Abstract of the Dissertation submitted by OLONINIIYI OLAMIDE KOLAWOLE

Title: Genetic characterization of Lassa virus strains isolated from 2012 to 2016 in southeastern Nigeria.

Japanese title: 2012 年から 2016 年にかけてナイジェリア南東部で検出されたラッサウイルスの遺伝学的解析

PLOS Neglected Tropical Diseases, In press.

Department of Infection Research
Nagasaki University Graduate School of Biomedical Sciences

Supervisor: Professor Jiro Yasuda, D.V.M, Ph.D

Introduction
Lassa virus (LASV) is the causative agent of Lassa fever (LF), a viral hemorrhagic fever with manifestations that range from asymptomatic to an acute, severe form associated with significant mortality (up to 50% in hospitalized patients). LASV is primarily transmitted via contact with infected tissue, blood, or excreta of Mastomys natalensis, which is the reservoir host for the virus. LASV is an enveloped, bi-segmented RNA virus of family Arenaviridae and genus Mammarenavirus, which consists of different species of arenaviruses that infect mammals. The RNA genome contains large (L) and small (S) segments that encode two proteins each with an ambisense strategy. The matrix RING finger protein (Z) and RNA-dependent RNA polymerase (L) are encoded by the L segment, whereas the nucleoprotein (NP) and glycoprotein precursor (GPC) are encoded by the S segment. Phylogenetic analyses have shown that LASV nucleotide sequences cluster into six lineages (I - VI) based on geographical locations. The delineation of LASV into distinct lineages reflects its diversity. Sequence diversity is higher (as high as 24.6%) between lineages, whereas they are more conserved (9.6%) within each lineage. This inter-lineage heterogeneity has practical implications for developing vaccines and effective diagnostic assays, especially nucleic acid detection tests such as reverse transcription polymerase chain reaction (RT-PCR). A significant proportion of LASV lineage II strains isolated in southern Nigeria are mainly from Irrua, and one sequence isolated from southeastern Nigeria. However, the isolation of new sequences from different areas will lead to better characterization of LASV diversity both within and between lineages, which will be important for the prevention and control of this disease; in addition, it will provide new insights into the pattern of LASV dissemination and evolution.

Materials and Methods
We determined the viral sequences from positive cases of LF reported at tertiary hospitals in Ebonyi and Enugu between 2012 and 2016. Samples were collected from suspected cases of LF among patients who presented to these tertiary hospitals. A diagnosis of acute LF was made based on clinical signs and symptoms of fever, headache, and/or hemorrhage, contact history with a suspected or confirmed LF patient, as well as the discretion of the managing physician. LASV was then detected from extracted RNA using conventional RT-PCR. The L and S segments of the LASV genome were sequenced using Sanger’s method from LASV-positive samples. Phylogenetic analysis was performed with the Bayesian Markov chain Monte Carlo (MCMC) method using complete coding sequences of our newly determined LASV viral genes, together with full length sequences available in GenBank.

Results

RT-PCR showed that 29 out of 123 suspected cases were positive for the virus among which 11 complete coding sequences of viral genes were determined. Phylogenetic analysis of the complete coding sequences of the four viral proteins revealed that lineage II strains are broadly divided into two genetic clades—clades A and B—that diverged from a common ancestor 195 years ago. This demarcation into clades A and B was present in the tree of all the genes and was well supported with posterior probabilities of 1. Clade A, consisting of strains from Ebonyi and Enugu, was more conserved at the nucleotide and amino acid level than the other from Irrua. However, the four viral proteins were evolving at similar rates in both clades. The evolutionary rate of clade A and clade B strains, respectively, across all the genes were 8.0 and 7.6 substitutions × 10^{-4} site^{-1} year^{-1} for GPC, 9.9 and 10.2 substitutions × 10^{-4} site^{-1} year^{-1} for NP, 7.8 and 8.1 substitutions × 10^{-4} site^{-1} year^{-1} for L, and 14.5 and 13.5 substitutions × 10^{-4} site^{-1} year^{-1} for the Z gene. These results suggested that the viruses of these clades have been distinctively evolving in geographically separate parts of southern Nigeria. Furthermore, the epidemiological data of the 2014 outbreak highlighted the role of human-to-human transmission in this outbreak, which was supported by phylogenetic analysis showing that 13 of the 16 sequences clustered together.

Discussion

In this study, we examined in detail the molecular epidemiology of LASV in southeastern Nigeria by sequencing the LASV genome from separate sites in the region. Using our new sequences, we conclusively demonstrated that lineage II strains can be broadly divided into two clades. The division of lineage II strains into two clades shows a pattern of clustering of LASV strains based on geographical locations with clade A strains occupying the southeastern part in Abakaliki and Enugu, whereas clade B strains are in the south-central part, including Onitsha, Ekpoma, and Irrua. As for clade A strains being more conserved while evolving at a similar rate to clade B, this may be due to the introduction and maintenance of the virus in reservoir hosts, where they evolve independently without interacting with each other. The barrier to the interaction between the reservoir hosts in separate foci could be the river Niger and its tributaries, with clade A strains being on the eastern side of the river, whereas clade B were restricted on the western side. These results provide new insights into the evolution of LASV in southern Nigeria and have important implications for vaccine development, diagnostic assay design, and LF outbreak management.