). These “C strains” could be classified into at least 24 different types according to the number of T>C substitutions and their positions (Figure 5B). Interestingly, apart from the well-known stretch with 13 serial T>C mutations at genome position 5512–5631, stretches with 11 and 12 serial T>C mutations were also found in the intergenic region between VP40 and GP in strains G4415.1 (GenBank No. KR105247) and LIBR10051 (GenBank No. KT725368) (Table S3), respectively. To study the phylogenetic association of these “C strains”, we mapped them onto the tree (Figure 5B and Supplementary Figure S2). However, these 49 “C strains” scattered across the entire tree, with little phylogenetic association (Figure 5B), indicating multiple independent origins of these T>C substitutions. However, some strains possessing the same T>C substitution type were clustered together (Supplementary Figure S2). Based on current evidence, human-to-human transmission is the most plausible reason for these cases.

In particular, ten strains possessed the 13 T>C mutations within positions 5512–5631 (Figure 5C). The prototype strain with the 13 T>C mutations was J0169 (KP759706), sequenced from the Magazine Wharf area, Freetown, Sierra Leone in November 2014 [17]. It has been reported that J0169-like EBOV continued to circulate in this region and had infected at least three additional patients by July 2015. However, the strain sequenced from patient 18797 on days 14 and 17







**Figure 6.** The schematic diagram for the dynamic intra-host substitution and inter-host transmission of EBOV. Based on the dynamic adaptation of Ebola virus in the patients we described herein, we further proposed the schematic model for the relationship of the human to human transmission and the dynamic adaptation of the Ebola viruses. Panel A shows the disease process of one survived patient A from the preclinical period to symptom presentation, and then to recovery. During this process, the dominant viruses in the patient are the T strain which possess the former nucleic acids in the 35 sites (Table 2), e.g. T in the 13 T>C stretch (5512–5631). There also will be emergence of the C strains during the recovery process, which possess the latter nucleic acids in the 35 sites (Table 2), e.g. C in the 13 T>C stretch (5512–5631) due to unknown reasons as we indicated in our patients. T strain virus dominates in the acute detoxification period of patients and certain patients died during this period. Thus in the general human to human transmission of Ebola virus (patient A to patient C), the transmitted viruses are T strain virus, which takes the dominates in the 1078 EBOV genomes of the West Africa outbreak publicly available from GenBank. However, we cannot exclude the possibility of the sporadic transmission of C strains in humans with close contact (A to B). Patient B, who was infected by the C strain of Ebola virus, may have a mixed quasi-species of the virus during the diseases process. This possibility requires further exploration.

possessed the 13 T>C mutations and clustered together with other such “C strains”. Surprisingly, on day 9 of patient 18797, the virus possessed Ts at all of the 13 positions, and did not fall within this cluster. This was also observed with patient 18791 (Figure 5C). Virus samples collected on day 12 from patient 18791 was located in a different cluster as the virus sequenced from the day 20 sample.

### 3.5. Co-current substitutions within the publicly released EBOV genomes

We further identified eight EBOV strains sharing the previously described 13 T>C mutations within positions 5512–5631 (the last column in Table 2). We called all the other viruses (1070 strains) as “T strains (5512–5631)”. Interestingly, we found that almost all 35 coincident substitutions that happened in our longitudinally sequenced EBOV genomes (Table 2) had already mutated in these eight “C strains (5512–5631)”. In contrast, none of the other EBOV “T strains (5512–5631)” without the 13 T>C mutations at position 5512–5631 have these substitutions. These data suggest that all substitutions distributed within the whole genome of EBOV, including the short stretch of T>C mutations within position 5512–5631, had a coincident evolution trend during the recovery process of the patients.

## 4. Discussion

In this study, we sequenced longitudinally-collected blood samples from EBOV patients based on an in-fieldNGS platform at the SLE-CHN Bio-safety Lab during the 2014–2016 EBOV epidemic. Although more a thousand of EBOV genomes have been sequenced during the 2014–2016 EBOV epidemic in West Africa, most of the viruses were from different patients during the early days of the acute phase during

the infection [1–5,7,8,13,17]. Phylogenetic analysis showed that the 18 genomes sequenced in this study were not grouped together. This suggested that in the late stages of the 2014–2016 EBOV outbreak, the virus became diversified and various minor EBOV lineages had been co-circulating in Sierra Leone.

