CHARACTERIZATION AND DISTRIBUTION OF LASSA VIRUS IN THE
NATURAL HOST RESEVOIR

AN ABSTRACT

SUBMITTED ON THE 13TH DAY OF DECEMBER 2021

TO THE GRADUATE PROGRAM IN BIOMEDICAL SCIENCES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE GRADUATE SCHOOL
OF TULANE UNIVERSITY
FOR THE DEGREE OF DOCTOR
OF
PHILOSOPHY BY

Allison R. Smither

APPROVED:

Robert F. Garry, Ph.D., Director
Monica E. Embers, Ph.D.

John S. Schieffelin, M.D., M.S.P.H
Jeffrey G. Shaffer, Ph.D.

Cindy A. Morris, Ph.D.
ABSTRACT

Lassa virus (LASV) is an ambisense mammarenavirus that causes Lassa fever, which burdens an estimated 100,000-300,000 people every year in West Africa. The natural host reservoir of the virus, the multimammate rat *Mastomys natalensis*, transmits the virus to humans via its bodily secretions. Despite that *M. natalensis* is the most broadly distributed rodent in sub-Saharan Africa, the Eastern Province of Sierra Leone has some of the highest incidences of Lassa fever throughout the continent. However, in recent years the number of patients presented or admitted to hospitals with suspicion of Lassa fever has decreased dramatically. Thus, we wanted to determine the presence of LASV within the host reservoir in the Kenema district, while also creating a diagnostic algorithm to identify active LASV infections within the host reservoir for further downstream sequencing analysis. To that end, we performed a cross-sectional study of small mammals in the Eastern Province of Sierra Leone in which we collected serum and tissues collected from over 500 small mammals between November 2018 and July 2019 and screened them for three biomarkers of LASV infection: antigen presence with a LASV antigen rapid test (RDT), viral presence by qRT-PCR, and IgG antibodies. RDTs deployed at the time of necropsy detected antigen in over 30% of *Mastomys spp.* specimens collected. Yet, positive antigen status itself as detected by RDT could not prove whether a specimen had an active LASV infection and was much less specific than qRT-PCR at detecting virus. Meanwhile, IgG antibody tests of collected serum revealed similar levels of LASV exposure in small
mammals in line with previous studies. Principal component-style analyses of the three biomarkers were used to create a definition of an active LASV infection in a specimen, which we defined as an antigen positive specimen with virus detected in two or more tissues by qRT-PCR at levels greater than 200 copies/µL RNA. Finally, tissue samples from actively infected specimens were sequenced and compared to human sequences to further elucidate the unique evolutionary history of LASV in Sierra Leone. Taken together, the results indicate high levels of LASV are present in Kenema district within the host reservoir, that RDTs can serve as a useful tool to screen collected specimens for potential active LASV infection, and that the evolution of LASV in Sierra Leone is more dictated by geography than viral transmission chains between rodents and humans. The ability to screen small mammal specimens for potential LASV presence will be useful not only for rodent control practices but also to prioritize specific specimens for further downstream analyses to further Lassa fever knowledge and therapies.
CHARACTERIZATION AND DISTRIBUTION OF LASSA VIRUS IN THE NATURAL HOST RESEVOIR

A DISSERTATION

SUBMITTED ON THE 13TH DAY OF DECEMBER 2021

TO THE GRADUATE PROGRAM IN BIOMEDICAL SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE GRADUATE SCHOOL

OF TULANE UNIVERSITY

FOR THE DEGREE OF DOCTOR

OF PHILOSOPHY

BY

Allison R. Smither

APPROVED:

Robert F. Garry, Ph.D., Director

Monica E. Embers, Ph.D.

Cindy A. Morris, Ph.D.

John S. Schieffelin, M.D., M.S.P.H

Jeffrey G. Shaffer, Ph.D.
ACKNOWLEDGEMENTS

I first off want to thank Dr. Garry and Dr. Schieffelin for the intellectual freedom to design and pursue this project, as well as their expert guidance along the way. I would also like to thank the other members of my committee: Dr. Shaffer for the immense help with the biostatistics involved with this project, Dr. Embers for the insightful questions and avenues explored during the proposal process, and Dr. Morris for clinical laboratory advice, particularly at the start of the COVID-19 pandemic. In addition to my committee, I am thankful for the support of Zalgen Labs, in particular Dr. Matt Boisen, for the technical support during assay development, and the lab of Dr. Kristian Andersen at Scripps Research, especially Dr. Refugio Robles-Sikisaka and Dr. Raphaëlle Klitting, for the heavy lifting behind the sequencing and phylogenetic analyses of this project. I am also grateful for the past and present members of the Garry and Schieffelin labs, and the friends I’ve made in the BMS program at Tulane who have all been there to weather the ups and downs of research. Also thank you to Dr. Zac Pursell and Dr. Jim Jackson in the biochemistry department, with whom I had many helpful conversations about troubleshooting assays.

This research was made possible by a multi-national collaboration between members of the Viral Hemorrhagic Fever Consortium, spearheaded at Kenema Government Hospital (KGH) in Sierra Leone. I thank every member of the KGH Lassa fever program across the admin, laboratory, outreach, and ecology teams for their assistance and dedication to Lassa fever research. I am deeply indebted to the ecology and
outreach teams who spent many nights camped in villages to collect specimens for the study: senior ecology officer James Koninga, junior ecology officer Franklyn Kanneh, ecology assistant Momoh Foday, senior outreach officer Lansana Kanneh, outreach officers Mohammed Yillah and Ibrihiam Mustapha, as well as driver Chris White. I sincerely hope to be working alongside them and everyone else from this outstanding group for years to come.

Finally, I want to say thank you for the unwavering love and support from my family over the years. To my sister Susan, who has always been close by to lend a helping hand. To my dad, who has been my biggest cheerleader since day 1. To my mom, Dr. Anna Robichaux Smither, and granny Dr. Frances Ainsworth Robichaux, to inspire and encourage my intellectual pursuits wherever in the world they lead me. Lastly, to my husband David and dog Daisy for your steadfast patience and commitment now and always.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .............................................................................................................................. i

LIST OF TABLES ........................................................................................................................................ viii

LIST OF FIGURES ...................................................................................................................................... ix

LIST OF COMMONLY USED ABBREVIATIONS .................................................................................... x

1. INTRODUCTION .................................................................................................................................... 1

1.1. Lassa Fever and Lassa Virus .............................................................................................................. 1

1.1.1. Lassa Fever ..................................................................................................................................... 1

1.1.2. Lassa Virus ..................................................................................................................................... 2

1.1.3. Lassa Virus Epidemiology ................................................................................................................. 3

1.2. Reservoirs of Lassa Virus .................................................................................................................. 4

1.2.1. Mastomys natalensis ....................................................................................................................... 4

1.2.2. Other Rodents ................................................................................................................................. 4

1.2.3. Additional Hosts ............................................................................................................................. 5

1.3. Epizoology Studies of Lassa Virus .................................................................................................... 6

1.3.1. Guinea ............................................................................................................................................. 6

1.3.2. Sierra Leone and Liberia ............................................................................................................... 7

1.3.3. Nigeria ........................................................................................................................................... 9
1.3.4 Other West African Countries ................................................................. 10

1.4 Dynamics of Lassa Virus in Mastomys natalensis ........................................ 12

1.4.1 Spatial and Temporal Dynamics ............................................................... 12

1.4.2 Transmission Modalities of Lassa Virus in the Host Reservoir ................. 14

1.4.3 Stopping the Cycle of Transmission ........................................................... 19

1.5 Evolutionary History of Old-World Mammarenaviruses in Rodents .......... 20

1.5.1 The Distinct Niche of Lassa Virus in West Africa ..................................... 20

1.5.2 Unraveling the Co-speciation Hypothesis ................................................. 21

1.5.3 LASV Sequencing and Phylogenetics ......................................................... 25

1.6 LASV Diagnostics ....................................................................................... 29

1.6.1 Nucleic Acid Amplification Tests ............................................................. 29

1.6.2 Antigen Tests ......................................................................................... 30

1.6.3 Antibody Tests ....................................................................................... 31

2. RESEARCH OBJECTIVE .............................................................................. 33

3. ASSESSMENT OF LASSA VIRUS PREVALENCE IN THE NATURAL HOST
RESERVOIR WITHIN IN THE KENEMA DISTRICT OF SIERRA LEONE ............ 34

3.1 Materials and Methods ............................................................................... 34

3.1.1 Ethics and biosafety Statement .............................................................. 34

3.1.2 Study design ........................................................................................... 34

3.1.3 Small mammal trapping ........................................................................ 37
4. CHARACTERIZATION OF LASSA VIRUS ANTIBODIES CIRCULATING IN
STUDY POPULATION AND DEVELOPMENT OF A DIAGNOSTIC ALGORITHM
TO SCREEN RODENTS FOR ACTIVE LASSA VIRUS INFECTIONS ................. 66

4.1 Materials and Methods ............................................................................................................ 66

4.1.1 Adaptation of human LASV IgG ELISA for detection of rodent LASV IgG. 66

4.1.2 Rodent LASV IgG ELISA .................................................................................................. 67

4.1.3 Statistical Analysis – Antibody ELISA .............................................................................. 68

4.1.4 Development of a diagnostic algorithm to screen rodents for active LASV
infection ........................................................................................................................................... 69

4.2 Results ..................................................................................................................................... 71

4.2.1 Modification of a commercially available LASV immunoassay ................................. 71

4.2.2 IgG antibody presence by ELISA ...................................................................................... 71

4.2.3 Differences in GP and NP IgG antibody response ............................................................. 73

4.2.4 Correlation between antigen status or qRT-PCR status and antibody status .. 76

4.2.5 Contributions of biomarkers to variance within the study population ............. 81

4.3 Discussion .................................................................................................................................. 83

4.3.1 Modification of a commercially available LASV immunoassay ................................. 83

4.3.2 IgG antibody presence by ELISA ...................................................................................... 83

4.3.3 Differences in GP and NP IgG antibody response ............................................................. 85

4.3.4 Correlation between antigen status or qRT-PCR status and antibody status .. 86
4.3.5 Contributions of biomarkers to variance within the study population ........ 87

5. VIRAL SEQUENCING OF ACTIVELY INFECTED RODENT SPECIMENS AND
PHYLOGEOGRAPHIC ANALYSIS OF LASV DIVERSITY WITHIN SIERRA
LEONE IN HUMANS AND RODENTS * ................................................................. 89

5.1 Materials and Methods .................................................................................. 89

5.1.1 Library generation and sequencing ............................................................... 89

5.1.2 Assembly of full-length LASV genomes ....................................................... 90

5.1.3 Sequence curation and alignment ................................................................. 91

5.1.4 Phylogenetic analysis and phylogeographic analysis .................................... 92

5.2 Results ............................................................................................................. 93

5.2.1 Sequences obtained from tested specimens ............................................... 93

5.2.2 Phylogenetic and phylogeographic trees of sequenced LASV specimens ..... 97

5.3 Discussion ....................................................................................................... 103

5.3.1 Sequences obtained from tested specimens ............................................... 103

5.3.2 Phylogenetic and phylogeographic trees analysis of sequenced LASV
specimens ............................................................................................................. 104

6. CONCLUSIONS AND FUTURE DIRECTIONS ............................................. 106

7. REFERENCES .................................................................................................... 112
LIST OF TABLES

Table 1: Names and GPS coordinates, in decimal degrees, of villages trapped in study. 35
Table 2: Cytochrome b primers used in genotyping experiments. 41
Table 3: Characteristics of specimens in study. 46
Table 4: Counts of Species by cytochrome b identification. 47
Table 5: Summary of trapping activity by village. 49
Table 6: Composition of specimens captured by villages and their distributions. 50
Table 7: Morphometric characteristics of specimens grouped by antigen RDT result. 52
Table 8: qRT-PCR status of tissues from antigen positive specimens. 56
Table 9: Frequency table of antigen and qRT-PCR status and its associated sensitivity and specificity. 58
Table 10: Factors and variables used to compute FAMD and HCPC analyses. 70
Table 11: Morphometric characteristics of specimens grouped by IgG ELISA result. 72
Table 12: Frequency table of antigen and IgG antibody status and its associated sensitivity and specificity. 77
Table 13: qRT-PCR vs IgG antibody status. 79
Table 14: LASV genome sequences obtained from actively infected specimens. 96
LIST OF FIGURES

Figure 1: Overview of LASV transmission routes.........................................................18
Figure 2: Location and reservoir of known Old-World mammarenaviruses throughout sub-Saharan Africa............................................................................................................ 22
Figure 3: GPS location of villages trapped in study within Sierra Leone. ................. 36
Figure 4: Possible test results of the Zalgen ReLASV Pan-Lassa Antigen Rapid Test.... 39
Figure 5: Traits of antigen presence among study population................................. 53
Figure 6: Viral levels in tissues of specimens............................................................... 55
Figure 7: Antigen strength vs viral levels. ................................................................. 59
Figure 8: Traits of IgG antibody presence among study population. ....................... 74
Figure 9: Differences between GP and NP IgG antibody responses. ......................... 75
Figure 10: Correlations between antibody presence and antigen status. ............... 78
Figure 11: Correlations between antibody presence and viral status by qRT-PCR..... 80
Figure 12: FAMD and HCPC analysis to assess variance within study population....... 82
Figure 13: Viral levels of specimens attempted for sequencing and genome coverage of curated LASV sequences. ................................................................. 95
Figure 14: Phylogenetic trees of S and L LASV segments. .......................................... 98
Figure 15: Phylogeographic analysis of LASV in Sierra Leone............................... 100
LIST OF COMMONLY USED ABBREVIATIONS

Ag  Antigen
IgG  Immunoglobulin G
IgM  Immunoglobulin M
KGH  Kenema Government Hospital
LASV  Lassa virus
RDT  Rapid Diagnostic Test
qRT-PCR/  Quantitative reverse-transcription polymerase chain reaction
qPCR
Spp.  Species
1. INTRODUCTION


1.1. Lassa Fever and Lassa Virus

1.1.1. Lassa Fever

Lassa virus (LASV) causes Lassa fever, a severe viral hemorrhagic fever spread primarily by the rodent *Mastomys natalensis* (*M. natalensis*) in West Africa. With an estimated 100,000–300,000 cases and more than 5000 deaths yearly, it is the most consequential rodent-borne virus worldwide (Monath 1975). Lassa fever was first identified in 1969, when missionary nurses in northeastern Nigeria contracted a hereinto unknown illness from a patient presenting to a local hospital.

Approximately 80% of individuals who contract Lassa fever will experience mild symptoms, such as fever, headache, and general malaise. The remaining 20% develop severe symptoms, including edema, vomiting, hemorrhage, and shock. Up to 2/3 of individuals who develop severe Lassa fever will die, including up to 95% of pregnant women (Branco, Boisen et al. 2011, Shaffer, Schieffelin et al. 2019). With no vaccine or other specific treatment measures currently available, rodent exclusion methods are currently the most effective preventive measure against Lassa fever.
LASV is transmitted through inhalation of aerosolized rodent urine in droplets or dust particles, as well as via ingestion of food or water, or contact with fomites contaminated with rodent urine or droppings. Upon inhalation, consumption, or handling of infectious material, the virus enters the body through contact with mucous membranes, cuts in the skin, or gastrointestinal inclusion, where it infects tissue resident macrophages (Walker, McCormick et al. 1982, Kenyon, McKee et al. 1992). For reasons not yet fully understood, the infection is either contained or disseminated throughout the body, causing fulminant Lassa fever.

1.1.2. Lassa Virus

Lassa virus (LASV), formally *Lassa mammarenavirus*, is an ambisense enveloped RNA virus belonging to the *Arenaviridae* family, order *Bunyavirales*. The LASV genome consists of four genes on two segments: Glycoprotein precursor (GP) and Nucleoprotein (NP) genes on the positive sense small segment, and L and Z genes on the negative sense large segment. The NP gene encodes the nucleoprotein, which encases the genome within the viral particle (Hastie, Liu et al. 2011). The GP gene is cleaved into two envelope proteins, GP1 and GP2, which form a heterotrimer complex that allows for binding with the host receptor alpha-dystroglycan (αDG) and entry into the cell, respectively (Hastie, Zandonatti et al. 2017). The L gene on the large segment encodes the RNA-dependent RNA polymerase, while the Z gene encodes the matrix protein, a small zinc finger protein which is involved in viral budding (Eichler, Strecker et al. 2004, Peng, Xu et al. 2020).

LASV and lymphocytic choriomeningitis virus (LCMV) use an endocytic entry pathway independent of clathrin, caveolin, dynamin and actin (Vela, Zhang et al. 2007). Once within the cell the viruses are rapidly delivered to endosomes via vesicular
trafficking, albeit one that is largely independent of the small GTPases Rab5 and Rab7. Upon contact with the endosome, pH-dependent membrane fusion occurs and is mediated by the envelope glycoprotein, which at the lower pH of the endosome binds the lysosome protein LAMP1 which results in membrane fusion and escape from the endosome (Cosset, Marianneau et al. 2009, Nunberg and York 2012). Once the viral particle enters a cell, the viral polymerase will first transcribe both the NP and L genes, then functions as a replicase to create an antigenome from which the GP and Z genes are then transcribed.

1.1.3 Lassa Virus Epidemiology

The natural ecology of the rodent reservoir *M. natalensis* is critical to the understanding of Lassa fever epidemiology. Of hundreds of rodent species present throughout sub-Saharan Africa, *M. natalensis* is the most broadly distributed (Wilson 1993). Despite the broad range of *M. natalensis* across sub-Saharan Africa, LASV is only found in West Africa. Most cases of Lassa fever are in either Nigeria or the countries of the Mano River Union: Guinea, Liberia, and Sierra Leone. Epizootology studies undertaken throughout West Africa indicate high levels (upward of 30%) of LASV presence in *M. natalensis*. While Lassa fever cases can and do occur year-round, cases of Lassa fever peak during the West African dry season, spanning from November-April (Lecompte, Fichet-Calvet et al. 2006). Seasonality and fecundity of *M. natalensis* with respect to the temporal dynamics of LASV in rodents provide insight into human Lassa fever, as well as strategies for Lassa fever prevention.
1.2 Reservoirs of Lassa Virus

1.2.1 *Mastomys natalensis*

The 1972 outbreak of Lassa fever in the Eastern province of Sierra Leone saw cases stemming from individual communities rather than from nosocomial infections, further suggesting the possibility of a non-human reservoir (Fraser, Campbell et al. 1974). A thorough search of local vertebrate animals was conducted in two of the villages where the cases originated (Monath, Newhouse et al. 1974). Snap-frozen organs from collected animals, predominantly rodents and bats, were crushed, and the extracted supernatant was used to test for potential cytopathic effect (CPE). The only specimens testing positive for CPE were from *Mastomys natalensis*. This finding was later confirmed during a subsequent outbreak investigation of Nigerian rodents where *M. natalensis* was the predominant species where CPE was observed from organ extracts of collected animals (Wulff, Fabiyi et al. 1975). Multiple studies have concluded that *M. natalensis* is the predominant reservoir of LASV throughout its entire geographic distribution in West Africa (Lecompte, Fichet-Calvet et al. 2006, Safronetz, Sogoba et al. 2013, Kouadio, Nowak et al. 2015, Leski, Stockelman et al. 2015).

1.2.2 Other Rodents

Anti-LASV IgG antibodies have been detected in multiple other rodent species: *Praomys spp.*, *Mus spp.*, *Rattus rattus*, and *Lemniscomys striatus*, providing evidence of occasional spillover and transient infection (Demby, Inapogui et al. 2001, Fichet-Calvet, Becker-Ziaja et al. 2014, Olayemi, Cadar et al. 2016, Olayemi, Obadare et al. 2016). New evidence suggests three other species of rodents could be potential hosts to LASV. LASV has been isolated in multiple specimens of *Hylomyscus pamfi* in Nigeria and *Mastomys*
*erythroleucus* in both Nigeria and Guinea over a period of several years by qRT-PCR (Olayemi, Cadar et al. 2016, Agbonlahor, Erah et al. 2017). In villages where the virus is isolated from *M. erythroleucus*, virus could not be obtained from *M. natalensis*. LASV and LASV-like viruses have also been isolated in Pygmy mice (*Mus baoulei*) in Benin and Ghana (Kronmann, Nimo-Paintsil et al. 2013, Yadouleton, Agolinou et al. 2019). These newly isolated viral sequences from “alternative” reservoirs do not cluster with viruses in the traditional four/five lineages, revealing a potentially new lineage of LASV previously unseen in circulation. Whether these species have always been reservoirs to LASV or if viral mutations allow for host-switching and persistence in a new population is unknown.

It is plausible that *M. erythroleucus* is not a newly established host reservoir of LASV. Members of the *Mastomys* genus are typically indistinguishable by morphological characteristics and thus can only be confirmed to the species level by molecular genotyping. The distinct species of *Mastomys* were initially characterized by chromosomal karyotyping or hemoglobin electrophoresis, whereas the foundational reports of LASV in *Mastomys* categorized all *Mastomys* caught as *M. natalensis* (Monath 1975, Robbins, Krebs et al. 1983, McCormick, Webb et al. 1987). It will be important to appropriately identify all potential reservoirs to the species level to elucidate the potential distribution of LASV hosts and determine specific ecological niches specific to each potential host.

