



## OPEN ACCESS

## EDITED BY

Junki Maruyama,  
University of Texas Medical Branch at  
Galveston, United States

## REVIEWED BY

Shuzo Urata,  
Nagasaki University, Japan  
Jinwoo Lee,  
University of Maryland, United States

## \*CORRESPONDENCE

Cheng-Qiang He  
✉ hchqiang@sdnu.edu.cn  
Nai-Zheng Ding  
✉ nzding@sdnu.edu.cn

RECEIVED 03 April 2024

ACCEPTED 29 April 2024

PUBLISHED 20 May 2024

## CITATION

He C-Q, Kong C, He M, Chen G-X, Liu S-M  
and Ding N-Z (2024) Intrasegmental  
recombination as an evolutionary force of  
Lassa fever virus.  
*Front. Microbiol.* 15:1411537.  
doi: 10.3389/fmicb.2024.1411537

## COPYRIGHT

© 2024 He, Kong, He, Chen, Liu and Ding.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Intrasegmental recombination as an evolutionary force of Lassa fever virus

Cheng-Qiang He\*, Chao Kong, Mei He, Guan-Xiang Chen,  
Shu-Min Liu and Nai-Zheng Ding\*

Shandong Provincial Key Laboratory of Animal Resistance Biology, College of Life Science, Shandong Normal University, Jinan, China

Lassa fever (LF), caused by Lassa virus (LASV), is one of the most dangerous diseases to public health. Homologous recombination (HR) is a basic genetic power driving biological evolution. However, as a negative-stranded RNA virus, it is unknown whether HR occurs between LASVs and its influence on the outbreak of LF. In this study, after analyzing 575 S and 433 L segments of LASV collected in Africa, we found that LASV can achieve HR in both of its segments. Interestingly, although the length of S segment is less than half of the L segment, the proportion of LASVs with S recombinants is significantly higher than that with L recombinants. These results suggest that HR may be a feature of LASV, which can be set by natural selection to produce beneficial or eliminate harmful mutations for the virus, so it plays a role in LASV evolution during the outbreak of LF.

## KEYWORDS

Lassa fever, LASV, homologous recombinant, virus evolution, negative-stranded RNA virus

## Introduction

Lassa fever (LF) is a type of viral hemorrhagic fever with high mortality, which has been listed by the World Health Organization (WHO) as one of the diseases posing the greatest risk to public health.<sup>1</sup> It is estimated that LF causes tens of thousands of hospitalizations and thousands of deaths every year. Case fatality rates among hospitalized LF patients can be up to 50%, although numerous sub-clinical infections may also occur (Troup et al., 1970; McCormick and Fisher-Hoch, 2002). Its pathogen, Lassa virus (LASV), is classified as (BSL)-4 agent, which causes endemic diseases in much of West Africa, primarily including Sierra Leone, Guinea, Liberia, and Nigeria (Andersen et al., 2015). The rodent *Mastomys natalensis* functions as a reservoir and maintains persistent infections of the virus (Lecompte et al., 2006). Although most patients are infected by exposure to excreta from this animal, human-to-human transmission has also been reported (McCormick and Fisher-Hoch, 2002). Given its potential to cause a public health emergency and the absence of efficacious drugs and vaccines, the WHO Research and Development (R&D) Blueprint Committee suggested that there is an urgent need for accelerated R&D for the virus (Mehand et al., 2018).

1 <https://www.who.int/activities/prioritizing-diseases-for-research-and-development-in-emergency-contexts>

LASV belongs to the *Mammarenavirus* genus of the *Arenaviridae* family (Pennington and Lee, 2022). Its genome consists of two segments of negative-sense single-stranded RNA, namely S and L segments (Torriani et al., 2017). Their lengths are approximately 3,400 and 7,300 nucleotides, respectively. They ambisense strategy is used to encode four proteins, namely, large (L) protein (RNA-dependent RNA polymerase, RdRp), zinc-binding (Z) protein, nucleoprotein (NP), and glycoprotein complex (GPC) (Garry, 2022). The L protein is encoded at the 3' end of the L segment in the genome-complementary sense, whereas the Z protein is encoded at the 5' end of the L RNA in the genomic sense. In a similar fashion, the *np* gene encodes NP at the 3' end of the S RNA, whereas GPC is encoded at the 5' end of the S RNA. Of them, glycoprotein precursor undergoes proteolytic cleavage to yield three subunits, namely, GP1, GP2, and the stable signal peptide (SSP) (Pennington and Lee, 2022). SSP/GP1/GP2 lies on the virion surface and mediates host cell attachment and membrane fusion and entry (Hastie and Saphire, 2018). As the antigen on the viral surface, GP1 is also the response target of protective antibody of host and becomes a focus for vaccine design efforts (Sommerstein et al., 2015; Robinson et al., 2016).

As an RNA virus, LASV has wide genetic diversity. Phylogenetic analyses showed that LASVs have differentiated into seven genotype lineages associated with geographical locations (lineages I, II, III, IV, V, VI, and VII). Sequence diversity is as high as 24.6% between lineages (Bowen et al., 2000). Moreover, each lineage may also include different sublineages (Oloninji et al., 2018). The LASV circulating in Nigeria has the most extensive genetic diversity, including lineages I, II, and III, whereas strains from Guinea, Sierra Leone, and Liberia seemed to be more closely related and belong exclusively to lineage IV (Fichet-Calvet et al., 2016). Lineage V is mainly found in Mali (Gunther et al., 2001). Lineage VII was reported from Togo in 2016 (Whitmer et al., 2018). The regionalization of genetic diversity may be the consequence of the adaptation of the virus to the local niches. Understanding the molecular mechanism of genetic diversity is important for controlling LASV.

Due to the lack of proofreading function, RdRp of LASV can result in a high mutation rate in progeny virus genome during replication. The evolution rate of LASV was as high as  $10^{-3}$ /site/year (Andersen et al., 2015). However, high mutation rates are generally considered deleterious for RNA virus due to its replication mode, which causes a continuous decrease in the mean fitness of its population (Chao, 1990). As a form of similar to sexual reproduction (Simon-Loriere and Holmes, 2011), Genetic recombination can compensate for the adverse effects of the high mutation rate (Bodine et al., 1990; Naito and Pawlowska, 2016). Recombination in RNA viruses can be defined as an exchange of genetic material between at least two separate viral genomes. Two types of recombination are known as reassortment and homologous recombination (HR). The former is defined as the exchange of genomic segments between viruses in a segmented genome, while the latter involves the exchange of homologous regions of genome isolated by one or several breakpoints (Simon-Loriere and Holmes, 2011). HR also plays an important role in the evolution of some positive-stranded RNA viruses (Kirkegaard and Baltimore, 1986; Nagy and Simon, 1997). Virulent variants of some viruses have been generated by homologous recombination (Worobey et al., 1999; Anderson et al., 2000; Pita et al., 2001). Nevertheless, HR has been reported to be rare in negative-stranded RNA virus because of its genome packaged with

nucleoproteins (Chare et al., 2003). Despite this, there is evidence of recombination in negative-stranded RNA viruses, such as influenza virus (He et al., 2008, 2019), Newcastle disease viruses (Zhang et al., 2010), rabies virus (Ding et al., 2017), Zaire ebolavirus (Wittmann et al., 2007), and severe fever with thrombocytopenia syndrome bunyavirus (He and Ding, 2012). So far, the occurrence and evolutionary influence of HR have not been known in LASV evolution, although several groups have attempted to find whether HR affects the evolution of this virus (Chare et al., 2003; Andersen et al., 2015).