Previous phylogenetic studies indicated that SNPs carried by different EBOV lineages were important molecular markers to study the virus transmission among humans [1,4,13]. Short stretches of the T>C substitutions appeared sporadically without phylogenetic association as shown in our analysis (Figure 5B). Although the exact reason is yet unclear, this was speculated to be the result of adenosine deaminases acting on RNA (ADARs) [16]. Thus, it is understandable that the T>C substitutions can occur in both the coding regions and the non-coding regions. In the present study, we found several related iSNVs across the EBOV genome occurring in the late acute phase of EBOV infection, including a short stretch of 13 T>C mutations within positions 5512–5631. However, these mutations were not found in samples collected during the early infection phase. In one patient, we even observed the intermediate stage for these T>C mutations. These data suggest that the 13 serial T>C mutations were emerged coincidentally, and that the mutations could arise independently in different EBOV patients. Since only a few T>C mutations are able to change the transcription level of adjacent genes [15], such serial T>C mutations in our study, though some in non-coding regions, might be functional during EBOV infection in humans, which should be a subject of further investigation. Meanwhile, further analysis should be performed to investigate whether all these mutations can be found on the same nucleic acid strands (reads).

Phylogenetic analysis showed that earlier samples without the 13 T>C mutations were not grouped together, whereas later samples with the 13 T>C mutations were clustered together and formed a separate cluster. We

**Figure 5.** Phylogenetic analysis of the newly characterized and public full-length EBOV genome sequences. A. Sequences characterized in the present study are marked in red. B. All of the strains with serial T>C mutations are marked in blue. The numbers of serial T>C mutations are given. C. Detailed phylogenetic relationships of the strains with 13 serial T>C mutations. In this panel, sequences in the same color are obtained from the same patient at different time points. Sequences marked with black stars are reported in this paper.

also identified a total of 49 strains with serial T>C substitutions that belonged to 24 different T>C mutation types, with different number of T>C substitutions and/or within different genomic regions. Generally, strains possessing different T>C mutation types were not clustered together, but scattered across the tree. Based on current evidence, there was no direct phylogenetic association between these different mutation types. This once again revealed that these various T>C mutation types could arise independently. However, some of the T>C substitutions were indeed shared by a few strains, which were clustered together in the tree. Thus, the transmission hypothesis cannot be fully rejected yet [17].

This study is the first showing that linked and coincident mutations observed in EBOV could arise independently during human infections. It can be rationally proposed that the “T strains” are the dominant circulating EBOVs during the epidemic (Figure 6). Most of the cases are infected via contact with the patients during the early stages of acute infection. During the late stages of the infection, the “C strains” emerge, although the virus titer has become lower during this period. Transmission may also occur rarely between close contacts at this stage of disease. This may explain why the “C strains” could be found in such few patients in the acute infection phase. Meanwhile, the T>C mutation happened at different time points of the patients after EBOV infection, which may be related with disease progress or prognosis. For patient 18791, the corresponding sites for these T>C were still T, while the T>C substitution was completed on day 5 of the infection of patient 18889. The fever duration and the viremia duration of patient 18791 were longer than 18889. However, the relationship of the occurrence time of these mutations with disease progress or prognosis should be further investigated.

Overall, our data indicate that short stretches of T>C substitutions are part of the linked and convergent evolution during the infection process of EVD patients, shedding light on the dynamic intra-host adaptation of EBOV during the 2014–2016 EBOV epidemic.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bsheat.2019.02.001>.

## Acknowledgements

Hundreds of Chinese and Sierra Leonean staff on-site took part in the sample collection, transportation, detection, and data analysis. We especially acknowledge the excellent work of four previous detection teams of ChinaCDC (mobile and fixed labs) and the essential support (materials and reagents) of the National Institute for Viral Disease Control and Prevention, China CDC. We also express our deep condolences to the family and associates of Dr. Abdul Kamara of the Ministry of Health and Sanitation, Sierra Leone, who contributed a lot to this project but succumbed after the successful control of Ebola epidemic in Sierra Leone.

