Characterization of novel LCMV strains in Southern Iraq

Master's thesis

Alburkat Hussein

Department of Virology

Faculty of Medicine,

University of Helsinki

Viral Zoonoses Research Unit

Helsinki, Finland

and

University of Helsinki

Faculty of Biological and Environmental Sciences

Molecular and Integrative Biosciences Programme

Degree Programme in Molecular Biosciences, General Microbiology, Virology

Helsinki, Finland

UNIVERSITY OF HELSINKI
FACULTY OF BIOLOGICAL AND ENVIRONMENTAL SCIENCES

Helsinki

2019

Tiedekunta – Fakultet – Faculty Faculty of Biological and Environmental Sciences		Koulutus ohjelma – Utbildingsprogram _ Degree Programme Degree Programme in Molecular Biosciences
Tekijä – Författare – Author Alburkat Hussein		
Työn nimi – Arbetets titel – Title Characterization of novel LCMV strains in Southern Iraq		
Oppiaine – Läroämne – Subject General microbiology-virology		
Työn laji – Arbetets art – Level Master's Thesis	Aika – Datum – Month and year January 2019	Sivumäärä – Sidoantal – Number of pages 55

Tiivistelmä – Referat – Abstract

LCMV Lymphocytic choriomeningitis virus is a rodent-borne pathogen belongs to *Arenaviridae* family. Most of the studies have referred *Mus musculus* as the main reservoir of the LCMV. It has been detected in pet rodents, laboratory rodents, and wild mice. Humans be infected with LCMV through the ingestion or inhalation of sources contaminated with rodent feces, urine, or both. LCMV infection can be asymptomatic, present with mild symptoms, or it can cause aseptic meningoencephalitis (AME) and teratogenic effects in infants. However, clinical cases of LCMV infection have been rarely reported, and there is only fragmental knowledge on the presence and prevalence of LCMV infections around the world. Likewise, the genetic characteristics of the circulating LCMV strains and impact of LCMV on public health have remained poorly characterized. This study was performed in the Southern Iraq, due to the lack of comprehensive information about LCMV in this area.

There were three main aims in this thesis. First, to assess the prevalence of LCMV among the healthy human population in the Nasiriyah region, southern Iraq. Second, to assess whether LCMV infections can be associated with neurological manifestations. Third, to characterize the genetic variation and evolutionary history of LCMV strains circulating in southern Iraq.

Serum and CSF samples were collected from patients and healthy people in Nasiriyah governorate in the Southern Iraq. Serum samples were screened for LCMV using Immunofluorescence assay (IFA) to detect IgG and IgM antibodies. Real-time PCR was used to detect LCMV genome. In order to confirm the PCR positive samples, we sequenced these samples by Next-generation sequencing.

The serological assay results showed 12.22% IgG prevalence of LCMV among healthy people and 7.36% IgG prevalence among patients with neurological symptoms. The IgM prevalence was 1.25% among the patients with acute infections. From symptomatic patients, we sequenced partial L-segments of two new LCMV strains. The phylogenetic tree constructed on the basis of all known LCMV strains suggested that these new LCMV strains from Iraq are genetically distant from the previously known LCMV strains and form a novel sub-cluster within LCMV species.

This study is the first survey of LCMV in the Southern Iraq. LCMV appears to be a rather common infection in Iraq. I reported new strains of LCMV that are circulating in the study site and most likely is the causative agent of the central nervous system-associated clinical manifestations in these patients. For future work, I'm aiming the detection of other *Arenaviruses* spreading in the Southern Iraq.

Avainsanat – Nyckelord – Keywords
LCMV, IFA, RT-PCR, Amplicon PCR, Metagenomics, Next-generation sequencing.

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Sarah Butcher, Haider Al-Hello, and Teemu Smura.

Säilytyspaikka – Förvaringställe – Where deposited
E-thesis

Muita tietoa, Övriga uppgifter, Additional information

Contents

Abbreviations

AMAV	Amapari virus
BIBDAV	Boid inclusion body disease-associated *Arenavirus*
ER	Endoplasmic reticulum
FITC	Fluorescein isothiocyanate
FU	Fluorouracil
GP	Glycoprotein
GPC	Glycoprotein complex
IFA	immunofluorescence assay
IFA	Immunofluorescence assay
IGR	intergenic region
JUNV	Junin virus
LASV	Lassa virus
LATV	Latina virus
LCMV	Lymphocytic choriomeningitis virus
MACV	Machupo virus
MCMC	Markov Chain Monte Carlo
NP	Nuclear protein
PARV	Paraná virus
PEDs	Pairwise evolutionary distances
PICV	Pichindé virus
RBS	Resuspension buffer
RdRp	RNA-dependent RNA polymerase
RNPs	Ribonucleoprotein
S1P	site- 1 protease
SKI-1	Subtilisin kexin isozyme-1
SSP	Stable signal peptides
TAE	Tris base, acetic acid and EDTA

TCRV	Tacaribe virus
vRNP	viral ribonucleoprotein
WTA	Complete Whole Transcriptome Amplification
αDG	Alpha Glycoprotein dystroglycan
βDG	Beta Glycoprotein dystroglycan

1. Introduction

Lymphocytic choriomeningitis virus (LCMV) is a *Mammarenavirus* causing a zoonotic neurological disease called lymphocytic choriomeningitis (LCM) in humans. The main host of LCMV is the common house mouse, *Mus musculus*. Infection in humans occurs when they inhale or ingest contaminated mice secretions. LCMV can infect humans of all ages including the fetus (Bonthius, 2012).

The majority of the epidemiological studies related to LCMV have been done on samples originating in Europe, United States of America, Japan and China (Blasdell *et al.*, 2016; Knust *et al.*, 2011; Knust *et al.*, 2014; Leibler, Zakhour, Gadhoke, & Gaeta, 2016; Takagi, Ohsawa, Morita, Sato, & Ohsawa, 2012). The available reports do not address the presence or prevalence of LCMV in Iraq. In addition, the effect of LCMV on public health is unknown (Zhang *et al.*, 2018) as is the prevalence of the virus, its disease burden in humans (Albarino *et al.*, 2010), and its genetic diversity in Iraq. In this study I addressed the prevalence and the characterization of new strains of LCMV; probably, were circulating between for long time in Southern Iraq.

1.1. Taxonomy of *Arenaviridae*

The name *Arenavirus* is derived from the Latin word *arenosus* which refers to the sandy appearance that was observed by electronic microscopy of ultrathin sections of viral particles. Until recently, the family *Arenaviridae* included only one genus *Arenavirus* with twenty two species (Emonet, de la Torre, Domingo, & Sevilla, 2009). In the latest International Committee on Taxonomy of Viruses (ICTV) the family *Arenaviridae* includes three genera (Piet Maes *et al.*, 2018) *Mammarenavirus*, *Reptarenavirus* and *Hartmanivirus*. The genus *Mammarenavirus* is comprised of 36 species which infect mammals including LCMV (Piet Maes *et al.*, 2018). Recently, reports from China have recently detected new *Mammarenaviruses* in small mammals (Blasdell *et al.*, 2016; Takagi *et al.*, 2012). *The genus Reptarenavirus*, includes five verified species and two putative new species, and the genus *Hartmanivirus*, which infect snakes, includes 1 verified and 3 putative new species (Jussi Hepojoki *et al.*, 2018).

The reorganisation of the *Arenaviridae* was based on the discovery, genome sequencing and genetic analysis of a number of new viruses, including those isolated from snakes (J. Hepojoki *et al.*, 2015). The new phylogenetic tree in Figure 1 shows these taxonomic relationships.

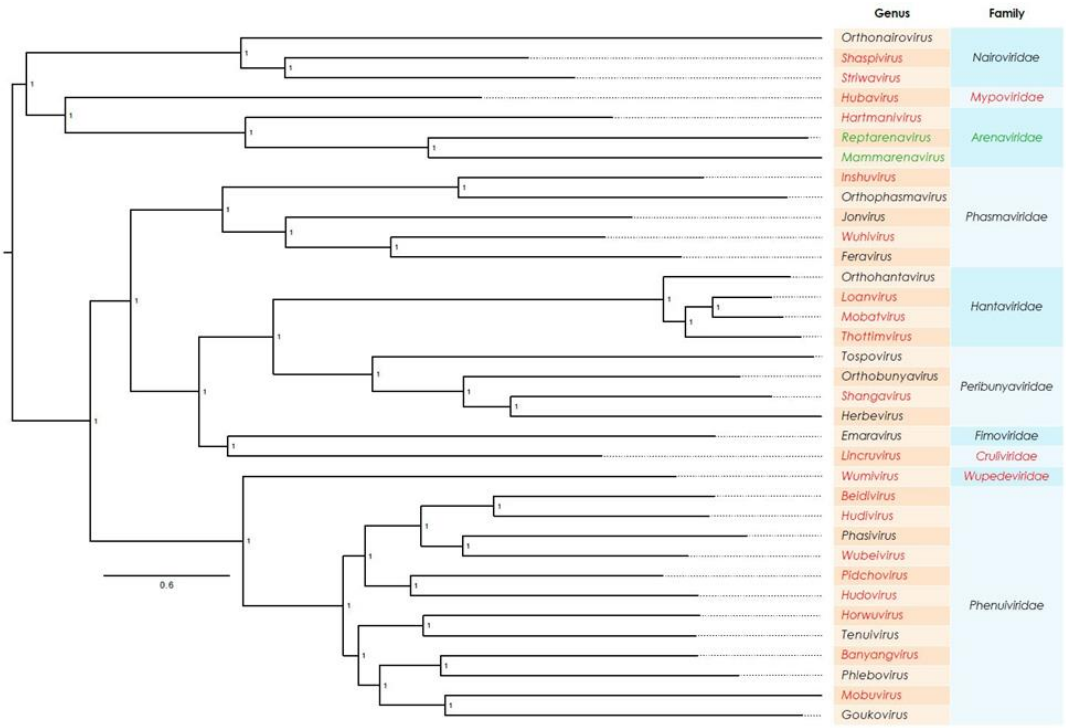

Figure1: The figure shows a Bayesian MCMC tree estimated using a Bayesian Markov Chain Monte Carlo method implemented in BEAST, using the WAG amino acid model of amino acid substitutions. Maximum clade credibility trees were determined using TreeAnnotator with a burn-in of 10% of the sampled trees. The Markov chain Monte Carlo analyses were run until effective sample sizes above 200 were obtained. The dataset used consists of full length products of coding regions of the S, M, and L segments (nucleocapsid protein, glycoprotein precursor, and RNA-dependent RNA polymerase, respectively) or in the case of *Arenavirus* S and L segments, concatenated in one alignment. The dashed boxes show the proposed virus families as calculated with DEmARC version 1.0. The DEmARC method is an approach for partitioning the genetic diversity of a virus family /genus within a hierarchically organized framework, providing quantitative support for both the delineated classification levels and the inferred taxa by devising the number and threshold values on genetic divergence at each taxonomic level. Genetic divergence is quantified by pairwise evolutionary distances (PEDs) estimated by maximum likelihood method on amino acid alignment. Numbers at main nodes of the tree represent the Bayesian Posterior probabilities. The number below the bar (0.6) corresponds to the amino acid distance between sequences. The picture used from (P. Maes *et al.*, 2018) with copyright permission number 4514191027338.

