**Inference from phylogeography and molecular epidemiology of Lassa virus is limited by sampling and sequencing bias in endemic regions.**

**Authors**

Hayley Free1a§, David Simons1§, Isobella Honeyborne2, Linzy Elton2, Najmul Haider1, Rashid Ansumana3, Richard Kock1, Francine Ntoumi4, Alimuddin Zumla2,5, Timothy D McHugh2, Liã Bárbara Arruda2b#

**Affiliations**

1 The Royal Veterinary College, University of London, Hatfield, UK.

2 Centre for Clinical Microbiology, Division of Infection and Immunity, University College London, London, UK

3 School of Community Health Sciences, Njala University, Bo, Sierra Leone

4 Fondation Congolaise pour la Recherche Médicale (FCRM), Brazzaville, Republic of Congo; Institute for Tropical Medicine, University of Tübingen, Germany

5 NIHR Biomedical Research Centre, UCLHospitals NHS Foundation Trust, London, UK

a Current affiliation Oxford Brookes University, Oxford, UK

b Current affiliation Wellcome Connecting Science, Hinxton, UK

§ Both authors contributed equality to this work

# Corresponding author

**Abstract**

The viral haemorrhagic infection caused by Lassa virus (LASV) is an important endemic zoonotic disease in West African with evidence for increasing outbreak sizes. The Natal multimammate mouse (*Mastomys natalensis*) is the predominant viral reservoir, although few studies have investigated the role of other animal species. To identify sequencing biases , all LASV nucleotide sequences and associated metadata (n = 2,298) available on GenBank were retrieved. Spatial modelling of sequencing effort highlighted the bias in locations of available sequences. Using available sequences phylogenetic analyses showed geographic clustering of LASV lineages, suggested isolated events of human-to-host transmission and the emergence of currently circulating strains of LASV from the year 1498 in Nigeria. Overall, the current study highlights significant geographic limitations in LASV surveillance, particularly, in non-human species. Further investigation of the non-human reservoir of this virus, alongside improved surveillance in other endemic countries, are required for further characterisation of the historic emergence and dispersal of LASV. Accurate assessment on viral circulation in non-human hosts is vital to guide public health interventions to prevent recurrent Lassa fever epidemics.

**Key-words**

Lassa virus; Phylogeography; Metadata

1. **Introduction**

Lassa fever (LF) is a viral haemorrhagic disease, caused by *Lassa mammarenavirus* (LASV) that causes several thousand deaths in West Africa annually (Asogun et al, 2019). The WHO reports that it is endemic in eight West African countries including Benin, Ghana, Guinea, Liberia, Mali, Sierra Leone, Togo and Nigeria [1].

There is limited epidemiological data on LF and making accurate estimates of its true burden remains challenging. Many individuals infected with LASV do not seek healthcare as up to 80% of infections are asymptomatic or present as mild illness [2]. [2,3] Identification of true cases is additionally confounded due to overlapping symptoms with other diseases such as malaria and Ebola (Nnaji 2021; Asogun et al, 2019; Ashcroft et al 2022) and lack of available diagnostic methods (Takah et al., 2019). Access to diagnostic assays varies spatially, increased availability at centers of excellence in Lassa fever treatment and research such as the Irrua Specialist Teaching Hospital, Nigeria and Kenema General Hospital, Sierra Leone results in a spatial bias of samples in these locations. Phylogenetic analysis and molecular dating of sequence clinical and research samples suggest a westward route of dispersal of LASV lineages, from the most recent common ancestor in Nigeria. [6–12].

The Natal multimammate mouse (*Mastomys natalensis*) is the primary reservoir of LASV, however, other rodents have been found to be infected; *Mastomys erythroleucus, Hylomyscus pamfi,* *Mus baoulei* and *Rattus rattus* [9,13–17]. Humans become infected with LASV upon contact with or inhalation of excretions from the rodent species [6,18]. Although human-to-human transmission has been reported – typically associated with nosocomial outbreaks – these are rare events when compared with spillover from rodent hosts [19].

LASV is a bisegmented ssRNA- virus of the family *Arenaviridae* [20,21]. Based on the genomic analysis of the large (L) and small segments (S) LASV has been classified into seven lineages which demonstrate spatial segregation across the endemic range [22]. The high nucleotide variability (25-32%) of these lineages introduces complexity into assays to detect LASV infection. Here, we compiled a comprehensive dataset of publicly available full-segment LASV sequences, spanning West Africa and host species, to inform our understanding of the phylogeny of LASV dispersal. We identified substantial variability in the origin of available sequences and completeness of records. We show strong geographic clustering among lineages supporting prior hypotheses of radiation from both Nigeria and a subsequent introduction into Liberia [23]. The synthesis of available metadata highlights important gaps in currently available data, including spatial bias in the sequencing of samples and should be used to inform the design of epidemiological programmes going forward. A better understanding of LASV phylogeography would improve and support effective implementation of measures to prevent an expected increase in the size of the endemic region due to projected climate, human population and land-use change.