1.2.3 Additional Hosts

Macaques and other non-human primates (NHPs) simulate human infection upon inoculation with LASV; nearly all species of NHP used in laboratory testing manifest fatal pathology even with low titers of LASV (Malhotra, Yen et al. 2013). Interestingly, IgG antibodies to LASV have been detected in otherwise healthy Mona monkeys
(Cercopithecus mona) and a single Anubis baboon (Papio anubis) in two different Nigerian states (Ogunro, Olugasa et al. 2019). A single Mona monkey was also found to be both antibody and antigen positive for LASV, though none of the animals were positive for IgM. Virus was unable to be detected by RT-PCR in this animal. Nearly all the IgG positive NHPs were found scavenging refuse near the Lagos lagoon, though the Lagos state has no previously reported cases or seroprevalence of LASV. Exactly how NHPs encountered LASV or a LASV-like virus, as well as their role, if any, in the sylvatic transmission cycle remains to be determined. To date, no other vectors (ticks, mosquitoes, etc.) have been identified as a source or intermediate of LASV. However, a recent study demonstrated IgG presence in insectivorous shrews (Crocidura spp.) (Olayemi, Oyeyiola et al. 2018). Sequencing and sero-epidemiological studies focusing on non-rodent mammals will be needed to better elucidate the diversity of reservoir and spillover species.

1.3 Epizoonology Studies of Lassa Virus

1.3.1 Guinea

Much of the research about the spatial and temporal dynamics of LASV in rodents has been conducted in Guinea. Nationwide trapping studies surveying every ecozone were conducted between 1996–1997 and 2002–2005, with additional ad hoc trapping in endemic hot spots along the Guinea–Sierra Leone border and in the eastern half of the country. Antibodies to LASV are found in approximately 10% of M. natalensis nationwide, with up to 27% seroprevalence seen in villages in high prevalence regions (Demby, Inapogui et al. 2001, Fichet-Calvet, Becker-Ziaja et al. 2014). Despite high IgG antibody levels, only 5% of M. natalensis had measurable antigen levels (Demby, Inapogui et al. 2001). LASV has been detected via RT-PCR in 10–15% of M. natalensis, with higher focality in certain
endemic villages (Lecompte, Fichet-Calvet et al. 2006, Fichet-Calvet, Lecompte et al. 2007, Lalis, Leblois et al. 2012). Most specimens were not antibody and virus/antigen positive at the same time, suggesting immunological clearance and horizontal transmission of the virus. LASV in *M. erythroleucus* was detected in 6/14 specimens in coastal Guinea (Mandia Oula), and no *M. natalensis* were caught in the village where sampling occurred, indicating a new potential host of the virus (Olayemi, Cadar et al. 2016). Humans in this same village had a 35% seroprevalence rate of LASV despite little other evidence of LASV in the region (Lukashevich, Clegg et al. 1993). As the virus clusters phylogenetically with Lineage IV Josiah as seen in Sierra Leone and forest Guinea, it is likely that the virus was imported to the area and underwent host-switching to adapt to its new environment (Olayemi, Cadar et al. 2016).

1.3.2 Sierra Leone and Liberia

LASV was first isolated in cell culture from *M. natalensis* captured in the eastern province of Sierra Leone in 1972 during the first documented non-nosocomial outbreak of Lassa fever (Monath, Newhouse et al. 1974). This study showed that higher presence of *M. natalensis* corresponded with higher LASV seroprevalence in the human population. Intriguingly, rodent trapping in Zorzor, Liberia, during the 1972 outbreak did not yield any positive specimens (Monath 1975); subsequent rodent trapping studies have not been performed to date. Ensuing rodent studies performed in Sierra Leone have focused on the anthropological issues surrounding Lassa fever.

Two of the few epizooology studies in Sierra Leone took place in the 1980s. In the first investigation, case–control trapping found viremic *M. natalensis* were ten times as likely to be found in a patient’s house than in a control house (Keenlyside, McCormick et
A third of all *M. natalensis* collected in case houses had evidence of LASV viremia. However, all houses sampled had similar seroconversion rates to LASV, indicating potential movement of rodents from other non-trapped houses nearby to continue the chain of infection. The same study found higher antibody prevalence in individuals from households with a Lassa fever case than without (30 vs. 20%).

A second study examined LASV seroprevalence in villages across Sierra Leone (8–52% depending on the village and region) with rodent LASV viremia (0–81%; overall 11.2%), though no distinct correlation was seen with rodent viral presence and human seroprevalence (McCormick, Webb et al. 1987). LASV antibodies occurred in up to approximately 35% of all *M. natalensis*; concurrent viremia and antibody detection were found in about 5% of animals. No LASV was observed in *M. natalensis* located along the Atlantic coast of the southern province of Sierra Leone, despite serological evidence of LASV infection in humans, raising the possibility that there is an additional reservoir of LASV in this area. Within the northern and eastern provinces, two different species of *Mastomys* were seen. One species predominated in the northern plateau, an area that had lower seroprevalence and seroconversion rates than in the eastern province where the other species dominated. The two different *Mastomys* species were delineated by karyotyping, as genotyping at the time was not able to distinguish two separate species. With advances in genotyping and PCR techniques, a new undertaking of LASV distribution in rodents in the eastern province isolated virus from 20% of captured *M. natalensis* (Leski, Stockelman et al. 2015). Within the small study range, high sequence diversity within Lineage IV was seen, demonstrating local clustering of the virus within specific villages.
1.3.3 Nigeria

LASV was first successfully extracted from 8/151 rodents (5 *M. natalensis*, 2 *Mus musculus*, and 1 *Rattus rattus*) in the Benue-Plateau and North-Eastern States in 1975 (Wulff, Fabiyi et al. 1975). Since then, surveys have been orchestrated throughout Nigeria to measure LASV in the rodent population. In one survey conducted in Ekpoma, Edo State, 25% of specimens collected were *M. natalensis*, of which nearly half were reactive to LASV complement-fixing antibody (Okoror, Esumeh et al. 2005). An extraordinarily high presence of LASV in the rodent population would explain why Lassa fever is endemic in the Edo State. Other studies in the Edo State have shown lower prevalence in LASV in the rodent population. One study found only 16/274 *M. natalensis* positive for LASV, and another could not isolate viral sequences in any *M. natalensis* found (Olayemi, Cadar et al. 2016, Olayemi, Obadare et al. 2016). All LASV isolates in this region clustered to Lineage II.

A more extensive collection of 1500 specimens in Edo, Delta, and Bayelsa states found *M. natalensis* comprised only 2% of rodent species collected (Agbonlahor, Erah et al. 2017). All but two of the *M. natalensis* found in Edo and half of those found in Delta were LASV positive, with only one positive *M. natalensis* found in the non-endemic state of Bayelsa. The most recent and extensive trapping studies show a wide range of LASV or LASV-like virus seroprevalence in multiple rodent species within LASV endemic areas, ranging anywhere from 1 to 52% depending on the locality (Olayemi, Cadar et al. 2016, Olayemi, Oyeyiola et al. 2018). This evidence suggested that the rodent hosts of LASV may have expanded in Nigeria. It should also be noted that investigations that did not
examine other species for LASV in Nigeria may not have uncovered the pervasiveness of the virus in other rodent populations.

1.3.4 Other West African Countries

Though not traditionally considered endemic for Lassa fever, several West African countries in addition to the Mano River Union countries (Guinea, Sierra Leone, Liberia) and Nigeria have reported cases of Lassa fever over the past 10–15 years. Côte d’Ivoire has not experienced a confirmed case of Lassa fever during the prodromal phase, despite LASV circulating in the rodent population (Akoua-Koffi, Ter Meulen et al. 2006, Emmerich, Gunther et al. 2008). Extensive trapping surveys conducted between 2003 and 2005 found serological evidence of LASV or LASV-like virus in four different species of rodent including *M. natalensis* (Coulibaly-N’Golo, Allali et al. 2011). Two novel mammarenaviruses were also found in the southwest part of the country in *Hylomyscus spp.* and *Mus (Nannomys) setulosus*. The virus found in *Hylomyscus spp.* (named Gbagroube) closely resembled LASV Lineage II, while the other virus (Menekre) clustered more with East African mammarenaviruses Mopeia and Mobala. LASV RNA from *M. natalensis* was isolated and sequenced in 2013; those strains converged with strain AV, which was the first strain of LASV detected from a traveler in the area (Kouadio, Nowak et al. 2015, Manning, Forrester et al. 2015). Epizoology studies in Mali found an overall prevalence rate of LASV across all ecozones to be 6.8%. Prevalence rates of LASV in rodents in the southern region hover near 20%, with individual villages ranging from 0 to 52% (Safronetz, Sogoba et al. 2013). LASV in Mali clusters with lineage AV.

The first human serosurveys in Ghana indicated the overall LASV seroprevalence to be 5.2%, with individual villages as high as 16% (Nimo-Paintsil, Fichet-Calvet et al.
Rodents screened in the villages also showed no evidence of LASV via RT-PCR in *M. natalensis*, though two novel mammarenaviruses were discovered in two separate *Mus* species (Kronmann, Nimo-Paintsil et al. 2013). The first virus, found in the central part of the country near the border with Togo, clustered with Lineage I (Pinneo). The second, discovered in the northwest, grouped with LCMV. With clear serological evidence of LASV in the population, it is plausible that the virus either has a different reservoir in Ghana or that *M. natalensis* has circulating virus levels beneath the limit of detection with conventional PCR techniques. Despite the discovery of two novel mammarenaviruses near the borders of Togo and Burkina Faso, no rodent trapping studies designed to detect LASV have been performed in these countries to date. In Benin, which borders Nigeria to the east, three pygmy mice (*Mus baoulei*) captured in the central part of the country were PCR positive for LASV, clustering with strain AV (Yadouleton, Agolinou et al. 2019).

Senegal has had no confirmed cases of Lassa fever. A potential case of Lassa fever in eastern Senegal was reported in the early 1970s, but no evidence of LASV was found in human or rodent sera (Monath 1975). Rodent trapping performed in the 1980s revealed a very low overall LASV seroprevalence rate of about 1.2%, jumping up to 2.1% in *M. natalensis* (Saluzzo, Adam et al. 1988). Likewise, while the Central African Republic has never reported a case of Lassa fever, related Old-World mammarenaviruses have been isolated from *Arvicanthus spp.* and *Praomys spp.*, which have been demonstrated to cross-react with LASV and other related mammarenaviruses (Gonzalez, McCormick et al. 1983, Swanepoel, Leman et al. 1985). This cross-reactivity calls into question the specificity of the few seroprevalence studies undertaken in the rodent population.
1.4 Dynamics of Lassa Virus in *Mastomys natalensis*

1.4.1 Spatial and Temporal Dynamics

The ecology of Lassa fever is entwined with the environmental, geographic, and spatial factors affecting the distribution of *M. natalensis*. *Mastomys* biome preference has been inferred from studies conducted in Guinea (West Africa) and Tanzania (East Africa). *M. natalensis* density is highest in tropical savanna regions in both West and East Africa. West Africa also has a high density of *M. natalensis* in forested regions, while forested areas of East Africa have lower *M. natalensis* density (Demby, Inapogui et al. 2001, Lecompte, Fichet-Calvet et al. 2006, Makundi, Massawe et al. 2007, Borremans, Leirs et al. 2011). *Rattus* and *Mus* species tend to predominate over *M. natalensis* in coastal and urban areas (Keenlyside, McCormick et al. 1983, Demby, Inapogui et al. 2001). Higher human LASV seroprevalence is found in forested and savannah regions, mirroring general *M. natalensis* distribution in West Africa (Lukashevich, Clegg et al. 1993). An explanation for the dearth of *M. natalensis* along the coast could be competition by imported rodents. It is plausible that European-lineage rodents such as *Rattus* and *Mus* may have been brought over on ships during slave transit and trade expeditions where they settled in the nearby shores and outcompeted other African-lineage rodents in environments that matched their ideal niche. Preferences in building construction in relation to rodent burrowing preferences (e.g., concrete vs. mud floors; metal vs. thatched roofs) could also explain the absence of *M. natalensis* in urban areas (Bonwitt, Saez et al. 2017). These differences in building finishes and construction in rural vs. urban housing allow for increased rodent burrowing and hiding in rural communities thereby contributing to
infestations with potentially infected animals in rural areas (Bonner, Schmidt et al. 2007, Bonwitt, Saez et al. 2017).

At the village level, *M. natalensis* is preferentially found inside homes as opposed to open fields (Demartini, Green et al. 1975, Fichet-Calvet, Lecompte et al. 2007, Marien, Kourouma et al. 2018), especially during the dry season. *M. natalensis* population levels decrease from the outskirts of a village to its core, potentially due to lack of nearby vegetation in adjacent gardens or forest (Demby, Inapogui et al. 2001, Fichet-Calvet, Lecompte et al. 2007). During the rainy season, food and resources are abundant outside of homes in proximal fields. The burning and clearing of fields during the dry season can lead to aerosolization of excrement and driving mice to enter homes in search of food, all activities that increase the likelihood of individuals encountering an infected rodent. The prevalence of LASV in *M. natalensis* remains constant year-round in homes, which seemingly contradicts the cyclical nature of Lassa fever (Fichet-Calvet, Lecompte et al. 2007). The authors of this study suggested the possibility of non-seasonal breeding inside homes in which the mice breed year-round (as opposed to a peak reproductive period at the beginning of the rainy season), contributing to consistent viral levels within the rodents. When other rodents such as *Mus spp.* and *Rattus spp.* out-compete *M. natalensis* in homes, reduced human seroprevalence and rodent viral presence of LASV are seen (Monath, Newhouse et al. 1974, McCormick, Webb et al. 1987, Demby, Inapogui et al. 2001, Lecompte, Fichet-Calvet et al. 2006). Furthermore, infectious rodents tend to be clustered in a select few houses as opposed to evenly distributed across homes in a village (Demby, Inapogui et al. 2001, Marien, Lo Iacono et al. 2020). The in-home LASV prevalence rate should be further investigated to better understand the temporal dynamics of Lassa fever.
Lassa fever appears to fluctuate seasonally. Most Lassa fever cases occur during the West African dry season spanning roughly from November to April, peaking in January (Bausch, Demby et al. 2001, Fichet-Calvet, Lecompte et al. 2008). Statistical modeling has shown regions in West Africa most at risk of Lassa fever epidemics receive heavy amounts of rain, up to 3000 mm during the rainy season spanning from May–October. This correlates with observational studies of *M. natalensis* reproduction and fecundity in Guinea showing robust turnover in the rodent population from the start of the rainy season to its end, possibly an effect of stabilized food and water resources (Fichet-Calvet, Lecompte et al. 2008). Similar results have been observed in Tanzania, where *M. natalensis* population levels peak during the biannual rainy seasons (Borremans, Leirs et al. 2011, Mulungu, Ngowo et al. 2013). There is conflicting evidence of overall LASV presence in *M. natalensis* in the rainy season or during the dry season (Demby, Inapogui et al. 2001, Fichet-Calvet, Lecompte et al. 2007). Other mammarenaviruses, such as Lymphocytic Choriomeningitis Virus (LCMV) and Junín virus, have increased viral presence during the breeding season (Mills, Ellis et al. 1992, Tagliapietra, Rosa et al. 2009). The increase in virus during the breeding season is highly indicative of vertical transmission, as will be discussed below.

1.4.2 Transmission Modalities of Lassa Virus in the Host Reservoir

The extent of horizontal and vertical transmission of LASV and other closely related Old World (OW) mammarenaviruses in the host reservoir has been hotly debated. Both horizontal and vertical transmissions have been demonstrated in the laboratory. LASV-inoculated *Mus musculus* neonates develop low-level persistent infections, while inoculated adults develop transient infections. In this study, both neonates and adults
developed antibodies to the viruses shortly upon inoculation, which remained for the duration of the study (Walker, Wulff et al. 1975, Borremans, Vossen et al. 2015). Colonies of mice (M. musculus) persistently infected with LCMV via vertical transmission have been established (Kunz, Rojek et al. 2006, Sun, Vasek et al. 2014) where they develop lifelong viremia yet have low antibody titers against the virus (Oldstone 2002). While the LCMV model may parallel LASV infection in M. natalensis, this feature has not been experimentally validated. LASV has been found in placental and fetal tissues as well as breast milk of humans, further suggesting that vertical transmission may occur as part of the transmission cycle outside of the reservoir (Theiler, Rasmussen et al. 2008). Observations of wild-caught animals have noted that rodents are infrequently virus and antibody positive at the same time, pointing to vertical transmission (Demby, Inapogui et al. 2001, Borremans, Leirs et al. 2011, Marien, Borremans et al. 2017). In transiently infected animals, the virus is able to persist in secretions for months at a time (Walker, Wulff et al. 1975). Aerosols generated from mucous membranes are stable for short periods of time and may be a key component of intra-rodent horizontal transmission when rodents engage in fighting, nosing, grooming, and other social activities (Stephenson, Larson et al. 1984, Vanden Broecke, Borremans et al. 2018). From an evolutionary context, vertical transmission provides the most logical explanation of viral persistence over generations. Although laboratory data provides a solid foundation to build hypotheses upon, discrepancies are seen between laboratory and natural conditions and must be taken into consideration when drawing conclusions.

The longitudinal transmission dynamics of OW mammarenaviruses in M. natalensis remains understudied. During an evaluation of rodents from LASV endemic
Guinean villages, it was determined that older mice have higher LASV seroprevalence than young and juvenile (adolescent) mice, indicative of horizontal transmission (Demby, Inapogui et al. 2001, Fichet-Calvet, Becker-Ziaja et al. 2014). It was also revealed that LASV seroprevalence was lowest in juvenile mice and higher in young (infant) mice, indicating a potential transfer of protective antibodies from mother to fetus in utero or through breast milk. At the same time, LASV has been detected in young mice in other studies, which could be an indicator of vertical transmission (Demby, Inapogui et al. 2001, Fichet-Calvet, Lecompte et al. 2008). A study of related Mopeia virus revealed peak viral RNA presence in young rodents, indicative of either vertical transmission or horizontal transmission preferentially occurring in younger rodents (Borremans, Leirs et al. 2011). This virus may not necessarily be persistent for the life of the animal in nature. Taken together, it is probable that an equal combination of vertical and horizontal transmission enables continual persistence in the LASV host populations (Figure 1).

Assuming undetectable but still extant virus in the maternal population, LASV presence peaks and a healthy amount of antiviral antibody is observed during the rainy season where there is high turnover within the *M. natalensis* population. During this time, the virus is spread via vertical transmission to newborn mice, or via horizontal transmission to mice that have not received in utero antibodies. As transient maternal antibodies protecting young mice from infection wane as they mature during the dry season, a recrudescence of horizontally transmitted virus occurs from other animals in the population, or perhaps “reactivating” the virus within the now juvenile rodents. Human activity and habitat changes bring the mice in closer contact with humans as indicated by the spike of human cases of Lassa fever. As the dry season draws to a close and the rainy
season begins, nearly all mice will have been exposed to the virus and viral levels drop once again until new litters of mice are born and the cycle repeats. It is important to note that *M. natalensis* is capable of breeding year-round which could account for cases seen during the rainy season.
Rodents spread the virus amongst themselves through both horizontal and vertical transmission. Most human cases of Lassa fever derive from contact with rodent excrement. The remaining cases arise from direct or indirect contact with infected individuals. It is unknown whether other species come into contact with LASV via infected rodents or whether humans have the ability to transfer the virus back to rodents or other species.
1.4.3 Stopping the Cycle of Transmission

Certain behaviors may promote spread of the virus from rodents to humans. Consumption of rodent bushmeat, at times the sole source of protein in rural areas, is associated with higher prevalence of antibodies to Lassa in humans compared to communities where rodents are not eaten (Ter Meulen, Lukashevich et al. 1996). Indeed, a recent case of Lassa fever in Benin can be traced back to regular consumption of rodent bushmeat by the patient (Attinsounon, Ossibi Ibara et al. 2018). Activities such as agriculture, logging, and mining can contribute to forest clearing and thus increase rodent–human interactions in a more “ideal” environment for rodents, promoting conditions ripe for spillover. Much of the economies of the high Lassa fever countries of Guinea, Sierra Leone, and Liberia are dependent on mineral mining and subsistence agriculture (United States. Central Intelligence Agency. 1995). In these areas, education and proper rodent control is vital to prevent the further spread of disease.

There are multiple methods of rodent control used throughout West Africa. Both trapping rodents and dispersing rodent poison are effective yet expensive, making it difficult to implement long term without coordinated efforts from communities, local health authorities, and government backing (Mari Saez, Cherif Haidara et al. 2018). The most sustainable rodent control measures are preventing rodents entering homes and further rodent proofing all food sources. Cats and dogs are being explored as rodent control agents, though they may bring their own public health issues (Haun, Kamara et al. 2019). In the absence of a vaccine, rodent exclusion measures are necessary to protect those at risk (Sogoba (Khan, Goba et al. 2008, Sogoba, Feldmann et al. 2012, Olayiwola and Bakarey 2017, Roberts 2018).
1.5 Evolutionary History of Old-World Mammarenaviruses in Rodents

1.5.1 The Distinct Niche of Lassa Virus in West Africa

Members of the *Mastomys* genus must be genotyped to delineate the exact species present. Upon the widespread adaptation of PCR in laboratories, “DNA barcoding” using mitochondrial DNA (mtDNA) microsatellite genes has been used to aid phylogenetic inference of different rodent species and subspecies (Lecompte, Granjon et al. 2002, Lecompte, Aplin et al. 2008). Mitochondrial DNA is particularly useful for evolutionary and phylogeny studies as it mutates faster than the rest of the nuclear genome, enabling minute differences over relatively short periods of time to be quantified and mapped (Parson (Brown, George et al. 1979, Parson, Pegoraro et al. 2000). Phylogenetic analysis using the mtDNA gene cytochrome *b* has distinguished eight different *Mastomys* species, of which only *M. natalensis* and *M. erythroleucus* have been associated with OW mammarenaviruses (Olayemi, Cadar et al. 2016).