This work was supported by the Megaproject for Infectious Disease Research of China(2016ZX10004222-003), the research of Ebola pathogen from the National Natural Science Foundation of China (NSFC, 81590763), National Key Research and Development Program of China(2016YFC1200200 to Y. Shu), the Distinguished Young Scientist Program of the NSFC (81525017 to Y. Shu), the Excellent Young Scientist Program of the NSFC(81822040 to W.J. Liu), the Taishan Scholar Project of Shandong Province(ts201511056 to W. Shi). G.F. Gao is a primary principal investigator of the NSFC Innovative Research Group (81621091).

## Conflict of interest statement

The authors declare that there are no conflicts of interest.

## Author contributions

W.J. Liu, W. Shi, W. Zhu, C. Jin, S. Zou, J. Wang, Y. Ke, X. Li, and M. Liu performed the experiments. C. Chen, C. Geng, W. Xie, H. Jiang, I.L. Kamara, and A. Kamara provided the necessary materials. W.J. Liu, W. Shi, T. Hu, H. Fan, Y. Tong, X. Zhao, W. Chen, Y. Zhao, and D. Liu

performed the data analysis. G. Wong, M. Lebby, B. Kargbo, X. Qiu, Y. Wang, X. Liang, M. Liang, X. Dong, and G. Wu contributed to fruitful discussions and key ideas. W.J. Liu, W. Shi, G.F. Gao, and Y. Shu contributed to the overall concept, experimental design and hypothesis, and wrote the manuscript.