Defining distinct *Arenavirus* species is conducted either with the detection of the geographic distribution and the antigenic properties, such as glycoprotein (GP) or nuclear protein (NP), or genetic variations among L- and S-segments and their related genes (Radoshitzky *et al.*, 2015). However, a systematic comparison between serogroup and genome sequence has not been performed for the majority of recently discovered *Arenavirus* species. Figure 2 represents an example of the identification of *Arenaviridae*

members, genera and species. To date, more than 30 LCMV strains have been characterized and this number is increasing (Takagi *et al.*, 2017).

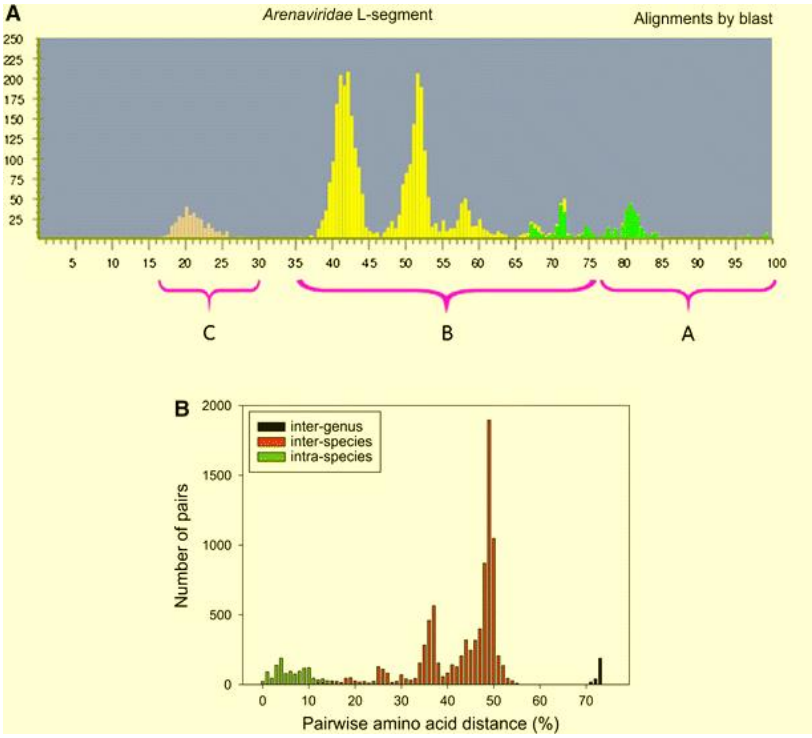

Figure 2: Pairwise Sequence Comparison (PASC) analysis of L segment sequences and amino acid distance analysis of NP sequences. (A) Distribution of pairwise identities among 87 complete sequences of the L segments of members of the family *Arenaviridae*. Regions A, B and C represent virus pairs from the same species (100 %-76 %), different species but the same genus (76 %-35 %), and different genera (16 %-30 %), respectively, based on the proposed identity values indicated in parentheses. The x-axis shows percent identity, and the y-axis shows the number of L segment sequence pairs. (B) Amino acid sequence distances were compared using the pairwise-distance algorithm in the MEGA 6 software package and shown as frequency histograms. This image is modified from (Radoshitzky *et al.*, 2015) with copyright permission number 4514180149734.

Areas like United States, Europe and Japan have had outbreaks of LCMV leading to the identification of new strains through the isolation of the virus from infected mice, hamsters and humans. Usually, the outbreaks have originated from lab animal facilities, organ donors and occasionally pet animals. There are very few recent reports of clinical infection transmitted from wildlife. Most likely this is due to lack of diagnostic methods and under-reporting in most parts of the world.

Those isolated strains show variable clinical signs depending on the interaction between the viral proteins, host receptors and MHC class I molecule (Takagi *et al.*, 2017). Generally, RNA viruses evolve quickly compared to DNA viruses. This is due to the

error-prone polymerases of most RNA viruses including *Arenaviridae* (Duffy, Shackelton, & Holmes, 2008). The high mutation rate enables rapid adaptation of RNA viruses to different ecological niches, such as adaptation to different host species. Within LCMV genome, mutations occur mostly in GP and polymerase genes affecting the sequences of the encoded proteins as well (Salvato, Borrow, Shimomaye, & Oldstone, 1991).

1.2. History of *Mammarenaviruses*

In 1930, LCMV was discovered in North America by three different groups (reviewed in (Emonet *et al.*, 2009). In 1934, the name LCMV was created by Armstrong and Lillie based on the signs and symptoms seen in mice and monkeys after intracerebral inoculation as mentioned in (Emonet *et al.*, 2009). Furthermore, in 1935 Rivers and Scott characterized LCMV as a human pathogen (Rivers & McNair Scott, 1935). Also, in 1935, *Mus musculus* was identified as a natural reservoir of LCMV (Traub, 1935). Thereafter, LCMV was determined as the causative agent of aseptic meningitis in human (Rivers & McNair Scott, 1935; Rivers & Scott, 1936; Scott & Rivers, 1936). Due to the global distribution of its natural host *Mus musculus*, also LCMV is considered to be globally distributed. The confirmed geographical distribution of LCMV is shown in Figure 3. A new agent Tacaribe virus (TCRV) was isolated in 1956 from Jamaican Artibeus bat and mosquitoes in Trinidad, West Indies (Downs, Anderson, Spence, Aitken, & Greenhall, 1963). After three years, Junin virus (JUNV) was identified as the causative agent of Argentinian hemorrhagic fever (Parodi *et al.*, 1958).

Attempts to discover and detect new viruses continued, until 1963, where Mettler established a new group called the "Tacaribe antigenic group". Mettler relied on the determination of serological relationships between TCRV and JUNV using a complement fixation test, as well as, demonstrating the differences between these viruses using a neutralization test (Mettler, Gianantonio, & Parodi, 1963). Hence, new viruses were discovered and included in the Tacaribe antigenic group. In 1963, Machupo virus (MACV) was isolated from a patient suffering from Machupo hemorrhagic fever in Bolivia (Johnson *et al.*, 1965). Thereafter, isolation of new viruses expanded to include new members as following: Amapari virus (AMAV) was isolated from rodents and mites of Amapa Territory, Brazil in 1966 (Murphy, Webb, Johnson, Whitfield, & Chappell, 1970), Latina virus (LATV), which has been mentioned for the first time by (Murphy *et al.*, 1970), Pichindé virus (PICV) was isolated from rodents in a village in Parana´ Colombia (Trapido & Sanmartin, 1971), Paraná virus (PARV) was isolated in Paraguay (Radoshitzky *et al.*, 2015), and Tamiami virus isolated from a rat in South Florida in 1969 (Calisher, Tzianabos, Lord, & Coleman, 1970). After that, Lassa virus (LASV) was identified in Nigeria as the causative agent of the death of two missionary-nurses (Frame, Baldwin, Gocke, & Troup, 1970).

Altogether, based on the morphological and serological properties among LCMV, TCRV and LASV, the *Arenavirus* group was established (Buckley & Casals, 1970). Subsequently, in 1976, the family *Arenaviridae* was established (Fenner, 1976).

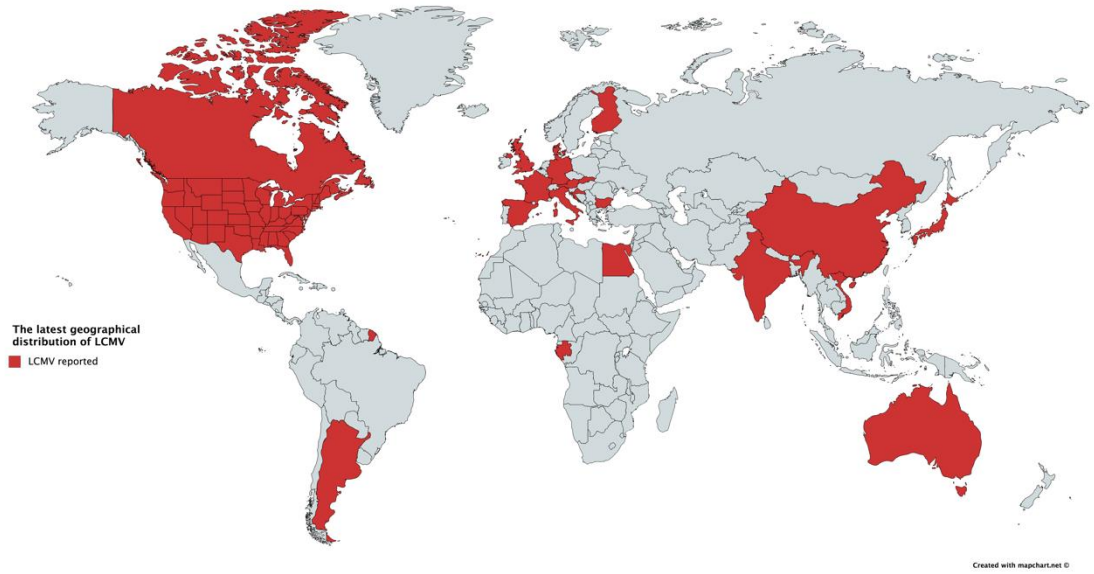

Figure 3: The latest geographical distribution LCMV. The viral distribution was obtained from literature search of *Mammarenavirus* identification either by genome sequencing or serology in different countries. LCMV distribution rate was gained from (Pontremoli, Forni, & Sironi, 2018). Map was created with mapchart.net.