1. **Methods**

**2.1 Data Collection and Processing**

LASV nucleotide and protein sequences were obtained from the National Centre for Biotechnology Information (NCBI) GenBank [24]. The search query run on 24 Sep 2021 was for “Lassa mammarenavirus” in the organism field of the NCBI nucleotide dataset. Data were obtained using the NCBI Entrez API [24] with analysis conducted using the “genbankr” package [25] and the R statistical programme [26]. Associated citations were searched to identify missing metadata for sequences including hosts and geographic location of samples. Sequences with large portions (10% missing compared to reference sequences, NC\_004296.1 and NC\_004297.1 for S and L segments respectively) of missing nucleotide data on the L- or S-segment or lacking associated metadata (collection year, host species, country, and geographical region of sampling) were excluded from phylogenetic analysis. Nucleotide sequences were aligned using the ‘map to reference’ tool on Geneious Prime 20201.2. Alignment, visual inspection and manual editing were performed, and entries that contained >100 continuous ambiguous nucleotide calls were excluded.

**2.3 Sequencing bias**

To understand the bias of sequenced samples at a sub-national level the origin of a sequenced sample was geocoded using the Google Geocoding API through the “ggmap” package [27]. Locations were associated with level-1 administrative regions and data were separated into human and rodent sources of samples to visualise the heterogeneity of spatial sampling. To measure sampling effort bias, the centroid of regions and number of samples obtained from both rodents and humans were used to produce a Generalised Additive Model, with number of sequences uploaded to GenBank as the response variable. In sensitivity analysis the number of reported cases for a region or country was added as a covariate.

* 1. **Phylogenetic Analysis**

Phylogenetic analysis was undertaken through Bayesian Markov Chain Monte Carlo (MCMC) method using BEAST.v1.10.4 [28]. In BEAUTi, the parameters were a substitution model as a generalised time reversible plus gamma site heterogeneity, with codon partition positions 1, 2, 3. A strict clock and a coalescent tree prior with a constant size population was used. Each analysis consisted of 20 million MCMC steps and trees were sampled every 20,000 generations. Sample collection dates from the metadata were used as tip dates to fit to a molecular clock, and country of sample collection was incorporated as a discrete state [10,29]. To assess the log files of the output TRACER.v.1.7.1 was used. Maximum-clade credibility trees were generated through TreeAnnotator v1.8.4 and visualised in FigTree.v1.4.4 [30].

1. **Results**
   1. **Compiled Dataset**

The initial dataset comprised 2,298 records (from samples obtained 1969-2019), including nucleotide sequences and associated metadata. Sequences lacking country information (n = 134) and incomplete gene sequences (n = 906) were removed from phylogenetic analyses. Therefore, 680 sequences of complete S segment and 578 sequences of partial L segment (L protein only) were used. Accession numbers of included and excluded sequences are available in Supplementary table 1.

* 1. **Descriptive Analysis**

Year of collection was available for 2,108 records, with the oldest sequence dating from 1969. Among these records, most sequences (n = 2,063) have been obtained since 2000. Human-derived LASV sequences comprised most of the available records (67%), other host species include *Mastomys natalensis* (29%) and *Mastomys sp.* (3%), while *Mastomys erythroleucus (n=18)*, *Mus baoulei (n=9)* and *Hylomyscus pamfi (n=10)* represent < 1% each. The species sampled was not documented in 107 records. The majority of sequences were produced from samples collected in Nigeria (56%), followed by Guinea (20%), Sierra Leone (14%), Liberia (4%) and Cote d’Ivoire (3%) with the remainder obtained from, Benin, Ghana, Mali and Togo (Figure 1).