In this study, we provide a series of significant evidence to show that HR has occurred in LASV. The recombination was found in both segments of LASV and could take place between viruses of the same or different lineages. Moreover, an LASV with recombinant S segment has become the dominant variant in response to 40% of LF cases in Nigeria. These results show that HR plays an important role in shaping LASV evolution and has the potential to result in a novel highly pathogenic LASV for human and/or host tropism change.

## Materials and methods

### Viruses and sequences

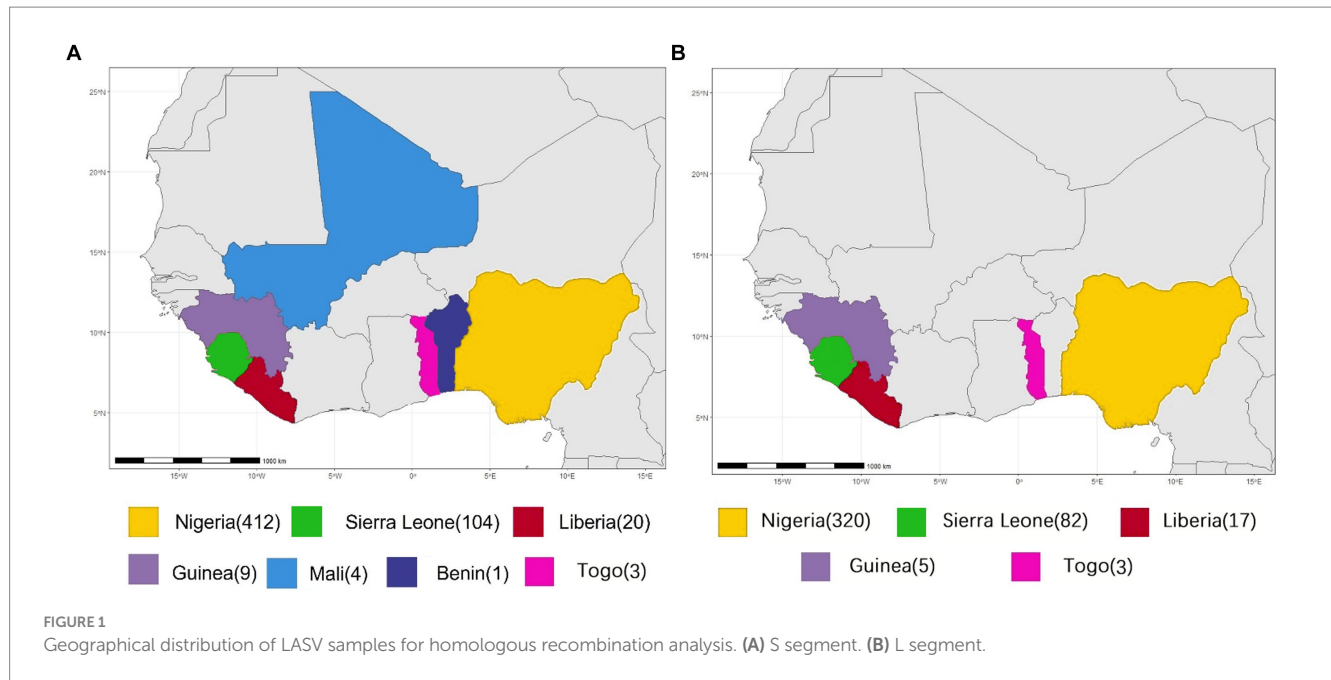
All available complete LASV S and L segment sequences were downloaded from GenBank. Most of these viruses are sequenced by the next-generation sequencing method. Since low coverage and co-infection might result in ambiguous bases, those sequences containing many "N" bases were considered as low quality and discarded in order to avoid unreliable recombination results as much as possible. Finally, 575 S and 433 L segments were used to analyze potential HR. Their geographical distributions are shown in Figure 1. These viruses were mainly collected from Nigeria, accounting for >70%.

### Reconstruction of phylogeny

These sequences were aligned with CLUSTAL W (Thompson et al., 1997). Nucleotide substitution models with the lowest Bayesian Information Criterion (BIC) scores are considered as the best substitution pattern. Phylogenetic histories were inferred using the Maximum-Likelihood (ML) or the Neighbor-Joining (NJ) methods, with the best nucleotide substitution model. The bootstrap method was used to assess the robustness of tree branches with 1,000 replicates. A branch with more than 70% bootstrap value was considered to have a recent common ancestor with robustness. Evolutionary analyses were conducted in MEGA11 (Tamura et al., 2021). Shimodaira-Hasegawa (SH) test method, which was implemented in TREE-PUZZLE program (Schmidt et al., 2002), was employed to test whether topology of phylogenetic trees was significantly different.

### Recombination analysis

The recombination analysis was performed referring to previous report (He et al., 2010). In brief, the recombination signals are visualized using the Simplot and Bootscan programs implemented in



the Simplot software package (Lole et al., 1999). Employing the FindSites subprogram of Simplot, potential breakpoint was estimated at parsimonious region with the maximization of chi-square values ( $\chi^2$ ) of the information sites. Fisher's exact test was used to statistically analyze the difference of information sites on both sides of the breakpoints. At the same time, the sequence alignment data were examined for potential recombination events using the RDP software package (Martin et al., 2005). If a recombinant signal can be detected by at least two recombination detection methods, the signal is considered robust.

To demonstrate the presence of recombination, the gold-standard approach is a set of incongruent phylogenetic trees inferred from the different genomic regions of the putative recombinant viruses (Chare et al., 2003; Boni et al., 2010). The ML method implemented in MEGA 11 was used to reconstruct the phylogeny of different regions defined by the recombination breakpoints of the recombinants (Tamura et al., 2021). The difference in significance of the topology of these phylogenetic trees was confirmed by the SH test (Schmidt et al., 2002). If different regions have different phylogenetic histories, the recombination signs of these recombination groups are considered as robustness.

## Results and discussion

In this study, recombination was analyzed in the 575S and 433L segments. 193S and 21L segments were found to have significant recombination signals (Table 1 LASV strains with evidence for potential recombination). This means that more than one-third of the LASVs circulating in Africa might carry recombinant S segment. Among the 193 viruses, 5 were single isolates while the other 188 composed of 6 recombinant groups. The largest group consisted of 168 members. Interestingly, although the length of L segment is more than twice that of S segment, the recombinants found in L segment were significantly lower than in S segment, which were less than 5%. In

addition, considering that HR might not be frequent in the virus, those isolates with the same recombination event should be descended from a common recombinant ancestor. These results suggested that HR is likely a trait of LASV that can be set by natural selection to create advantageous or purge deleterious mutations in this virus.