1.3. LCMV genome and virion structure

Arenaviruses are enveloped viruses with bi-segmented, negative-sense, single-stranded RNA genomes with a large (L)-segment and small (S)-segment (Figure 4 A). Although there is some variability in the genomic RNA segments lengths, L (ca 7.2-7.3 kb) and S (ca 3.4-3.5 kb), their overall organization is conserved (Ortiz-Riano, Cheng, Carlos de la Torre, & Martinez-Sobrido, 2013; Takagi *et al.*, 2012)

Each *Arenavirus* genome segment uses an ambisense coding strategy to synthesize two polypeptides in opposite orientations, separated by a noncoding intergenic region (IGR) with a predicted folding of a stable hairpin structure. The 5' and 3' ends of each segment are complementary to each other, thereby forming a so-called pan-handle structure. Probably, this pan-handle structure has a role in the control of RNA synthesis (Perez & de la Torre, 2003). The negative stranded genome is transcribed to positive strand mRNA. These viral mRNAs of *Mammarenaviruses* bear a cap structure on the 5' end but they lack a Poly (A) structure at the 3' end. (Foscaldi *et al.*, 2017)

The S-segment encodes the viral glycoprotein precursor (GPC; ca 75 kDa), and the nucleoprotein (NP; ca 63 kDa). The glycoprotein complex (GPC) is subsequently cleaved into three polypeptides: the stable single peptide (SSP), glycoprotein 1 (GP1) and glycoprotein 2 (GP2).

The L-segment encodes the viral RNA-dependent RNA polymerase (RdRp; ca 200 kDa), and a small RING finger protein Z (ca 11kDa) (Urata & de la Torre, 2011).

The virion structure is shown in (Figure 4B). The virion is enveloped and spherical or pleomorphic. Its Diameter ranges from 50-300 nm. The viral glycoproteins (in purple color) shown on the viral particle surface are embedded in the viral envelope which is shown in black. The matrix Z protein is located beneath the viral envelope directly (shown in green). The viral RNA is encapsidated by NP (shown in yellow) and associated with RdRp (shown in red) to form the ribonucleoparticles.

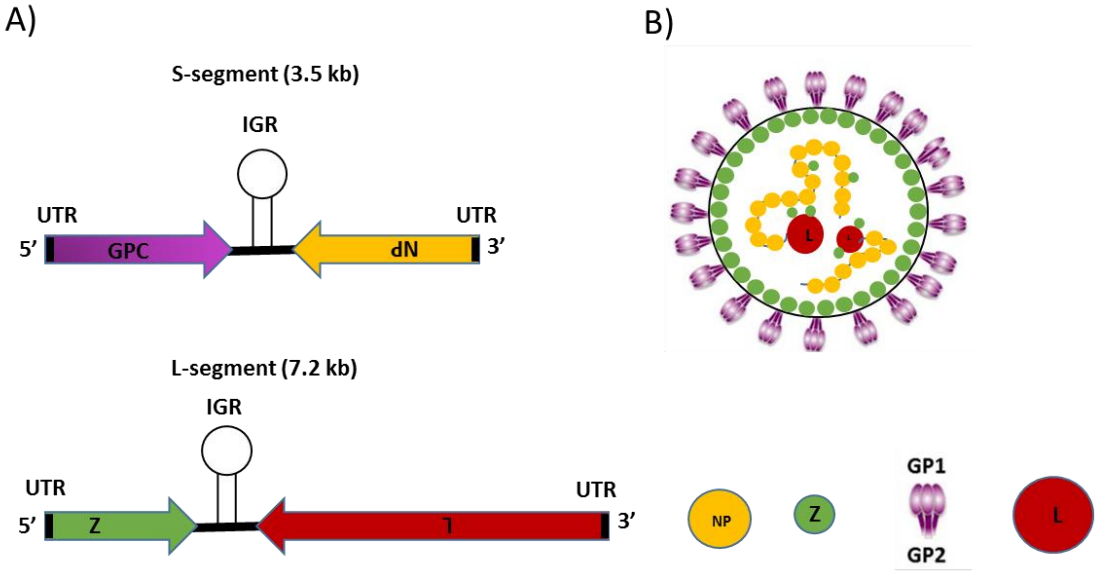

Figure 4: The genome and virion structure of *Mammarenavirus* LCMV. A) The genome consists of L-segment (7.2 kb) and S-segment (3.5 kb) encoding the viral proteins. B) The virion particle covered by GPC. The matrix Z protein from inside associated with the viral membrane. The association between NP and viral RNAs to form RNPs. RdRp (L) responsible for viral replication. The image is adapted from (Martinez-Sobrido & de la Torre, 2016).

1.4. Life cycle of LCMV

The life cycle of LCMV is outlined in Figure 5. The initial step in the LCMV life cycle is the attachment to the host cell. The major receptor is α subunit of the glycoprotein dystroglycan (αDG) (Spiropoulou, Kunz, Rollin, Campbell, & Oldstone, 2002).

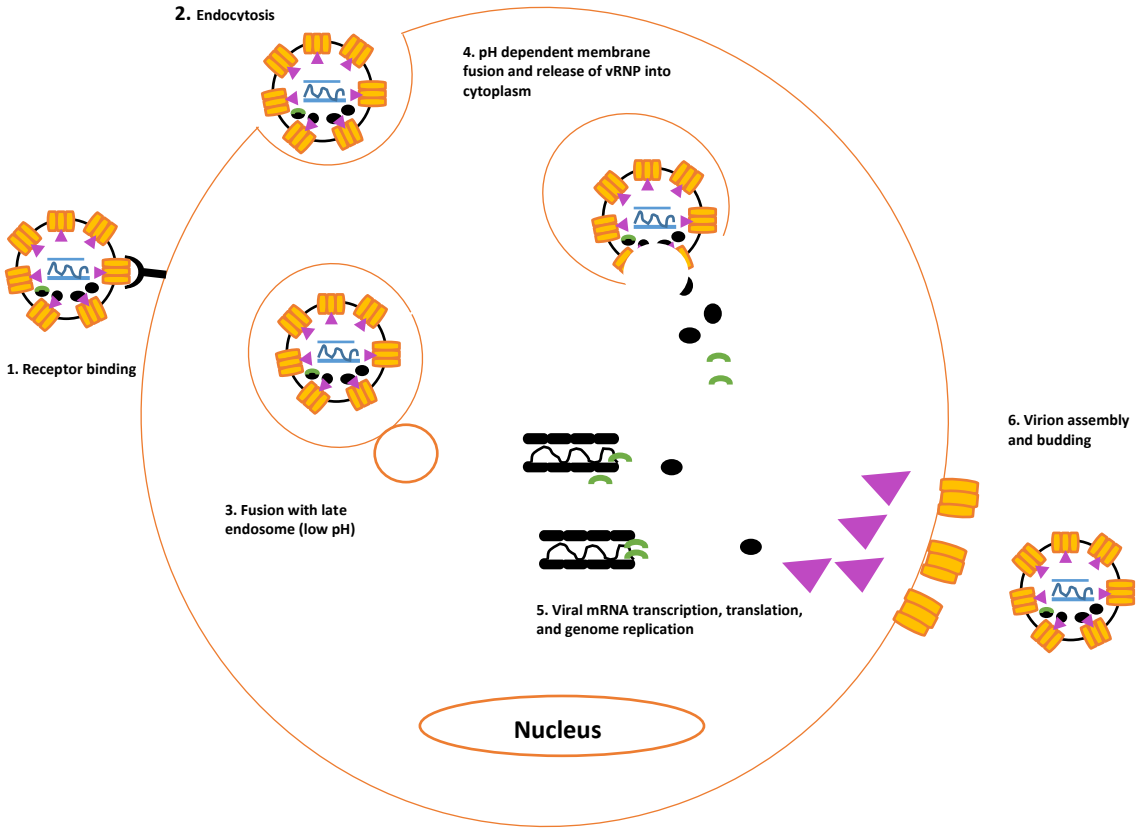

Figure 5: *Arenavirus* life cycle. This image is adapted from (Shao, Liang, & Ly, 2015).

Dystroglycan contains both αDG and βDG subunits, but only αDG is implicated in the viral life cycle (Kunz, Campbell, & Oldstone, 2003). The physiological function of αDG is to complete the link between the cytoskeleton and basal lamina by calcium-dependent binding activity associated with the high affinity to the extracellular matrix protein (Michele & Campbell, 2003).

The glycoprotein that is embedded in the viral envelope, is composed of an N-terminal stable signal peptides (SSP) and the GPC which is included in the SSP. In the

Endoplasmic reticulum (ER), the GPC cleavages into GP1 and GP2 subunits by signal peptidases and subtilisin kexin isozyme-1/site- 1 protease (SKI-1/S1P) (Beyer, Popplau, Garten, von Laer, & Lenz, 2003). Altogether, SSP, GP1 and GP2 are associated with the virus attachment and entry to the host cell. In detail, GP1 is responsible of virus binding to the cellular receptors on the host cell surface. Uptake of the virus into cells is mediated by GP2 and endocytosis. GP2 also mediates the fusion of the virus that occurs in low pH in late endosome (Lenz, ter Meulen, Klenk, Seidah, & Garten, 2001).

This leads to the release of the viral ribonucleoprotein (vRNP) that is comprised of NP, viral RNAs and L, into the cytoplasm. Subsequently, the protein expression and viral genome replication, which is initiated by RdRp complex, take place in the cytoplasm. Finally, the assembly of the virus and budding occurs in the plasma membrane exclusively utilizing the αDG. Z proteins is associated with budding of the progeny particles and then the virus is released to infect other cells (Shao *et al.*, 2015).