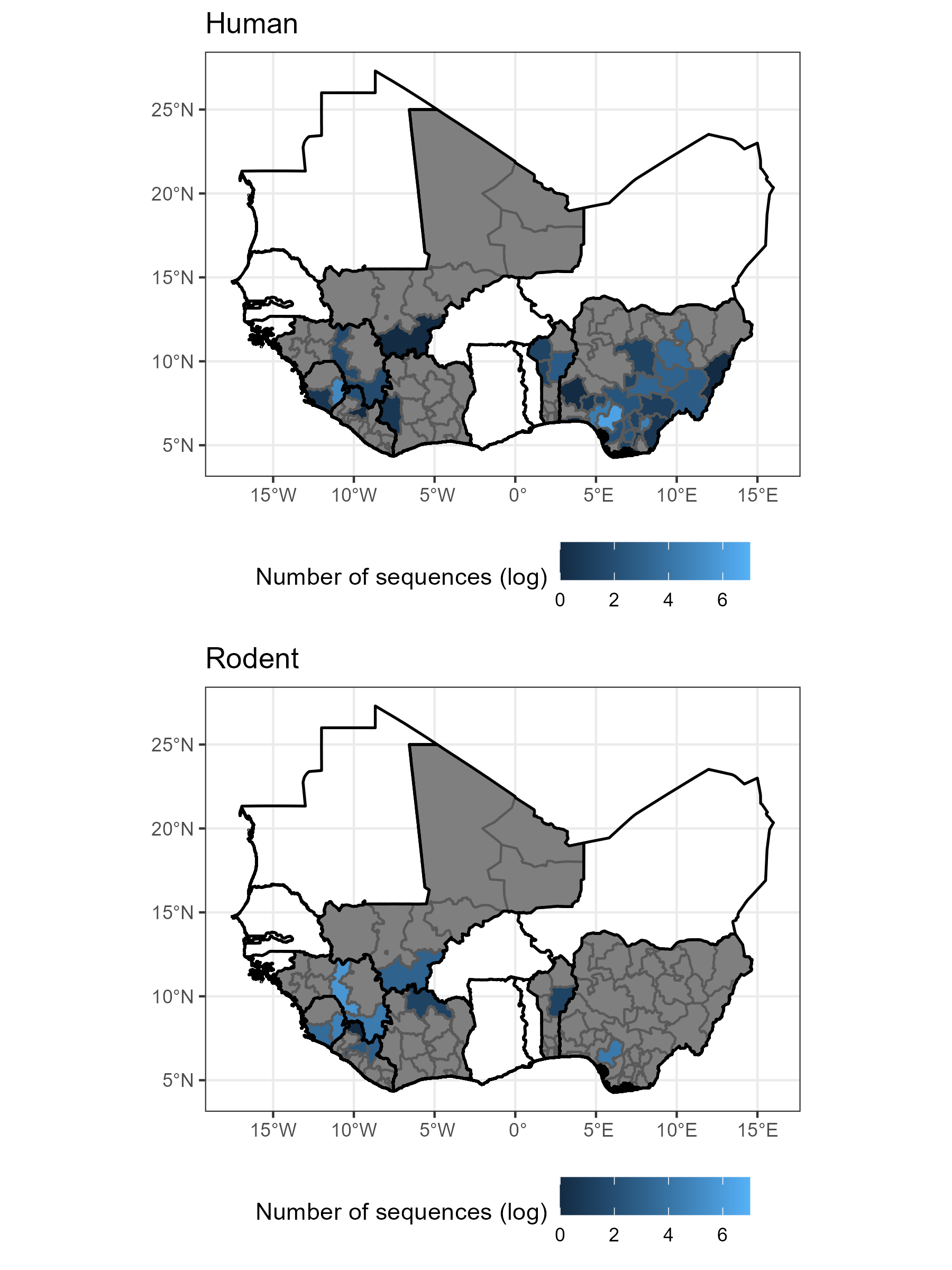


Figure 1 – The number of sequences, shown on a log scale, retrieved from NCBI GenBank with associated location and host for human samples (top, N = 1070) and rodent samples (bottom, N = 656). Sequences for human samples were clustered in Edo State, Nigeria and Eastern Province, Sierra Leone with 51 samples from the remaining endemic countries. Sequences from rodent samples were most commonly obtained from Faranah, Guinea and Eastern Province, Sierra Leone with 136 samples from the remaining endemic countries. Grey regions represent level-1 administrative areas with no sequences within countries that have at least one available sequence. White countries are West African countries with no available *Lassa mammarenavirus* sequences. See Supplementary Figure 3 for country names.

* 1. **Sequencing bias**

Combining both human and rodent derived samples, relative sequencing effort was found to be greatest in Southwest Nigeria, centered over Edo State and the border region of Guinea, Sierra Leone and Liberia centered over the Faranah. Nzérékoré regions of Guinea, Eastern Province of Sierra Leone and Nimba district of Liberia (Figure 2.). Adjusting for the reported number of cases in sensitivity analysis did not have an important effect on the geographic distribution of sequencing effort (Supplementary Figure 1.).