## HR has influence on the evolution of S segment

S segment encodes GPC and NP proteins of LASV. NP oligomerizes and encapsidates genomic RNA to form a nucleocapsid (Young and Howard, 1983). Although the core structural function of NP oligomers is used to package, protect, and control access to the viral genome (Ruigrok et al., 2011; Reguera et al., 2014), NP also suppresses host type I interferon response (Huang et al., 2018), suggesting that it plays a role in adapting to the innate immunity of the host. As the extracellular spike, GP1 of the surface glycoprotein mediates the receptor ( $\alpha$ -dystroglycan, an extracellular protein) recognition and viral entry into a host cell and bears the strong immune selection pressure from the host (Burri et al., 2012; Hastie and Saphire, 2018). Therefore, rapid evolution of S segment may be the inherent requirement for LASV to adapt to selection pressure from immunity of the host. In the samples we analyzed, more than 30% of viruses contained recombinant S segment. Interestingly, LASV966 recombinant family has caused more than 45% cases of LF in Nigeria.

The isolate LASV966 was collected from a patient in 2009 in Nigeria (Andersen et al., 2015). Its genome was sequenced using Illumina method and assembled *de novo* through Trinity v. r2011-11-26 and Novoalign v. v2.08. Previous study showed that its genome was sequenced with high quality (Andersen et al., 2015). This study found that LASV966 and its sister branch consisting of more than 160 members contained the same recombinant S segment (Figure 2; Supplementary Figure S1).

TABLE 1 LASV strains with evidence for potential recombination.

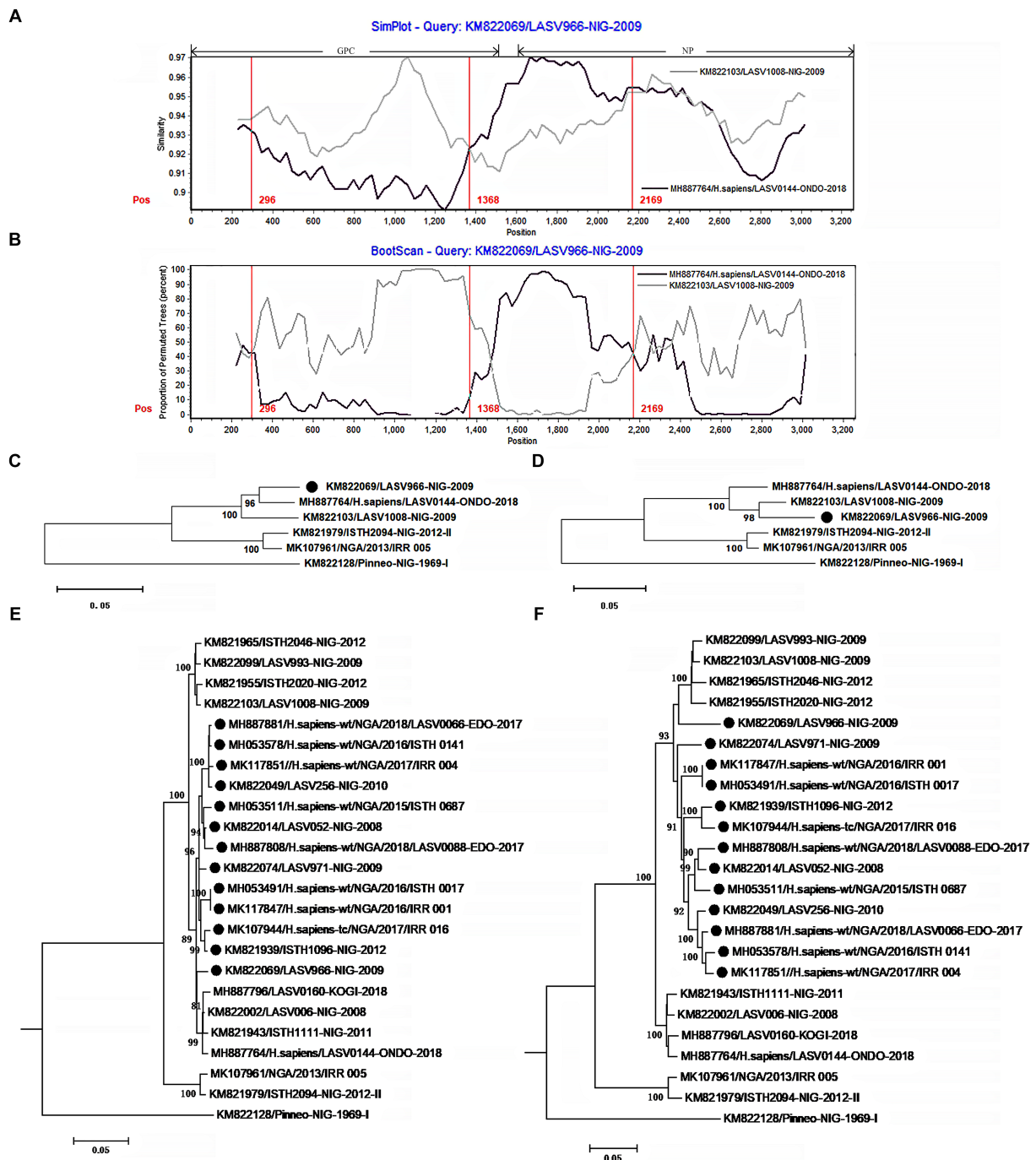
GenBank number	Recombinant	Num. of Sisters	Segment	Putative parents	Position of Breakpoints	Host	Country	Sequencing Technology
MH053560	ISTH_1150	0	S	ISTH_0543; ISTH_1250	2,907	Human	Nigeria	Illumina MiSeq
MK107957	IRR_030	0	S	IRR_027; 0167- ONDO	93, 556, 1,020, 1,218, 2,877, 3,004	Human	Nigeria	Illumina MiSeq
MK107947	IRR_003	0	S	IRR_005; ISTH_0621	1,412, 1,581	Human	Nigeria	Illumina MiSeq
MH887871	0169-ONDO	0	S	0106-ONDO; 0083-ONDO	2,151, 2,401, 3,067	Human	Nigeria	Illumina MiSeq
MK107970	IRR_004	0	S	IRR_002; ISTH_0543	3,018; 3,248	Human	Nigeria	Illumina MiSeq
OM735974	Barlie-00241	1	S	G1774-SLE; G1618-SLE	1,095; 2,124; 2,584; 2,939	<i>L. sikapusi</i>	Sierra Leone	Illumina
KM821877	G2906-SLE	2	S	G1774-SLE; G2392-SLE	2,515; 3,045	Human	Sierra Leone	Illumina
MK118000	IRR_073	3	S	ISTH2271-NIG; 0133-EDO	194; 3,001	Human	Nigeria	Illumina
MK107920	IRR_028	5	S	IRR_005; IRR_015	346; 559	Human	Nigeria	Illumina
MK107951	IRR_001	5	S	IRR_011; ISTH1096-NIG	688; 1,120; 2,778	Human	Nigeria	Illumina
KM822069	LASV966	167	S	0144-ONDO; 1,008-NIG	1,368; 2,169	Human	Nigeria	Illumina
MK117929	IRR_096	0	L	0192.2-ONDO; 0193-DELTA	2,819; 3,634	Human	Nigeria	Illumina
MK117903	IRR_052	18	L	IRR_071; Nig08-A37	378; 584	Human	Nigeria	Illumina
MH887906	LASV0169	0	L	IRR072; LASV0215	376; 740; 1,055; 6,425	Human	Nigeria	Illumina