In-depth microsatellite analysis of *M. natalensis* mtDNA has also been employed to further characterize differences between members of this species and its subspecies throughout Africa (Colangelo, Verheyen et al. 2013, Gryseels, Baird et al. 2017). This insight has shown that *M. natalensis* has different groups of a specimen characterized by matrilineal lineages, or clades, of its mitochondrial DNA, and sub-clades that vary slightly throughout the continent. The clades and sub-clades are generally constrained to a specific geographic area separated by natural boundaries such as lakes, mountains, and rivers. Clade A is predominantly found in West Africa entirely west of the Congo River and a small sub-clade found around Lake Victoria in Kenya. Clade B is restricted to Eastern Africa, bound by Great Rift Valley and Eastern Arc Mountains in Tanzania to the west and the Kalahari
Basin and Orange River in the south. Since the discovery of LASV, several OW mammarenaviruses have been found in different *M. natalensis* sub-clades across Africa. Only sub-clade A1 in West Africa has been associated with LASV; other closely related OW mammarenaviruses seem to be associated with its own specific sub-clade (Olayemi, Obadare et al. 2016, Gryseels, Baird et al. 2017). The sub-clades of *M. natalensis* may explain the spatial spread of highly related mammarenaviruses throughout Africa.

Even within the LASV-specific sub-clade of *M. natalensis*, certain populations of rodents are more likely to carry the virus than others. There is no evidence of *M. natalensis* migrating long distances along trade routes and major thoroughfares, nor even short distances outside of their commensal human habitats (Lalis and Wirth 2018). To establish a population, these rodents would have had to come along in food and storage containers during periods of human migration and re-settle in new areas. As such, LASV-sustaining populations of *M. natalensis* are seen within small geographic and inquiline niches, as evidenced by the low rates of gene flow and high co-sanguinity among LASV-positive animals (Lalis, Leblois et al. 2012, Lalis, Evin et al. 2015). The dependency on human transport to spread the rodents (and thus the virus) may explain why there is inter-village variability within a localized geographic area.

1.5.2 Unraveling the Co-speciation Hypothesis

In addition to the mammarenaviruses found exclusively in *M. natalensis*, other mammarenaviruses in closely related rodents have also been discovered (reviewed in (Gryseels, Rieger et al. 2015); see also (Figure 2). All rodents are either members of the same tribe *Praomyini* or a species of *Mus* (Lecompte, Aplin et al. 2008).
While LASV is exclusively found in West Africa (colored circles), the approximate range of *M. natalensis* (blue dashed line) and *M. erythroleucus* (orange dashed line) is shown to demonstrate the potential for expanded geographic distribution of LASV. Recently, it has been proposed that some of the novel mammarenaviruses in West Africa may be new lineages of LASV; this would greatly expand the number of potential reservoirs of LASV. Mammarenaviruses found in East Africa form a distinct clade within OW mammarenaviruses.
This widespread distribution leads to the original theory of co-speciation, where an individual rodent species becomes the host to a specific mammarenavirus over millions of years of evolution (Gonzalez, Georges et al. 1986, Bowen, Peters et al. 1997). Overlaying molecular clock analyses of OW mammarenaviruses with African murine radiation and the respective phylogenetic trees show considerable overlap (Bowen, Peters et al. 1997). It was believed that a common LCMV-like arenavirus progenitor may have infected a common rodent ancestor prior to its split into different muroid (rodent) subfamilies. From there, muroids began to spread from Asia to Europe, Africa, and the Americas and subsequently evolved (Lecompte, Aplin et al. 2008). According to this hypothesis, OW mammarenaviruses arose from the Murinae subfamily and New-World (NW) mammarenaviruses arose from the Sigmodontinae subfamily around 30 million years ago (Arata and Gratz 1975, Gonzalez 1999). Over time, mammarenaviruses evolved with specific rodent species: LCMV with Mus spp., NW mammarenaviruses with Calomys and Neotoma spp., and OW mammarenaviruses with the Stenocephalomyys complex (specifically Mastomys, Praomys, and Hylomyscus spp.) (Lecompte, Granjon et al. 2002, Schenk, Rowe et al. 2013). Fossil evidence indicates that the Mastomys genus arose during the Late Pliocene era approximately 2.5–3.5 million years ago in migration from Eurasia to Africa. This evolutionary event allowed for further co-speciation between specific rodent species and mammarenaviruses to develop throughout Africa (Lecompte, Aplin et al. 2008). Recent reports of additional OW mammarenaviruses in brown rats (Rattus norvegicus) in China, Cambodia, and Thailand fit the model of co-speciation from a common viral ancestor, yet also showed an expanded range of potential Murinae hosts beyond the Praomyini tribe (Li, Lin et al. 2015, Blasdell, Duong et al. 2016).
A thorough analysis of the ribosomal-dependent RNA polymerase of mammarenaviruses reveals the division between OW and NW mammarenaviruses occurred as recently as 45,000 years ago (ya), with OW mammarenaviruses originating between 23,000 and 1800 ya (Forni, Pontremoli et al. 2018). Such a short evolutionary history cannot not be explained by co-divergence. In this vein, deep sequencing of clinical isolates of LASV from Sierra Leone and Nigeria indicates the evolutionary origin of LASV in Nigeria 1,000–2,000 years ago before spreading across West Africa indicating further disconnect between speciation and viral evolution (Andersen, Shapiro et al. 2015). It is thought that colonial rule, common travel, and slave trade may have allowed LASV to spread westward from Nigeria to other regions of West Africa around 150–200 years ago (Lalis, Leblois et al. 2012, Manning, Forrester et al. 2015).

A persistent flaw of the co-speciation model has been revealed by the presence of NW Tacaribe virus in phyllostomid bats, mosquitos, and ticks in the Caribbean, demonstrating a wide taxonomic difference between potential mammarenavirus hosts (Downs, Anderson et al. 1963, Sayler, Barbet et al. 2014). The recent discovery of arenaviruses in reptiles and fish further proves the broad diversity of Arenavirus hosts. Indeed, this has since led to reclassification of Arenaviridae from a single Arenavirus genus into four different geneses: Mammarenaviruses encompassing OW and NW arenaviruses primarily found in rodents, Reptarenavirus and Hartmanivirus in reptiles, and Antennavirus in fish (Maes, Alkhovsky et al. 2018, Shi, Lin et al. 2018). Subsequently, the Arenaviridae family has since merged with members of the now-defunct Bunyaviridae family to form a new order Bunyavirales (Maes, Alkhovsky et al. 2018). This taxonomic change better reflects the common genome and structural similarities between family
members of this order. Genetic reassortment and recombination between these related viruses may have allowed for sylvatic transmission between species, allowing a novel host niche to evolve, emerge, and further diverge over short periods of time.

1.5.3 LASV Sequencing and Phylogenetics

Prior to the early 2000’s, only three LASV strains isolated from some of the first known cases of Lassa fever had been sequenced in full (Bowen, Rollin et al. 2000). Early insights into the variability of LASV strains circulating in West Africa were first characterized by Sanger (first-generation) sequencing of partial genome fragments from clinical Lassa fever samples (Bowen, Rollin et al. 2000). These initial characterizations indicated four different lineages of LASV: Lineages I-III in Nigeria, and Lineage IV in the Mano River Union countries of Guinea, Liberia, and Sierra Leone.

Commercialization and mainstream adaptation of next-generation sequencing (NGS) has allowed in-depth investigations of LASV genomes previously too costly and time-prohibitive to perform; over 1000 full-length or nearly full-length LASV sequences have been produced in the past 20 years (Klitting, Mehta et al. 2020). With these advances in sequencing technology and bioinformatic analyses, at least two new LASV lineages have recently been uncovered. Lineages I-IV (particularly II-IV) remain the best characterized, while Lineages V and VI have been recently discovered in other west African countries (Mali/Côte d’Ivoire and Togo/Benin, respectively) (Manning, Forrester et al. 2015, Whitmer, Strecker et al. 2018, Salu, James et al. 2019). Further epidemiological investigations of LASV throughout West Africa combined with clinical sequencing of diagnostic specimens continue to deepen our knowledge of LASV diversity and its unique evolutionary history.
Three of the largest and most detailed LASV genomic epidemiology studies of
LASV (Andersen, Shapiro et al. 2015, Siddle, Eromon et al. 2018, Kafetzopoulou, Pullan
et al. 2019) each employed a unique sequencing approach demonstrating the variety of
approaches to LASV sequencing. Andersen et al. and Siddle et al. both used the Illumina
sequencing platform to generate LASV sequences. The Illumina sequencing platform reads
short fragments of a sequencing library, which is constructed from the samples of interest,
in a massively parallel fashion. To prepare the sequencing library, cDNA is generated from
randomly primed, reverse transcribed RNA. The cDNA is then further processed to include
fragment-specific DNA barcodes for individual downstream identification, and ligates
adaptor sequences to the fragments that allow the fragments to hybridize to complementary
oligonucleotides (oligos) inside the sequencing flow cell. Once the completed sequence
library is placed into the flow cell and fragments bind to their complementary oligos, clonal
bridge amplification produces clusters containing many amplified copies of the fragments.
After cluster generation, the clusters are then simultaneously sequenced and read
(“sequencing by synthesis”) by incorporating fluorescently labeled deoxynucleotides
dNTPs into the fragment being sequenced. Fluorescent lasers excite the cluster to detect
the specific dNTP most recently added. Sequences are constructed by grouping
overlapping fragments together, called contigs, then aligning the contigs to a reference
genome.

While Siddle et al. used an metagenomic approach to allow for unbiased detection
of any pathogen present in a sample (Matranga, Gladden-Young et al. 2016), Andersen et
al. performed targeted hybrid capture amplification of unbiased library fragments to enrich
for LASV sequences prior to sequencing on a flow cell (Matranga, Andersen et al. 2014).
Hybrid capture amplification uses custom-designed oligonucleotide probes to specifically bait the sequences of interest, in this case LASV-specific fragments, and selectively amplify the captured sequences prior to addition to the flow cell. Performing hybrid capture during library construction allows for both viral sample enrichment and detection of low frequency variants present within a single virus sample that may not be possible with unbiased metagenomic approaches.

In contrast to the Illumina platform technology employed in Siddle et al. and Andersen et al., Kafetzopoulou et al. used the Oxford Nanopore minION platform to perform metagenomic sequencing of clinical LASV specimens in real time during a recent Lassa fever outbreak in Nigeria. The minION platform and other third-generation sequencing technologies perform real-time sequencing of metagenomic libraries on protein nanopores in a sequencing flow cell. Each nucleotide added to the growing sequence produces a specific change in the ion gradient flowing through the nanopore, which can then be detected by changes in current applied through the flow cell. The platform produces single molecule, gigabase long reads of either RNA or DNA from the sequencing libraries. Furthermore, the minION platform is designed to be highly portable and able to be deployed in low-resource settings for near instantaneous detection of viruses and other pathogens (Greninger, Naccache et al. 2015). Despite the relative ease of use and portability, third-generation sequencing technologies run the risk of higher error rates due to the nature of long reads of single molecules. With potential complications in mind, newer amplification and computational techniques have been developed to compensate for possible sequencing error (Cressiot, Greive et al. 2018, Rang, Kloosterman et al. 2018).
Once viral samples are sequenced and assembled, sequences can be explored in numerous ways. Beyond deducing changes to protein functionality from mutations found in the sequenced sample as compared to a reference sample, a phylogenetic tree can be assembled to trace the sample’s evolutionary history. The evolutionary history of a virus sample is inferred by determining the time to most common recent ancestor (tMRCA), the real or theoretical sequence from which other viral sequences evolved. tMRCA can be deduced from molecular clock substitution models, which calculate rates of nucleotide substitution per nucleotide per year as assessed in exploratory laboratory studies, and/or Bayesian inferences, in which viral samples with known collection dates are used to inform and supplement molecular clock substitution models of actual viral evolution (Zhou and Teo 2016, Suchard, Lemey et al. 2018). Transmission chains and transmission dynamics can then be examined from detailed phylogenetic trees, which can hint at reoccurring spillover or human-to-human transmission chains. Individual transmission events can be explored in even greater detail by examining mutations occurring within a single sample, called intra-host single nucleotide variations (iSNVs). Accumulation of iSNVs within a host may help to elucidate transmission bottlenecks or explain why certain mutations become fixed in a population (Andersen, Shapiro et al. 2015, Grubaugh, Gangavarapu et al. 2019). Furthermore, phylogeographic analysis allows for observation of viral evolution with respect to geography. Indeed, LASV lineages can be further broken down to multiple sub-lineages roughly bound by geographic parameters within the area of parent lineage circulation (Ehichioya, Dellicour et al. 2019, Klitting, Kafetzopoulou et al. 2021). Within the context of this dissertation, our team uses phylogeographic analysis to observe the
demarcation of LASV evolution and diversity of both humans and rodents within geographic confines.

1.6 LASV Diagnostics

Timely laboratory diagnosis of LASV is an essential component of Lassa fever case management. Although viral culture of a patient blood sample provides the most accurate assessment of LASV presence, it takes several days to perform and can only be performed under BSL-4 conditions which prohibits its use in LASV-endemic areas. To that end, the most commonly used LASV diagnostic assays in LASV-endemic areas are nucleic acid amplification tests (NAAT), antigen detection tests, and antibody detection tests (Mazzola and Kelly-Cirino 2019). As each method has its own advantages and disadvantages, multiple diagnostic methods are often performed to distinguish an active LASV infection from a previous exposure.

1.6.1 Nucleic Acid Amplification Tests

NAAT such as polymerase chain reaction (PCR) and loop-mediated isothermal amplification (LAMP) enable detection of LASV at low levels but are constrained by the high sequence diversity found among and within LASV lineages. PCR assays often target highly conserved sequences within the S segment (Demby, Chamberlain et al. 1994, Olschlager, Lelke et al. 2010, Trombley, Wachter et al. 2010). Degenerate primers and probes have been developed to allow for detection of the more variable L segment with comparable or better specificity to S segment assays (Pang, Li et al. 2014, Nikisins, Rieger et al. 2015); such degenerate primers are used in commercially available LASV qRT-PCR diagnostic kits like the Altona Diagnostics RealStar© Lassa Virus RT-PCR 2.0 Kit. While PCR assays deploying probe-based detection are often more specific than assays deploying
SYBR Green based detection, the sensitivity of the assay may be increased using sequence-independent intercalculating fluorescent dyes such as SYBR Green as opposed to probe-based detection (Boisen, Hartnett et al. 2018). Although LAMP assays have been developed as potential LASV diagnostics, they have not been deployed in clinical settings due to the wide strain variability present in LASV which require lineage-specific primers. (Fukuma, Kurosaki et al. 2011, Pemba, Kurosaki et al. 2019). Recently, CRISPR-Cas13 technology has been used to detect LASV in clinical samples at levels comparable to qRT-PCR (Barnes, Lachenauer et al. 2020). The technology, dubbed SHERLOCK (Specific High-sensitivity Enzymatic Reporter unLOCKing) first isothermally amplifies potential viral targets, then uses a Cas13 protein paired with virus-specific guide RNA designed to detect the amplified viral targets within the sample (Myhrvold, Freije et al. 2018). Like LAMP, however, lineage-specific primer sets and guide RNAs for use with SHERLOCK must be used due to the wide strain variability of LASV throughout its geographic distribution.

1.6.2 Antigen Tests

Although PCR is an effective method to detect low levels of LASV, it requires specialized equipment and highly trained personnel to perform the assay. In contrast, antigen diagnostic tests are considerably less complex and take less time to perform than PCR testing. Antigen diagnostics are particularly robust in diagnosing active LASV infection, as protein sequences typically mutate less than genetic sequences would mutate. NP antigen can be detected either by an enzyme-linked immunosorbent assay (ELISA) or rapid diagnostic test (RDT) (Bausch, Rollin et al. 2000, Saijo, Georges-Courbot et al. 2007, Boisen, Hartnett et al. 2018). Unlike ELISA which must be performed in a laboratory
setting, RDTs can be performed in low resource settings by community health workers and give results in under 30 minutes with similar results to ELISA (Boisen, Hartnett et al. 2018, Boisen, Uyigue et al. 2020). Antigen diagnostic results are often coupled with an IgM ELISA to determine active infection.

For this project, our research team chose to use an RDT in-field to assess presence of LASV antigen in rodents. The Zalgen Pan-Lassa Antigen RDT is a lateral flow dipstick immunoassay designed to detect the presence of LASV NP antigen in serum, plasma, or whole blood, using as little as one drop of blood (~30 µL). The dipstick consists of a sample test pad containing proprietary recombinant rabbit anti-LASV NP polyclonal antibodies conjugated to gold nanoparticles, which is attached to a nitrocellulose membrane containing one stripe of the LASV antibodies and one stripe of anti-rabbit IgG antibodies. The sample is placed onto the sample test pad of the dipstick, then placed in a small test tube containing sample buffer to initiate lateral flow across the nitrocellulose membrane. If NP antigen is present in the sample, it will form an antigen-antibody conjugate complex to the gold-antibody complex present in the sample test pad. As lateral flow moves the sample up the membrane, the first (test) stripe of LASV antibodies will capture the complex and change color based on the amount of sample present. The excess gold-antibody complex will be captured by the second anti-rabbit IgG control stripe to indicate a valid result. The results can be read after 15-25 minutes. The use of an anti-rabbit IgG as a secondary antibody allows the RDT to detect antigen in all Muridae species.

1.6.3 Antibody Tests

Enzyme-linked immunosorbent assay (ELISA) can be used to detect IgG and IgM to LASV. Indirect immunofluorescence assays (IFA) have also been used to detect LASV
antibodies, however they are impractical for clinical diagnoses as they require BSL-4 facilities to perform. IFA for IgG surveillance is still used, namely for animal surveillance studies (Fichet-Calvet, Becker-Ziaja et al. 2014, Mari Saez, Cherif Haidara et al. 2018). IgM detection by ELISA or IFA as a stand-alone test for LASV presence is weak, with sensitivity hovering around 30% (Gabriel, Adomeh et al. 2018). Therefore, IgM testing is often performed alongside antigen and PCR testing to confirm presence of LASV. IgG testing is mostly restricted to LASV surveillance, though it can be used within a clinical context to monitor the course of a LASV infection with a Lassa fever patient (Shaffer, Schieffelin et al. 2019). A limitation of both ELISA and IFA is that the secondary antibody used to detect bound antibody to plated antigen or virus must be specific to the species being examined thereby limiting the range of species that can be tested by these formats. A species-neutral method for detection of LASV antibodies is the double-antigen binding assay (DABA), a variation of an indirect ELISA. DABA substitutes a species-specific secondary detection antibody with a horseradish-peroxidase linked antigen against the antibody captured by the plated antigen to allow antibodies from any species to be detected (Akpogheneta, Dicks et al. 2021). As accurate assessment of IgG antibody is vital to LASV surveillance, species-neutral assays like DABA may allow for broader LASV surveillance beyond humans and rodents.
2. RESEARCH OBJECTIVE

While LASV and Lassa fever occurs throughout West Africa, the Eastern province of Sierra Leone has some of the highest incidences of Lassa fever in the world (Shaffer, Grant et al. 2014). However, in recent years the number of patients presenting to Kenema Government Hospital (KGH) with suspicion of Lassa fever has decreased dramatically (Shaffer, Schieffelin et al. 2021). Likewise, published serological studies of rodents in Sierra Leone have not been performed since the 1980s (Keenlyside, McCormick et al. 1983, McCormick, Webb et al. 1987). To that end, our group performed a cross-sectional serological study of small mammals in the Kenema district of Sierra Leone to i) uncover the true incidence rate of LASV in rodents in the Kenema district ii) establish a diagnostic algorithm driven by field-deployable diagnostics to screen for small mammals, namely rodents, with active LASV infections and iii) sequence LASV from rodents with active LASV infections to further visualize LASV diversity in Sierra Leone. We hypothesize that evidence of LASV infection, indicated by serum antigen and/or IgG antibody presence, will be present in 20-30% of the sampled rodent specimens in the Kenema district, in line with previous epizooLOGY studies in Sierra Leone. Furthermore, we predict that roughly half of the specimens with LASV antigen presence will be seen to have active LASV infections. Finally, we expect to see high levels of LASV diversity within Sierra Leone based on geographic zones.
3. ASSESSMENT OF LASSA VIRUS PREVALENCE IN THE NATURAL HOST RESEVOR WITHIN IN THE KENEMA DISTRICT OF SIERRA LEONE

3.1 Materials and Methods

3.1.1 Ethics and biosafety Statement

The Tulane University Institutional Animal Care and Use Committee reviewed and approved this study (Protocol ID 1034). Verbal consent to rodent trapping was obtained from both the village chief and the occupants of each house where trapping was performed prior to placing traps. Trapping and processing of collected specimens at field BSL-3 conditions were followed as described (James N. Mills, James E. Childs et al. 1995). Tissue samples shipped to the United States were chemically inactivated to allow for manipulation in a standard BSL-2 environment. Serum samples and tissue samples processed in the laboratory at KGH were handled under BSL-2+ conditions.

3.1.2 Study design

A cross-sectional rodent trapping survey was conducted in Kenema District, Eastern Province, Sierra Leone between November 2018-July 2019. Each of the villages (Table 1; Figure 3) in the study was trapped for two consecutive nights, one time. The villages selected for trapping have high human seroprevalence of LASV or have had patients admitted to the Lassa Fever ward at KGH to maximize the likelihood of observing an active infection in rodents. Geospatial visualization was performed with ArcGIS Online (ESRI, CA, USA).
Table 1: Names and GPS coordinates, in decimal degrees, of villages trapped in study.