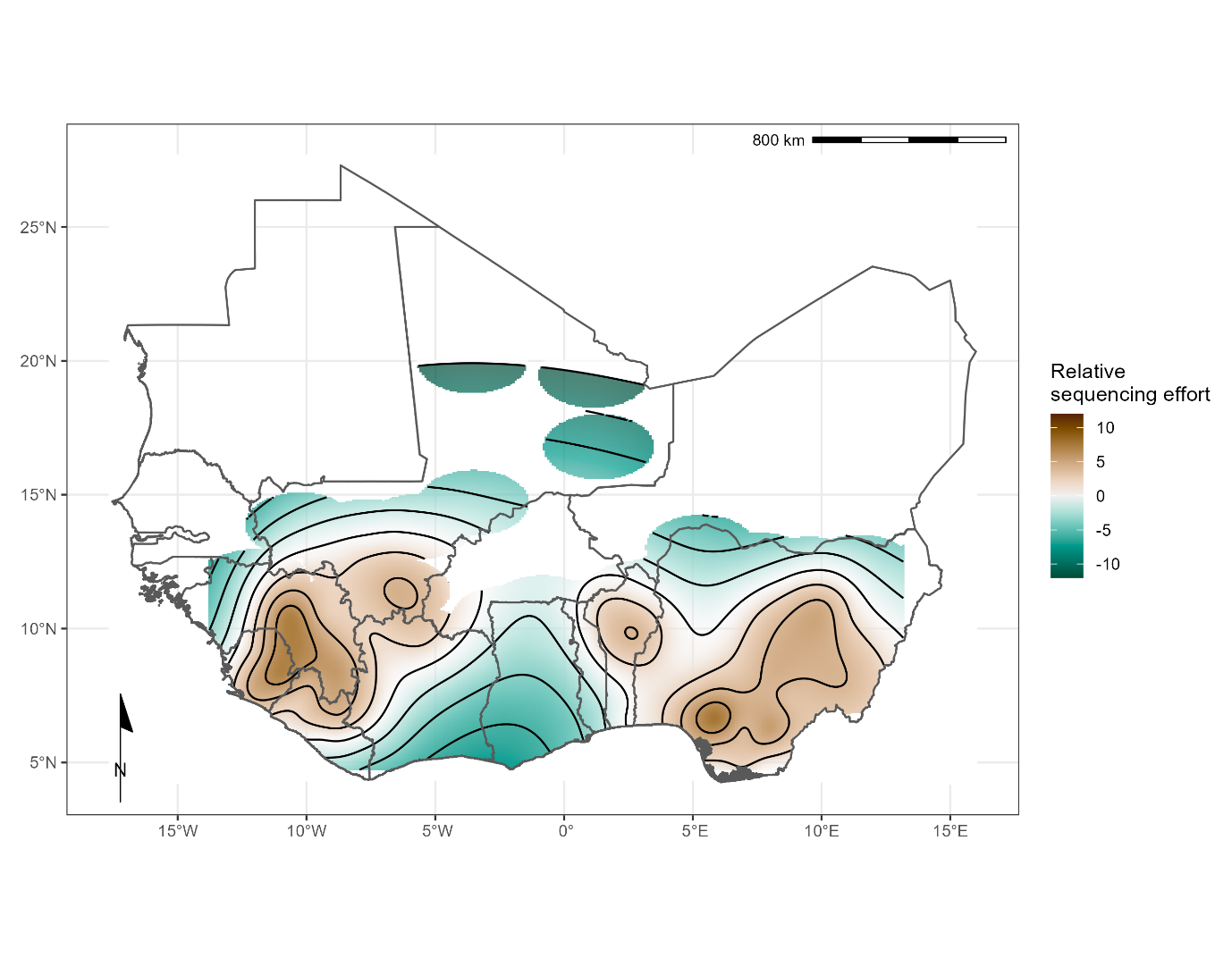


Figure 2 – Modelled relative sequencing effort derived from both human and rodent samples. Greatest sequencing effort coincides with areas where sampling in humans (Edo, Nigeria and Kenema, Sierra Leone) and rodents (Faranah, Guinea) have historically been focussed.

**3.4 Phylogenetic Analysis**

Sequences for each segment of LASV showed clustering according to previously documented lineages I-VII alongside geographical clustering with lineages I-III and VI present in Nigeria, lV in Liberia, Guinea and Sierra Leone, V in Mali and VII in Togo (Figure 3). In this analysis only L segment sequences of lineage V from Cote d’Ivoire were included due to quality control exclusion criteria. The phylogeny of the L segment indicates an older emergence of LASV in the human population, with the most recent common ancestor (MRCA) predicted in the year 828 in Nigeria, inference based on the S segment indicates the emergence in the year 1350 (Table 1).

A picture containing diagram

Description automatically generated

Figure 3 – Time-calibrated phylogenetic analysis of Lassa mammarenavirus (LASV) small (S) segment (left; n = 650) and large (L) segment (right; n = 573). Predicted emergence date (year) is indicated on the axis. Country of collection is indicated by branch colour (key upper left) and host species of collection is indicated by tip shape colour (key lower left). This analysis used Bayesian Markov Chain Monte Carlo method of inference. Lineage is indicated on the right-hand side of each tree; lineage II is collapsed and consisted of sequences from humans only.

Table 1 - The most recent common ancestor (MRCA) according to host and country of collection of Lassa mammarenavirus (LASV) S and L segments. Samples were collected between 1969-2018.

|  |  |  |  |
| --- | --- | --- | --- |
| **Host species** | **Country** | **S segment MRCA** | **L segment MRCA** |
| Homo sapiens (n=1181) | Benin | 1995 | 1989 |
| Guinea | 1895 | 1871 |
| Liberia | 1895 | 1627 |
| Nigeria | 1681 | 1498 |
| Sierra Leone | 1901 | 1874 |
| Togo | 2016 | 2014 |
| Hylomyscus pamfi (n=2) | Nigeria | 1681 | 1498 |
| Mastomys erythroleucus (n=18) | Guinea | 1975 | 2010 |
| Nigeria | 2008 | 2006 |
| Mastomys natalensis (n=36) | Guinea | 1938 | 1997 |
| Mali | 1951 | 2007 |
| Sierra Leone | 1909 | 1979 |

There was a lack of sequence information from lineage I and VI, however, phylogeny suggests these lineages are basal to others in Nigeria (Figure 3). Lineage VII in Togo is most closely related to Nigerian isolates and potentially diverged between 500-900 years ago. The divergence of lineage III and IV is predicted to have occurred between the years 1332-1551. Introduction to countries west of Nigeria appears to be by dispersal initially to Liberia, followed by Guinea in the 1700s, followed by Sierra Leone and Mali approximately 100 years later. A lack of full segment sequences from lineage V limits calculation of divergence from the most recent common ancestor from lineage IV (approximately 200 years). Regional-level data were available for sequences obtained from Nigeria (Supplementary Figure 2A-B). The lineages circulating in Nigeria also tend to form regional clusters, with lineage II dispersed in the southern region and III across the central region of the country (data not shown).