1.5. Transmission

Rodents are considered as the natural reservoirs of LCMV. Almost all the identified members of *Mammarenavirus* can cause either an acute or persistent infection in humans. The main transmission pathway of LCMV member LCMV is the contact with urine or saliva from the infected rodents moving within the humans' houses. The transmission routes from rodents to humans occur either by direct contact with the infected materials to open wounds or mucosa or indirectly by exposure to contaminated water and food.

The common house mouse *Mus musculus* is not only a reservoir of the virus but is considered to be the main natural host of LCMV. The European house mouse *Mus musculus domesticus* has also been reported as a host for LCMV in Africa (N *et al.*, 2015).

The main key in the transmission of LCMV through the mouse populations is the ability of the virus to pass vertically from one generation to the next by intrauterine infection (Bonthius, 2012). In humans, there is some evidence for congenital transmission from mother to fetus, but there is no evidence of horizontal human to human transmission. In case of congenital transmission, the fetus is exposed to the infection if the mother aquires a primary LCMV infection during pregnancy. The infection may be transferred congenitally to the fetus or possibly during the delivery (i.e. during intrapartum period. (Bonthius, 2012). There is also an increased risk of infection during organ transplantation (Fischer *et al.*, 2006).

1.6. Clinical picture

1.6.1. Signs and symptoms

The signs and symptoms of LCMV infection ranges from asymptomatic to severe symptoms depending on general health status of the patient during the infection (Bonthius & Karacay, 2002). Immunocompromised individuals are especially susceptible, leading more commonly to deaths as shown in outbreaks in the USA and Australia (Urata & de la Torre, 2011). Also, the clinical course of LCMV infection depends on whether the infection occurs during the postnatal or prenatal life (Bonthius, 2009). The LCMV infection symptoms during both childhood and adulthood are often mild. Symptoms in the first phase of infection include fever which may be accompanied by headache, myalgia, malaise, anorexia, nausea, and vomiting. Severe symptoms develop up to two weeks after full recovery from the first mild phase of the infection. The severe symptoms include aseptic meningitis, headache, photophobia, fever, vomiting, and rigidity (Barton & Hyndman, 2000).

Regarding the LCMV infection and its implications to the fetus, (Bonthius *et al.*, 2007) referred to the occurrence of brain and retinal damages as a result of acquisition of LCMV infection during pregnancy. Nevertheless, the link between the congenital LCMV infection and the congenital neurological and ophthalmological diseases mostly is not considered. This link between LCMV and these diseases in fetus, child and adults should be considered due to the high increasing seroprevalence of LCMV among mice and humans (Mets, Barton, Khan, & Ksiazek, 2000).

So far, among the *Mammarenavirus*es, only LCMV is known to cause human illness in Asia and many cases maybe undiagnosed because mild symptoms are similar to common cold (Blasdell *et al.*, 2016).

1.6.2. Epidemiology

LCMV infections are related strongly to the wild mice, thus wherever the mice are existing, also LCMV is present (Bonthius, 2012). An epidemiologic study on the prevalence of LCMV in Baltimore (Childs, Glass, Korch, Ksiazek, & Leduc, 1992), indicated that 9% of the mice population were infected with LCMV. Furthermore, LCMV is considered the highest distribution rate among *Arenavirus*es, due to its association with the global rodents species *Mus musculus* (Charrel, de Lamballerie, & Emonet, 2008).

The lack of resources in poor countries is also a logical reason for the dangerous and rapid spread of infectious diseases . This makes it difficult for people to go to health care professionals, and therefore leads to the spread of undiagnosed diseases. The lack of scientific experience, training or laboratory support by clinicians is also reflected in the increase of infectious diseases (Mueller *et al.*, 2014).

Recommendations to study LCMV are based on the evidence of increased risk to public health especially for pregnant women's and fetus who are at risk of contracting meningitis (Kallio-Kokko *et al.*, 2006).

1.6.3. The pathogenicity of LCMV

The connection between LCMV and menigitis disease comes from the LCMV pathogenicity of the meninges. LCMV infection usually occurs through an inhalation of aerosols that enter the lungs, which is the primary site of infection (Bonthius, 2012). At the first stages, the infection appears as interstitial lung infiltrates and lung edema. After that, the virus enter the blood stream, from which it can infect other organs in the body, including the meninges. Ultimately, the meningitis symptoms appear as a result of inflammatory response (Bonthius, 2012).

1.6.4. Diagnostics and detection

The diagnostic methods to detect LCMV rely on the detection of antibodies against LCMV (IgM or seroconversion from IgG- to IgG+) or the detection of viral nucleic acids using PCR combined with sequencing. LCMV can also be cultured. However, it replicates slowly in the cell culture and does not induce cytopathic effects. Also, there are general methods for *Arenaviruses* diagnostics.

The common detection methods range from isolation of the virus, ELISA to RT-PCR. Detection of known diagnostic monoclonal antibodies to *Arenavirus* is useful to detect infected tissues. This could be as a starter method to determine viral antigens as well as a deferential method between *Arenavirus*es (Buchmeier, Lewicki, Tomori, & Oldstone, 1981).

For infectious pathogens like *Arenaviruses*, immunofluorescence assay (IFA) also was described as a useful assay to detect antibodies in serum samples that bind to the virus infected cells. The binding occurs during an incubation time in a temperature, which is equivalent to the body temperature. Thus, by IFA one can distinguish antibodies by their interaction with virus under fluorescence microscope (Bausch *et al.*, 2000).

In addition, serologic diagnosis for IgM and IgG are recommended but they should be sensitive, specific and reliable to avoid misdiagnosis. IgM-specific ELISA is a necessary assay to detect infection of acute cases because IgM is the first antibody to appear in response to initial antigen exposure. While IgG-specific ELISA is considered an efficacious method for both diagnostics and epidemiological studies (Bausch *et al.*, 2000; Ksiazek *et al.*, 1999)

1.6.5. Treatments and prevention

So far, there is no efficient *Arenavirus* vaccine. There are some attempts to produce a live attenuated *Arenavirus* by developing a reverse genetic system to rescue recombinant *Arenavirus*es from human 293T cells and Vero cells (Ortiz-Riano *et al.*, 2013). Although this system was used successfully in recent study, it was restricted only to the rodent cells. In this study, FDA-approved cell lines were used to rescue a recombinant virus from human cells 293T and Vero cells in order to facilitate vaccine development. However, the transfection of the virus was not highly efficient.

Better understanding of the LCMV molecular and cell biology may lead to the identification of potential drug targets for antiviral development. The Z protein required for virus budding maybe one such target that facilitate the development of antiviral drug (Urata & de la Torre, 2011).

The available antiviral treatment currently is the purine nucleoside analog ribavirin (1-β-D-ribofuranosyl-1-*H*-1,2,4-triazole-3-carboxamide), although it has undesirable side effects (Urata & de la Torre, 2011). The purine nucleoside analogue ribavirin was reported as a mutagenic agent to LCMV through cell culture experiments (Moreno *et al.*, 2011). They demonstrated the inhibitory effect of ribavirin on LCMV replication in BHK-21 cells at low MOI (multiplicity of infection). Subsequently, they found ribavirin caused high mutation of LCMV in the treated viral population compared to the untreated ones.

The base analogue 5-fluorouracil (FU) is also effective as a mutagenic agent which leads to virus extinction by interfering with viral replication. This was shown by infecting BHK-21 cells with LCMV in the presence of the FU. According to *in silico* simulations, this interfering activity with seems to be an effective system to suppress the viral infectivity by enhanced mutagenesis (Martin, Abia, Domingo, & Grande-Perez, 2010).

Combination of such inhibitors could be explored to prevent or cure LCMV infection (Moreno *et al.*, 2011).

1.7. Sequencing strategies

This study was performed to assess the presence of LCMV *Mammarenavirus* in Southern Iraq. After RT-PCR detection of LCMV, the serum samples were pretreated with ApoH. The ApoH is a plasma protein able to bind to viruses and protect them of the possible inhibitors including proteins and antibodies. Using ApoH protein results in concentrated virus, and its viability and infectivity are maintained (Stefas *et al.*, 2001). ApoH is also considered as a highly sensitive tool to detect and isolate viruses from various sources, such as serum which includes different molecules that inhibit the viruses (Adlhoch *et al.*, 2011).

For the partial sequencing of L-segment, I used Nextera XT DNA library preparation kit which allows the fragmentation and tagging of the genomic DNA to adapter sequences using a unique engineered transposon. This process allows dual-indexed sequencing of the pooled libraries on a MiSeq sequencing platform. Figure 6 shows the library preparation process. Phylogenetic analysis was performed to evaluate the evolutionary rate of LCMV. Using MEGA 6, the phylogenetic tree was produced after alignment against LCMV strains obtained from GenBank. The main steps of sequences analyses are presented in (Figure 8).

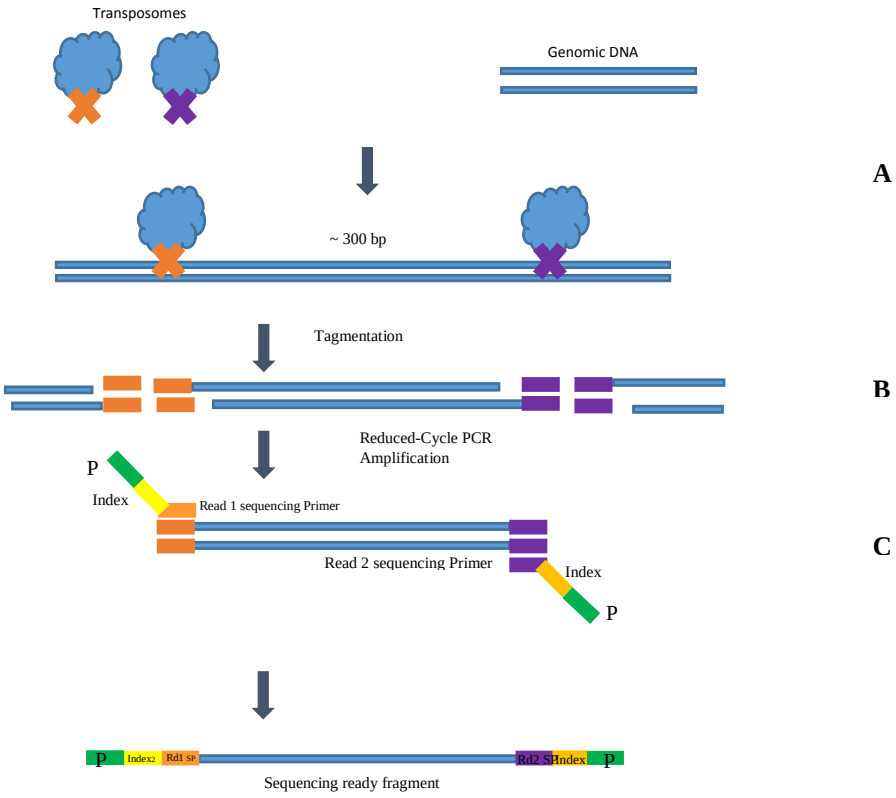

Figure 6: The library preparation workflow. A) Represents the transposomes attachment to the genomic DNA. B) Represents the tagmentation of the genomic DNA. C) Represents the indexes attachment to the amplified DNA.