<table>
<thead>
<tr>
<th>Village</th>
<th>Latitude ('N')</th>
<th>Longitude ('W')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bambawo</td>
<td>8.008983</td>
<td>-11.13190</td>
</tr>
<tr>
<td>Bowohun</td>
<td>8.105333</td>
<td>-11.21270</td>
</tr>
<tr>
<td>Dandabu *</td>
<td>7.538883</td>
<td>-11.18082</td>
</tr>
<tr>
<td>Goura</td>
<td>8.174800</td>
<td>-11.08528</td>
</tr>
<tr>
<td>Guabu **</td>
<td>8.175033</td>
<td>-11.20037</td>
</tr>
<tr>
<td>Jamboma</td>
<td>8.121583</td>
<td>-11.21472</td>
</tr>
<tr>
<td>Joru</td>
<td>7.692162</td>
<td>-11.05662</td>
</tr>
<tr>
<td>Kamboma</td>
<td>8.356900</td>
<td>-11.43385</td>
</tr>
<tr>
<td>Koi</td>
<td>8.040883</td>
<td>-11.05437</td>
</tr>
<tr>
<td>Kormolu</td>
<td>7.995233</td>
<td>-11.08348</td>
</tr>
<tr>
<td>Koyama Ngeima</td>
<td>8.234950</td>
<td>-11.18197</td>
</tr>
<tr>
<td>Kpalu</td>
<td>7.907617</td>
<td>-11.08485</td>
</tr>
<tr>
<td>Kptema</td>
<td>8.273750</td>
<td>-11.07435</td>
</tr>
<tr>
<td>Largo</td>
<td>8.042244</td>
<td>-11.10564</td>
</tr>
<tr>
<td>Lawana **</td>
<td>8.126683</td>
<td>-11.18898</td>
</tr>
<tr>
<td>Maleh</td>
<td>7.898251</td>
<td>-11.04634</td>
</tr>
<tr>
<td>Mano-Ngeiya</td>
<td>8.087167</td>
<td>-11.09905</td>
</tr>
<tr>
<td>Ngeihun</td>
<td>8.174800</td>
<td>-11.08528</td>
</tr>
<tr>
<td>Njagor</td>
<td>8.244633</td>
<td>-11.13432</td>
</tr>
<tr>
<td>Nyahahun</td>
<td>7.907617</td>
<td>-11.08485</td>
</tr>
<tr>
<td>Pandembu</td>
<td>8.214267</td>
<td>-11.06597</td>
</tr>
<tr>
<td>Panguma (vaama)</td>
<td>8.186067</td>
<td>-11.12788</td>
</tr>
<tr>
<td>Pujehun</td>
<td>8.151694</td>
<td>-11.09213</td>
</tr>
<tr>
<td>Saahun</td>
<td>7.769467</td>
<td>-11.11262</td>
</tr>
<tr>
<td>Saleima</td>
<td>8.121933</td>
<td>-11.19347</td>
</tr>
<tr>
<td>Tongola</td>
<td>8.211017</td>
<td>-11.05090</td>
</tr>
</tbody>
</table>

* Village Excluded From Study
** No Animals Captured
Figure 3: GPS location of villages trapped in study within Sierra Leone.
3.1.3 Small mammal trapping

Trapping and specimen collection was performed by the KGH Ecology Team. Two Sherman traps (Sherman Live Trap Co., FL, USA) baited with a mixture of oats, fish, and peanut butter were set in opposing corners of each room of the home or dwelling enrolled in the study. Traps were set in the early evening and checked the next morning at dawn for presence of animals. The traps were removed from the rooms during the day, then set again in the evening.

3.1.4 Necropsy

Specimen processing took place away from inhabited dwellings. Off-target captures (e.g. Frogs, Toads) and any endangered species were immediately released. Captured target animals were sedated via isoflurane inhalation, and morphometric data (sex, weight, genus, female reproductive status, total length, tail length, right ear and right foot length) were collected. The animals were then euthanized via cardiac puncture under isoflurane sedation. Whole blood collected during the euthanasia process was placed into 3 mL serum vacutainer tubes (BD Biosciences, CA, USA), spun at 1000 x g for 15 minutes in a car battery-powered centrifuge, and the generated serum then transferred to a 1.5mL microcentrifuge tube. 30 μL of the generated serum was used to perform an antigen rapid diagnostic test, and another 140 μL of the generated serum was placed into Buffer AVL (Qiagen, Hilden, DE) for viral screening. The remaining generated serum, if any, was kept for antibody analysis. Liver, lung, kidney, and spleen tissue samples were harvested and stored in TRIzol® LS (Invitrogen, MA, USA). All serum and tissue samples were kept on ice until transport to the laboratory at KGH, where samples were then stored at -20°C. Inactivated tissue samples were shipped to the United States for downstream analyses.
3.1.5 In-field antigen rapid diagnostic test

The ReLASV Pan-Lassa Antigen Rapid Test (RDT) (Zalgen, MD, USA) was deployed during field collections in accordance with the manufacturer’s instructions. In brief, 30 μL (the approximate volume of one drop of blood) of fresh serum was placed onto the sample test pad of the RDT dipstick. Four drops of sample buffer were added to a 5mL conical tube, and the dipstick was oriented sample test pad down into the conical tube. The test was read 20 minutes after the dipstick was placed in the conical tube, and a photograph was taken of the dipstick. The test was scored based on the color of the test line. A negative result (no color) was scored as 0, and a positive result was scored on a scale of 1 (faint pink) to 5 (dark red) (Figure 4). A photograph of the test was taken for documentation.

3.1.6 RNA extraction from serum and tissue samples

Viral RNA was extracted from serum samples stored in Buffer AVL with the Viral RNA Mini Kit (Qiagen) in accordance with the manufacturer’s instructions; however, 10 μg/mL of linear acrylamide (VWR, PA, USA) was used in place of carrier RNA to prevent interference with downstream viral sequencing. Total RNA was extracted from tissue samples by guanidinium thiocyanate-phenol-chloroform phase extraction (Chomczynski and Sacchi 1987) and cleaned with the RNeasy Plus Mini Kit (Qiagen). 30 mg pieces of tissue were homogenized in 600 μL of TRIzol® (Invitrogen, MA, USA) using a Bullet Blender Storm Pro tissue homogenizer with zirconium oxide beads (NextAdvance, NY, USA).
Figure 4: Possible test results of the Zalgen ReLASV Pan-Lassa Antigen Rapid Test.

**ReLASV® Pan-Lassa Antigen Rapid Test**

A negative test result will have no color on the test stripe and has an RDT score of 0. A positive test result will have color on the test stripe, with color intensity and RDT Score corresponding to antigen strength.
Phase extraction was performed with 520 μL of homogenate, 60 μL of Nuclease – free H2O and 120 μL of Chloroform in PhaseMaker tubes (Invitrogen). Samples were vigorously shaken for 15 seconds, then spun at 12,000 x g for 5 minutes at 4°C. 350 μL of the aqueous layer was moved to a fresh 1.5 mL microfuge tube, and the RNA was precipitated with 350 μL of 100% Ethanol. The sample was then placed onto an RNeasy Mini Spin column, cleaned, and eluted in accordance with the RNeasy Plus Mini Kit instructions. Isolated RNA was stored at -80°C.

3.1.7 DNA extraction from tissue samples

15 mg pieces of tissue samples were homogenized in 600 μL of Buffer RLT Plus from the Qiagen RNeasy Plus Mini Kit with a Bullet Blender Storm Pro tissue homogenizer as described in RNA extraction. The homogenate was then spun for 3 minutes at 8,000 x g to dissolve homogenate foam. 520 μL of the homogenate was then placed onto a gDNA eliminator spin column from the RNeasy Plus Mini kit and spun for 30 seconds at 8,000 x g. Wash steps of the gDNA eliminator column with Buffer AW1 and AW2 were performed as described in the Viral RNA Mini Kit. 100 μL of DNA was eluted from the column with Buffer AE (Qiagen). Isolated DNA was stored at -20°C.

3.1.8 Cytochrome b PCR

DNA extracted from tissue samples was used to identify collected specimens to the species level with the cytochrome b gene as previously described (Lecompte, Granjon et al. 2002). Primers (Table 2) were used at a working concentration of 10 μM for a final concentration of 0.5 μM.
Table 2: Cytochrome b primers used in genotyping experiments (Lecompte et al., 2002).

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L14723_F</td>
<td>ACC AAT GAC ATG AAA AAT CAT CGT T</td>
</tr>
<tr>
<td>L15408_I</td>
<td>ATA GAC AAA ATC CCA TTC CA</td>
</tr>
<tr>
<td>L15146_I</td>
<td>CAT GAG GAC AAA TAT CAT TCT GAG</td>
</tr>
<tr>
<td>H15149_I</td>
<td>CTC AGA ATG ATA TTT GTC CTC</td>
</tr>
<tr>
<td>L15513_I</td>
<td>CTA GGA GAC CCA GAC AAC TA</td>
</tr>
<tr>
<td>H15553_I</td>
<td>TAG GCA AAT AGG AAA TAT CAT TCT GGT</td>
</tr>
<tr>
<td>H15915_R</td>
<td>TCT CCA TTT CTG GTT TAC AAG AC</td>
</tr>
</tbody>
</table>
For each specimen tested, 3 μL of isolated specimen DNA was combined with 22 μL of Master Mix, consisting of 1.25 μL of each primer, 12.5 μL of Q5 High-Fidelity 2x Master Mix (New England Biolabs, MA, USA) and 0.75 μL of Nuclease-free H2O for a final volume of 25 μL.

The samples were run in 96 well plates on a Veriti 96-well Fast Thermal Cycler (Applied Biosystems, MA, USA). The cycling conditions were as follows: 98°C for 30 seconds, 40 cycles of 98°C for 10 seconds, 50°C for 30 seconds, and 72°C for 40 seconds, and a final extension at 72°C for 2 minutes. 5 μL of PCR product was cleaned with the Exo-CIP Rapid PCR Cleanup Kit (New England Biolabs) to remove excess nucleotides prior to Sanger sequencing. DNA concentration was then assessed with the Nanodrop One/One spectrophotometer (ThermoFisher).

3.1.9 Sanger sequencing and alignment

25 ng of DNA was combined with 2.5 μL of 10 μM primer L14723_F and brought up to a final volume of 15 μL in Nuclease – free H2O. Sanger sequencing was performed by Genewiz (Genewiz, NJ, USA). The trace files of the Sanger sequences were examined for signal integrity and FASTA files of the examined sequences were downloaded. The FASTA files were cleaned and analyzed using Biopython, a bioinformatic package for Python (Cock, Antao et al. 2009). The first 30 nucleotides of the sequence were trimmed and only next 400 bp of the sequence were kept. The trimmed sequence reads were kept as the representative cytochrome \( b \) gene fragment of the specimen. The cytochrome \( b \) gene fragments from each specimen were compared to all other eukaryotic cytochrome \( b \) genes by the BLAST algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul, Gish et al. 1990). The BLAST results were downloaded as an XML file and parsed using Biopython.
The species identified from the top scoring hit of the BLAST results was considered the species of the specimen.

3.1.10 Quantitative Reverse-Transcription PCR (qRT-PCR)

One-step qRT-PCR was performed using previously validated primers (Nikisins, Rieger et al. 2015) on all serum samples, and tissues from antigen positive specimens with the SuperScript III Platinum One-Step qRT-PCR Kit (Invitrogen, MA, USA). 1 μL each of forward primer (5’ CCA CCA TYT TRT GCA TRT GCC A 3’), reverse primer (5’ GCA CAT GTN TCH TAY AGY ATG GAY CA 3’) and probe (5’ FAM-AAR TGG GGY CCD ATG ATG TGY CCW TT-BBQ) were added to the reaction mix. The primers were used at a working concentration of 10 μM and the probe 2 μM for final concentrations of 0.5 μM and 0.1μM respectively. 5 μL of RNA were used for a total volume of 20 μL. The qRT-PCR reaction was run on a Roche Lightcycler 480 in white 96 well plates or 8-well strips with the following cycling parameters: 48°C for 30 minutes, 95°C for 2 minutes, 45 cycles of 95°C for 15 seconds, 55°C for 45 seconds, and 72°C for 15 seconds, and a final machine cooldown of 40°C for 30 seconds. All samples and controls were run in duplicate.

3.1.11 qRT-PCR Controls

Amplicon-specific single-strand RNA (ssRNA) was run with every qRT-PCR reaction to generate a standard curve. ssRNA was generated from a gBlock fragment corresponding to LASV Josiah RNA (BEI Resources, MD, USA) using the T7 RiboMAX Express Large Scale RNA Production System (Promega, WI, USA) in accordance with the manufacturer’s instructions. The ssRNA was quantified with the QuBit Broad Range RNA Quantification Kit (Thermofisher, MA), then diluted to 10⁶ copies/20 μL reaction in a 20 ng/μL carrier RNA solution. The ssRNA was stored in 13 μL aliquots at -80°C until use.
At the time of use, a ten-fold, six-point standard curve ranging from 10⁶ - 10⁰ copies/20μL reaction was made in nuclease-free H₂O, and 5 μL of each standard was plated. A non-template control of nuclease-free H₂O and an additional positive control of LASV Josiah RNA (BEI Resources, MD, USA) diluted 1:10 in nuclease-free H₂O were used. 18S RNA was detected with a SYBR Green Based Assay (IDT, IA, USA) as an extraction quality control.

3.1.15 Statistical Analysis – Data cleaning and general statistics

Rodent collection data were initially captured in paper records prior recording in Microsoft Excel (Microsoft, WA, USA). The final dataset, with animal information, location, and test results was compiled and cleaned in RStudio v.1.4.1717 with the packages tidyverse, gt (RStudio Team 2020), ggpubr (https://CRAN.R-project.org/package=ggpubr), ggprism (Dawson 2021), gtsummary (Sjoberg, Whiting et al. 2021), epiR (https://CRAN.R-project.org/package=epiR), and rstatix (https://CRAN.R-project.org/package=rstatix). Statistical comparisons between tests and figures used in this publication were also generated in RStudio.

3.1.16 Statistical Analysis – qRT-PCR

Absolute quantification of viral RNA levels for all samples was calculated with the Roche Lightcycler 480 software from the ssRNA standard curve by the second derivative maximum method. Viral copies per μL of RNA were then calculated for the samples from the standard curve. The limit of detection (LOD) for each tissue type was initially assessed with RNA isolated from negative control M. coucha specimens (n = 5, see chapter 2 methods) to an ssRNA standard to generate Receiver Operator Characteristic (ROC) curves in GraphPad Prism software v.9.0 (GraphPad, CA, USA). However, LOD tests indicated
theoretical detection below zero copies per reaction; in addition, we were unable to obtain enough negative control specimens to create accurate ROC curves. Thus, the LOD was set at 1.5x the mean number of viral copies of the $10^1$ copies/reaction standard to account for potential differences in PCR inhibition between different species. Ten RNA copies per reaction is broadly considered the limit of detection by qRT-PCR (Forootan, Sjoback et al. 2017). To determine the exact LOD for each tissue type, a standard curve diluted with extracted RNA from negative control tissue was compared to a standard curve diluted with nuclease-free H$_2$O. The LOD of each tissue type was then set to reflect the number of viral copies needed to equate $10^1$ copies/reaction of the water-diluted standard curve. The cutoffs for each tissue were as follows: Serum (AVL): 22.8 copies/µL RNA; Liver: 129.6 copies/µL RNA; Lung: 11.4 copies/µL RNA; Kidney: 53.2 copies/µL RNA; Spleen: 5.16 copies/µL RNA. RNA from samples determined to have more than 150 viral copies/µL RNA (~Cycle threshold [Ct] value of 30) were shipped to the Andersen lab at Scripps Research to attempt viral sequencing.

3.2 Results

3.2.1 Traits and distribution of study population

Our final study population contains a total of 534 small mammal specimens, representing six different genera. Approximately 94% of the specimens were identified in the field as either *Mastomys spp.* or *Rattus spp.* (Table 3). Of the *Mastomys spp.* specimens available for cytochrome $b$ identification, all but two were identified as *M. natalensis*; the remaining two specimens were identified as *M. erythroleucus* (Table 4).
Table 3: Characteristics of specimens in study.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mastomys, $N = 300^1$</th>
<th>Rattus, $N = 200^1$</th>
<th>Praomys, $N = 18^1$</th>
<th>Crocidura, $N = 8^1$</th>
<th>Hylomyscus, $N = 7^1$</th>
<th>Graphiurus, $N = 1^1$</th>
<th>Overall, $N = 534^1$</th>
<th>p-value\footnote{\textsuperscript{2}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.281</td>
</tr>
<tr>
<td>F</td>
<td>163 (54%)</td>
<td>102 (51%)</td>
<td>14 (78%)</td>
<td>4 (50%)</td>
<td>4 (57%)</td>
<td>1 (100%)</td>
<td>288 (54%)</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>137 (46%)</td>
<td>98 (49%)</td>
<td>4 (22%)</td>
<td>4 (50%)</td>
<td>3 (43%)</td>
<td>0 (0%)</td>
<td>246 (46%)</td>
<td></td>
</tr>
<tr>
<td>Weight (Age proxy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.040</td>
</tr>
<tr>
<td>&gt; 40g (Adult)</td>
<td>111 (37%)</td>
<td>94 (47%)</td>
<td>7 (39%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>212 (40%)</td>
<td></td>
</tr>
<tr>
<td>25-40g (Sub-Adult)</td>
<td>97 (32%)</td>
<td>67 (34%)</td>
<td>8 (44%)</td>
<td>4 (50%)</td>
<td>1 (14%)</td>
<td>0 (0%)</td>
<td>177 (33%)</td>
<td></td>
</tr>
<tr>
<td>&lt; 25g (Juvenile)</td>
<td>92 (31%)</td>
<td>39 (20%)</td>
<td>3 (17%)</td>
<td>4 (50%)</td>
<td>6 (86%)</td>
<td>1 (100%)</td>
<td>145 (27%)</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{1} n (\%)
\textsuperscript{2} Fisher’s Exact Test for Count Data with simulated p-value (based on 2000 replicates)

Counts and percentage of study population for each collected specimen genera by sex and weight (a proxy for age of the specimen; described in (Leirs 1992). F = Female, M = Male. The p-value for Weight (Age proxy) was calculated only with *Mastomys* *spp.*, *Rattus* *spp.*, and *Praomys* *spp.*
Table 4: Counts of Species by cytochrome b identification.

<table>
<thead>
<tr>
<th>Species</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Mastomys natalensis</em></td>
<td>181</td>
</tr>
<tr>
<td><em>Rattus rattus</em></td>
<td>93</td>
</tr>
<tr>
<td><em>Praomys rostratus</em></td>
<td>12</td>
</tr>
<tr>
<td><em>Hylomyscus simus</em></td>
<td>5</td>
</tr>
<tr>
<td><em>Crocidura olivieri</em></td>
<td>3</td>
</tr>
<tr>
<td><em>Mastomys erythroleucus</em></td>
<td>2</td>
</tr>
<tr>
<td><em>Graphiurus kelleni</em></td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>297</strong></td>
</tr>
</tbody>
</table>

Cytochrome *b* identification was only able to be performed on a subset of samples available at time of publication.
All the available *Rattus* specimens for cytochrome *b* identification were identified as *Rattus rattus*, the common black rat. Similar numbers of male and female specimens were captured (p = 0.288 by Fisher’s Exact Test) (Table 3). The number of specimens grouped by weight, a general proxy for age of a specimen (Leirs 1992), varied among *Mastomys spp.*, *Rattus spp.*, and *Praomys spp.* specimens (p = 0.042 by Fisher’s Exact Test). *Crocidura spp.*, *Hylomyscus spp.*, and *Graphiurus spp.* species were not included in the weight-age analysis due to the limited number of specimens collected.

Among the 26 villages trapped during the study, specimens were captured in all but two of the villages trapped across a total of 17,943 trap-nights (traps set per night) (Table 5). The number of captured specimens in a village was only weakly correlated with ratio of trap-nights/capture in a village (R² = 0.218, p = 0.005 by Linear Regression). The composition of captured specimens varied between villages, however, the relative amounts of *Mastomys spp.* and *Rattus spp.* were evenly distributed among the villages when normalized to percent composition of captured specimens per village (Table 6A, 6B) (p = 0.235 and p = 0.152 respectively by the Shapiro-Wilk Test for normality). Other captured specimens were not evenly distributed (p = 0.0122 by the Shapiro-Wilk Test for normality). Across all villages, *Mastomys spp.* accounted for an average of 56% of captured specimens in a village, *Rattus spp.* accounted for an average of 37% of captured specimens in a village, and all other species accounted for an average of 7% of captured specimens in a village.
Table 5: Summary of trapping activity by village.