1. **Discussion**

Our analyses of 2,298 LASV sequences obtained from GenBank further informs the dispersal of LASV lineages in West Africa. There are several important findings from this study. First, most sequence data was reported from only three countries: Nigeria (56%), Guinea (20%) and Sierra Leone (14%), highlighting the need for further research and developing increased surveillance, sequencing and reporting capacity in other countries. This bias has been mapped as relative sequencing effort to identify regions where *Lassa mammarenavirus* is considered endemic to support efforts to counteract current sequencing deficits. Second, geographic clustering of LASV lineages, suggested isolated events of host-to-human transmission and the emergence of the first Lassa fever cases dating from 1498 in Nigeria. Third, there was comparatively limited data from non-human hosts and only 69/703 sequences encompassed complete genes. Altogether, the data indicate limited surveillance approaches among animal species, and further investments are required to make available reliable data for accurately defining the space-temporal pathway of *Lassa mammarenavirus*. Strengthening surveillance and research capacities on the non-human host are vital for preventing Lassa fever outbreaks.

The phylogenetic analysis of LASV according to host species appears to maintain the trend of spatial evolution, as opposed to intra-host viral evolution (Figure 3). For instance, LASV sequences from *M. erytholeucus* sampled in Nigeria and Guinea clustered within lineages III and IV, respectively. Interestingly, these isolates appear to occur after the most recent common ancestor of humans and *M. natalensis* in the same country (Table 1), suggesting introduction of LASV to *M. erythroleucus* was secondary. Sequences from *M. natalensis* in Sierra Leone exhibit minimal clustering, and were interspersed with sequences from humans, potentially representing frequent introductory events with spillback into rodent populations from human sources (reverse zoonosis). The most recent common ancestor of LASV sequences from *M. natalensis* in Sierra Leone suggest a relatively later emergence of the virus in this country. Our findings corroborate those of Andersen et al., that within Sierra Leone LASV appears to have emerged in human hosts before rodents (Andersen et al., 2015). However, this data must be caveated by the limited information from rodent species available for analysis.

There is a lower coverage of rodent derived LASV sequences, with those from the primary reservoir *M. natalensis* forming fewer than one third of the sequences (n = 609), with substantially lower sampling of other possible rodent hosts, including other *Mastomys* species. Rodent sampling has not increased on the same trajectory as human samples despite increased sampling effort apparent from 2008 [9,15,31]. There is substantial heterogeneity in the locations in which rodent and human samples are available. Despite a significant number of rodent samples being obtained from Guinea few human sequences are available from these locations. The inverse is true of Nigeria where most human samples and fewer than 80 rodent sequences have been obtained, and all of these from a single state. The number of reported cases was not found to be importantly associated with the number of available sequences. This is suggestive of both under-reporting of human cases and the consolidation of research efforts into few locations. The paucity of full segment sequences from rodents, from limited geographic locations, limits our understanding of viral radiation in rodent hosts, particularly from species which are not considered the primary reservoir, e.g. *H. pamfi.* In the current literature, despite the initial report of LASV in *H. pamfi* in 2016, the most recent common ancestor appears in the late 1600s [9]. It is therefore possible lineage VI and/or *H. pamfi* as a reservoir of LASV has gone undetected due to lack of sufficient sampling.

There were several limitations to interpreting our data. A high number of sequences (70%) from Nigerian and Sierra Leonean samples correlate with the location of Lassa fever research programs, representing spatial ascertainment bias [32–34]. From 2016 there was a substantial increase in the number of LASV sequences in the repository, this reflects increasing research effort, availability of sequencing platforms and increased data gathering during Lassa fever epidemics, such as in the 2018 Nigeria Lassa fever outbreak – the largest known to date [35]. There are notably fewer recorded sequences of LASV from Benin, Togo, and Ghana, potentially suggesting a potential a gap in surveillance and research capacity in these locations or a lack of circulating LASV. Despite almost 40% of the original dataset being removed due to incomplete sequences (n=906) or missing data on country of sample collection (n=134) phylogenetic analysis on included samples demonstrated geographic clustering of LASV lineages, supporting previous research [8–10,29,36–39]. LASV surveillance and case finding in Nigeria has improved since the establishment of the Nigerian Center for Disease Control in 2011, leading to an overrepresentation of LASV sequences from this region on GenBank [40]. Therefore, it was possible to evaluate the data at a regional-level corroborating previous finding that lineages II and III show clustering aligning with the progression of their ancestry from North-East to South-West within Nigeria [7,41].