With the help of Simplot program, the S segment sequences of LASV966 and its parent lineage representatives LASV0144 and LASV1008 were compared (Figure 2A). The translation starting site of GPC is defined as nucleotide 1. Two crossover sites were found around the positions 1,368 and 2,169. The recombinant LASV966 and the putative minor parent LASV0144 have higher sequence similarity between the two crossover sites. However, outside of the two crossover sites, the sequence similarity between LASV966 and LASV1008 was higher than that of LASV0144. Bootscan analysis showed that there might be three crossover sites on the S segment of LASV966, and the new crossover sites were located near position 300 (Figure 2B). The maximum  $\chi^2$  values of the three putative breakpoints were 0.6, 20.08, and 13.09, respectively. It was found that the first breakpoint had no statistical significance support ( $p > 0.05$  of Fisher's exact test). For the second and third breakpoints, their information sites had statistically significant differences ( $p < 0.0001$  and  $p < 0.0005$ , respectively). Employing the recombination detection program (RDP 4.0), six recombination detection methods, RDP ( $p < 0.005$ ), GENECONV ( $p < 0.05$ ), BootScan ( $p < 0.05$ ), Maxchi ( $p < 0.0001$ ), Chimera ( $p < 0.005$ ), and 3Seq ( $p < 0.0001$ ) found significant evidence of recombination. The most compelling evidence for recombination was the occurrence of

incongruent phylogenetic trees in different genomic regions (Nelson and Holmes, 2007), which was found in LASV966 S segment (Figures 2C,D). These results suggested that S segment of LASV966 had multi-origin. The most likely parent was sought in GenBank using the basic local alignment search tool (BLAST), but putative parent with 100% sequence identity could not be found, indicating that S segment of LASV966 might originated from a recombinant ancestor that had experienced multiple recombination events before 2009.

Running RDP, we scanned the sequence alignment data of S segments and found another 167 recombinants identical to LASV966, and they accounted for more than 40% (168 of 412) of the Nigerian LASV in our sample. These viruses have been circulating in Nigeria in 2008, indicating that the group with the recombinant S segment should have become the dominant pathogen responsible for LF in this country.

To determine these recombinants, 13 representatives were used to reconstruct their phylogenetic histories (Figures 2E,F). In the phylogenetic tree inferred from the regions 1–300 and 1,368–2,169, they were clustered into the LASV0144 lineage. However, in the phylogenetic tree inferred from the regions 300–1,367 and 2,169–3,052, the 13 isolates and the LASV1008 lineage constituted a



**FIGURE 2**  
 The evidence for recombination in the S segment of LASV966 and its sisters circulating in Nigeria. (A) Comparison of S segments of the recombinant LASV966 and its putative parents LASV1008 and LASV0144. LASV966 is used as the query. The y-axis gives the percentage of identity between LASV966 and its two parents. Different colors on the bold line represent the distribution of genes in the S segment. (B) Bootscan plot of S segment sequences of LASV966 and its putative parents. The y-axis gives the percentage of permuted trees using a sliding window. The vertical lines show the recombination breakpoints with the maximum  $\chi^2$  of information sites. (C,D) Phylogenetic trees inferred from different regions of S segment. (E,F) The phylogenetic history of the S recombinant lineage. (E) The phylogenetic history of the 13 S recombinants inferred from the regions 1 to 300 and 1,368 to 2,169. (F) The phylogenetic history of the 13 S recombinants inferred from the regions 300–1,368 and 2,169–3,025. The S recombinants are marked with “●”.

mono-phylogenetic group. Moreover, there was a statistically significant discrepancy in the topology of the two trees (Shimodaira–Hasegawa test,  $p < 0.05$ ), which also constituted a powerful evidence

for the recombination event. It was also noticed that 147 of the 168 apparent S recombinants had the signal of multiple origins (Supplementary Figure S1).

In addition to the LASV966 family, another 19S recombinants have also been found in the LASV circulating in Nigeria (Supplementary Figures S2–S9). Of them, 5 recombinants were single, and the other 14 formed three groups. Moreover, except for the recombinant ISTH\_1150 (Supplementary Figure S5), there are at least two potential recombination breakpoints in these S recombinants.

In S segment of the isolate IRR 030, there were up to six obvious recombination breakpoints (Supplementary Figure S2). This virus has been purified by cell culture, and virus purification was performed by cell culture to obtain a high-titer viral stock which was suitable for subsequent experiments. This method ensures the removal of contaminants from other viruses, thereby providing a purified virus sample for accurate characterization of viral properties. In addition, there is an N base in the nucleotide sequence of its segment S during Illumina MiSeq sequencing. Moreover, there is no recombination signal in its segment L. These results suggested that these recombination signals should be more likely natural, although further sequencing verification is necessary.

It was noticed that these recombination events mainly occurred between viruses within the same lineage. This is understandable because viruses within lineages have a higher chance of meeting in the same host than viruses outside lineages. Interestingly, recombination might take place between different lineages (Supplementary Figure S3). IRR 003 has also been purified by cell culture (Ehichioya et al., 2019). In its S segment, there are two significant recombination breakpoints located at the 3' end of GPC and NP genes. In S segment of IRR 003, most of the genetic material were inherited from the lineage II, but there were approximately 170 bases inherited from lineage III.

In this study, the sample also included 104 isolates from Sierra Leone. Of them, five recombinants were found with two different recombinant S segments. Three human isolates contain one of them (Supplementary Figure S10), while two rodent isolates contain the other (Supplementary Figure S11). All these five viruses had been sequenced in high quality, and there were no ambiguous bases in their S segment.

The three human isolates with the recombinant *np* gene were collected in 2011, 2012, and 2013, respectively. In addition, no putative parent strains can be found in the GenBank database, which are completely consistent with their nucleotide sequences. These results indicated that the recombination event should have occurred before 2011. Although the recombinant did not become the dominant strain of Sierra Leone, it repeatedly infected humans, indicating that it could have been widely circulating in rodents of this country.