<table>
<thead>
<tr>
<th>Village</th>
<th>Total trap-nights</th>
<th>Total captures</th>
<th>Trap-nights/Capture</th>
<th>R²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>17943(^1)</td>
<td>534(^1)</td>
<td>49.35(^2)</td>
<td>0.218(^3)</td>
<td>0.005(^3)</td>
</tr>
<tr>
<td>Bambawo</td>
<td>344</td>
<td>3</td>
<td>114.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bowohun</td>
<td>456</td>
<td>13</td>
<td>35.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guabu</td>
<td>256</td>
<td>0</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jamborna</td>
<td>578</td>
<td>10</td>
<td>57.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joru</td>
<td>486</td>
<td>14</td>
<td>34.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kamborna</td>
<td>768</td>
<td>28</td>
<td>28.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koi</td>
<td>552</td>
<td>15</td>
<td>36.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kormolu</td>
<td>720</td>
<td>11</td>
<td>65.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koyama Ngiena</td>
<td>486</td>
<td>30</td>
<td>16.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kpalu</td>
<td>1397</td>
<td>27</td>
<td>51.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kpetema</td>
<td>762</td>
<td>6</td>
<td>127.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Largo</td>
<td>1046</td>
<td>69</td>
<td>15.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lawana</td>
<td>184</td>
<td>0</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maleh</td>
<td>694</td>
<td>29</td>
<td>23.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mano Ngieya</td>
<td>518</td>
<td>40</td>
<td>12.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ngeihun</td>
<td>1458</td>
<td>8</td>
<td>182.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ngouma</td>
<td>392</td>
<td>14</td>
<td>28.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Njagor</td>
<td>470</td>
<td>22</td>
<td>21.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyahahun</td>
<td>722</td>
<td>15</td>
<td>48.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pandembu</td>
<td>1458</td>
<td>37</td>
<td>39.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panguma (vaama)</td>
<td>1106</td>
<td>28</td>
<td>39.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pujejun</td>
<td>510</td>
<td>12</td>
<td>42.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saahun</td>
<td>796</td>
<td>70</td>
<td>11.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saleima</td>
<td>546</td>
<td>13</td>
<td>42.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tongola</td>
<td>1218</td>
<td>20</td>
<td>60.90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) Sum of all villages  
\(^2\) Average trap-nights/capture across all villages  
\(^3\) Linear regression of Trap-nights/Capture vs Total captures

Total number of trap-nights, total number of captures, and trap-nights per capture for each village in the study, and statistics for linear regression of Trap-nights/Capture vs. Total captures.
Table 6: Composition of specimens captured by villages and their distributions.

A

<table>
<thead>
<tr>
<th>Village</th>
<th>Mastomys, N = 300(^1)</th>
<th>Rattus, N = 200(^2)</th>
<th>Praomys, N = 18(^3)</th>
<th>Cricidura, N = 8(^4)</th>
<th>Hylomyscus, N = 7(^5)</th>
<th>Graphiurus, N = 1(^6)</th>
<th>Overall, N = 534(^7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bambawo</td>
<td>1 (33%)</td>
<td>2 (67%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>3 (100%)</td>
</tr>
<tr>
<td>Bwokohun</td>
<td>5 (38%)</td>
<td>8 (62%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>13 (100%)</td>
</tr>
<tr>
<td>Gouma</td>
<td>6 (43%)</td>
<td>4 (29%)</td>
<td>4 (29%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>14 (100%)</td>
</tr>
<tr>
<td>Jamboma</td>
<td>13 (46%)</td>
<td>15 (54%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>28 (100%)</td>
</tr>
<tr>
<td>Joru</td>
<td>6 (40%)</td>
<td>7 (47%)</td>
<td>2 (13%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>15 (100%)</td>
</tr>
<tr>
<td>Kamboma</td>
<td>9 (82%)</td>
<td>1 (9%)</td>
<td>0 (0%)</td>
<td>1 (9.1%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>11 (100%)</td>
</tr>
<tr>
<td>Koi</td>
<td>5 (17%)</td>
<td>25 (83%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>30 (100%)</td>
</tr>
<tr>
<td>Komolou</td>
<td>19 (70%)</td>
<td>6 (22%)</td>
<td>2 (7.4%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>27 (100%)</td>
</tr>
<tr>
<td>Koyama</td>
<td>3 (50%)</td>
<td>2 (33%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (17%)</td>
<td>0 (0%)</td>
<td>6 (100%)</td>
</tr>
<tr>
<td>Kgulu</td>
<td>52 (75%)</td>
<td>12 (17%)</td>
<td>5 (7.2%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>69 (100%)</td>
</tr>
<tr>
<td>Kpetera</td>
<td>5 (17%)</td>
<td>23 (79%)</td>
<td>1 (3.4%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>29 (100%)</td>
</tr>
<tr>
<td>Laro</td>
<td>29 (72%)</td>
<td>10 (25%)</td>
<td>1 (2.5%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>40 (100%)</td>
</tr>
<tr>
<td>Malah</td>
<td>6 (75%)</td>
<td>1 (13%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (12%)</td>
<td>0 (0%)</td>
<td>8 (100%)</td>
</tr>
<tr>
<td>Mano-Neinya</td>
<td>11 (79%)</td>
<td>2 (14%)</td>
<td>1 (7.1%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>14 (100%)</td>
</tr>
<tr>
<td>Njehun</td>
<td>12 (55%)</td>
<td>6 (27%)</td>
<td>1 (4.5%)</td>
<td>0 (0%)</td>
<td>3 (14%)</td>
<td>0 (0%)</td>
<td>22 (100%)</td>
</tr>
<tr>
<td>Njigor</td>
<td>3 (30%)</td>
<td>7 (70%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>10 (100%)</td>
</tr>
<tr>
<td>Nyahun</td>
<td>5 (33%)</td>
<td>8 (53%)</td>
<td>2 (13%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>15 (100%)</td>
</tr>
<tr>
<td>Pandembu</td>
<td>10 (27%)</td>
<td>24 (65%)</td>
<td>0 (0%)</td>
<td>1 (2.7%)</td>
<td>1 (2.7%)</td>
<td>1 (2.7%)</td>
<td>37 (100%)</td>
</tr>
<tr>
<td>Panguma</td>
<td>20 (71%)</td>
<td>7 (25%)</td>
<td>0 (0%)</td>
<td>1 (3.6%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>28 (100%)</td>
</tr>
<tr>
<td>Pujehun</td>
<td>1 (8.3%)</td>
<td>10 (83%)</td>
<td>0 (0%)</td>
<td>1 (8.3%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>12 (100%)</td>
</tr>
<tr>
<td>Saahun</td>
<td>53 (76%)</td>
<td>15 (21%)</td>
<td>1 (1.4%)</td>
<td>1 (1.4%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>70 (100%)</td>
</tr>
<tr>
<td>Seimba</td>
<td>8 (62%)</td>
<td>5 (38%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>13 (100%)</td>
</tr>
<tr>
<td>Tongola</td>
<td>18 (90%)</td>
<td>0 (0%)</td>
<td>2 (10%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>20 (100%)</td>
</tr>
</tbody>
</table>

\(^1\) n (% composition of village)

B

<table>
<thead>
<tr>
<th>Species</th>
<th>W</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mastomys</td>
<td>0.945</td>
<td>0.235</td>
</tr>
<tr>
<td>Rattus</td>
<td>0.937</td>
<td>0.152</td>
</tr>
<tr>
<td>Other</td>
<td>0.884</td>
<td>0.012</td>
</tr>
</tbody>
</table>

Table A: The number and percent composition of each type of specimen collected in each village where specimens were captured. Table B: Results of the Shapiro-Wilk Test (W) for normality to assess distribution of specimen composition between villages. W = 1 indicates perfectly even distribution, W = 0 indicates extreme variation of distribution.
3.2.2 Antigen presence via rapid diagnostic test (RDT)

To determine LASV antigen (Ag) presence in the study population we deployed the Zalgen ReLASV Pan-Lassa Antigen Rapid Test (RDT) at the time of specimen processing in-field, using serum from the collected specimen. Of the 523 specimens tested with the RDT, 99 (19%) tested antigen positive (Table 7). 92 (93%) of the antigen positive specimens were identified as *Mastomys* spp.; the remaining antigen positive specimens were identified as *Praomys* spp. and *Rattus* spp. No antigen positive specimens were identified as *Hylomyscus* spp., *Crocidura* spp., or *Graphiurus* spp. Interestingly, the ratio of antigen positive specimens to the total number of collected specimens for *Mastomys* spp. and *Praomys* spp. is not significant (p = 0.414 by Pairwise Fisher’s Exact Test) (Figure 5A). No significant differences are observed in the ratio of antigen positive specimens to antigen negative specimens by sex (p = 0.737 by Fisher’s Exact Test) or by weight group, a proxy for the specimen age (p = 0.329 by Fisher’s Exact Test) (Figure 5B, 5C). Additionally, a broad range of antigen strength, as represented by a semi-quantitative RDT score, are observed among the antigen positive specimens (Figure 5D).

We identified 85 of the 99 antigen positive specimens to the species level, including all the *Praomys* spp. and *Rattus* spp. antigen positive specimens (Figure 4E). Of the 78 antigen-positive *Mastomys* spp. specimens identified to the species level, one was identified as *M. erythroleucus* and the remaining 77 specimens were identified as *M. natalensis*. All the antigen positive *Praomys* spp. were identified as *P. rostarus*, and all the antigen positive *Rattus* spp. were identified as *R. rattus*. 
Table 7: Morphometric characteristics of specimens grouped by antigen RDT result.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Ag-, N = 424</th>
<th>Ag+, N = 99</th>
<th>Overall, N = 523</th>
<th>p-value&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mastomys</td>
<td>207 (69%)</td>
<td>92 (31%)</td>
<td>299 (100%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Rattus</td>
<td>188 (98%)</td>
<td>4 (2.1%)</td>
<td>192 (100%)</td>
<td></td>
</tr>
<tr>
<td>Praomys</td>
<td>14 (82%)</td>
<td>3 (18%)</td>
<td>17 (100%)</td>
<td></td>
</tr>
<tr>
<td>Crocidura</td>
<td>7 (100%)</td>
<td>0 (0%)</td>
<td>7 (100%)</td>
<td></td>
</tr>
<tr>
<td>Hylomuscus</td>
<td>7 (100%)</td>
<td>0 (0%)</td>
<td>7 (100%)</td>
<td></td>
</tr>
<tr>
<td>Graphiurus</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>1 (100%)</td>
<td></td>
</tr>
<tr>
<td>Weight (Age proxy)</td>
<td></td>
<td></td>
<td></td>
<td>0.371</td>
</tr>
<tr>
<td>&gt; 40g (Adult)</td>
<td>171 (83%)</td>
<td>35 (17%)</td>
<td>206 (100%)</td>
<td></td>
</tr>
<tr>
<td>25-40g (Sub-Adult)</td>
<td>135 (78%)</td>
<td>39 (22%)</td>
<td>174 (100%)</td>
<td></td>
</tr>
<tr>
<td>&lt; 25g (Juvenile)</td>
<td>118 (83%)</td>
<td>25 (17%)</td>
<td>143 (100%)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td>0.738</td>
</tr>
<tr>
<td>F</td>
<td>227 (80%)</td>
<td>55 (20%)</td>
<td>282 (100%)</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>197 (82%)</td>
<td>44 (18%)</td>
<td>241 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup>n (%)
<sup>2</sup>Fisher’s exact test

Counts and percentage of antigen RDT test results in study population by genera, weight (a proxy for age of the specimen; described in (Leirs 1992)) and sex. The table is visualized in Figure 4A-4C. Ag<sup>-</sup> = antigen negative, Ag<sup>+</sup> = antigen positive, F = Female, M = Male. This table is visualized in Figure 4A-4C.
Figure 5: Traits of antigen presence among study population.

A-C: Percentage of antigen positive/antigen negative specimens by characteristic. A: Genera (Pairwise Fisher’s Exact Test, **** = p < 0.0001, * = p < 0.05), B: Sex (Fisher’s Exact Test), C: Weight (a proxy for age of the specimen; described in (Leirs 1992) (Fisher’s Exact Test). Ag- = antigen negative, Ag+ = antigen positive. D: Histogram of RDT score distribution among antigen positive specimens. A higher score means more antigen was present in the sample. E: Histogram of species with a positive RDT test (antigen positive).
3.2.3 Viral presence as detected by qRT-PCR and its relation to antigen status

We screened serum from each individual specimen for viral presence by a previously established qRT-PCR assay (Nikisins, Rieger et al. 2015). Surprisingly, only one serum sample screened (n = 158) had a viral level above the negative cutoff (22.8 copies/μL RNA) (Figure 5A). We then focused our efforts on screening tissues from specimens testing antigen positive, as tissues from antigen negative specimens were unavailable at the time of publication. Here, we found 41% of the specimens testing antigen positive had at least one tissue with viral levels above the negative cutoff (Table 8). Of the qRT-PCR+ specimens, detectable viral levels were most often seen in the lung (78% of specimens) and spleen (74% specimens). Detectable viral levels were less frequently seen in kidney (37% of specimens) and liver (32% of specimens). The median viral level across all tested tissues was highest in lung tissue (Figure 6A).

Among qRT-PCR+ tissues, viral levels were classified as either a “strong positive” (≥ 200 copies/μL RNA) or a “weak positive” (< 200 copies/μL RNA). No liver tissue samples and only one kidney tissue sample were classified as weak positive, while lung and spleen tissues had approximately equal numbers of tissues classified as strong positive or weak positive (Figure 6B). Among tissues classified as strong positive, lung and kidney tissues had the highest median viral levels (Figure 6C). However, the viral levels between tissues of a single specimen remained relatively consistent; typically clustered within one log of each other (Figure 6D).
Figure 6: Viral levels in tissues of specimens.

A: Mean viral levels of each tissue tested by qRT-PCR. Tissue-specific cutoff concentrations are indicated in the methods. The collective mean for each tissue was compared to one another (Dunn’s Post-hoc Pairwise Rank Test, **** = p < 0.0001, ** = p < 0.01). Serum qRT-PCR results were excluded from further analyses., B-C: Mean viral levels of each tissue testing qRT-PCR positive, classified as a strong or weak qRT-PCR positive. The collective mean for each tissue was compared to one another (Dunn’s Post-hoc Pairwise Rank Test, ** = p < 0.01, * = p < 0.05),. D: Mean viral levels of every tissue of antigen positive specimen tested.
Table 8: qRT-PCR status of tissues from antigen positive specimens.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Negative, N = 54&lt;sup&gt;†&lt;/sup&gt;</th>
<th>Positive, N = 38&lt;sup&gt;†&lt;/sup&gt;</th>
<th>Overall, N = 92&lt;sup&gt;†&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum</td>
<td>0 / 32 (0%)</td>
<td>1 / 19 (5.3%)</td>
<td>1 / 51 (2.0%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>22</td>
<td>19</td>
<td>41</td>
</tr>
<tr>
<td>Liver</td>
<td>0 / 42 (0%)</td>
<td>11 / 34 (32%)</td>
<td>11 / 76 (14%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>12</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Lung</td>
<td>0 / 45 (0%)</td>
<td>28 / 36 (78%)</td>
<td>28 / 81 (35%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>9</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Kidney</td>
<td>0 / 48 (0%)</td>
<td>13 / 35 (37%)</td>
<td>13 / 83 (16%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>6</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Spleen</td>
<td>0 / 41 (0%)</td>
<td>23 / 31 (74%)</td>
<td>23 / 72 (32%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>13</td>
<td>7</td>
<td>20</td>
</tr>
</tbody>
</table>

<sup>†</sup><em>n / N (%)</em>

Counts and percentage qRT-PCR<sup>+</sup> for each tissue tested from antigen positive (Ag<sup>+</sup>) specimens. Negative = Ag<sup>+</sup> specimens testing qRT-PCR<sup>−</sup> in all tissues, Positive = Ag<sup>+</sup> specimens testing qRT-PCR<sup>+</sup> in at least one tissue, Overall = all Ag<sup>+</sup> specimens tested by qRT-PCR.
A clear link is observed between antigen presence and viral presence as assessed by qRT-PCR (Fisher’s Exact Test, \( p = 0.000239 \)) (Table 9). The sensitivity and specificity of the RDT compared to qRT-PCR from the samples tested is 41.3% and 100% respectively (Table 9). However, we observed little to no correlation between average viral level among all tissues in the specimen and antigen strength, as well as viral level of the individual tissue and antigen strength (Figure 7A, 7B). Overall, LASV antigen status of a specimen is mildly indicative of whether a specimen will also have detectable viral levels of LASV by qRT-PCR.
Table 9: Frequency table of antigen and qRT-PCR status and its associated sensitivity and specificity.

<table>
<thead>
<tr>
<th></th>
<th>Ag -</th>
<th>Ag +</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR -</td>
<td>108</td>
<td>54</td>
<td>162</td>
</tr>
<tr>
<td>PCR +</td>
<td>0</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>Column Total</td>
<td>108</td>
<td>92</td>
<td>200</td>
</tr>
</tbody>
</table>

$p < 0.001$ by Fisher's Exact Test

<table>
<thead>
<tr>
<th></th>
<th>est</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>100.0%</td>
<td>96.6%-100.0%</td>
</tr>
<tr>
<td>Specificity</td>
<td>41.3%</td>
<td>31.1%-52.1%</td>
</tr>
</tbody>
</table>
A: Mean viral level across all tissues of antigen positive specimens vs. antigen strength of specimens as indicated by RDT score. A specimen was considered qRT-PCR positive if at least one tissue tested qRT-PCR positive. B: Mean viral level of each tissue of qRT-PCR positive specimens vs. antigen strength of specimens as indicated by RDT score. A specimen was considered a strong qRT-PCR positive if at least one tissue tested had a mean viral level $\geq 200$ copies/μL.
3.3 Discussion

3.3.1 Traits and distribution of study population

Herein, we have performed one of the largest epizoology surveys in Sierra Leone since the 1980’s (McCormick, Webb et al. 1987). We designed our cross-sectional study with the intent to capture as many small mammal specimens as possible in a short timeframe. With the collected specimens, we sought to i) uncover the true incidence rate of LASV in small mammals in the Kenema district, ii) establish a diagnostic algorithm driven by field-deployable diagnostics to screen for small mammals with active LASV infection, and iii) sequence LASV from small mammals with active LASV infections to visualize LASV diversity in Sierra Leone. We partly chose our collection sites relative to the location of the main research site at Kenema Government Hospital, and partly due to the high incidence of LASV and Lassa fever in the villages sampled. Homes selected for trapping were clustered together in a village, and a variety of locations within the village were trapped to ensure a wide variety of small mammals, mainly rodent species, would be captured. The broad, non-targeted approach to in-home trapping yielded many specimens, with an even distribution of numbers and species collected across villages (Table 6). Two of the villages trapped in the study, Guaba and Lawana, did not yield any captured specimens (Table 5). This result is unsurprising, given that those two villages had the fewest number of trap-nights across the entire study. However, the number of trap-nights was only weakly correlated with number of specimens captured. Other factors not examined in this study may explain discrepancies of rodent abundance and composition within and between villages.
Almost 95% of specimens captured were identified as either *Mastomys spp.* (56%) or *Rattus spp.* (38%) upon collection (Table 3), with a nearly identical composition of rodent species collected in a previous non-serological trapping survey performed in Sierra Leone (Bonwitt, Saez et al. 2017). This study, however, noted potential under-sampling of *Rattus spp.* due to the size of the Sherman traps used in the collection process. Large *Rattus spp.* may be unable to fit in the traps and thus go uncaptured.

To further identify specimens to the species level, we sequenced cytochrome *b* PCR fragments on 305 available specimens (Table 4). While nearly all (98.8%) of the *Mastomys spp.* specimens were identified as *M. natalensis*, two (1.2%) of the *Mastomys spp.* specimens analyzed were identified as *M. erythroleucus*. All *Rattus spp.* specimens analyzed were identified as *R. rattus*. The composition of our study population is in line with previous studies in Sierra Leone (Keenlyside, McCormick et al. 1983, McCormick, Webb et al. 1987, Bonwitt, Saez et al. 2017). As the foregoing studies primarily focused on Eastern Sierra Leone, as did our study, future trapping studies are needed to characterize small mammal diversity in other regions of the country, which could provide insight into the relative paucity of LASV infections in these regions (Olayemi, Obadare et al. 2017).

3.3.2 Antigen presence via rapid diagnostic test (RDT)

One of the motivations of this study was to evaluate the performance of the Zalgen ReLASV Pan-Lassa RDT in collected rodents as part of a diagnostic algorithm to detect active LASV infections in small mammals. Of the 523 specimens examined for LASV antigen in our study, 99 (18.5%) tested antigen positive as determined by RDT (Table 7). Equal numbers of male and female specimens tested antigen positive, and equal numbers of specimens grouped by weight categories as a proxy for age tested antigen positive (Table
7, Figure 4B-4C). Antigen strength among specimens as indicated by RDT score (1-5) were broadly distributed (Figure 5D). Few previous studies have assessed the level of LASV antigen in small mammals. In the two known studies where circulating LASV antigen probed, LASV antigen was detected in less than 5% of specimens tested (Wulff, Fabiyi et al. 1975, Demby, Inapogui et al. 2001). In contrast, the even distribution among sex and age of specimens testing antigen positive in our study indicates robust antigen detection by the RDT in the host reservoir.