A substantial number (n = 906) of the sequences retrieved corresponded to short fragments derived from PCR products used for diagnostic purposes rather than for viral genomic surveillance. LASV is a segmented virus, and it was not possible to identify complete genome sequences since both S and L segments are reported separately on the sequence’s repository. The molecular clock analyses from L protein indicated an earlier emergence of LASV when compared to S segment analysis (828 and 1350 respectively), possibility because the viral RNA polymerase (L protein) is less affected by selective pressure than the S segment [6,36,42].

Nevertheless, our study has synthesised available data on LASV sequences available on GenBank to investigate the location and period of sampling to reconstruct the viral lineages dispersal across the endemic region. We corroborated the strong lineage and geographic clustering of LASV samples, supporting the role of the rodent reservoir to sustain the endemic cycle. Despite the regionalisation of LF being pivotally driven by rodent-to-human transmission, there is still scarce LASV genomic data from animal species, suggesting limited surveillance approaches investigating the disease reservoir. The intensification of surveillance programmes targeting animal species will not only improve the understanding of the relationships host-pathogen, but also provide instrumental data to support public policies to respond more efficiently to public health emergencies.

**Supplementary material**

Supplementary table 1 presents the GenBank accession number of analysed sequences, including available data about host, country, region, year, sequence length, genome segment (L or S) and predicted MRCA.

**Funding**

Pan-African Network for Rapid Research, Response and Preparedness for Infectious Diseases Epidemics – PANDORA-ID-NET ([www.pandora-id.net](http://www.pandora-id.net)), funded through the European and Developing Countries Clinical Trials Partnership (EDCTP) (grant number RIA2016E-1609). DS is supported by a PhD studentship from the UK Biotechnology and Biological Sciences Research Council [BB/M009513/1]. AZ is in receipt of UK NIHR Senior Investigator Award.

**Author contributions**

Conceptualisation: LBA; Methodology: HF, DS and LBA; Formal Analyses: HF, DS and LBA; Investigation: HF, DS and LBA; Supervision: LBA; Data Curation: HF and DS; Writing – original draft preparation: HF, DS, LBA; Writing – Review and Editing: IH, LE, NH, RA, RK, FN AZ and TMcH; Funding acquisition: AZ and FN.

**Data availability and reproducibility**

All data used in these analyses are publicly available from GenBank. The accession numbers of records used and code to reproduce the metadata analyses are available as an archived Git release on Zenodo (<https://doi.org/10.5281/zenodo.6340162>)

**Conflict of interests**

The authors declare no conflict of interests

**References**

1. World Health Organisation Lassa Fever Available online: https://www.who.int/health-topics/lassa-fever#tab=tab\_1 (accessed on 22 February 2022).

2. McCormick, J.B.; Webb, P.A.; Krebs, J.W.; Johnson, K.M.; Smith, E.S. A Prospective Study of the Epidemiology and Ecology of Lassa Fever. *J Infect Dis* **1987**, *155*, 437–444, doi:10.1093/infdis/155.3.437.

3. Basinski, A.J.; Fichet-Calvet, E.; Sjodin, A.R.; Varrelman, T.J.; Remien, C.H.; Layman, N.C.; Bird, B.H.; Wolking, D.J.; Monagin, C.; Ghersi, B.M.; et al. Bridging the Gap: Using Reservoir Ecology and Human Serosurveys to Estimate Lassa Virus Spillover in West Africa. *PLoS Comput Biol* **2021**, *17*, e1008811, doi:10.1371/journal.pcbi.1008811.

4. Nnaji, N.D.; Onyeaka, H.; Reuben, R.C.; Uwishema, O.; Olovo, C.V.; Anyogu, A. The Deuce-Ace of Lassa Fever, Ebola Virus Disease and COVID-19 Simultaneous Infections and Epidemics in West Africa: Clinical and Public Health Implications. *Tropical Medicine and Health* **2021**, *49*, 102, doi:10.1186/s41182-021-00390-4.

5. Takah, N.F.; Brangel, P.; Shrestha, P.; Peeling, R. Sensitivity and Specificity of Diagnostic Tests for Lassa Fever: A Systematic Review. *BMC Infectious Diseases* **2019**, *19*, 647, doi:10.1186/s12879-019-4242-6.

6. Andersen, K.G.; Shapiro, B.J.; Matranga, C.B.; Sealfon, R.; Lin, A.E.; Moses, L.M.; Folarin, O.A.; Goba, A.; Odia, I.; Ehiane, P.E.; et al. Clinical Sequencing Uncovers Origins and Evolution of Lassa Virus. *Cell* **2015**, doi:10.1016/j.cell.2015.07.020.

7. Bowen, M.D.; Rollin, P.E.; Ksiazek, T.G.; Hustad, H.L.; Bausch, D.G.; Demby, A.H.; Bajani, M.D.; Peters, C.J.; Nichol, S.T. Genetic Diversity among Lassa Virus Strains. *Journal of Virology* **2000**, *74*, 6992–7004, doi:10.1128/jvi.74.15.6992-7004.2000.