1.7.1. De novo assembly of sequence reads

In conjunction with the massive development of next generation sequencing, whole genome sequencing has become easier and more cost effective than before. Many

assembly programs are available to align short sequence reads (typically 75-300 bp) resulted from the sequencing instruments like the MiSeq platform from Illumina which was used in this thesis work. De novo assemblers are useful because they do not require a reference genome. Using an improved de novo assembly leads to more reliable results (Lischer & Shimizu, 2017).

The de novo assembly pipeline includes all the main steps to alignment the sequences from the sequencing instrument leading to precise whole genome sequencing. Figure 7 shows the principle of De novo assembly.

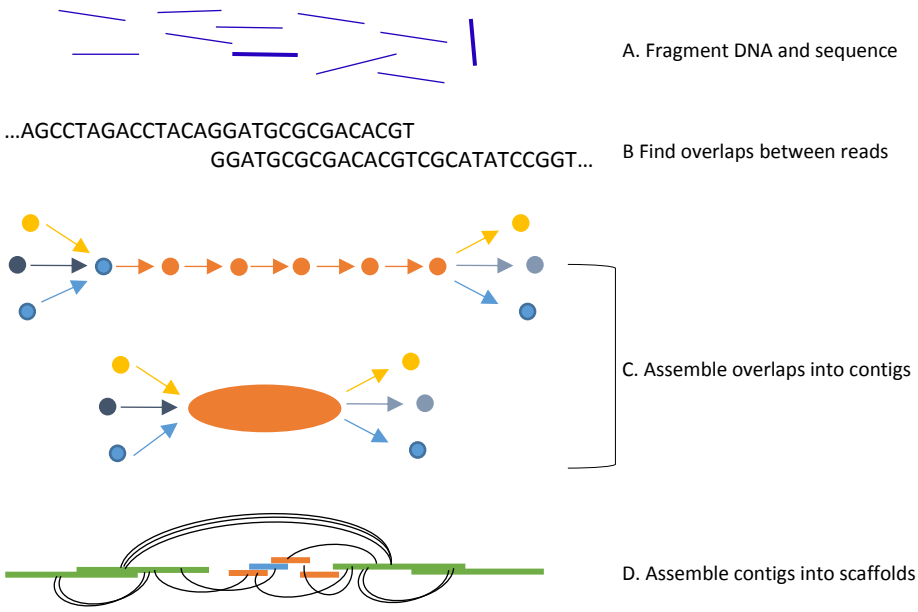

Figure 7: Schematic representation of the principle of De novo assembly. The image is modified from (Baker, 2012) with copyright permission number 4514210782125.

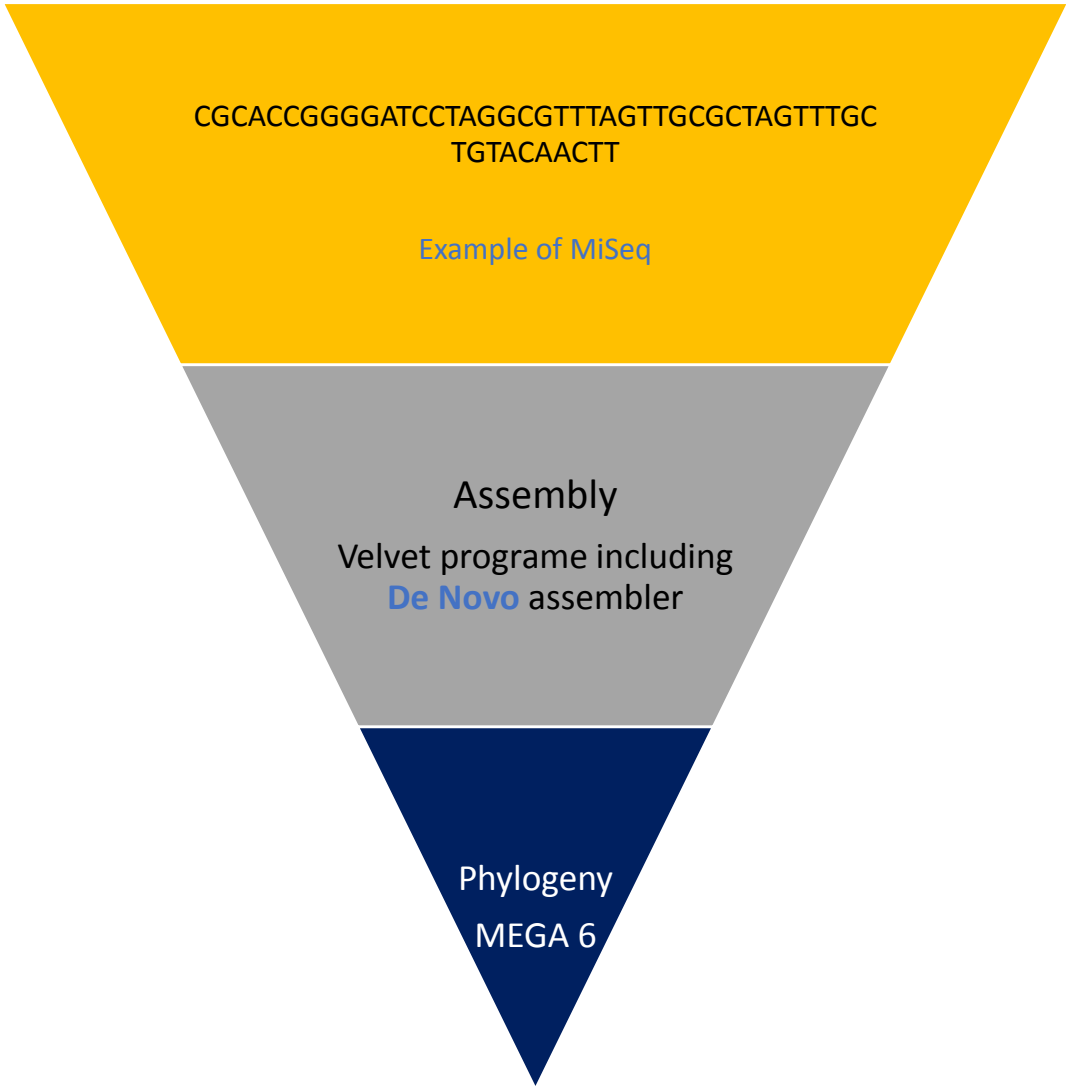

Figure 8: This flow chart is showing the workflow after performing the sequencing of the libraries using MiSeq platform of illumina. The outcomes of MiSeq are Fastq files including the sequences of the conducted samples. The assembly was performed using de novo assembler which is included in Mira and Velvet programs. The phylogenetic tree is the last step of the work that was generated using Mega 6 program.

2. Aims of the study

There were three main aims in this thesis: First, to assess the prevalence of LCMV among the healthy human population in the Nasiriyah region, Southern Iraq. Second, to assess whether LCMV infections can be associated with neurological manifestations. Third, to characterize the genetic variation and evolutionary history of LCMV strains circulating in southern Iraq.

3. Materials and Methods

3.1. Study site and samples collection

A total of 297 samples were collected from people suspected of being infected with LCMV as well as healthy people as control, to investigate the seroprevalence of the LCMV virus. The samples were collected with ethical permits from Nasiriyah capital of the Dhi Qar Governorate in southern Iraq including both rural and urban areas (Figure 9). There were both serum (N = 289) and CSF (N = 8) samples.

Figure 9: Map of Iraq showing the location of Nasiriyah from which the samples were collected. The star mark refers to the capital city Baghdad. The image is modified from 123rf.com.

3.2. Immunofluorescence assay (IFA)

An immunofluorescence assay was utilized to detect IgG and IgM antibodies from the serum samples. IFA is a commonly used assay to determine the seroprevalence of the viruses. Initially, Vero E6 cells were cultivated at 37 °C / 5% CO2 atmosphere for four

days in Eagle's Minimum essential medium that included 10% FBS, L-glutamine, penicillin and streptomycin. LCMV slides were prepared by infecting Vero E6 cells (5 cell culture flasks, 75 cm^2) from monkey kidney cells with LCMV (*Armstrong* strain) to use it as antigen then the cells were incubated at 37 °C for 1 h. After that, 20 ml of Eagle's Minimum essential cell culture medium was added to the cells. The appearance of the cells was monitored 7-8 days post infection and then these cells were washed with PBS and detached with trypsin-EDTA. Afterward, 45 ml of PBS was used to wash the cells 5 times and then the cells were centrifuged at 1800 rpm for 3 minutes. The cells were resuspended with 25 ml of PBS and divided into 10 well slides with 25µl/well. The laminar flow cabinet was used to dry the slides overnight. Subsequently, 99.98 % acetone was added to fix the cells at +4 °C for 7 min and the slides were stored at -70°C until use.