## References

- [1] Y.G. Tong, W.F. Shi, D. Liu, J. Qian, L. Liang, X.C. Bo, J. Liu, H.G. Ren, H. Fan, M. Ni, Y. Sun, Y. Jin, Y. Teng, Z. Li, D. Kargbo, F. Dafaie, A. Kanu, C.C. Chen, Z.H. Lan, H. Jiang, Y. Luo, H.J. Lu, X.G. Zhang, F. Yang, Y. Hu, Y.X. Cao, Y.Q. Deng, H.X. Su, Y. Sun, W.S. Liu, Z. Wang, C.Y. Wang, Z.Y. Bu, Z.D. Guo, L.B. Zhang, W.M. Nie, C.Q. Bai, C.H. Sun, X.P. An, P.S. Xu, X.L. Zhang, Y. Huang, Z.Q. Mi, D. Yu, H.W. Yao, Y. Feng, Z.P. Xia, X.X. Zheng, S.T. Yang, B. Lu, et al., Genetic diversity and evolutionary dynamics of Ebola virus in Sierra Leone, *Nature* 524 (2015) 93–96.
- [2] J. Quick, N.J. Loman, S. Duraffour, J.T. Simpson, E. Severi, L. Cowley, J.A. Bore, R. Koundouno, G. Dudas, A. Mikhail, N. Ouedraogo, B. Afrough, A. Bah, J.L. Baum, B. Becker-Ziaja, J.P. Boettcher, M. Cabeza-Cabrero, A. Camino-Sanchez, L.L. Carter, J. Doerrbecker, T. Enkirch, I. Garcia-Dorival, N. Hetzelt, J. Hinzmann, T. Holm, L.E. Kafetzopoulou, M. Koropogui, A. Kosgey, E. Kuisma, C.H. Logue, A. Mazzarelli, S. Meisel, M. Mertens, J. Michel, D. Ngabo, K. Nitzsche, E. Pallasch, L.V. Patrono, J. Portmann, J.G. Repits, N.Y. Rickett, A. Sachse, K. Singethan, I. Vitoriano, R.L. Yemanaberhan, E.G. Zekeng, T. Racine, A. Bello, A.A. Sall, O. Faye, et al., Real-time, portable genome sequencing for Ebola surveillance, *Nature* 530 (2016) 228–232.
- [3] J.T. Ladner, M.R. Wiley, S. Mate, G. Dudas, K. Prieto, S. Lovett, E.R. Nagle, B. Beitzel, M.L. Gilbert, L. Fakoli, J.W. Diclaro, R.J. Schoepp, J. Fair, J.H. Kuhn, L.E. Hensley, D.J. Park, P.C. Sabeti, A. Rambaut, M. Sanchez-Lockhart, F.K. Bolay, J.R. Kugelman, G. Palacios, Evolution and spread of Ebola virus in Liberia, 2014–2015, *Cell Host Microbe* 18 (2015) 659–669.
- [4] D.J. Park, G. Dudas, S. Wohl, A. Goba, S.L. Whitmer, K.G. Andersen, R.S. Sealfon, J.T. Ladner, J.R. Kugelman, C.B. Matranga, S.M. Winnicki, J. Qu, S.K. Gire, A. Gladden-Young, S. Jalloh, D. Nosamiefan, N.L. Yozwiak, L.M. Moses, P.P. Jiang, A.E. Lin, S.F. Schaffner, B. Bird, J. Towner, M. Mamoh, M. Gbakie, L. Kanneh, D. Kargbo, J.L. Massally, F.K. Kamara, E. Konuwa, J. Sellu, A.A. Jalloh, I. Mustapha, M. Foday, M. Yillah, B.R. Erickson, T. Sealy, D. Blau, C. Paddock, A. Brault, B. Amman, J. Basile, S. Bearden, J. Belser, E. Bergeron, S. Campbell, A. Chakrabarti, K. Dodd, M. Flint, A. Gibbons, et al., Ebola virus epidemiology, transmission, and evolution during seven months in Sierra Leone, *Cell* 161 (2015) 1516–1526.
- [5] S.K. Gire, A. Goba, K.G. Andersen, R.S. Sealfon, D.J. Park, L. Kanneh, S. Jalloh, M. Momoh, M. Fullah, G. Dudas, S. Wohl, L.M. Moses, N.L. Yozwiak, S. Winnicki, C.B. Matranga, C.M. Malboeuf, J. Qu, A.D. Gladden, S.F. Schaffner, X. Wang, P.P. Jiang, M. Nekoui, A. Colubri, M.R. Coomber, M. Fonnies, A. Moigboi, M. Gbakie, F.K. Kamara, V. Tucker, E. Konuwa, S. Saffa, J. Sellu, A.A. Jalloh, A. Kovoma, J. Koninga, I. Mustapha, K. Kargbo, M. Foday, M. Yillah, F. Kanneh, W. Robert, J.L. Massally, S.B. Chapman, J. Bochicchio, C. Murphy, C. Nusbaum, S. Young, B.W. Birren, D.S. Grant, J.S. Scheffelin, et al., Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak, *Science* 345 (2014) 1369–1372.