8. Manning, J.T.; Forrester, N.; Paessler, S. Lassa Virus Isolates from Mali and the Ivory Coast Represent an Emerging Fifth Lineage. *Frontiers in Microbiology* **2015**, doi:10.3389/fmicb.2015.01037.

9. Olayemi, A.; Cadar, D.; Magassouba, N.; Obadare, A.; Kourouma, F.; Oyeyiola, A.; Fasogbon, S.; Igbokwe, J.; Rieger, T.; Bockholt, S.; et al. New Hosts of The Lassa Virus. *Scientific Reports* **2016**, doi:10.1038/srep25280.

10. Olayemi, A.; Adesina, A.S.; Strecker, T.; Magassouba, N.; Fichet-Calvet, E. Determining Ancestry between Rodent-and Human-Derived Virus Sequences in Endemic Foci: Towards a More Integral Molecular Epidemiology of Lassa Fever within West Africa. *Biology* **2020**, doi:10.3390/biology9020026.

11. Whitmer, S.L.M.; Strecker, T.; Cadar, D.; Dienes, H.P.; Faber, K.; Patel, K.; Brown, S.M.; Davis, W.G.; Klena, J.D.; Rollin, P.E.; et al. New Lineage of Lassa Virus, Togo, 2016. *Emerging Infectious Diseases* **2018**, *24*, 599–602, doi:10.3201/eid2403.171905.

12. Okoro, O.A.; Bamgboye, E.; Dan-Nwafor, C.; Umeokonkwo, C.; Ilori, E.; Yashe, R.; Balogun, M.; Nguku, P.; Ihekweazu, C. Descriptive Epidemiology of Lassa Fever in Nigeria, 2012-2017. *Pan Afr Med J* **2020**, *37*, 15, doi:10.11604/pamj.2020.37.15.21160.

13. Bangura, U.; Buanie, J.; Lamin, J.; Davis, C.; Bongo, G.N.; Dawson, M.; Ansumana, R.; Sondufu, D.; Thomson, E.C.; Sahr, F.; et al. Lassa Virus Circulation in Small Mammal Populations in Bo District, Sierra Leone. *BIOLOGY-BASEL* 2021, *10*.

14. Forni, D.; Sironi, M. Population Structure of Lassa Mammarenavirus in West Africa. *Viruses* **2020**, *12*, 437.

15. Lecompte, E.; Fichet-Calvet, E.; Daffis, S.; Koulémou, K.; Sylla, O.; Kourouma, F.; Doré, A.; Soropogui, B.; Aniskin, V.; Allali, B.; et al. Mastomys Natalensis and Lassa Fever, West Africa. *Emerging Infectious Diseases* **2006**, doi:10.3201/eid1212.060812.

16. Wulff, H.; Fabiyi, A.; Monath, T.P. Recent Isolations of Lassa Virus from Nigerian Rodents. *Bull World Health Organ* **1975**, *52*, 609–613, doi:PMC2366652.

17. Yadouleton, A.; Agolinou, A.; Kourouma, F.; Saizonou, R.; Pahlmann, M.; Bedié, S.K.; Bankolé, H.; Becker-Ziaja, B.; Gbaguidi, F.; Thielebein, A.; et al. Lassa Virus in Pygmy Mice, Benin, 2016-2017. *Emerging Infectious Diseases* **2019**, doi:10.3201/eid2510.180523.

18. Oti, V.B. A Reemerging Lassa Virus: Aspects of Its Structure, Replication, Pathogenicity and Diagnosis. In *Current Topics in Tropical Emerging Diseases and Travel Medicine*; Alfonso J. Rodriguez-Morales, Ed.; BoD – Books on Demand, 2018.

19. Lo Iacono, G.; Cunningham, A.A.; Fichet-Calvet, E.; Garry, R.F.; Grant, D.S.; Khan, S.H.; Leach, M.; Moses, L.M.; Schieffelin, J.S.; Shaffer, J.G.; et al. Using Modelling to Disentangle the Relative Contributions of Zoonotic and Anthroponotic Transmission: The Case of Lassa Fever. *PLoS Neglected Tropical Diseases* **2015**, doi:10.1371/journal.pntd.0003398.

20. Hallam, S.J.; Koma, T.; Maruyama, J.; Paessler, S. Review of Mammarenavirus Biology and Replication. *Frontiers in Microbiology* **2018**, *9*, 1–8, doi:10.3389/fmicb.2018.01751.

21. Günther, S.; Lenz, O. Lassa Virus. *Critical Reviews in Clinical Laboratory Sciences* 2004.

22. Welch, S.R.; Scholte, F.E.M.; Albariño, C.G.; Kainulainen, M.H.; Coleman-McCray, J.D.; Guerrero, L.W.; Chakrabarti, A.K.; Klena, J.D.; Nichol, S.T.; Spengler, J.R.; et al. The S Genome Segment Is Sufficient to Maintain Pathogenicity in Intra-Clade Lassa Virus Reassortants in a Guinea Pig Model. *Frontiers in Cellular and Infection Microbiology* **2018**, *8*.