Before initiating the IFA processing, the slides were brought to room temperature. Serum samples along with the LCMV positive control (serum from confirmed positive patient for LCMV) were diluted 1:40 in PBS. Thereafter, 25 µl of the diluted samples were applied to the wells on the slides and then incubated at 37 °C for 30 minutes. PBS was used to wash the slides three times for 5 minutes each wash in room temperature and then they were washed with distilled water one time for 5 minutes. Anti-human conjugate IgG FITC as well as anti-human conjugate IgM FITC were diluted 1:100 in PBS and then 25 µl of these conjugates were added to the wells. Then, the slides were incubated again for 30 minutes. Thereafter, the slides were washed again with PBS three times for 5 minutes each wash and one time with distilled water for 5 minutes. Ultimately, 3 drops of 5 µl of mounting medium (50% glycerol) were added to the slides and covered by coverslips. The slides were stored at 4 °C until they were scanned with a fluorescence microscope

IgM was detected similarly except that GullSORB was added to the sample to inactivate the IgG and the first incubation was extended to 3 h.

3.3. RNA extraction

Viral RNA were extracted from serum and CSF samples (140 µl/sample) using of QIAamp Viral RNA Mini kit (Qiagen, Hilden, Germany). The extraction was conducted according to the manufacturer's instructions.

3.4. One step RT-PCR

The RT-PCR reaction was accomplished with the SuperScript II One-Step RT-PCR system with Platinum Taq High Fidelity (Invitrogen). The cDNA synthesis from serum and CSF samples was done in a single tube reaction in 25 µL. The reaction mixture was prepared in a clean-lab environment to avoid contamination. The reaction mixture included 12.50 µl of SuperScript II buffer (250 mM Tris-HCl, pH 8.3 at room

temperature; 375 mM KCl; 15 mM $MgCl_2$), 0.25 µl of $MgSO_4$ (5mM), 0.40 µl of the primers, 0.80 of enzyme mix and 10 µl of 2 ng RNA sample. Primer sequences are given in (Table 1). Sterile water was included in the PCR as negative control. The RT-PCR cycling program had the following parameters: 30 min at 50 °C to generate cDNA, 2 min at 94 °C to denature the cDNA, followed by 45 cycles of 20 s at 94 °C, 30 s at 55 °C to anneal the new copies of the cDNA, 60 s at 72 °C. Finally the reaction was incubated at 72 °C for 5 min. The size of the PCR products were validated in 2% agarose gel-electrophoresis including Tris base, acetic acid and EDTA (TAE) buffer and 2µl of 10mg/ml Gel Red nucleic acid stain to visualize the PCR products.

Table 1. Primers used for the cDNA synthesis in RT-PCR. The primers arranged in 5'-3' direction according to their ID numbers. F* refers to forward primer direction. R* refers to reverse primer direction.

Oligo ID	Oligo name sequence 5' – 3'	Orientation
180208X44A05 1/4	LVL_3359D_Y+ AGA ATC AGT GAA AGG GAA AGC AAY TC	F*
180208X44A06 2/4	LVL_3359G_Y+ AGA ATT AGT GAA AGG GAG AGT AAY TC	F
180208X44A07 3/4	LVL_3754A_R- CAC ATC ATT GGT CCC CAT TTA CTA TGR TC	R*
180208X44A08 4/4	LVL_3754D_R- CAC ATC ATT GGT GGG GAT TTA CTG TGR TC	R

3.5. Detection of *Arenaviruses* by amplicon PCR and sequencing

Amplicon PCR targeted the L gene of *Arenaviruses* using two primers sequences. The primers were overhanged with adapters to facilitate the comptability with the illumina indexes. The produced amplicon composed of 400 bp as maximum. Black sequence is the adapter and the red sequence is the interest region of LCMV genome as follow: DNA adapt_LVL_3359D_Y+ TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGA ATC AGT GAA AGG GAA AGC AAY TC, DNA adapt_LVL_3359G_Y+ TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGA ATT AGT GAA AGG GAG AGT AAY TC, DNA adapt_LVL_3754A_R- GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAC ATC ATT GGT CCC CAT TTA CTA TGR TC, DNA adapt_LVL_3754D_R- GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAC ATC ATT GGT GGG GAT TTA CTG TGR TC

As recommended by Illumina, the resulted amplicon of the targeted region L-segment of LCMV contained approximately 50-bp overlapping sequence in the middle of the target

amplicon. The purpose is to permit the overlaping of the bases at the end of each read. See figure (9) shows the principle work of amplicon PCR.

Regardless the adapters, the designed primers have a melting temprature (Tm) about 60°C, as determined using the available online tool (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/).

The thermal cycler was programmed to the following conditions: 30 min at 50 °C, 2 min at 94 °C to denature the DNA template to be two separated strands, followed by 45 cycles of 20 seconds at 94 °C, 30 seconds at 55 °C. This cooling in temperature permit the primer linking to the DNA template. The temperature was raised again to 72 °C for 60 seconds to permit the polymerase activating, thus adding new nucleotides. The final incubation was conducted at 72 °C for 5 min.

Additionally, the amplicon PCR products were validated using Agarose gel 1.5% including Gel Red stain.

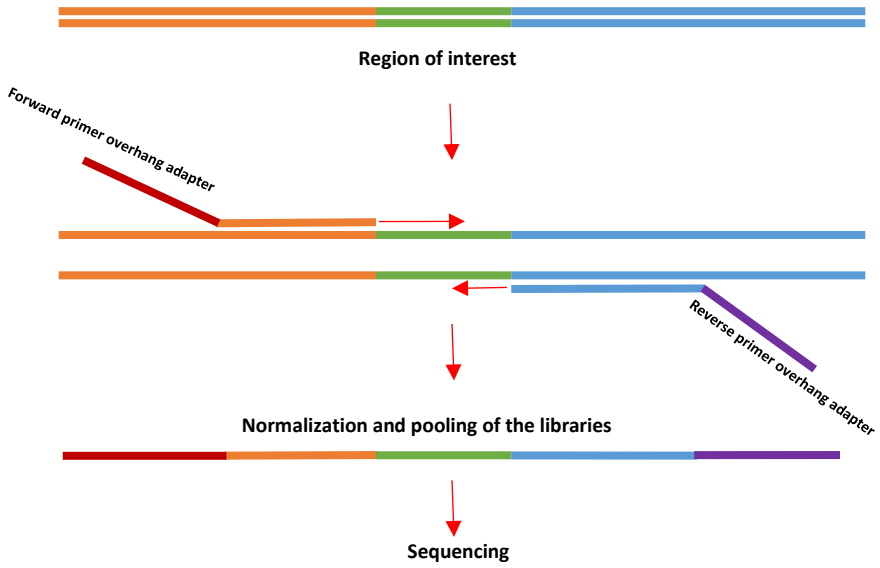

Figure 10: The construction of amplicon PCR using primers deisgned with overhang adapters. The primers are complementary to the upstreamand downstream of the region of interest, where the adapters are complementary of the Illumina indexes. Libraries are normalized and pooled before the performance of sequencing on MiSeq system.

After the identification of short sequences (Maximum 400 bp) of LCMV L-segment by amplicon PCR and MiSeq, the samples were underwent to partial genome sequencing of L-segment of LCMV.

3.6. Library preparation

Firstly, the target was to identify only short sequences of LCMV *Mammarenavirus*. Therefore, I used amplicon PCR technique which allows to attach illumina sequencing adapters at the end of the used primers using Nextera XT index Kit. This pathway facilitated the recognition of the joined pairs by MiSeq machine. Figure 11 illustrates the workflow of the library preparation. However, the libraries were washed twice Ampure XP beads. One time using to purify the amplicon of free primers and primer dimer species, and second time after adding the indexes to clean up the final library before quantitation.

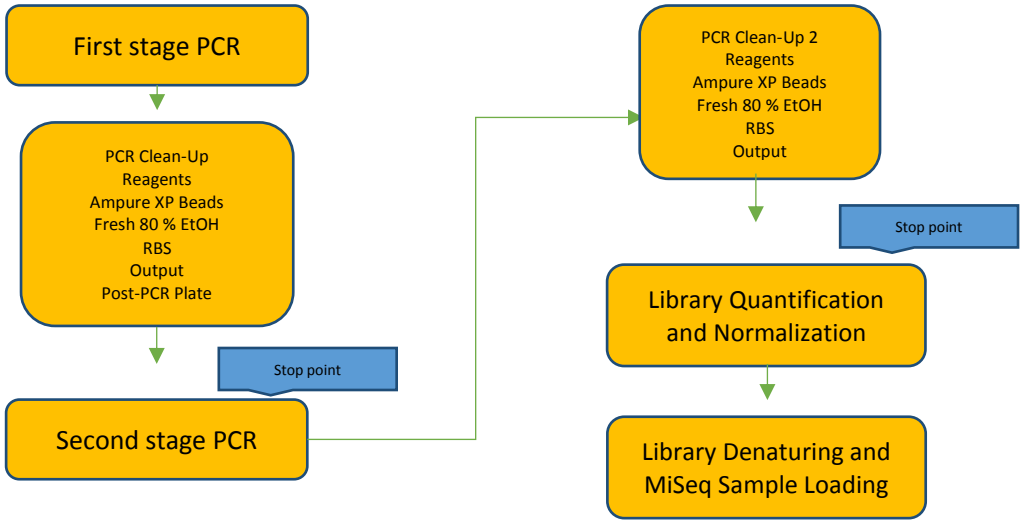

Figure 11: The workflow of library preparation. Safe stop points included in the work are mentioned in blue beetween steps.

3.6.1. DNA library quantification by qPCR

DNA library quantification was performed by SYBR green based qPCR using the New England Biolabs library quantification kit. First, the libraries were diluted to 1:1000, 1:10 000 and 1:100 000 in 10 mM Tris-HCL, pH 8.0. The master mixture was prepared including 12.4 µl 2x qPCR Master Mix, 10x primer premix, 50x ROX, 3.6 µl of RNAse free-water and 4µl of sample or standard, accordingly. Subsequently, the qPCR cycling cnditions were programmed as follow: one cycle of initial denaturation at 95 °C for 1 minutes followed by 35 cycles at 95 °C for 15 seconds and 63°C for 45 seconds, respectively Finally, the libraries were denatured with NaOH before the introducing for MiSeq. The prepared libraries were pooled into 4 nM solution.