Of the 99 antigen positive specimens, 92 specimens were identified as *Mastomys* spp. All the antigen positive *Mastomys* spp. were identified as *M. natalensis* except for one specimen identified as *M. erythroleucus* (Table 7, Figure 5E). While our survey did not collect enough *M. erythroleucus* specimens to infer its potential as a host reservoir, other studies have suggested *M. erythroleucus* as a novel LASV reservoir in Nigeria and Guinea (Olayemi, Cadar et al. 2016). Three other antigen positive specimens were identified as *Rattus* spp., and the four remaining antigen positive specimens were identified as *Praomys* spp. (Table 7). LASV antigen or virus has occasionally been detected in *Rattus* spp., most likely due to a secondary spillover event as opposed to viral maintenance amongst the species (Wulff, Fabiyi et al. 1975, Agbonlahor, Erah et al. 2017). LASV antigen or viral presence has not been previously detected in *Praomys* spp., though LASV IgG has been seen in *P. rostarurs* (Olayemi, Cadar et al. 2016). The ratio of specimens testing antigen positive relative to the total number of specimens collected was the same for *Mastomys* spp. and *Praomys* spp. (Figure 5A). Given the close relation of *Praomys* spp. to other identified LASV reservoirs, this further demonstrates the ability of LASV to infect a broad range of the *Praomyini* tribe (Lecompte, Granjon et al. 2002, Nicolas, Mikula et al. 2021).
As *Praomys spp.* is primarily a tree-associated rodent (Ademola, Vanden Broecke et al. 2021), future research should place traps in and around trees both within village centers and on village outskirts to further investigate the dynamics of LASV in *Praomys spp.*

3.3.3 Viral presence as detected by qRT-PCR and its relation to antigen status

qRT-PCR is often considered the “gold standard” of viral diagnostics; thus, we wanted to compare detection of LASV by antigen RDT to qRT-PCR. Our initial screening strategy was to test viral RNA from serum samples of each specimen by qRT-PCR to determine specificity and sensitivity of the RDT for LASV detection as compared to qRT-PCR. Total RNA from liver, lung, kidney, and spleen tissue samples would then only be extracted and screened from antigen negative specimens testing qRT-PCR positive by serum. Total RNA of the aforenoted tissue samples would also be extracted from all specimens testing antigen positive regardless of the serum qRT-PCR result. This strategy proved unsuccessful, as only 1 serum sample tested positive for LASV by qRT-PCR (Figure 6A, Table 8). The lack of viremia indicates LASV is a well-controlled intracellular infection in the rodent host (Meyer and Ly 2016). Indeed, LASV presence in rodents is frequently assessed by dried whole blood spots (DBS) to test for IgG presence by IFA and viral presence by PCR (Fichet-Calvet, Becker-Ziaja et al. 2014, Borremans, Vossen et al. 2015, Gryseels, Rieger et al. 2015). Although the RDT has been validated for LASV antigen detection with whole blood, serum, and plasma, it was unknown whether antigen could be detected from DBS. In retrospect, antigen detection by DBS should have been experimentally validated prior to field deployment of the RDT. If DBS validation of antigen detection did not work, we should have saved the cells remaining after serum separation for virus detection alongside the serum.
To that end, this study only examined tissue samples from specimens that tested antigen positive; antigen negative tissue was not readily available in quantities needed for qRT-PCR analysis in the US due to shipping disruptions caused by the COVID-19 pandemic. Nonetheless, 38/92 specimens (41%) that tested antigen positive also tested positive by qRT-PCR in at least one tissue (Table 9). Lung was the tissue most likely to test positive by qRT-PCR (28/36 samples tested; 78%), followed by spleen (23/31 samples tested; 71%), kidney (13/35 samples tested; 37%) and liver (11/31 samples tested; 32%). Viral levels of different tissues within a specimen were typically contained within one log of each other (Figure 6D). In addition, lung tissue tended to have the highest mean viral levels, followed by kidney, spleen, and liver (Figure 6A, 6B). Other experimental animal studies of LASV infection have also observed high LASV levels in lung and kidney tissues (Walker, Wulff et al. 1975, Baillet, Reynard et al. 2021). We further classified qRT-PCR results as a weak positive (< 200 copies/μL RNA) or a strong positive (≥ 200 copies/μL RNA) (Figure 6B). Even with the stratification, lung tissues still had the highest mean viral loads (Figure 6C). However, within the constraints of this study, we cannot elucidate whether LASV detected in the lungs is found in capillary endothelial cells, resident lung macrophages, or other white blood cells found within the lungs. Furthermore, we may have had more specimens and tissues test positive if we also performed qRT-PCR with two or more primer/probe sets to detect different parts of the LASV genome.

We next wanted to observe if there was correlation between viral levels and antigen strength. Although antigen can be detected prior to viral detection by qRT-PCR, as well as linger in the blood after viral clearance, we still hypothesized antigen strength would moderately correspond with viral levels as previously observed (Boisen, Hartnett et al. 2021).
2018, Boisen, Uyigue et al. 2020). Curiously, we saw no correlation between viral levels and antigen strength among all qRT-PCR samples tested, both positive and negative, despite a clear link between antigen presence and viral presence (Figure 7A, Table 9). The lack of correlation was also observed within each individual organ and among qRT-PCR positive samples classified as a weak positive or a strong positive (Figure 7B). The overall specificity and sensitivity of the RDT assay as compared to qRT-PCR was 41% and 100% respectively (Table 9). However, antigen negative tissues will need to be examined to assess the true specificity and sensitivity of the RDT in comparison to qRT-PCR. Our unbiased, single timepoint sampling method means specimens were collected in various states of viral infection, which could explain the lack of correlation between viral levels and antigen strength.
4. CHARACTERIZATION OF LASSA VIRUS ANTIBODIES CIRCULATING IN STUDY POPULATION AND DEVELOPMENT OF A DIAGNOSTIC ALGORITHM TO SCREEN RODENTS FOR ACTIVE LASSA VIRUS INFECTIONS

4.1 Materials and Methods

4.1.1 Adaptation of human LASV IgG ELISA for detection of rodent LASV IgG

Commercial Human LASV IgG ELISA kits purchased from Zalgen Labs were modified to allow for detection of mice and rat LASV IgG antibodies. Anti-Rat and Anti-Mouse IgG horseradish peroxidase (HRP) secondary antibodies were used in place of the anti-human secondary antibody solution provided in the kits, diluted in antibody conjugate solution provided by Zalgen. Species-specific positive and negative controls were used in place of the provided lyophilized human controls. All other kit components, including the antigen coated ELISA plates, sample diluent, wash solution (PBS-Tween 20), TMB detection substrate, and stop solution (2% methylsulfuric acid) from the kits were used in accordance with the manufacturer’s instructions.

To create species-specific positive controls, serum from mice and rats were immunized with antigens to either LASV GP (Lineage IV Lassa rGP, Zalgen Labs) or NP (A mixture of Lineage II, II, and IV Lassa N-terminal (1-340) rNP, Zalgen Labs). Immunization and bleeds were performed by either Prosci Inc. (Prosci Incorporated, CA, USA) or Thermofisher. Five BALB/C mice (Prosci) and two Sprague-Dawley Rats (Thermofisher) were immunized with either Pf-GP or NP antigen with Complete Freund’s
Adjuvant and boosted on days 21 and 35 with Incomplete Freund’s Adjuvant. Terminal bleeds were performed on day 50. Equal amounts of terminal bleed from each animal were mixed for a species-specific positive control. Mouse GP serum was diluted 1:25 and Mouse NP serum was diluted 1:35 in PBS prior to use in ELISA assays.

*Mastomys spp.* rodents, identified by cytochrome *b* PCR as *M. coucha*, (Safronetz, Rosenke et al. 2021) were obtained from a local python breeder (Bailey & Bailey Reptiles, LA, USA) to provide negative mouse serum and tissue controls. Additional *M. natalensis* serum was kindly provided by Heinz Feldman of Rocky Mountain Laboratories (NIAID, Montana, USA) to compare with *M. coucha* bleeds. Serum from individual Sprague-Dawley rats (*Rattus rattus*) was purchased (Biomed Chemical, VA, USA) for use as rat negative control serum.

Prior to ELISA testing, specimens were classified as either mouse-related specimens (*Mastomys spp.*, *Praomys spp.*, and *Hylomyscus spp.*) or rat-related specimens (*Rattus spp.*) to ensure the appropriate ELISA reagents were utilized for the respective specimens. (Lecompte, Granjon et al. 2002). *Crocidura spp.* were not tested in this study as anti-shrew antibodies are not commercially available.

4.1.2 Rodent LASV IgG ELISA

The serum from individual specimens and serum from the negative controls was diluted 1:100 in sample diluent. A three-fold, six-point standard curve with an initial 1:100 dilution was created from the pre-diluted positive control serum in sample diluent. 100 μL of diluted serum/controls were placed on the antigen coated ELISA plate and incubated at room temperature for 30 minutes. The plate was washed four times x 300 μL/well with PBS-Tween wash buffer using an automatic plate washer. 100 μL/well of a species-specific
Horseradish peroxidase-labeled secondary antibody solution (Donkey anti-Rat IgG (H+L) HRP [Thermofisher] diluted 1:2000 in Zalgen Antibody Conjugate; Goat anti-Mouse IgG(γ), HRP [Seracare, MA, USA] diluted 1:2500 in Zalgen Antibody Conjugate) were placed on the wells and incubated at room temperature for 30 minutes. The 4 x 300 μL wash step was repeated. 100 μL/well of TMB Substrate was added to the plate and incubated at room temperature for ten minutes while protected from light. Finally, 100 μL/well of stop solution was added and the plate absorbance was read at 450nm.

4.1.3 Statistical Analysis – Antibody ELISA

To interpolate relative sample concentrations from the standard curve, a four-parameter (4PL) logistic regression model was created in GraphPad Prism software v.9.0 (Graphpad, CA, USA) as previously described (Boisen, Hartnett et al. 2018, Boisen, Uyigue et al. 2020). The initial concentration of the positive controls on the plate were considered to be 1000 arbitrary units (U)/mL. An ROC curve for 95% specificity was initially created and applied to negative control serum for both mouse (n = 40) and rat (n = 20) ELISAs. However, due to discrepancies between laboratory and wild animals, the number of animals available for ROC testing, and laboratory temperature and humidity differences between the United States and KGH, the ROC-established cutoff was determined to be inadequate. Instead, the negative cutoff was changed to that of twice the mean of the average negative sample run across the experimental plates. Cutoffs for the IgG ELISAs were as follows: Mouse GP: 28.2 U/mL (n = 6); Mouse NP: 45.9 U/mL (n = 6); Rat GP: 5.24 U/mL (n = 4); Rat NP: 1.41 U/mL (n = 4).
4.1.4 Development of a diagnostic algorithm to screen rodents for active LASV infection

Factor analysis of mixed data (FAMD) and hierarchical clustering of principal components (HCPC) were used to create a definition of an animal actively infected with LASV for use as a diagnostic screening tool. From the final dataset of all study specimens (n = 534), we pulled eight variables that described morphometric identifiers, biomarker status, antigen strength and number of qRT-PCR positive organs (Table 10). Primary categorical variables were the three biomarkers encoded as factors (i.e. `Ag+`, `Ag-`, `Ag not tested`). The variable `qPCR` was stratified based on the strength of the positive result. Morphometric information was encoded as supplementary categorical information. Antigen strength, as determined by RDT score, and the number of qPCR-positive organs were encoded as quantitative variables. Missing quantitative data were singularly imputed with the R package MissMDA (Josse and Husson 2016) (parameters: ncp = 10). The R package FactoMineR (Lê, Josse et al. 2008) was then used to perform FAMD and HCPC, while the R package factoextra (https://cran.r-project.org/package=factoextra) was used to extract and display the results of FAMD and HCPC. HCPC was cut in four clusters and was constructed using Euclidean distance and unweighted pair-means to increase cophenetic correlation between the clusters (Legendre, Legendre et al. 2012).
Table 10: Factors and variables used to compute FAMD and HCPC analyses.

<table>
<thead>
<tr>
<th>FAMD Factors and Variables</th>
<th>N = 534</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary categorical variables: biomarker status</strong></td>
<td></td>
</tr>
<tr>
<td>Ag +/-</td>
<td></td>
</tr>
<tr>
<td>Ag +</td>
<td>99</td>
</tr>
<tr>
<td>Ag -</td>
<td>424</td>
</tr>
<tr>
<td>Ag not tested</td>
<td>11</td>
</tr>
<tr>
<td>IgG +/-</td>
<td></td>
</tr>
<tr>
<td>IgG +</td>
<td>52</td>
</tr>
<tr>
<td>IgG -</td>
<td>323</td>
</tr>
<tr>
<td>IgG not tested</td>
<td>159</td>
</tr>
<tr>
<td><strong>qPCR status</strong></td>
<td></td>
</tr>
<tr>
<td>Strong qPCR +</td>
<td>14</td>
</tr>
<tr>
<td>Weak qPCR +</td>
<td>24</td>
</tr>
<tr>
<td>qPCR -</td>
<td>162</td>
</tr>
<tr>
<td>qPCR not tested</td>
<td>334</td>
</tr>
<tr>
<td><strong>Supplementary categorical variables: morphometrics</strong></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>288</td>
</tr>
<tr>
<td>M</td>
<td>246</td>
</tr>
<tr>
<td><strong>Genus</strong></td>
<td></td>
</tr>
<tr>
<td>Mastomys</td>
<td>300</td>
</tr>
<tr>
<td>Rattus</td>
<td>200</td>
</tr>
<tr>
<td>Praomys</td>
<td>18</td>
</tr>
<tr>
<td>Crocidura</td>
<td>8</td>
</tr>
<tr>
<td>Hylomyscus</td>
<td>7</td>
</tr>
<tr>
<td>Graphiurus</td>
<td>1</td>
</tr>
<tr>
<td><strong>Age proxy</strong></td>
<td></td>
</tr>
<tr>
<td>Juvenile</td>
<td>145</td>
</tr>
<tr>
<td>Sub-adult</td>
<td>177</td>
</tr>
<tr>
<td>Adult</td>
<td>212</td>
</tr>
<tr>
<td><strong>Quantitative variables</strong></td>
<td></td>
</tr>
<tr>
<td>Ag strength</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>0.59 (1.36)</td>
</tr>
<tr>
<td>Not tested</td>
<td>11</td>
</tr>
<tr>
<td># qPCR + Organs</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>0.37 (0.94)</td>
</tr>
<tr>
<td>Not tested</td>
<td>334</td>
</tr>
</tbody>
</table>

Table legend:
Select morphometric details and biomarker statuses were chosen to complete FAMD and HCPC analyses. There were three primary categorical variables, three supplementary categorical variables, and two quantitative variables. F = Female, M = Male.
4.2 Results

4.2.1 Modification of a commercially available LASV immunoassay

We adapted the Zalgen ReLASV Pan-Lassa Pre-fusion GP and NP ELISAs to detect rodent antibodies in place of human antibodies. In place of the anti-human HRP secondary antibody provided in the ELISA kits, we substituted the appropriate rodent secondary antibody. *Mastomys spp.*, *Praomys spp.*, and *Hylomyscus spp.* were tested using an anti-mouse secondary antibody. *Rattus spp.* were tested using an anti-rat secondary antibody. The single *Graphirurus spp.* specimen collected did not yield enough serum to perform both antigen and antibody testing. We replaced the human calibration standards with pooled serum from immunized rodents against either GP or NP specific to the assay, and replaced the negative serum control with pooled serum from immunologically naïve rodents. All other components provided in kit were used in accordance with the original assay specifications: sample diluent, secondary antibody conjugate solution, wash buffer, TMB substrate, and 2% methylsulfuric acid stop solution. The ELISA, with the aforenoted modifications, was performed in accordance with the original manufacturer’s instructions. IgG concentrations were calculated from a four-parameter standard curve using arbitrary units per mL of serum (U/mL).

4.2.2 IgG antibody presence by ELISA

Of the 543 specimens in the study, we were able to test 325 specimens for IgG by both GP and NP ELISA. A specimen was considered positive for IgG if it had a concentration of antibody above the cutoff specified for the assay for at least one of the proteins tested. 52 specimens (16%) were positive for IgG (Table 11).
Table 11: Morphometric characteristics of specimens grouped by IgG ELISA result.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>IgG -, N = 323 (^1)</th>
<th>IgG +, N = 52 (^1)</th>
<th>Overall, N = 375 (^1)</th>
<th>p-value(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genus</td>
<td></td>
<td></td>
<td></td>
<td>0.005</td>
</tr>
<tr>
<td>Mastomys</td>
<td>177 (81%)</td>
<td>42 (19%)</td>
<td>219 (100%)</td>
<td></td>
</tr>
<tr>
<td>Rattus</td>
<td>131 (93%)</td>
<td>10 (7.1%)</td>
<td>141 (100%)</td>
<td></td>
</tr>
<tr>
<td>Praomys</td>
<td>12 (100%)</td>
<td>0 (0%)</td>
<td>12 (100%)</td>
<td></td>
</tr>
<tr>
<td>Hylomyscus</td>
<td>3 (100%)</td>
<td>0 (0%)</td>
<td>3 (100%)</td>
<td></td>
</tr>
<tr>
<td>Crocidura</td>
<td>0 (NA%)</td>
<td>0 (NA%)</td>
<td>0 (NA%)</td>
<td></td>
</tr>
<tr>
<td>Graphiurus</td>
<td>0 (NA%)</td>
<td>0 (NA%)</td>
<td>0 (NA%)</td>
<td></td>
</tr>
<tr>
<td>Weight (Age proxy)</td>
<td></td>
<td></td>
<td></td>
<td>0.015</td>
</tr>
<tr>
<td>&gt; 40g (Adult)</td>
<td>136 (82%)</td>
<td>29 (18%)</td>
<td>165 (100%)</td>
<td></td>
</tr>
<tr>
<td>25-40g (Sub-Adult)</td>
<td>119 (86%)</td>
<td>20 (14%)</td>
<td>139 (100%)</td>
<td></td>
</tr>
<tr>
<td>&lt; 25g (Juvenile)</td>
<td>68 (96%)</td>
<td>3 (4.2%)</td>
<td>71 (100%)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td>0.368</td>
</tr>
<tr>
<td>F</td>
<td>174 (84%)</td>
<td>32 (16%)</td>
<td>206 (100%)</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>149 (88%)</td>
<td>20 (12%)</td>
<td>169 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) n (%)

\(^2\) Fisher's exact test

Counts and percentage of IgG ELISA test results in study population by genera, weight (a proxy for age of the specimen; described in (Leirs 1992) and sex. The table is visualized in Figure 7. IgG\(^{-}\) = antibody negative, IgG\(^{+}\) = antibody positive, F = Female, M = Male.
42 specimens (80.8%) that tested positive for IgG were identified as *Mastomys spp.*, and the remaining 10 specimens (19.2%) testing positive for IgG were identified as *Rattus spp.* None of the *Praomys spp.* or *Hylomyscus spp.* identified specimens tested were positive for IgG. 19% of the *Mastomys spp.* tested were positive for IgG, while 7.1% of the *Rattus spp.* tested were positive for IgG (p = 0.007 by Fisher’s Exact Test) (Figure 8A). There were no significant sex differences between specimens testing IgG positive and specimens testing IgG negative (p = 0.368 by Fisher’s Exact Test) (Figure 8B). Significant differences were seen in specimens by weight group, a proxy for the specimen age. Specimens weighing < 25g, a proxy for juvenile specimens, were less likely to be antibody positive than specimens weighing > 40g, a proxy for adult specimens (p = 0.019 by Fisher’s Exact Test) (Figure 8C).

### 4.2.3 Differences in GP and NP IgG antibody response

36 (69.3%) of the specimens tested positive for both GP and NP IgG, 9 (19.2%) of the specimens tested positive for GP only, and 7 (13.5%) tested positive for NP only (Figure 9A). Interestingly, of the 9 GP-only positive specimens, all but two of the specimens were identified as *Rattus spp.* Similarly, only one of the NP-only positive specimens was identified as *Rattus spp.* Among the 36 specimens that tested positive for both GP and NP, 30 (83.3%) had higher NP IgG concentrations than GP IgG concentrations (Figure 9A). Indeed, linear regression analysis indicated NP IgG concentrations were weakly correlated with GP IgG concentrations (R² = 0.272, p = 0.0009 by T-regression) (Figure 9D).
Figure 8: Traits of IgG antibody presence among study population.

A-C: Percentage of antibody positive/antibody negative specimens by characteristic. A: Genera (Pairwise Fisher’s Exact Test, ** = p < 0.01, ns = not significant), B: Sex (Fisher’s Exact Test, ns = not significant), C: Weight (a proxy for age of the specimen; described in (Leirs 1992) (Fisher’s Exact Test, * = p < 0.05, ns = not significant). IgG⁻ = antibody negative, IgG⁺ = antibody positive.
Figure 9: Differences between GP and NP IgG antibody responses.

A: Mean GP and NP IgG antibody levels of every tissue of every antibody positive specimen tested. A specimen was considered antibody positive if at least one IgG ELISA (GP or NP) had a positive result. B-C: Percentage of specimens with higher mean levels of GP IgG/only GP IgG positive vs those with higher mean levels of NP IgG/only NP IgG positive by genus (B) and antigen status (C) (Fisher’s Exact Test, **** = p > 0.0001, ns = not significant). D: Mean levels of GP IgG antibody vs mean levels of NP IgG antibody (Linear regression, p > 0.0009 by T-regression). E: Mean levels of GP IgG vs NP IgG across all specimens, grouped by genus (Unpaired two-tailed Student’s t-test, ** = p < 0.01, ns = not significant).
In addition, the mean NP IgG concentration of all specimens that tested NP positive was found to be higher than the mean GP IgG concentration of all *Mastomys spp.* specimens that tested GP positive, though the trend was not statistically significant in *Rattus spp.* specimens (*Mastomys spp.:* p = 0.001, *Rattus spp.:* p = 0.689; both by Fisher’s Exact Test) (Figure 9E). The only significant, observable difference between the specimens with GP IgG concentrations higher than NP IgG concentrations, including specimens that tested positive for GP only, was that the higher GP IgG concentration specimens were more likely to be *Rattus spp.* than *Mastomys spp.* (p < 0.00001 by Fisher’s Exact Test) (Figure 9B). Specimens with higher GP IgG concentrations than NP IgG concentrations were also less likely to test antigen positive by RDT, but the difference was not statistically significant (p = 0.069 by Fisher’s Exact Test) (Figure 9C).