- [6] T. Hoenen, D. Safronetz, A. Groseth, K.R. Wollenberg, O.A. Koita, B. Diarra, I.S. Fall, F.C. Haidara, F. Diallo, M. Sanogo, Y.S. Sarro, A. Kone, A.C. Togo, A. Traore, M. Kodio, A. Dosseh, K. Rosenke, E. de Wit, F. Feldmann, H. Ebihara, V.J. Munster, K.C. Zoon, H. Feldmann, S. Sow, Mutation rate and genotype variation of Ebola virus from Mali case sequences, *Science* 348 (2015) 117–119.
- [7] J.R. Kugelman, M.R. Wiley, S. Mate, J.T. Ladner, B. Beitzel, L. Fakoli, F. Taweh, K. Prieto, J.W. Diclaro, T. Minogue, R.J. Schoepp, K.E. Schaefer, J. Pettitt, S. Bateman, J. Fair, J.H. Kuhn, L. Hensley, D.J. Park, P.C. Sabeti, M. Sanchez-Lockhart, F.K. Bolay, G. Palacios, Diseases USAMRIID, National Institutes of Health, Integrated Research Facility–Frederick Ebola Response T, Monitoring of Ebola virus Makona evolution through establishment of advanced genomic capability in Liberia, *Emerg. Infect. Dis.* 21 (2015) 1135–1143.
- [8] E. Simon-Loriere, O. Faye, O. Faye, L. Koivogui, N. Magassouba, S. Keita, J.M. Thiéberge, L. Diancourt, C. Bouchier, M. Vandenbogaert, V. Caro, G. Fall, J.P. Buchmann, C.B. Matranga, P.C. Sabeti, J.C. Manuguerra, E.C. Holmes, A.A. Sall, Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic, *Nature* 524 (2015) 102–104.
- [9] G. Dudas, L.M. Carvalho, T. Bedford, A.J. Tatem, G. Baele, N.R. Faria, D.J. Park, J.T. Ladner, A. Arias, D. Asogun, F. Bielejec, S.L. Caddy, M. Cotten, J. D’Ambrosio, S. Dellicour, A. Di Caro, J.W. Diclaro, S. Duraffour, M.J. Elmore, L.S. Fakoli, O. Faye, M.L. Gilbert, S.M. Geva, S. Gire, A. Gladden-Young, A. Gnirke, A. Goba, D.S. Grant, B.L. Haagmans, J.A. Hiscox, U. Jah, J.R. Kugelman, D. Liu, J. Lu, C.M. Malboeuf, S. Mate, D.A. Matthews, C.B. Matranga, L.W. Meredith, J. Qu, J. Quick, S.D. Pas, M.V.T. Phan, G. Pollakis, C.B. Reusken, M. Sanchez-Lockhart, S.F. Schaffner, J.S. Schieffelin, R.S. Sealfon, E. Simon-Loriere, et al., Virus genomes reveal factors that spread and sustained the Ebola epidemic, *Nature* 544 (2017) 309–315.
- [10] D. Ka, G. Fall, V.C. Diallo, O. Faye, L.D. Fortes, O. Faye, E.I. Bah, K.M. Diallo, F. Balique, C.T. Ndour, M. Seydi, A.A. Sall, Ebola virus imported from Guinea to Senegal, 2014, *Emerg. Infect. Dis.* 23 (2017) 1026–1028.
- [11] Team WHOER, Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections, *N. Engl. J. Med.* 371 (2014) 1481–1495.
- [12] S. Baize, D. Pannetier, L. Oestereich, T. Rieger, L. Koivogui, N. Magassouba, B. Soropogui, M.S. Sow, S. Keita, H. De Clerck, A. Tiffany, G. Dominguez, M. Loua, A. Traore, M. Kolie, E.R. Malano, E. Heleze, A. Bocquin, S. Mely, H. Raoul, V. Caro, D.