3.7. Genome sequencing based on Metagenomics approach (ApoH-MiSeq)

The process was started with the centrifugation of the serum and CSF samples for 3 min at 17000g. 150 µl of the supernatant was taken from each sample. Thereafter, the supernatants were filtrated using 0.8 µm (PES) filter at 17000g for 1 minute. The filtration was done to reduce the number of host cells and potential contaminating bacteria. 130 µL of supernatant containing RNA was digested using 7 µL of 20x buffer (1M Tris, 100 mM $CaCl_2$ and 30 mM $MgCL_2$, pH 8) and nuclease treatment was performed by adding 2 µL of benzonase and 1 µL of micrococcal nuclease. Nuclease treatment reduces free RNA and DNA that are outside of viral capsid or membrane.

The content was incubated for 2h at 37˚C. To stop the reaction, 7 µL of 10 nM EDTA was added. The virus was further enriched using 10 µl of ApoH which was added for each sample following by incubation for 30 min at 8°C under appropriate agitation. I proceeded to RNA extraction according to the manufacturer's instruction using QIAamp Viral RNA Mini kit (Qiagen, Hilden, Germany).

Semi-random amplification was performed using a modified WTA2 (Complete Whole Transcriptome Amplification Kit) protocol according to the manufacturer's instructions (Sigma Aldrich). In this protocol, the primers designed to contain random part and universal part, the random part attaches to RNA/DNA and the universal part is used for amplification. I purified the PCR products using GeneJET PCR purification kit from Thermo Fisher Scientific using the manufacturer's instructions. Subsequently, I have measured the PCR products concentration using Quant-iT™ Broad-Range dsDNA Assay Kit. The measurement of PCR products concentration was done by Qubit device. The library preparation was conducted using an adjusted protocol for the Nextera XT DNA sample preparation kit from Illumina. The workflow of library preparation is characterized in Figure 6.

3.8. Sequence determination, molecular characterization and phylogenetic analysis

Sequences analysis included multiple steps based on the sequence library recorded from the Miseq sequencing. Reads were demultiplexed, adapter sequences were removed, and sample FASTQ files were produced with the MiSeq reporter. Bases with low Q score (less than 30) together with short sequences (less than 20) bp were removed using Trimmomatic program (Bolger, Lohse, & Usadel, 2014).

De novo assembler embedded in the Velvet program (Zerbino & Birney, 2008) was utilized to perform the assembly of the sequences. The sequences were produced from

two directions and then overlapping consensus sequences were combined for editing and analysis. Clustal W algorithm (Thompson, Higgins, & Gibson, 1994) was used to align the consensus sequences along with the LCMV strains obtained from GenBank. Clustal W aligment tool is accessible online: https://embnet.vital-it.ch/software/ClustalW.html.

Sequences of the conducted strains were compared to the published LCMV strains in GenBank by Blast (Altschul, Gish, Miller, Myers, & Lipman, 1990) which is available online: https://blast.ncbi.nlm.nih.gov/Blast.cgi. The comparison was restricted to the L-segment only.

Measuring of the evolutionary rate and constructing of the phylogenetic tree was performed using MEGA6 (Koichiro Tamura, Stecher, Peterson, Filipski, & Kumar, 2013), where multiple sequence alignments were performed using Clustal W. The best model to estimate the number of nucleotide substitutions was Tamura T92+G+I (K. Tamura, 1992) with pairwise gap deletion. The phylogenetic tree was constructed using maximum-likelihood method with 100 bootstrap replicates. Bootstrap values greater than 70 are considered as significant support for clusters in phylogenetic trees.

4. Results

4.1. IFA results revealed the seroprevalence of LCMV in Southern Iraq

IFA was performed to investigate the seroprevalence of LCMV *Mammarenavirus* among the communities present in the study site. The observation of LCMV positive samples was done by fluorescence microscope, where the the cells stained with FITC-conjugated secondary antibody, see Figure 12. Seropositive samples were detected among both healthy individuals and patients with suspected virus infections. Samples collected from healthy individuals had higher IgG seroprevalence compared to the suspected individuals, 12.22% (95% CI 6-20%) and 7.36% (95% CI 3-9) (Chi-squared test P=0.6) respectively (Figure 13 A).

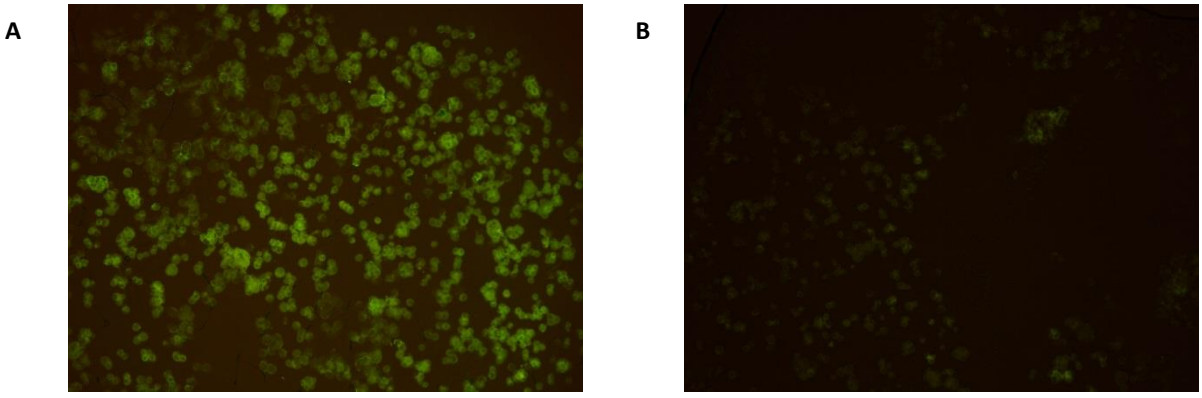

Figure 12: Observation of the LCMV particles in IFA. A) Represents positive LCMV particles. Primary antibodies were detected with specific secondary antibody: anti-human IgG-FITC (green). The cells were distributed equivalently among the whole well, thus negative and positives cells were shown closely. B) Represents negative control where no antibodies against LCMV particles were detected.

There were no distinguishable difference in seropositivity among individuals living in the urban compared to individuals living in rural regions. IFA results revealed that 12.50% (95% CI 3-13%) of the urban districts residents were seropositive but the rural regions residents showed a slightly lower seropositivity rate of 12.12% (Chi-squared test P = 0.5) (Figure 13 B). Among the gender group, 7.22 % (95% CI 3-10%) of females were seropositive in the suspected samples, compared to 6.10 % (95% CI 2-8) of the males (Chi-squared test P = 0.8). In contrast, the seropositivity of healthy females 5.74% (95%

CI 2-12%) was slightly less than healthy males 6.8% (95% CI 3-14%) (Chi-squared test P = 0.02). in the same category (Figure 13 C).

Table 2: The total IgG and IgM seroprevalence in serum samples.

IFA group	Positive	Negative
IgG	23	266
IgM	2	197

Moreover, most of the age groups of healthy individuals, except the 81-90 years old, contained seropositive individuals. On the other hand, only age groups 21-40 and 61-80 ages had seropositive samples among the suspected individuals (Figure 13 D). The suspected samples were screened for both IgM and IgG, while healthy samples were screened only for IgG. There were only two IgM positive samples, see (Table 2).

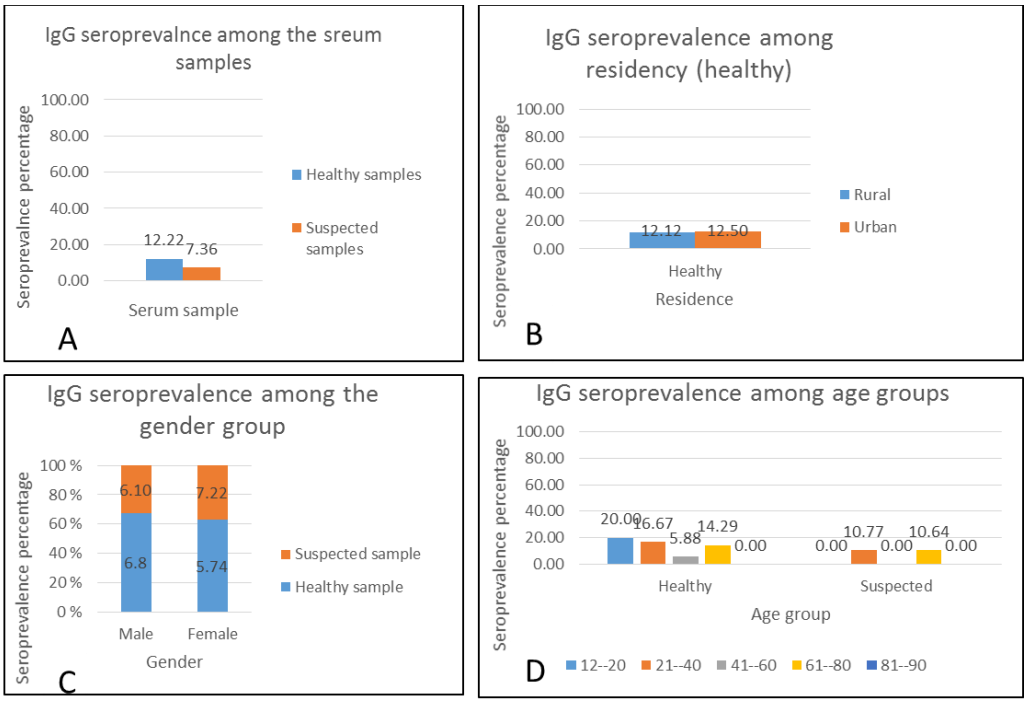

Figure 13: The results of immunofluorescence assay of serum samples. A: represents the seroprevalence among the healthy and suspected serum samples. B: represents the seroprevalence of LCMV in healthy people according to their residence. C: represents the seroprevalence of LCMV in the sense of gender group (female/male). D: represents the seroprevalence of LCMV among the age groups.

4.2. Detection of LCMV *Mammarenavirus* by RT-PCR and amplicon PCR

To identify the presence of infectious agent *Arenavirus*, primers targeting only *Arenaviruses* were used in RT-PCR. These primers facilitated the indication of *Arenaviruses* from the extracted RNA after converting to cDNA. Indeed, 22 samples (Table 4) were detected in gel-electrophoresis as positive for *Arenaviruses*.