23. Klitting, R.; Kafetzopoulou, L.E.; Thiery, W.; Dudas, G.; Gryseels, S.; Kotamarthi, A.; Vrancken, B.; Gangavarapu, K.; Momoh, M.; Sandi, J.D.; et al. *Predicting the Evolution of Lassa Virus Endemic Area and Population at Risk over the next Decades*; Microbiology, 2021;

24. National Center for Biotechnology Information National Center for Biotechnology Information Available online: https://www.ncbi.nlm.nih.gov/ (accessed on 3 February 2022).

25. Becker, G.; Lawrence, M. Genbankr: Parsing GenBank Files into Semantically Useful Objects 2021.

26. R Core Team R: A Language and Environment for Statistical Computing 2021.

27. Kahle, D.; Wickham, H. Ggmap: Spatial Visualization with Ggplot2. *The R Journal* **2013**, *5*, 144–161, doi:10.32614/RJ-2013-014.

28. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10. *Virus Evolution* **2018**, doi:10.1093/ve/vey016.

29. Olayemi, A.; Fichet-Calvet, E. Systematics, Ecology, and Host Switching: Attributes Affecting Emergence of the Lassa Virus in Rodents across Western Africa. *Viruses* 2020.

30. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* **2018**, doi:10.1093/sysbio/syy032.

31. Lecompte, E.; Brouat, C.; Duplantier, J.M.; Galan, M.; Granjon, L.; Loiseau, A.; Mouline, K.; Cosson, J.F. Molecular Identification of Four Cryptic Species of Mastomys (Rodentia, Murinae). *Biochemical Systematics and Ecology* **2005**, doi:10.1016/j.bse.2004.12.015.

32. Townsend Peterson, A.; Moses, L.M.; Bausch, D.G. Mapping Transmission Risk of Lassa Fever in West Africa: The Importance of Quality Control, Sampling Bias, and Error Weighting. *PLoS ONE* **2014**, doi:10.1371/journal.pone.0100711.

33. Ehichioya, D.U.; Hass, M.; Ölschläger, S.; Becker-Ziaja, B.; Onyebuchi Chukwu, C.O.; Coker, J.; Nasidi, A.; Ogugua, O.O.; Günther, S.; Omilabu, S.A. Lassa Fever, Nigeria, 2005-2008. *Emerging Infectious Diseases* 2010.

34. Khan, S.H.; Goba, A.; Chu, M.; Roth, C.; Healing, T.; Marx, A.; Fair, J.; Guttieri, M.C.; Ferro, P.; Imes, T.; et al. New Opportunities for Field Research on the Pathogenesis and Treatment of Lassa Fever. *Antiviral Research* **2008**, doi:10.1016/j.antiviral.2007.11.003.

35. Control, N.C. for D. *Lassa Fever Situation Report*; 2018;

36. Ibukun, F.I. Inter-Lineage Variation of Lassa Virus Glycoprotein Epitopes: A Challenge to Lassa Virus Vaccine Development. *Viruses* 2020.

37. Lalis, A.; Leblois, R.; Lecompte, E.; Denys, C.; ter Meulen, J.; Wirth, T. The Impact of Human Conflict on the Genetics of Mastomys Natalensis and Lassa Virus in West Africa. *PLoS ONE* **2013**, *7*, doi:10.1371/journal.pone.0037068.

38. Wiley, M.R.; Fakoli, L.; Letizia, A.G.; Welch, S.R.; Ladner, J.T.; Prieto, K.; Reyes, D.; Espy, N.; Chitty, J.A.; Pratt, C.B.; et al. Lassa Virus Circulating in Liberia: A Retrospective Genomic Characterisation. *The Lancet Infectious Diseases* **2019**, doi:10.1016/S1473-3099(19)30486-4.

39. Yadouleton, A.; Picard, C.; Rieger, T.; Loko, F.; Cadar, D.; Kouthon, E.C.; Job, E.O.; Bankolé, H.; Oestereich, L.; Gbaguidi, F.; et al. Lassa Fever in Benin: Description of the 2014 and 2016 Epidemics and Genetic Characterization of a New Lassa Virus. *Emerging Microbes & Infections* **2020**, 1–23.

40. Naidoo, D.; Ihekweazu, C. Nigeria’s Efforts to Strengthen Laboratory Diagnostics – Why Access to Reliable and Affordable Diagnostics Is Key to Building Resilient Laboratory Systems. *African Journal of Laboratory Medicine* **2020**, *9*, doi:10.4102/ajlm.v9i2.1019.

41. Ehichioya, D.U.; Hass, M.; Becker-Ziaja, B.; Ehimuan, J.; Asogun, D.A.; Fichet-Calvet, E.; Kleinsteuber, K.; Lelke, M.; Ter Meulen, J.; Akpede, G.O.; et al. Current Molecular Epidemiology of Lassa Virus in Nigeria. *Journal of Clinical Microbiology* **2011**, doi:10.1128/JCM.01891-10.

42. Hastie, K.M.; Saphire, E.O. Lassa Virus Glycoprotein: Stopping a Moving Target. *Current Opinion in Virology* 2018.