## HR also affects the evolution of L segment

L segment encodes Z and L proteins. The Z protein coordinates the packaging, budding of infectious particles, and virus egress (Urata and Yasuda, 2012). L protein is a multifunctional enzyme with RNA-dependent RNA polymerase, host-cap-snatching endonuclease, and regulatory domains (Vieth et al., 2004; Brunotte et al., 2011; Vogel et al., 2019). As the inner proteins, these two proteins may have undergone less natural selection pressure. Therefore, the selection orientation of recombinant L might be relatively lower. In fact, only 21 isolates were found to carry recombinant L segment among the 433 viruses analyzed in this study (Figure 3; Supplementary Figures S12, S13). The rate of recombinant is less than 5% of the sample. This is much

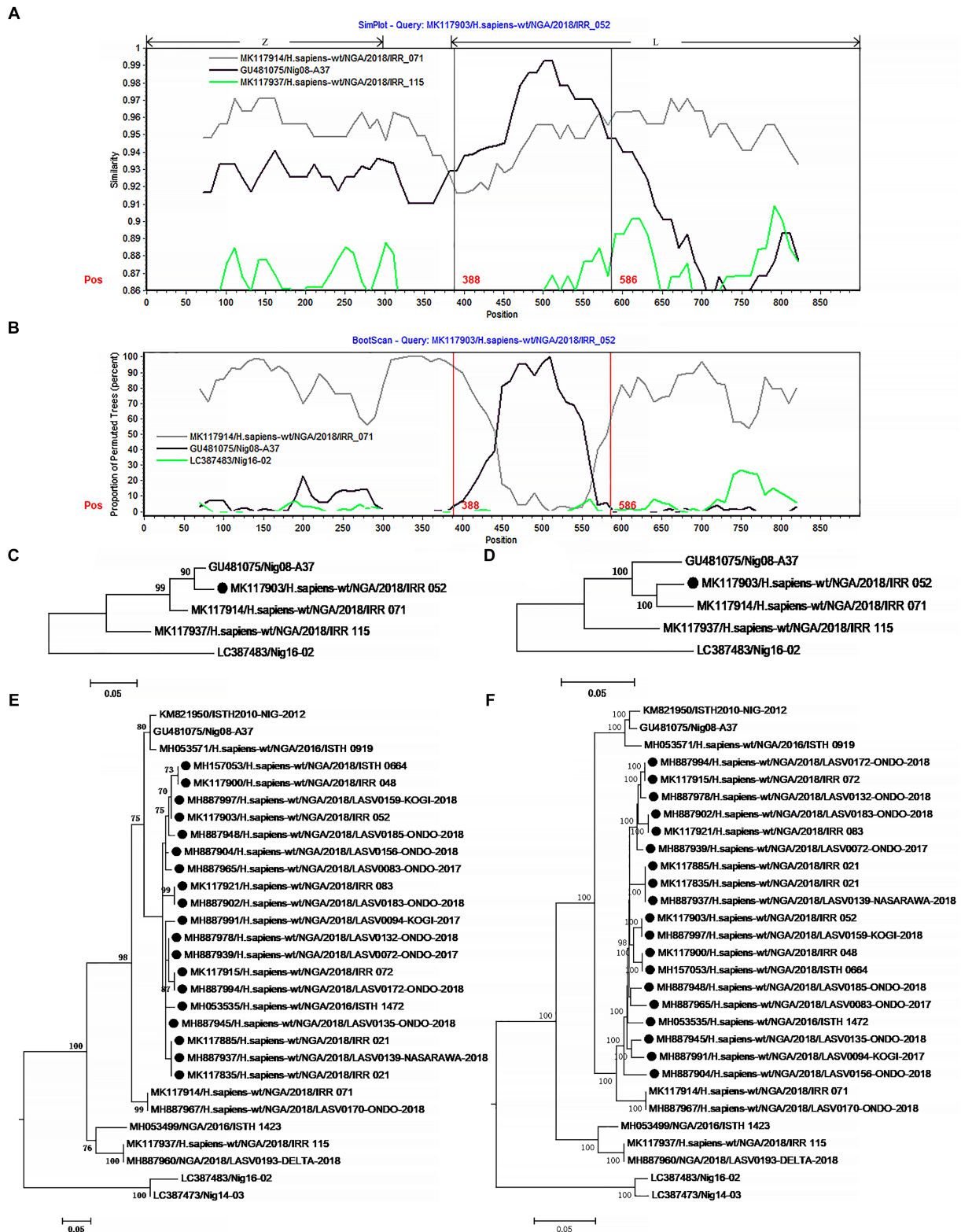
lower than the 33% recombinant ratio of the S segment. Of the 21 L recombinants, 19 might originate from the same common ancestor which underwent the HR. In addition, the single sequence of the recombinant IRR 0096 contains many ambiguous bases which are evenly distributed in its L segment.

The 19L recombinants shared more than 96.3% sequence similarity with each other. For recombination analysis, IRR052 was used as the recombinant representative strain. Two isolates, IRR071 and Nig08-A37, could be supposed as the putative parents of the recombinant. Combined with the results of the  $\chi^2$  analysis and Fisher's exact test, two breakpoints were located at the positions 388 and 586 with maximum  $\chi^2$  values (5.4 and 9.6, respectively) when the start codon of the Z gene was defined as the first nucleotide. These two breakpoints fall in the 3' end of the L gene (Figure 3A). Fisher's exact test showed that the information sites on both sides of these two breakpoints were significantly different ( $p < 0.05$  and  $p < 0.01$ , respectively). We compared the similarity between the first 900 nucleotides of the L segment of IRR052 and its two putative parents (Figure 3A). Between the two breakpoints, IRR052 shared 96.9% sequence identity with the putative minor parent Nig08-A37 (versus 93.4% with IRR071). However, before and after these two breakpoints, IRR052 only shared 90.9% similarity with Nig08-A37 (versus 95.4% with IRR071). Bootscan result supported the hypothesis that IRR052 was descended from the two lineages, IRR071 and Nig08-A37 (Figure 3B). The recombination signal was also recognized by RDP, and its five programs, namely, GENECONV ( $p < 0.05$ ), BootScan ( $p < 0.01$ ), Maxchi ( $p < 0.01$ ), Chimera ( $p < 0.05$ ), and 3Seq ( $p < 0.05$ ) gave statistically significant evidence of recombination. Phylogenetic reconstruction showed that IRR052 and Nig08-A37 constituted a monophylogenetic group with 90% bootstrap value (Figure 3C) between the two breakpoints. On the contrary, IRR052 and IRR071 fell into the same monophylogenetic group with 100% bootstrap value (Figure 3D). Shimodaira–Hasegawa test showed that the difference of phylogenetic histories inferred from different regions were statistically significant ( $p < 0.001$ ).

When RDP was run to scan the alignment data of the 433 L segments, it was found that another 19 L segments had the same recombination signal as IRR052. Of them, 18 were determined by the incongruent phylogenetic histories of different regions (Figures 3E,F) ( $p < 0.001$  of Shimodaira–Hasegawa test).