4.2.4 Correlation between antigen status or qRT-PCR status and antibody status

Simultaneous presence of LASV antigen and LASV IgG antibody are strongly linked in the sample population, where 21 of the 373 specimens that tested positive for IgG also tested positive for antigen (p < 0.0001 by Fisher’s Exact Test) (Table 12, Figure 10A). However, no correlation was observed between antigen strength and antibody concentration, by both mean antibody concentration of GP and NP together (R² = 0.0009 by linear regression) and individual protein antibody concentration (GP R² = 0.0001, NP R² = 0.0032; both by linear regression) (Figure 10B, 10C). Simultaneous presence of LASV detected by qRT-PCR and LASV IgG antibody was also observed in 9 specimens where both tests were performed, though the link was not statistically significant (p = 0.128 by Fisher’s Exact Test when including tested serum, p = 1 when not including testing serum) (Table 13A, 13B; Figure 11A, 11B).
Table 12: Frequency table of antigen and IgG antibody status and its associated sensitivity and specificity.

<table>
<thead>
<tr>
<th></th>
<th>IgG -</th>
<th>IgG +</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag -</td>
<td>268</td>
<td>31</td>
<td>299</td>
</tr>
<tr>
<td>Ag +</td>
<td>53</td>
<td>21</td>
<td>74</td>
</tr>
<tr>
<td>Column Total</td>
<td>321</td>
<td>52</td>
<td>373</td>
</tr>
</tbody>
</table>

p = <0.0001 by Fisher’s Exact Test

<table>
<thead>
<tr>
<th></th>
<th>est</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>83.5%</td>
<td>79.0%-87.4%</td>
</tr>
<tr>
<td>Specificity</td>
<td>40.4%</td>
<td>27.0%-54.9%</td>
</tr>
</tbody>
</table>
Figure 10: Correlations between antibody presence and antigen status.

A: Visualization of frequency table from Table 12 (Fisher’s Exact Test, *** = p < 0.001).

B: Mean antibody level, averaged from GP and NP IgG, vs antigen strength as indicated by RDT score (Linear Regression).

C: Mean antibody level vs antigen strength by antibody protein (Linear Regression).
Table 13: qRT-PCR vs IgG antibody status.

<table>
<thead>
<tr>
<th></th>
<th>IgG -</th>
<th>IgG +</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>qPCR -</td>
<td>95</td>
<td>18</td>
<td>113</td>
</tr>
<tr>
<td>qPCR +</td>
<td>23</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td>Column Total</td>
<td>118</td>
<td>27</td>
<td>145</td>
</tr>
<tr>
<td>p = 0.128 by Fisher’s Exact Test</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A: Frequency table of qRT-PCR (qPCR) status vs IgG antibody status of all tissues tested, including serum.

<table>
<thead>
<tr>
<th></th>
<th>IgG -</th>
<th>IgG +</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>qPCR -</td>
<td>28</td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td>qPCR +</td>
<td>23</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td>Column Total</td>
<td>51</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>p = 1 by Fisher’s Exact Test</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B: Frequency table of qRT-PCR status vs IgG antibody status of all tissues tested excluding serum.

C: Table of organs testing qRT-PCR positive in IgG antibody positive specimens.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Negative, N = 18</th>
<th>Positive, N = 9</th>
<th>Overall, N = 27</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum</td>
<td>0 / 14 (0%)</td>
<td>0 / 5 (0%)</td>
<td>0 / 19 (0%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Liver</td>
<td>0 / 8 (0%)</td>
<td>5 / 9 (56%)</td>
<td>5 / 17 (29%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Lung</td>
<td>0 / 10 (0%)</td>
<td>9 / 9 (100%)</td>
<td>9 / 19 (47%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Kidney</td>
<td>0 / 10 (0%)</td>
<td>3 / 6 (50%)</td>
<td>3 / 16 (19%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>8</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Spleen</td>
<td>0 / 9 (0%)</td>
<td>5 / 7 (71%)</td>
<td>5 / 16 (31%)</td>
</tr>
<tr>
<td>Not tested</td>
<td>9</td>
<td>2</td>
<td>11</td>
</tr>
</tbody>
</table>

\(^1 n / N (%)\)
Figure 11: Correlations between antibody presence and viral status by qRT-PCR.

A-B: Visualization of frequency tables from Table 13 for (A) all specimens tested by qRT-PCR, including serum, (B) all specimens tested by qRT-PCR not including serum (Fisher’s Exact Test, ns = not significant)., C: Mean viral level averaged across all tissues of an individual specimen vs mean antibody level averaged from GP and NP IgG (Linear Regression)., D: Mean viral level vs mean antibody level averaged from GP and NP IgG, by tissue (Linear Regression).
Virus was detected by qPCR in all the lung samples (9/9) from the qPCR+/IgG+ specimens (Table 13C). 5/7 spleen samples, 5/9 lung samples, and 3/6 kidney samples had detectable viral levels by qPCR, though there was no significant correlation between viral levels and serum antibody concentration for any tissue (p = 0.433 by Kruskal-Wallis Test) (Figure 11C, 11D).

4.2.5 Contributions of biomarkers to variance within the study population

Factor analysis of mixed data (FAMD) was applied to evaluate how select factors contribute to the variance among the study population. FAMD was applied to all specimens in the study population (n = 534), which resulted in five dimensions across eight factors (Table 10, Figure 12A): three primary categorical variables (qPCR status, IgG status, Ag status), three supplementary categorical variables (Sex, Genus, Age), and two quantitative variables (number of qPCR+ organs, RDT score). The first dimension alone accounted for 31.4% of the variance within the study population and was driven primarily by qPCR status and antigen status (Figure 12A-12C). IgG antibody status did not begin to significantly contribute to variance of the dataset until the third dimension (Figure 12D).

Individual specimens oriented by their individual position within the first two dimensions of the FAMD matrix separate into distinct groups (Figure 12E). Hierarchical clustering of principal components (HCPC) performed from the FAMD matrix of the top two dimensions grouped the individual specimens into four clusters (Figure 12G).
Figure 12: FAMD and HCPC analysis to assess variance within study population.

A: Scree plot of FAMD results indicating the number of dimensions needed to explain the variance among the selected factors of the dataset., B: Factor map of all variables contributing to the variance of the first two dimensions., C: Correlation plot of the weights of quantitative variables to the variance of the first two dimensions., D: Factor map of all variables contributing to the variance of the second and third dimensions., E: Grouping of specimens in the dataset as explained by variance of the first two dimensions of FAMD., F: Specimens in the dataset grouped according to qRT-PCR (qPCR) status in the first two dimensions of FAMD., G: Clusters of specimens generated by HCPC analysis of the first two dimensions of FAMD.
Likewise, similar size clusters were obtained when individual specimens within the first two dimensions of the FAMD matrix were grouped by qPCR status alone (Figure 12F). Based off the results of the HCPC analysis and specimen grouping by qPCR status, we propose cluster 1 of HCPC, which corresponds to the cluster of “strong PCR positive” specimens to be actively infected with LASV.

4.3 Discussion

4.3.1 Modification of a commercially available LASV immunoassay

Commercial and in-house LASV IgG and IgM ELISAs have been developed for human research and diagnostic use (Bausch, Rollin et al. 2000, Emmerich, Thome-Bolduan et al. 2006, Boisen, Hartnett et al. 2018, Gabriel, Adomeh et al. 2018, Boisen, Uyigue et al. 2020). To that end, we adapted the Zalgen ReLASV Pre-fusion GP and NP ELISA kits for IgG and IgM antibody detection to detect LASV antibodies in rodents with minor modifications to the kits. At the outset, we adapted the ELISA kits to test for both IgG and IgM antibodies in rodents. However, immunized control specimens had minimal IgM responses to injected antigen and required ~100x more serum to obtain a signal response equivalent to that of IgG controls. Furthermore, initial testing of IgM ELISAs indicated significant background and non-specific binding even with increased amounts of blocking reagents in the sample diluent and secondary antibody conjugate. Therefore, we focused specifically on IgG antibody responses.

4.3.2 IgG antibody presence by ELISA

Of the 375 specimens examined for LASV GP and NP IgG antibodies, 52 tested positive for at least one of the two proteins for an overall positivity rate of 13.9% (Table 11). Only Mastomys spp. and Rattus spp. specimens tested IgG positive. Notably, antibody
presence was more often observed in larger, older specimens than smaller, younger specimens (Table 11, Figure 8C). This trend indicates horizontal transmission, with increasing exposure to the virus over the lifespan of the animal (Fichet-Calvet, Becker-Ziaja et al. 2014). However, we were unable to test 50% of the juvenile specimens due to low blood volumes upon collection. In contrast, we tested 77.8% of adult specimens for IgG antibodies (Table 11). The lack of serum samples available from juvenile specimens may bias the trend of increased LASV exposure in adults. Future trapping studies should ensure collection of DBS and/or serum samples from an even distribution of specimen ages sampled to better draw conclusions of IgG presence as related to specimen age.

We did not detect IgG in *Hylomyscus* spp. or *Praomys* spp. specimens. None of the seven *Hylomyscus* spp. specimens caught tested antigen or antibody positive (Table 7, Table 11). Five *Hylomyscus* spp. specimens typed to the species level by cytochrome *b* PCR were identified as *H. simius* (Table 4). A different *Hylomyscus* spp. member, *H. pamfi*, has been documented as a novel LASV reservoir in Nigeria (both IgG+ and qPCR+) and a host of a related LASV-like virus in Cote d’Ivoire (Coulibaly-N’Golo, Allali et al. 2011, Olayemi, Cadar et al. 2016). 12 *Praomys* spp. specimens typed to the species level by cytochrome *b* PCR were identified as *P. rostarus* (Table 4). Indeed, *Praomys* spp., specifically *P. daltoni*, have tested positive for LASV antibody (Olayemi, Oyeyiola et al. 2018). Although an equal proportion of *Praomys* spp. specimens and *Mastomys* spp. specimens tested positive for LASV antigen, none of the *Praomys* spp. specimens examined tested positive for LASV antibody. Given that studies have used anti-mouse IgG secondary antibodies to detect LASV IgG in both *Hylomyscus* spp. and *Praomys* spp., it is unlikely that we did not detect IgG antibodies in these species due to non-reactivity of the
anti-mouse IgG secondary antibody. Rather, it is more plausible that LASV has distinct ecological niches that restricts both competent host replication and interactions between primary LASV reservoirs and other rodent species.

4.3.3 Differences in GP and NP IgG antibody response

We also wanted to observe and discern any differences between GP IgG and NP IgG response. Most specimens that tested antibody positive tested positive for both GP and NP IgG, and the NP IgG response was stronger than the GP IgG response (Figure 9A). However, a handful of specimens either tested IgG positive for GP only or had a stronger GP IgG response than NP IgG response. As NP is the more immunogenic of the two proteins, we wanted to see in what ways specimens with the single or stronger GP IgG response (GP > NP) were different from the other antibody positive specimens. Indeed, GP > NP specimens were more likely to be *Rattus spp.* and less likely to be antigen positive, although the latter was not statistically significant (p = < 0.00001 and p = 0.129 respectively by Fisher’s Exact Test) (Figure 9B, 9C). Likewise, no *Rattus spp.* specimens tested both antigen positive and antibody positive and tended to have a weaker antibody response than *Mastomys spp.* (Table 7, Table 11, Figure 9D). The high number of *Rattus spp.* specimens in the GP > NP population likely contributed to the decreased likelihood of antigen presence in the GP > NP population. These results could be expected, given differential immunological responses between mice and rats have been noted for numerous other viruses (Lebrec and Burleson 1994, Bonthius and Perlman 2007, Debing, Mishra et al. 2016, Boukhalova, Yim et al. 2018). Although not discernable from this study, it is likely that *Rattus spp.* transmission routes, viral shedding, and immunological responses to LASV infection differ from that of *Mastomys spp.* and other Praomyini tribe members.
4.3.4 Correlation between antigen status or qRT-PCR status and antibody status

A linkage between antigen and antibody presence is expected in a reservoir host if there is a sustained viral presence to continually produce viral antigen (Walker, Wulff et al. 1975, Walker, McCormick et al. 1982). Antigen may also be present towards the end of or after an infection once IgG antibodies begin to be produced. To that end, we observed a strong link between simultaneous antigen and antibody presence in Mastomys spp. specimens compared to only antigen presence or only antibody presence (p = < 0.001 by Fisher’s Exact Test) (Table 12, Figure 10A). However, antigen strength and antibody strength were not correlated in Ag+/IgG+ specimens (Figure 10B, 10C). Interpretation of these results makes sense in the context of capturing specimens at various states of infection at a single time point. Yet, our results contrast with the findings of Demby et. al in Guinea, which pointed to vertical transmission as the primary transmission modality (Demby, Inapogui et al. 2001). His research team demonstrated that antigen and antibody presence in Mastomys spp. specimens were mutually exclusive. Nonetheless, the strong link between simultaneous antigen and antibody presence in adult and sub-adult specimens indicates at least some horizontal transmission within the primary host reservoir.

We did not see any statistically significant linkage between simultaneous viral presence detected by qRT-PCR and antibody presence (Table 13A, 13B; Figure 11A, 11B), as observed in other mammarenavirus trapping studies (Demby, Inapogui et al. 2001, Borremans, Leirs et al. 2011, Marien, Borremans et al. 2017). Despite the lack of correlation between viral presence and antibody presence, all qPCR+/IgG+ specimens had detectable viral levels in lung tissue, while only half of the qPCR+/IgG+ specimens had detectable viral levels in kidney tissue (Table 13C). However, no correlation was observed...
between tissue viral levels and serum antibody concentration (Figure 11C). The dearth of qPCR+/IgG+ specimens relative to the study population further points to horizontal transmission in the host reservoir. Likewise, laboratory colonies of mice vertically infected with LCMV have lifelong viremia with little to no antibody response (Oldstone 2002). Nevertheless, the concurrent presence of virus and antibody provides insight into viral replication patterns, where both horizontal and vertical transmission play a role to maintain LASV within the host reservoir(s).

4.3.5 Contributions of biomarkers to variance within the study population

Factor analysis of mixed data (FAMD) and hierarchical clustering of principal components (HCPC) were performed to algorithmically classify specimens with active infections from the other study specimens. We specifically chose to perform FAMD as it combines principal component analysis of quantitative data with multiple correspondence analysis of categorical qualitative data. Although we did not have enough specimens with all three laboratory tests (antigen, qPCR, and antibody) performed to directly assess respective quantities of viral levels and antibody levels, FAMD allowed us to assess antigen strength and the number of organs testing qPCR-positive as continuous variables rather than discrete factor levels to strengthen the analysis. The variance in the study population, as described by eight variables listed in Table 10, could be explained in five dimensions (Figure 12A). Furthermore, the first dimension of FAMD described 31.4% of the variance in the study population (Figure 12A, 12B). qPCR status, number of organs testing qPCR-positive, antigen strength, and antigen status contributed the most to the variance of the first dimensions (Figure 12B, 12C). Interestingly, antibody status did not play a significant role in the variance of the first two dimensions, in line with the previous
observation of little linkage between qPCR status and antibody status (Figures 12D and 11A).

We then performed hierarchical clustering of principal components (HCPC) to predict grouped clusters of specimens from the interpretation of FAMD. HCPC classified the specimens into four distinct clusters, which roughly corresponded to those classified strictly by qPCR status (Figure 12G, 12F). Indeed, cluster 4 of the HCPC model and the “Strong qPCR++” cluster of FAMD classification are identical to one another. Furthermore, all the “strong qPCR++” specimens tested qPCR+ in two or more organs and were antigen positive (Figure 6D). Thus, we propose the definition of an actively infected specimen as a specimen that tests antigen positive and has a viral level greater than 200 copies/μL RNA (a “Strong qPCR-positive”) as detected by qRT-PCR in at least two tissues.

Overall, the Zalgen ReLASV Pan-Lassa Antigen Rapid Test serves as a useful tool to help prioritize specimens for viral detection in tissues by qRT-PCR. The high sensitivity of the RDT as compared to qRT-PCR (100%) compensates for its low specificity (41%) in detecting qPCR-positive specimens. With respect the proposed diagnostic criteria, presence of IgG does not necessarily preclude an active infection in a specimen. In fact, there was no significant difference in the numbers of IgG+ actively infected specimens and IgG- actively infected specimens. The ability to quickly assess captured specimens for active LASV infections will help reduce the amount of time and resources needed to screen large numbers of samples, which will allow more detailed downstream analyses to be performed sooner.
5. VIRAL SEQUENCING OF ACTIVELY INFECTED RODENT SPECIMENS AND PHYLOGEOGRAPHIC ANALYSIS OF LASV DIVERSITY WITHIN SIERRA LEONE IN HUMANS AND RODENTS *

Performed in conjunction with Refugio Robles-Sikisaka, Ph.D. and Raphaëlle Klitting, Ph.D in the lab of Kristian Andersen, Ph.D. at Scripps Research in San Diego, CA

5.1 Materials and Methods

5.1.1 Library generation and sequencing

Tissue samples from actively infected rodents were sent to Scripps Research for viral sequencing. Double stranded cDNA was prepared from RNA in tissue samples using randomly primed reverse transcription with the SuperScript® IV First-Strand cDNA Synthesis kit (Thermofisher, MA, USA) and the NEBNext Ultra II Non-directional RNA Second Strand Synthesis Module (New England Biolabs, MA, USA). The cDNAs were purified with AMPureXP beads (BeckmanCoulter, IN, USA) and quantified using the Qubit High Sensitivity (HS) DNA assay kit (Thermofisher). Sequencing libraries were generated with the SureSelect XT HS2 DNA Library Preparation Kit (Agilent, CA, USA). Libraries were quantified using the Qubit HS DNA kit and Tapestation D5000 tape (Agilent) then pooled using similar amounts of DNA based on input RNA concentration that differed by no more than 3 Cycle threshold (Ct) units as determined by qRT-PCR screening. Hybrid-capture was then performed on the libraries to enrich the viral content of sequencing libraries. The hybrid-capture panel consists of 120-mer oligonucleotides tiling the whole LASV genome, with one probe every 30 bp. The panel was generated based on a set of 14 LASV genomes, including one sequence for each of the LASV lineages
described to date (I to VII). Pooled libraries were added to the hybrid-capture baits and hybridized over 90 minutes. After capture, libraries were amplified, washed, quantified, normalized and pooled in equimolar amounts at 2 nM before sequencing on the MiSeq platform.

5.1.2 Assembly of full-length LASV genomes

Low quality reads and bases, and Illumina adapters were removed with Trimmomatic v0.39 (Bolger, Lohse et al. 2014) (parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:4:25:7:1:keepBothReads SLIDINGWINDOW:5:20 MINLEN:50). All reads were de novo assembled using SPAdes v3.15.2 (Prjibelski, Antipov et al. 2020) (parameters: -t 2 -k 55,77,127 --only-assembler --meta). Contigs generated from the assembly were filtered to remove low complexity sequences using PRINSEQ v0.20.4 (Schmieder and Edwards 2011). PROmer v3.0 (Delcher, Phillippy et al. 2002) was used to scaffold contigs based on the closest available full LASV segment identified with BLASTn from a database of full LASV S and L segments. All sequencing reads from individual samples were aligned back to its own LASV scaffold using NextGenMap (Sedlazeck, Rescheneder et al. 2013) and a consensus sequence was called using iVar (Grubaugh, Gangavarapu et al. 2019). Regions where contig coverage was less than 10x were called as ‘N’. A reference-based alignment was also generated for each of the generated sequences by aligning trimmed reads to the closest available full LASV segment, calling consensus using iVar then re-aligning reads to this consensus. Finally, consensus sequences resulting from either the de novo assembly or the reference alignment were manually inspected, checked and compared to control sequences for accuracy and the presence of intact open reading frames (ORFs) using Aliview (Larsson 2014). The most complete and accurate consensus was
kept as the final consensus. We generated a total of 20 full and 1 partial (> 50% of the genome covered) S segments, and 9 full and 11 partial L segments from 15 different rodent specimens. Select specimens had multiple tissue samples sequenced for future downstream analysis of intra-host variants.

5.1.3 Sequence curation and alignment

All publicly available sequences for LASV were downloaded from the NCBI Nucleotide database (https://www.ncbi.nlm.nih.gov/nuccore) (keywords: “lassa NOT mopeia NOT natalensis”). The sequences downloaded from NCBI were then combined with recently generated sequences from Sierra Leone from the Center for Viral Systems Biology consortium (https://cvisb.org/data/) that have been prepared and assembled as described here. From the downloaded sequences, we excluded i) sequences without a timestamp, ii) sequences from laboratory strains (i.e., adapted for tissue culture or passaged multiple times, recombinant viruses, or strains obtained from antiviral/vaccine experiments), iii) sequences from identified hospital epidemics or sequences for which the location corresponded to the site of hospitalization, and iv) sequences where no GPS coordinates were provided. If multiple time points from a single individual patient were available, only a single time point was kept. Likewise, duplicate sequences were removed when more than one sequence was available for a single viral strain. The remaining downloaded sequences, together with the sequences generated in this study, were manually trimmed to their coding regions and arranged in sense orientation both for the S segment (NP-NNN-GPC) and the L segment (L-NNN-Z) using AliView. Multiple sequence alignment was then performed with MAFFT v1.7 (Katoh and Standley 2013). Upon manual inspection of the clean and arranged alignments, low quality sequences and very
short sequences that had a combined ORF length of < 500nt were discarded. The final alignments with all curated sequences included 89 S segment sequences and 61 L segment sequences.