- Cadar, M. Gabriel, M. Pahlmann, D. Tappe, J. Schmidt-Chanasit, B. Impouma, A.K. Diallo, P. Formenty, M. Van Herp, S. Gunther, Emergence of Zaire Ebola virus disease in Guinea, *N. Engl. J. Med.* 371 (2014) 1418–1425.
- [13] E.C. Holmes, G. Dudas, A. Rambaut, K.G. Andersen, The evolution of Ebola virus: Insights from the 2013–2016 epidemic, *Nature* 538 (2016) 193–200.
- [14] X. Jiang, Z. Zhang, D. Zhuang, M.J. Carr, R. Zhang, Q. Lv, W. Shi, Non-coding regions of the Ebola virus genome contain indispensable phylogenetic and evolutionary information, *Sci. China Life Sci.* 58 (2015) 682–686.
- [15] M. Ni, C. Chen, J. Qian, H.X. Xiao, W.F. Shi, Y. Luo, H.Y. Wang, Z. Li, J. Wu, P.S. Xu, S.H. Chen, G. Wong, Y. Bi, Z.P. Xia, W. Li, H.J. Lu, J. Ma, Y.G. Tong, H. Zeng, S.Q. Wang, G.F. Gao, X.C. Bo, D. Liu, Intra-host dynamics of Ebola virus during 2014, *Nat. Microbiol.* 1 (2016) 16151.
- [16] J.F. Gelinac, G. Clerzius, E. Shaw, A. Gagnon, Enhancement of replication of RNA viruses by ADAR1 via RNA editing and inhibition of RNA-activated protein kinase, *J. Virol.* 85 (2011) 8460–8466.
- [17] S.L. Smits, S.D. Pas, C.B. Reusken, B.L. Haagmans, P. Pertile, C. Cancedda, K. Dierberg, I. Wurie, A. Kamara, D. Kargbo, S.L. Caddy, A. Arias, L. Thorne, J. Lu, U. Jah, I. Goodfellow, M.P. Koopmans, Genotypic anomaly in Ebola virus strains circulating in Magazine Wharf area, Freetown, Sierra Leone, 2015, *Euro Surveill.* 20 (2015), <https://doi.org/10.2807/1560-7917.ES.2015.20.40.30035>.
- [18] K.J. Emmett, A. Lee, H. Khiabani, R. Rabadan, High-resolution genomic surveillance of 2014 Ebolavirus using shared subclonal variants, *PLoS Curr.* 7 (2015), <https://doi.org/10.1371/currents.outbreaks.c7fd7946ba606c982668a96bcb43c90>.
- [19] R.A. Urbanowicz, C.P. McClure, A. Sakuntabhai, A.A. Sall, G. Kobinger, M.A. Muller, E.C. Holmes, F.A. Rey, E. Simon-Loriere, J.K. Ball, Human adaptation of Ebola virus during the West African outbreak, *Cell* 167 (2016) 1079–1087 (e1075).
- [20] W.E. Diehl, A.E. Lin, N.D. Grubaugh, L.M. Carvalho, K. Kim, P.P. Kyaw, S.M. McCauley, E. Donnard, A. Kucukural, P. McDonel, S.F. Schaffner, M. Garber, A. Rambaut, K.G. Andersen, P.C. Sabeti, J. Luban, Ebola virus glycoprotein with increased infectivity dominated the 2013–2016 epidemic, *Cell* 167 (2016) 1088–1098 (e1086).
- [21] E. Dietzel, G. Schudt, V. Kraehling, M. Matrosovich, S. Becker, Functional characterization of adaptive mutations during the West African Ebola virus outbreak, *J. Virol.* 91 (2017) e01913–e01916.
- [22] S.L. Whitmer, C. Albarino, S.S. Shepard, G. Dudas, M. Sheth, S.C. Brown, D. Cannon, B.R. Erickson, A. Gibbons, A. Schuh, T. Sealy, E. Ervin, M. Frace, T.M. Uyeki, S.T. Nichol, U. Stroher, Preliminary evaluation of the effect of investigational Ebola virus disease treatments on viral genome sequences, *J. Infect. Dis.* 214 (2016) S333–S341.
- [23] G.F. Gao, Y. Feng, On the ground in Sierra Leone, *Science* 346 (2014) 666.
- [24] Q. Wang, Y. Zhang, H.Y. Wang, H.J. Du, K. Nie, J.D. Song, K. Xiao, W.W. Lei, J.Q. Guo, H.J. Wei, K. Cai, Y.H. Wang, J. Wu, G. Bangura, I.L. Kamara, X.P. Dong, Detection and analysis of Ebola virus in Sierra Leone-China Friendship Biosafety Laboratory from March 11 to April 20, 2015, *Biomed. Environ. Sci.* 29 (2016) 443–447.
- [25] W.J. Liu, On the ground in Western Africa: from the outbreak to the elapse of Ebola, *Protein Cell* (2016), <https://doi.org/10.1007/s13238-016-0305-2>.
- [26] M.A. de La Vega, G. Caleo, J. Audet, X. Qiu, R.A. Kozak, J.I. Brooks, S. Kern, A. Wolz, A. Sprecher, J. Greig, K. Lokuge, D.K. Kargbo, B. Kargbo, A. Di Caro, A. Grolla, D. Kobasa, J.E. Strong, G. Ippolito, M. Van Herp, G.P. Kobinger, Ebola viral load at diagnosis associates with patient outcome and outbreak evolution, *J. Clin. Invest.* 125 (2015) 4421–4428.