All these results show that many LASV families with recombinant segment are circulating in Africa. It was noticed that some of recombinant segments were single. At present, we cannot determine whether they are natural or non-natural recombinants due to co-infected samples. To confirm whether they are natural recombinants, further purification and sequencing may be necessary.

The occurrence of recombination requires two viruses to infect a host cell at the same time. Persistent infection may provide more opportunities for such co-infection. Interestingly, previous studies indicated that LASV infection can downregulate its cell-surface receptor, alpha-DG, which inhibits co-infection (Rojek et al., 2007). This downregulation could increase viral diversity, potentially influencing LASV recombination and evolutionary dynamics. Reassortment has been found in this virus (Andersen et al., 2015), suggesting that co-infection of LASV can occur in the same cell. Although LASV is an acute infection in humans, it can cause the persistent infection in rodent host (Walker et al., 1975; Marien et al., 2017). Thus, these recombination events are more likely to occur in *Mastomys natalensis* than in humans.



**FIGURE 3**  
 The evidence for recombination in the L segment of IRR 052 and its sisters circulating in Nigeria. **(A)** Sequence similarity plot of L segment (positions 1–900) of IRR 052 and its putative parents IRR 071 and Nig08-A37. IRR 052 is used as the query. The y-axis indicates the percentage of similarity between IRR 052 and its parents. Different colors on the bold line represent the different genes in the L segment. **(B)** Bootscan plot of L segment sequences of IRR 052 and its putative parents. The y-axis gives the percentage of permuted trees. The vertical lines indicate the recombination breakpoints. **(C,D)** are the phylogenetic trees inferred from different regions of L segment. **(C)** The phylogenetic tree inferred from the region 388–586. **(D)** The phylogenetic tree inferred from the regions 1–388 + 586–6,500. **(E,F)** Incongruent phylogenetic histories of the L recombinant lineage. **(E)** The phylogenetic history of the 19 L recombinants inferred from the fragment 388–586. **(F)** The phylogenetic history of the 19 L recombinants inferred from the region 586–6,500. The recombinant isolates are marked with “●”.

In addition, although HR is a rare event in the virus, interestingly, recombinants are not unusual in the current prevalent LASV. The discovery of these recombinants also suggests that co-infection of LASV might be common in its reservoir host.

Although most recombination breakpoints are randomly distributed in S and L segments, several recombinants have a similar breakpoint near site 3,000 of S segment, including three single sequences and three groups, namely, IRR\_030, 0169-ONDO, IRR\_004, IRR\_073 group, Barlie-00241 group, and G2906 group. This suggested that this site might be a so-called recombination hotspot. According to the template shift hypothesis of RNA virus recombination, recombination occurred in RNA regions, where RNA could potentially form a secondary structure and play a role in recombination as activator (Nagy and Simon, 1997). This allows the two parental RNAs of different origins to form a complex and thereby forces recombination to occur (Romanova et al., 1986; Tolskaya et al., 1987). We analyzed the secondary structure of 100 nucleotides around this site of recombinant and found that there is a potentially hairpin structure with a minimum free energy of approximately  $-20$  kcal/mol (Supplementary Figure S14). Similar secondary structures have been observed in recombination studies of other viruses. These secondary structures may promote template switching by stalling the RNA polymerase during replication and facilitating the transfer of the polymerase and the nascent nucleic acid molecule onto the acceptor RNA. Consequently, these structures are considered as a factor driving the occurrence of recombination (Chare and Holmes, 2006; Simon-Loriere and Holmes, 2011). These findings further support our hypothesis that RNA secondary structures may play a crucial role in the process of viral recombination. The presence of such secondary structures provides an interaction platform for RNAs of different origins, facilitating template switching and gene recombination.

HR can sometimes play a major role in evolution, emergence, and epidemiology of virus. It has been found that recombination has been associated with the expansion of viral host range (Gibbs and Weiller, 1999), increases in virulence (Khatchikian et al., 1989), the evasion of host immunity (Malim and Eberman, 2001), and the evolution of resistance to antivirals (Nora et al., 2007). In this study, the recombinant rate of S segment was found to be significantly higher than that of L segment, which may be related to the different functions of the proteins encoded by these two segments. The proteins encoded by the S fragment bears the main pressure of host immune selection, which may be one of the reasons for the difference in recombination rates between the two fragments. GP1 and GP2 encoded by S segment are crucial for host cell attachment and virus entry. NP possesses several mechanisms as an anti-innate immune response at the 3' end, and one of them is known as DEDDh motif (Hastie et al., 2012), which aids in immune evasion. We found that all recombinant offspring have the motif. Interestingly, we found that an isolate lacking the motif was located at the ancestral position of one parent lineage (Supplementary Figure S15), suggesting that recombination may have given its offspring the potential to evade host immune pressure.

In conclusion, this study found that HR might be a basic genetic mechanism driving the rapid evolution of LASV. In Africa, LASVs with intrasegmental recombinant have an important responsibility in the outbreak of LF. Given that HR may not be a high probability event in LASV, it may be necessary to clarify whether the high proportion of recombinants in infected people is due to their higher human

tropism or there is such a high proportion of recombinants in its rodent reservoirs.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

## Author contributions

C-QH: Conceptualization, Formal analysis, Funding acquisition, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. CK: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. MH: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – review & editing. G-XC: Data curation, Methodology, Software, Writing – review & editing. S-ML: Data curation, Methodology, Software, Writing – review & editing. N-ZD: Conceptualization, Data curation, Funding acquisition, Resources, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research is supported by the Natural Science Foundation of Shandong Province (No. ZR2022MC105) to C-QH.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2024.1411537/full#supplementary-material>