5.1.4 Phylogenetic analysis and phylogeographic analysis

GPS coordinates of the curated alignments were examined to assign sequences to one of five geographic zones: Three in Kenema district (K1-K3), Bo district (B), and Makeni (M), the capital of the Bombali district. Past transitions between zones K1-3, B and M were inferred using the TreeTime migration model (Sagulenko, Puller et al. 2018) based on a maximum-likelihood tree reconstructed using IQ-TREE v1.6.10 (Nguyen, Schmidt et al. 2014) (Best-fit model: GTR+F+I+G4 chosen according to Bayesian Information Criterion). Spatially-explicit phylogeographic reconstructions were performed using the relaxed random walk (RRW) diffusion model (Lemey, Rambaut et al. 2010) implemented in BEAST v1.10 (Suchard, Lemey et al. 2018), which was coupled with the BEAGLE 3 library (Ayres, Cummings et al. 2019) to improve computational performance. The nucleotide substitution process was modeled according to a GTR+Γ parameterization (https://agris.fao.org/agris-search/search.do?recordID=US201301755037) and branch-specific evolutionary rates according to a relaxed molecular clock with an underlying log-normal distribution. A distinct BEAST analysis was run for each segment (S and L) in which we sampled Markov chain Monte-Carlo (MCMC) chains every $10^4$ generations. Tracer 1.7 (Rambaut, Drummond et al. 2018) was used to identify the number of sampled trees to discard as burn-in, inspect model convergence, and ensure estimated sampling size values associated with estimated parameters were all $> 200$. TreeAnnotator 1.10 (Suchard,
Lemey et al. 2018) was used to obtain a maximum clade credibility (MCC) tree for each individual BEAST analysis of the S and L segments.

5.2 Results

5.2.1 Sequences obtained from tested specimens

We attempted to sequence LASV from select lung, kidney, and spleen tissue samples for 21 unique specimens: 12 of the 13 actively infected specimens identified in Section 4.2.5, 7 specimens with a strong antigen presence as assessed by RDT score but low viral level as detected by qRT-PCR, 1 specimen with antigen presence but no virus detected, 1 specimen with no antigen presence and no virus detected, and 1 specimen with no antigen presence and where viral levels were not assessed. We successfully obtained at least one genome segment from all 12 actively infectious specimens sent for sequencing (Figure 13A). Furthermore, none of the ten additional specimens sent for sequencing yielded any LASV sequence fragments (Figure 13A).

From the 12 actively infected specimens in which LASV sequencing was attempted, we obtained a total of 41 LASV genome sequences from 18 different tissue samples: 20 full S genome segments and 1 partial S genome segment, and 9 full L genome segments and 11 partial L genome segments (Table 14). LASV genome sequences were most frequently obtained from lung tissue and kidney tissue, though three genome sequences were obtained from spleen tissue. Viral sequencing was not attempted with liver tissue due to the high concentration of sequencing inhibitors present within the RNA samples. L genome contig coverage at the 3’ end of the genome tended to be lower than at both the 5’ end of the L genome segment and the entirety of the S genome segment.
coverage (Figure 13B). However, > 100x genome coverage was obtained for 18/20 of the full S genome segments, suitable for iSNV analyses (data not shown).
Figure 13: Viral levels of specimens attempted for sequencing and genome coverage of curated LASV sequences.

A: Mean viral levels of individual specimens where LASV sequencing was attempted. B: LASV genome coverage. Coverage above 10x was achieved for 100% of the S and 99% of the L segment, and coverage above 100x was achieved (shown by a red line required for iSNV analysis) for 98% of the S and 92% of the L segment.
Table 14: LASV genome sequences obtained from actively infected specimens.

<table>
<thead>
<tr>
<th>Animal ID</th>
<th>Tissue</th>
<th>S Segment</th>
<th>L Segment</th>
<th>Mean viral copies/ µL RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>Lung</td>
<td>Full</td>
<td>Full</td>
<td>$1.73 \times 10^5$</td>
</tr>
<tr>
<td>33</td>
<td>Kidney</td>
<td>Full</td>
<td>Full</td>
<td>$4.39 \times 10^5$</td>
</tr>
<tr>
<td>117</td>
<td>Lung</td>
<td>Full</td>
<td>Full</td>
<td>$1.02 \times 10^5$</td>
</tr>
<tr>
<td>117</td>
<td>Kidney</td>
<td>Full</td>
<td>Full</td>
<td>$2.83 \times 10^5$</td>
</tr>
<tr>
<td>226</td>
<td>AVL</td>
<td>Full</td>
<td>Full</td>
<td>$1.25 \times 10^4$</td>
</tr>
<tr>
<td>226</td>
<td>Kidney</td>
<td>Full</td>
<td>Full</td>
<td>$2.07 \times 10^4$</td>
</tr>
<tr>
<td>245</td>
<td>Lung</td>
<td>Full</td>
<td>Partial</td>
<td>$3.20 \times 10^5$</td>
</tr>
<tr>
<td>292</td>
<td>Lung</td>
<td>Full</td>
<td>Partial</td>
<td>$1.99 \times 10^5$</td>
</tr>
<tr>
<td>292</td>
<td>Kidney</td>
<td>Full</td>
<td>Partial</td>
<td>$1.45 \times 10^5$</td>
</tr>
<tr>
<td>303</td>
<td>Spleen</td>
<td>Full</td>
<td>Partial</td>
<td>$7.53 \times 10^3$</td>
</tr>
<tr>
<td>313</td>
<td>Lung</td>
<td>Full</td>
<td>Partial</td>
<td>$1.00 \times 10^5$</td>
</tr>
<tr>
<td>318</td>
<td>Lung</td>
<td>Full</td>
<td>Partial</td>
<td>$8.87 \times 10^4$</td>
</tr>
<tr>
<td>407</td>
<td>Lung</td>
<td>Full</td>
<td>Partial</td>
<td>$1.69 \times 10^5$</td>
</tr>
<tr>
<td>407</td>
<td>Kidney</td>
<td>Full</td>
<td>Full</td>
<td>$8.86 \times 10^4$</td>
</tr>
<tr>
<td>408</td>
<td>Spleen</td>
<td>Partial</td>
<td>None</td>
<td>$1.06 \times 10^4$</td>
</tr>
<tr>
<td>488</td>
<td>Lung</td>
<td>Full</td>
<td>Partial</td>
<td>$1.07 \times 10^5$</td>
</tr>
<tr>
<td>488</td>
<td>Spleen</td>
<td>Full</td>
<td>Partial</td>
<td>$1.92 \times 10^3$</td>
</tr>
<tr>
<td>504</td>
<td>Lung</td>
<td>Full</td>
<td>Partial</td>
<td>$2.30 \times 10^5$</td>
</tr>
<tr>
<td>511</td>
<td>Kidney</td>
<td>Full</td>
<td>Partial</td>
<td>$1.46 \times 10^5$</td>
</tr>
</tbody>
</table>
5.2.2 Phylogenetic and phylogeographic trees of sequenced LASV specimens

Full-length sequences obtained from the rodents were compared to all other known full-length sequences from both humans and rodents in Sierra Leone to evaluate changes in LASV evolutionary history. We created phylogenetic trees of the S and L segments using IQ-TREE to visualize the genetic relatedness of the newly sequenced rodent sequences compared to other known LASV sequences. Surprisingly, the most evolutionarily distant LASV genome originated from one of the newly sequenced rodent specimens in both the S and L segments (Figure 14A, 14B). Both S and L segments show general grouping into four different clusters.

To see if the four clusters identified in the general phylogenetic analysis were associated with geographic locations, we re-created the phylogenetic trees with migration models, which are mutation models designed to analyze migration patterns given known geographic sampling locations and time of collection. For ease of analysis, sequences were identified as being part of one of five discrete zone clusters: three clusters within the Kenema district (K1-K3), one cluster for the Bo district (B), and one cluster for Makeni (M). The reconstructed trees showed that sequences tended to cluster roughly by geographic zone (Figure 15A-C). The one human sample from Makeni clustered with specimens from geographic zone K2, the heart of the Kenema district. Sequences from the K3 geographic zone also with the K2 geographic zone. Sequences from the K1 geographic zone and the B zone largely remained grouped together, particularly with the L segment.
Figure 14: Phylogenetic trees of S and L LASV segments.

A
A: S segment phylogenetic tree of all known full-length S sequences in Sierra Leone, curated as described in 5.1.3, combined with full-length and partial-length S segment sequences from obtained from rodent specimens in this study. Distinct clusters are encased in blue boxes. Old Mastomys = Sequences from *Mastomys spp.* previously collected. New Mastomys = Sequences from *Mastomys spp.* collected for this study. Tips are annotated with the date (MM/YYYY) of sample collection. B: Same as A but for L segment.
Figure 15: Phylogeographic analysis of LASV in Sierra Leone.
Figure 15 (continued): Phylogeographic analysis of LASV in Sierra Leone.
Figure 15 (continued): Phylogeographic analysis of LASV in Sierra Leone.

A-B: Phylogenetic trees colored by geographic zone A = S Segment, B = L Segment. C:
Map of respective geographic locations within Sierra Leone.
5.3 Discussion

5.3.1 Sequences obtained from tested specimens

LASV genome sequences were obtained from all 12 of the actively infected specimens selected for viral sequencing (Table 14, Figure 13A), which indicates that the proposed definition of an actively infected specimen is accurate. To further verify our proposed definition of an actively infected specimen, we chose to attempt to sequence an additional seven specimens that tested positive for both antigen presence and viral presence, but which viral presence was low (< 200 copies/μL RNA). Indeed, none of those seven specimens yielded LASV sequences (Figure 13A). Moreover, LASV sequences from three other specimens that had neither antigen presence nor viral presence were unable to be obtained. Although specimens without antigen presence remain to be examined for viral presence, we do not expect those specimens to yield observable quantities of virus by qRT-PCR such that the virus can be investigated in downstream analyses.

On the whole, depth of contig coverage on the 3’ end of the L genome segment tended to be lower than the depth of contig coverage on both the 5’ end of the L genome segment and the entirety of the S genome segment (Figure 13B). As the 3’ end of the L segment is highly variable even within lineages (Leski, Stockelman et al. 2015), the relative lack of coverage is likely due to less-than-optimal oligonucleotide probe hybridization during the hybrid capture step. The less-than-optimal binding of the oligonucleotide hybridization probes to the 3’ end of the L segment could also explain why more partial L genome segments were obtained than full L genome segments.
5.3.2 Phylogenetic and phylogeographic trees analysis of sequenced LASV specimens

Full-length sequences obtained from trapped rodents as part of this study were compared to all other known full-length sequences from both humans and rodents in Sierra Leone to evaluate changes in LASV evolutionary history. The close relation of rodent sequences sampled at the same time indicates rodent-to-rodent transmission, as expected. While paired rodent-human specimens are exceedingly difficult to amass, further analyses of these transmission chains could help to elucidate potential bottlenecks in rodent-human transmission.

The most evolutionarily distant LASV genome originated from one of the newly sequenced rodent specimens in both LASV genome segments (Figure 14A, 14B). The most closely related specimen to these two rodent specimens dates to 1976. Concerted trapping efforts should take place where these specimens were collected to further refine this evolutionary node of LASV within Sierra Leone.

Phylogenetic trees reconstructed and grouped by geographic zones showed that sequences tended to cluster roughly by geographic zone (Figure 15A-C). Even within zones, there is still a high degree of sequence variability. LASV, and mammarenaviruses in general, exhibit a high degree of variability within members of the same species (Sevilla and de la Torre 2006, Leski, Stockelman et al. 2015). Furthermore, the rodent host reservoir reproduces multiple times over its short lifespan, which could allow for the virus to accumulate mutations in many unique specimens. This reproductive strategy, coupled over generations, could produce many variants of similarly mutated viruses. Thus, it is most
likely that LASV evolves within unique geographic confines with potential clades being restricted by movement potential of the host reservoir.
6. CONCLUSIONS AND FUTURE DIRECTIONS

Individuals living in endemic hotspots of Lassa fever have recurrent exposure to LASV via spillover from the primary host reservoir *Mastomys natalensis*. As the Kenema district of Sierra Leone has the highest historical incidence of Lassa fever in the world, the Lassa Fever ward and Lassa fever research program at KGH is uniquely poised to diagnose and treat Lassa fever cases as well as execute critical laboratory studies of Lassa virus diagnostics. Additional Lassa fever outreach, epidemiology and survivor studies, and ecology studies of LASV are performed under the umbrella of the Lassa fever research program with dedicated programs for each of the respective units.

When a Lassa fever patient admits to the Lassa Fever ward at KGH, the KGH outreach and ecology teams travel to the home village of the patient to initiate a case investigation into the potential cause and source of the original LASV infection. To that end, the ecology team routinely performs a case-control study where traps are set in the home of the admitted patient and two or three homes adjacent to the patient’s home. However, the team previously had no way to determine whether captured rodents from the study were infected with LASV; traditional detection methods of LASV require specialized laboratory equipment that cannot be transported to the field. Deployment of a rapid diagnostic test on collected rodent serum at the time of necropsy in the field allows the ecology team to quickly assess captured rodents for potential presence of LASV.

To that end, we performed a cross-sectional serological survey of small mammals in the Kenema district of Sierra Leone, in part to assess the prevalence of LASV within the
endemic rodent population but also to determine the usefulness of the Zalgen ReLASV Pan-Lassa RDT in detection of active LASV infection in rodents. Despite low admissions to the Lassa Fever ward at KGH in recent years, we show evidence of current or previous LASV infection in 15-20% of specimens captured as part of our study. Indeed, we have seen evidence of LASV infection by antigen presence in four different species: *M. natalensis*, *M. erythroleucus*, *P. rostarus*, and *R. rattus* specimens. Furthermore, the ratio of antigen presence in *P. rostarus* relative to the number of specimens caught is equal to the ratio of antigen presence observed in *M. natalensis* specimens, which suggests the possibility of a novel host reservoir. In contrast, *R. rattus* specimens had significantly less antigen and antibody presence of LASV compared to *M. natalensis* specimens, which is indicative of secondary horizontal transmission between the two species. Surprisingly, active LASV infection, indicated by $\geq 200$ LASV copies/μL RNA by qRT-PCR in two or more tissues of a specimen, was seen in less than 5% of all *Mastomys* spp. specimens. All specimens with active LASV infections tested antigen positive, though no correlation was seen between antigen strength and viral levels in these specimens.

Therefore, we conclude that the RDT by itself is not a quality indicator that a captured specimen has an active LASV infection. Rather, the RDT should be used as a triage tool, with specimens testing antigen positive prioritized for further molecular screening by qRT-PCR. Tissue samples from positive specimens with high viral levels can then be further processed for viral sequencing and viral isolation. The RDT is also useful as a method for general LASV surveillance as another tool to assess LASV prevalence within the rodent population of a community. While serosurveillance for LASV IgG is critical to understand the true incidence of LASV in a population, it is not necessary for
specimen “triage” as presence of IgG does not preclude an active LASV infection in a specimen.

Despite the limitations of the RDT as a standalone diagnostic for active LASV infection in rodents, the ability to quickly screen whether a rodent has evidence of a LASV infection has numerous implications. Rodent serosurveillance provides a natural complement to human serosurveillance; the two together can be used to develop models that better predict risk of spillover. Knowledge of which species of small mammals, particularly rodents, harbor LASV and the level of LASV naturally present in a community allows for robust and targeted rodent exclusion measures to be swiftly implemented. Given proper protective equipment and training on handling captured rodents, the potential for community-led rodent LASV testing would strengthen and support Lassa Fever outreach and education. Finally, in-field diagnostic testing enables scientists to prioritize specimens for further in-depth LASV testing and sequencing to better understand the reservoir host response of LASV, saving both time and resources in the lab.

Evolutionary insights of LASV have added to our knowledge of closely related mammarenavirus distribution amongst rodents throughout the continent, as well as LASV diversity within individual countries. With our targeted approach to sequencing, we have shown high levels of LASV diversity in Sierra Leone, roughly bounded by distinct geographic zones as opposed to date of sampling. It remains to be determined what geographic or demographic features within the geographic zones analyzed contribute to LASV diversity seen within Sierra Leone.

Our cross-sectional study of small mammals in Sierra Leone provides much needed information of current LASV prevalence and diversity in the host reservoir population.
These data, communicated informally among colleagues, have already helped shape country-wide serological surveys of human LASV prevalence in Sierra Leone. Secondly, the development of a rodent-specific diagnostic algorithm will help streamline laboratory analysis of LASV-infected rodents for future research studies. Moreover, the LASV sequences obtained from actively infected specimens allowed for in-depth viral sequencing analyses that has furthered our knowledge of the unique evolutionary history and phylogeography of LASV in Sierra Leone. Furthermore, the mass specimen collections undertaken for this study have also allowed for the creation of a large biobank of serum and tissue samples that can be analyzed for other viral pathogens.

Despite the strengths and benefits of the study, there are some notable limitations to its scope. Due to the COVID-19 pandemic, secondary trapping events in select villages to evaluate spatial and temporal trends of LASV in the rodent population were unable to be executed. Additionally, home-resolution level GPS coordinates were not obtained for multiple villages in the study, which further prevented detailed spatial analyses to be performed. While we did not examine erudition of rodent habitats, reproduction and fecundity, movement patterns, and spatial preferences of rodents, the study of these factors are essential components of LASV ecology. Our method of age estimation by weight groups has largely been superseded by dried eye lens weight, which we did not collect (Fichet-Calvet, Lecompte et al. 2008). As mentioned in the Chapter 1 discussion, antigen detection by DBS should have been experimentally validated in the lab prior to study collections. If antigen could not have been detected by DBS, an alternative blood collection method would have been deployed to thoroughly examine blood for presence of antigen,
antibodies, and virus. Furthermore, we did not examine tissue samples for virus by qRT-PCR in specimens that tested antigen negative.

With respect to future analyses, the tissue samples from these antigen negative specimens will be examined by qRT-PCR. All statistical analyses will also be re-performed with the updated qRT-PCR results prior to publication of the data in a peer-reviewed journal. We will also further expand on LASV phylogeography within Sierra Leone, as well as investigate intra-host variants with rodent LASV sequences that could elucidate transmission bottlenecks between rodents and humans. New small mammal collections underway in Sierra Leone are actively incorporating the established diagnostic algorithm to screen rodent specimens for active LASV infections in order to perform viral isolation under BSL-4 conditions, as well as sampling villages at multiple timepoints to allow for seasonal and spatial analyses of LASV in the host reservoir population.

Our proposed definition of an actively infected specimen is defined as a specimen having a viral level of > 200 LASV copies/μL RNA. Meanwhile, hybrid capture and amplicon-based sequencing approaches typically require a viral level > 500 copies/μL RNA (Zeller, Gangavarapu et al. 2021). Both hybrid capture and amplicon-based sequencing methods yield high genome coverage (< 100x) to allow for in-depth phylogenetic analyses not necessarily possible with traditional metagenomic sequencing; metagenomic sequencing can detect virus with as little as 100 copies/μL RNA (Greninger, Naccache et al. 2015). We chose to perform hybrid capture sequencing to allow for in-depth investigations of iSNVs in rodent tissues; the results from these analyses are forthcoming. We may also perform minION sequencing from future specimens with viral
levels between 200-500 LASV copies/μL RNA to generate as many consensus sequences from rodents as possible.

Lassa fever has a multifactorial etiology intricately woven between man and nature. LASV will continue to evolve over time, potentially creating more new reservoir hosts or changes in pathogenicity. Climate change and increasing globalization will also likely enlarge the Lassa fever zone across West African and perhaps even across continental borders. An encroachment of this zone into larger population centers could result in a major public health emergency. Community hygiene practices, along with expanded surveillance and observation of both humans and rodents, are needed to trace and prevent outbreaks. Research parsing out the ecological niches of LASV and further understanding the effects of viral diversity of LASV offer more effective solutions for concerted outreach efforts and successful prevention of Lassa fever.
7. REFERENCES


Bouquet, S. Somasekar, J. M. Linnen, R. Dodd, P. Mulembakani, B. S. Schneider, J. J.
identification of viral pathogens in clinical samples by real-time nanopore sequencing

Grubaugh, N. D., K. Gangavarapu, J. Quick, N. L. Matteson, J. G. De Jesus, B. J. Main,
amplicon-based sequencing framework for accurately measuring intrahost virus diversity

Gryseels, S., S. J. Baird, B. Borremans, R. Makundi, H. Leirs and J. Gouy de Belloq
(2017). "When Viruses Don't Go Viral: The Importance of Host Phylogeographic

Gryseels, S., T. Rieger, L. Oestereich, B. Cuypers, B. Borremans, R. Makundi, H. Leirs,
S. Gunther and J. Gouy de Belloq (2015). "Gairo virus, a novel arenavirus of the
widespread Mastomys natalensis: Genetically divergent, but ecologically similar to Lassa

Hastie, K. M., T. Liu, S. Li, L. B. King, N. Ngo, M. A. Zandonatti, V. L. Woods, Jr., J. C.
de la Torre and E. O. Saphire (2011). "Crystal structure of the Lassa virus nucleoprotein-
RNA complex reveals a gating mechanism for RNA binding." Proc Natl Acad Sci U S A

Hastie, K. M., M. A. Zandonatti, L. M. Kleinfelter, M. L. Heinrich, M. M. Rowland, K.
"Structural basis for antibody-mediated neutralization of Lassa virus." Science
356(6341): 923-928.

Haun, B. K., V. Kamara, A. S. Dweh, K. Garalde-Machida, S. S. E. Forkay, M. Takaaze,
M. Namekar, T. A. S. Wong, A. E. R. Bell-Gam Woto, P. Humphreys, O. I. Weeks, M. P.
of Ebola virus exposure in dogs from affected communities in Liberia: A preliminary

James N. Mills, P. D., S. D. James E. Childs, D. V. M. Thomas G. Ksiazek, Ph.D., and
Virologic Testing. U. S. D. O. H. H. SERVICES, Centers for Disease Control and
Prevention.

Josse, J. and F. Husson (2016). "missMDA: A Package for Handling Missing Values in

Kafetzopoulou, L. E., S. T. Pullan, P. Lemey, M. A. Suchard, D. U. Ehichiyoa, M.