## References

- Andersen, K. G., Shapiro, B. J., Matranga, C. B., Sealfon, R., Lin, A. E., Moses, L. M., et al. (2015). Clinical sequencing uncovers origins and evolution of Lassa virus. *Cell* 162, 738–750. doi: 10.1016/j.cell.2015.07.020
- Anderson, J. P., Rodrigo, A. G., Learn, G. H., Madan, A., Delahunty, C., Coon, M., et al. (2000). Testing the hypothesis of a recombinant origin of human immunodeficiency virus type 1 subtype E. *J. Virol.* 74, 10752–10765. doi: 10.1128/JVI.74.22.10752-10765.2000
- Bodine, D. M., McDonagh, K. T., Brandt, S. J., Ney, P. A., Agricola, B., Byrne, E., et al. (1990). Development of a high-titer retrovirus producer cell line capable of gene transfer into rhesus monkey hematopoietic stem cells. *Proc. Natl. Acad. Sci. USA* 87, 3738–3742. doi: 10.1073/pnas.87.10.3738
- Boni, M. F., de Jong, M. D., van Doorn, H. R., and Holmes, E. C. (2010). Guidelines for identifying homologous recombination events in influenza A virus. *PLoS One* 5:e10434. doi: 10.1371/journal.pone.0010434
- Bowen, M. D., Rollin, P. E., Ksiazek, T. G., Hustad, H. L., Bausch, D. G., Demby, A. H., et al. (2000). Genetic diversity among Lassa virus strains. *J. Virol.* 74, 6992–7004. doi: 10.1128/jvi.74.15.6992-7004.2000
- Brunotte, L., Lelke, M., Hass, M., Kleinstueber, K., Becker-Ziaja, B., and Gunther, S. (2011). Domain structure of Lassa virus L protein. *J. Virol.* 85, 324–333. doi: 10.1128/JVI.00721-10
- Burri, D. J., da Palma, J. R., Kunz, S., and Pasquato, A. (2012). Envelope glycoprotein of arenaviruses. *Viruses* 4, 2162–2181. doi: 10.3390/v4102162
- Chao, L. (1990). Fitness of RNA virus decreased by Muller's ratchet. *Chin. Sci. Bull.* 348, 454–455. doi: 10.1038/348454a0
- Chare, E. R., Gould, E. A., and Holmes, E. C. (2003). Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *J. Gen. Virol.* 84, 2691–2703. doi: 10.1099/vir.0.19277-0
- Chare, E. R., and Holmes, E. C. (2006). A phylogenetic survey of recombination frequency in plant RNA viruses. *Arch. Virol.* 151, 933–946. doi: 10.1007/s00705-005-0675-x
- Ding, N. Z., Xu, D. S., Sun, Y. Y., He, H. B., and He, C. Q. (2017). A permanent host shift of rabies virus from Chiroptera to Carnivora associated with recombination. *Sci. Rep.* 7:289. doi: 10.1038/s41598-017-00395-2
- Ehichioya, D. U., Dellicour, S., Pahlmann, M., Rieger, T., Oestereich, L., Becker-Ziaja, B., et al. (2019). Phylogeography of Lassa virus in Nigeria. *J. Virol.* 93:929. doi: 10.1128/JVI.00929-19
- Fichet-Calvet, E., Olschlager, S., Strecker, T., Koivogui, L., Becker-Ziaja, B., Camara, A. B., et al. (2016). Spatial and temporal evolution of Lassa virus in the natural host population in upper Guinea. *Sci. Rep.* 6:21977. doi: 10.1038/srep21977
- Garry, R. F. (2022). Lassa fever – the road ahead. *Nat. Rev. Microbiol.* 21, 87–96. doi: 10.1038/s41579-022-00789-8
- Gibbs, M. J., and Weiller, G. F. (1999). Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. *Proc. Natl. Acad. Sci. USA* 96, 8022–8027. doi: 10.1073/pnas.96.14.8022
- Gunther, S., Weisner, B., Roth, A., Grewing, T., Asper, M., Drosten, C., et al. (2001). Lassa fever encephalopathy: Lassa virus in cerebrospinal fluid but not in serum. *J. Infect. Dis.* 184, 345–349. doi: 10.1086/322033
- Hastie, K. M., King, L. B., Zandonatti, M. A., and Saphire, E. O. (2012). Structural basis for the dsRNA specificity of the Lassa virus NP exonuclease. *PLoS One* 7:e44211. doi: 10.1371/journal.pone.0044211
- Hastie, K. M., and Saphire, E. O. (2018). Lassa virus glycoprotein: stopping a moving target. *Curr. Opin. Virol.* 31, 52–58. doi: 10.1016/j.coviro.2018.05.002
- He, C. Q., and Ding, N. Z. (2012). Discovery of severe fever with thrombocytopenia syndrome bunyavirus strains originating from intragenic recombination. *J. Virol.* 86, 12426–12430. doi: 10.1128/JVI.01317-12
- He, C. Q., Ding, N. Z., He, M., Li, S. N., Wang, X. M., He, H. B., et al. (2010). Intragenic recombination as a mechanism of genetic diversity in bluetongue virus. *J. Virol.* 84, 11487–11495. doi: 10.1128/JVI.00889-10
- He, C. Q., Han, G. Z., Wang, D., Liu, W., Li, G. R., Liu, X. P., et al. (2008). Homologous recombination evidence in human and swine influenza A viruses. *Virology* 380, 12–20. doi: 10.1016/j.virol.2008.07.014
- He, C. Q., He, M., He, H. B., Wang, H. M., and Ding, N. Z. (2019). The matrix segment of the "Spanish flu" virus originated from intragenic recombination between avian and human influenza A viruses. *Transbound. Emerg. Dis.* 66, 2188–2195. doi: 10.1111/tbed.13282
- Huang, Q., Shao, J., Liang, Y., and Ly, H. (2018). Assays to demonstrate the roles of Arenaviral nucleoproteins (NPs) in viral RNA synthesis and in suppressing type I interferon. *Methods Mol. Biol.* 1604, 189–200. doi: 10.1007/978-1-4939-6981-4\_13
- Khatchikian, D., Orlich, M., and Rott, R. (1989). Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. *Chin. Sci. Bull.* 340, 156–157.
- Kirkegaard, K., and Baltimore, D. (1986). The mechanism of RNA recombination in poliovirus. *Cell* 47, 433–443. doi: 10.1016/0092-8674(86)90600-8
- Lecompte, E., Fichet-Calvet, E., Daffis, S., Koulemou, K., Sylla, O., Kourouma, F., et al. (2006). Mastomys natalensis and Lassa fever, West Africa. *Emerg. Infect. Dis.* 12, 1971–1974. doi: 10.3201/eid1212.060812
- Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., et al. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73, 152–160. doi: 10.1128/JVI.73.1.152-160.1999
- Malim, M. H., and Emerman, M. (2001). HIV-1 sequence variation: drift, shift, and attenuation. *Cell* 104, 469–472. doi: 10.1016/s0092-8674(01)00234-3
- Marien, J., Borremans, B., Gryseels, S., Soropogui, B., De Bruyn, L., Bongo, G. N., et al. (2017). No measurable adverse effects of Lassa, Morogoro and Gairo arenaviruses on their rodent reservoir host in natural conditions. *Parasit. Vectors* 10:210. doi: 10.1186/s13071-017-2146-0
- Martin, D. P., Williamson, C., and Posada, D. (2005). RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* 21, 260–262. doi: 10.1093/bioinformatics/bth490
- McCormick, J. B., and Fisher-Hoch, S. P. (2002). Lassa fever. *Curr. Top. Microbiol. Immunol.* 262, 75–109. doi: 10.1007/978-3-642-56029-3\_4
- Mehand, M. S., Al-Shorbaji, F., Millett, P., and Murgue, B. (2018). The WHO R&D blueprint: 2018 review of emerging infectious diseases requiring urgent research and development efforts. *Antivir. Res.* 159, 63–67. doi: 10.1016/j.antiviral.2018.09.009
- Nagy, P. D., and Simon, A. E. (1997). New insights into the mechanisms of RNA recombination. *Virology* 235, 1–9. doi: 10.1006/viro.1997.8681
- Naito, M., and Pawlowska, T. E. (2016). Defying Muller's ratchet: ancient heritable Endobacteria escape extinction through retention of recombination and genome plasticity. *MBio* 7:257. doi: 10.1128/mBio.02057-15
- Nelson, M. I., and Holmes, E. C. (2007). The evolution of epidemic influenza. *Nat. Rev. Genet.* 8, 196–205. doi: 10.1038/nrg2053
- Nora, T., Charpentier, C., Tenaillon, O., Hoede, C., Clavel, F., and Hance, A. J. (2007). Contribution of recombination to the evolution of human immunodeficiency viruses expressing resistance to antiretroviral treatment. *J. Virol.* 81, 7620–7628. doi: 10.1128/JVI.00083-07
- Oloniniyi, O. K., Unigwe, U. S., Okada, S., Kimura, M., Koyano, S., Miyazaki, Y., et al. (2018). Genetic characterization of Lassa virus strains isolated from 2012 to 2016 in southeastern Nigeria. *PLoS Negl. Trop. Dis.* 12:e0006971. doi: 10.1371/journal.pntd.0006971
- Pennington, H. N., and Lee, J. (2022). Lassa virus glycoprotein complex review: insights into its unique fusion machinery. *Biosci. Rep.* 42:930. doi: 10.1042/BSR20211930
- Pita, J. S., Fondong, V. N., Sangare, A., Otim-Nape, G. W., Ogwal, S., and Fauquet, C. M. (2001). Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J. Gen. Virol.* 82, 655–665. doi: 10.1099/0022-1317-82-3-655
- Reguera, J., Cusack, S., and Kolakofsky, D. (2014). Segmented negative strand RNA virus nucleoprotein structure. *Curr. Opin. Virol.* 5, 7–15. doi: 10.1016/j.coviro.2014.01.003
- Robinson, J. E., Hastie, K. M., Cross, R. W., Yenni, R. E., Elliott, D. H., Rouelle, J. A., et al. (2016). Most neutralizing human monoclonal antibodies target novel epitopes requiring both Lassa virus glycoprotein subunits. *Nat. Commun.* 7:11544. doi: 10.1038/ncomms11544
- Rojek, J. M., Campbell, K. P., Oldstone, M. B., and Kunz, S. (2007). Old World arenavirus infection interferes with the expression of functional alpha-dystroglycan in the host cell. *Mol. Biol. Cell* 18, 4493–4507. doi: 10.1091/mbc.e07-04-0374
- Romanova, L. I., Blinov, V. M., Tolskaya, E. A., Viktorova, E. G., Kolesnikova, M. S., Guseva, E. A., et al. (1986). The primary structure of crossover regions of intertypic poliovirus recombinants: a model of recombination between RNA genomes. *Virology* 155, 202–213. doi: 10.1016/0042-6822(86)90180-7
- Ruigrok, R. W., Crepin, T., and Kolakofsky, D. (2011). Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr. Opin. Microbiol.* 14, 504–510. doi: 10.1016/j.mib.2011.07.011
- Schmidt, H. A., Strimmer, K., Vingron, M., and von Haeseler, A. (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18, 502–504. doi: 10.1093/bioinformatics/18.3.502
- Simon-Loriere, E., and Holmes, E. C. (2011). Why do RNA viruses recombine? *Nat. Rev. Microbiol.* 9, 617–626. doi: 10.1038/nrmicro2614
- Sommerstein, R., Flatz, L., Remy, M. M., Malinge, P., Magistrelli, G., Fischer, N., et al. (2015). Arenavirus glycan shield promotes neutralizing antibody evasion and protracted infection. *PLoS Pathog.* 11:e1005276. doi: 10.1371/journal.ppat.1005276
- Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120

- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882. doi: 10.1093/nar/25.24.4876
- Tolskaya, E. A., Romanova, L. I., Blinov, V. M., Viktorova, E. G., Sinyakov, A. N., Kolesnikova, M. S., et al. (1987). Studies on the recombination between RNA genomes of poliovirus: the primary structure and nonrandom distribution of crossover regions in the genomes of intertypic poliovirus recombinants. *Virology* 161, 54–61. doi: 10.1016/0042-6822(87)90170-X
- Torriani, G., Galan-Navarro, C., and Kunz, S. (2017). Lassa virus cell entry reveals new aspects of virus-host cell interaction. *J. Virol.* 91:1902. doi: 10.1128/JVI.01902-16
- Troup, J. M., White, H. A., Fom, A. L., and Carey, D. E. (1970). An outbreak of Lassa fever on the Jos plateau, Nigeria, in January-February 1970. A preliminary report. *Am. J. Trop. Med. Hyg.* 19, 695–696. doi: 10.4269/ajtmh.1970.19.695
- Urata, S., and Yasuda, J. (2012). Molecular mechanism of arenavirus assembly and budding. *Viruses* 4, 2049–2079. doi: 10.3390/v4102049
- Vieth, S., Torda, A. E., Asper, M., Schmitz, H., and Gunther, S. (2004). Sequence analysis of L RNA of Lassa virus. *Virology* 318, 153–168. doi: 10.1016/j.virol.2003.09.009
- Vogel, D., Rosenthal, M., Gogrefe, N., Reindl, S., and Gunther, S. (2019). Biochemical characterization of the Lassa virus L protein. *J. Biol. Chem.* 294, 8088–8100. doi: 10.1074/jbc.RA118.006973
- Walker, D. H., Wulff, H., Lange, J. V., and Murphy, F. A. (1975). Comparative pathology of Lassa virus infection in monkeys, Guinea-pigs, and *Mastomys natalensis*. *Bull. World Health Organ.* 52, 523–534
- Whitmer, S. L. M., Strecker, T., Cadar, D., Dienes, H. P., Faber, K., Patel, K., et al. (2018). New lineage of Lassa virus, Togo, 2016. *Emerg. Infect. Dis.* 24, 599–602. doi: 10.3201/eid2403.171905
- Wittmann, T. J., Biek, R., Hassanin, A., Rouquet, P., Reed, P., Yaba, P., et al. (2007). Isolates of Zaire ebolavirus from wild apes reveal genetic lineage and recombinants. *Proc. Natl. Acad. Sci. USA* 104, 17123–17127. doi: 10.1073/pnas.0704076104
- Worobey, M., Rambaut, A., and Holmes, E. C. (1999). Widespread intra-serotype recombination in natural populations of dengue virus. *Proc. Natl. Acad. Sci. USA* 96, 7352–7357. doi: 10.1073/pnas.96.13.7352
- Young, P. R., and Howard, C. R. (1983). Fine structure analysis of Pichinde virus nucleocapsids. *J. Gen. Virol.* 64, 833–842. doi: 10.1099/0022-1317-64-4-833
- Zhang, R., Wang, X., Su, J., Zhao, J., and Zhang, G. (2010). Isolation and analysis of two naturally-occurring multi-recombination Newcastle disease viruses in China. *Virus Res.* 151, 45–53. doi: 10.1016/j.virusres.2010.03.015