Table 4: The results of RT-PCR and amplicon PCR.

PCR	Negative	Positive
RT-PCR	94	22
Amplicon PCR	0	22

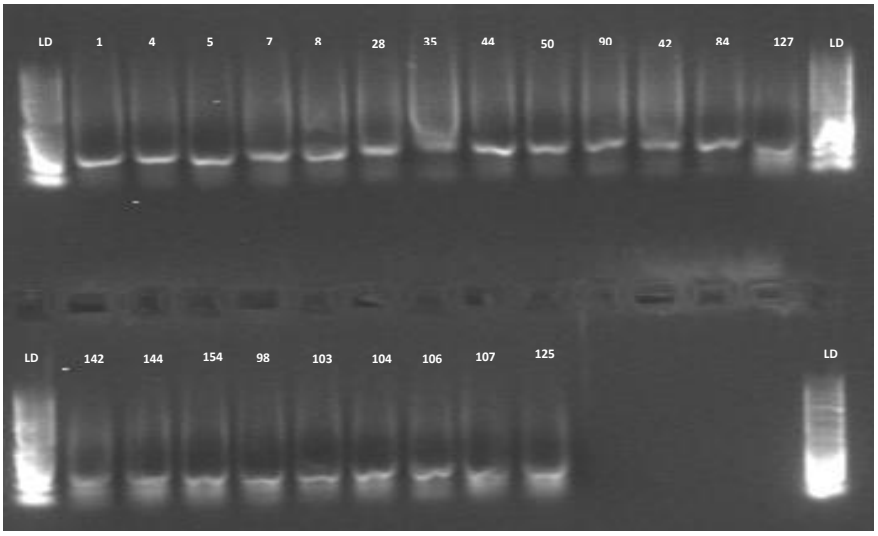

Figure 14. Gel electrophoresis of the 22 amplicon PCR products. LD represents the Ladder 50 bp. The numbers represent the samples.

4.3. Sequencing of short L-segment sequences indicated *Mammarenavirus*

In order to confirm the RT-PCR findings, the positive samples were first subjected to amplicon sequencing. The target region for this PCR was the same target as in RT-PCR, but the primers had, in addition to arenavirus-specific sequence, overhanging (Illumina Nextera) adapter sequences. In the agarose gel-electrophoresis the generated products' sizes were approximately 200-400 bp. Figure 14 shows comparison to the positive control of LCMV. A total of 22 samples showed amplicon PCR products.

Next step was to sequence the positive amplicon PCR products by NGS using Illumina MiSeq platform Figure 14). Libraries were built by adding the indexes to the 22 positive PCR products. Library concentrations were determined by real-time PCR using the primers of the indexes Figure 15. Based on the Real-time PCR results, all 22 libraries contained DNA. Then, the libraries were pooled into four groups and the concentration of the pools were measured using Qubit. According to the results, the concentration range was significantly high, 8.64 ng/µl, and should be diluted. Therefore, the concentration of the pooled libraries were diluted to 4nM as a final concentration.

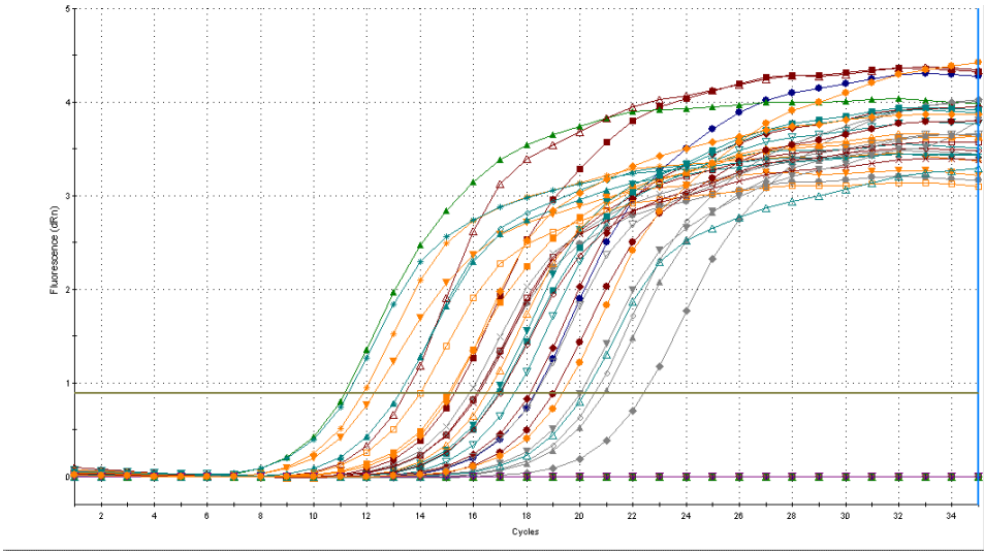

Figure 15: libraries plots. The figure contains six positive control (first three green lines and last three grey lines), one negative control and 22 libraries.

Thereafter, the pooled libraries were introduced to MiSeq sequncer to get short sequences of L-segment. The length of a single sequence was as maximum as 400 bp. The sequencing run had a good quality (Q) score indicating that most of the sequence reads passed the quality filter, see Figure 16.

BLAST analyses were performed of those clones shown similarities to LCMV and/or LASV. The identification results of the produced clones referred to 40-95% to *Mammarenavirus*es especially LCMV. Table 3 shows an example of blast results.

Table 3: Blast results of the some sequenced libraries.

	Max score	**Total score**	**Query cover**	**E value**	**Ident.**	**Accession**
Lymphocytic choriomeningitis virus strain 811316 segment L, complete sequence	233	233	77%	3e-57	81%	FJ607022.1
Lymphocytic choriomeningitis virus Z protein and L protein genes, complete cds	231	231	76%	9e-57	81%	AF004519.1
Lymphocytic choriomeningitis virus polymerase gene, partial cds	145	145	43%	1e-30	83%	M64451.1
Lassa virus L segment, strain AV, genomic RNA	86.1	86.1	66%	7e-13	73%	FR832710.1
Lassa virus strain AV RING finger protein (Z) and polymerase (L) genes, complete cds	86.1	86.1	66%	7e-13	73%	AY179171.1

4.4. Metagenomic Approach for sequencing

Since the BLAST results from partial L segment suggested that the LCMV strains from Iraq are highly divergent from the other LCMV strains (hampering the primer design for complete genome sequencing), a metagenomic approach directly from the serum samples was attempted. However, no additional virus sequences were detected with this approach.

Collectively, IFA screening of LCMV, detection of viral RNA by RT-PCR, amplification and using of sequencing adapters by amplicon PCR and sequencing analyses referred to the presence of LCMV. The short sequencing results of L-segment revealed that 12 samples were positive for LCMV.

Totally, all the PCR positive clones and *Arenavirus* sequences were obtained from serum samples rather than CSF (Table 5). CSF samples were not examined by IFA, but the serum samples. Furthermore, not all the IgM positive samples of IFA were positive in PCR.

Table 5: The LCMV screening resultsof and PCR performing for the samples.

Sample type	IFA IgG +/-	IFA IgM +/-	PCR +/-	MiSeq +/-
Serum	25/289	2/199	22	12
CSF	-	-	0	0

4.5. Phylogenetic analysis led to the Identification of two new strains of LCMV *Mammarenavirus*

A total of 56 sequences LCMV strains and one strain of Kodako virus as outgroup were obtained from GenBank. The phylogenetic tree was constructed on the basis of complete L-segment of available LCMV strains, partial L-segment of Kodako virus, and two strains sequenced from Iraq. Only two strains from Iraq were included in the analysis, since long enough sequence to contain phylogenetic information was obtained only from these two samples. Obviously, the new LCMV strains from Iraq clustered to LCMV species but they are distinct from the known LCMV strains. The high values of bootstrap (minimum as 70%) verifies the reliability of LCMV phylogenetic tree Figure 17.

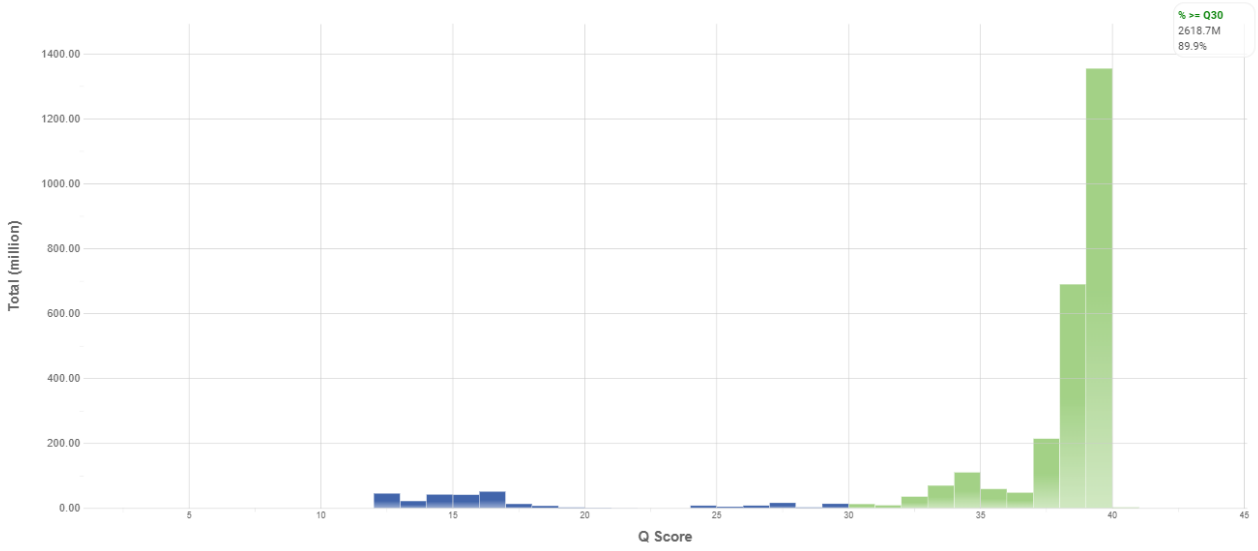

Figure 16: Illustrates the quality (Q) score distribution of the sequenced libraries was 89.9%.

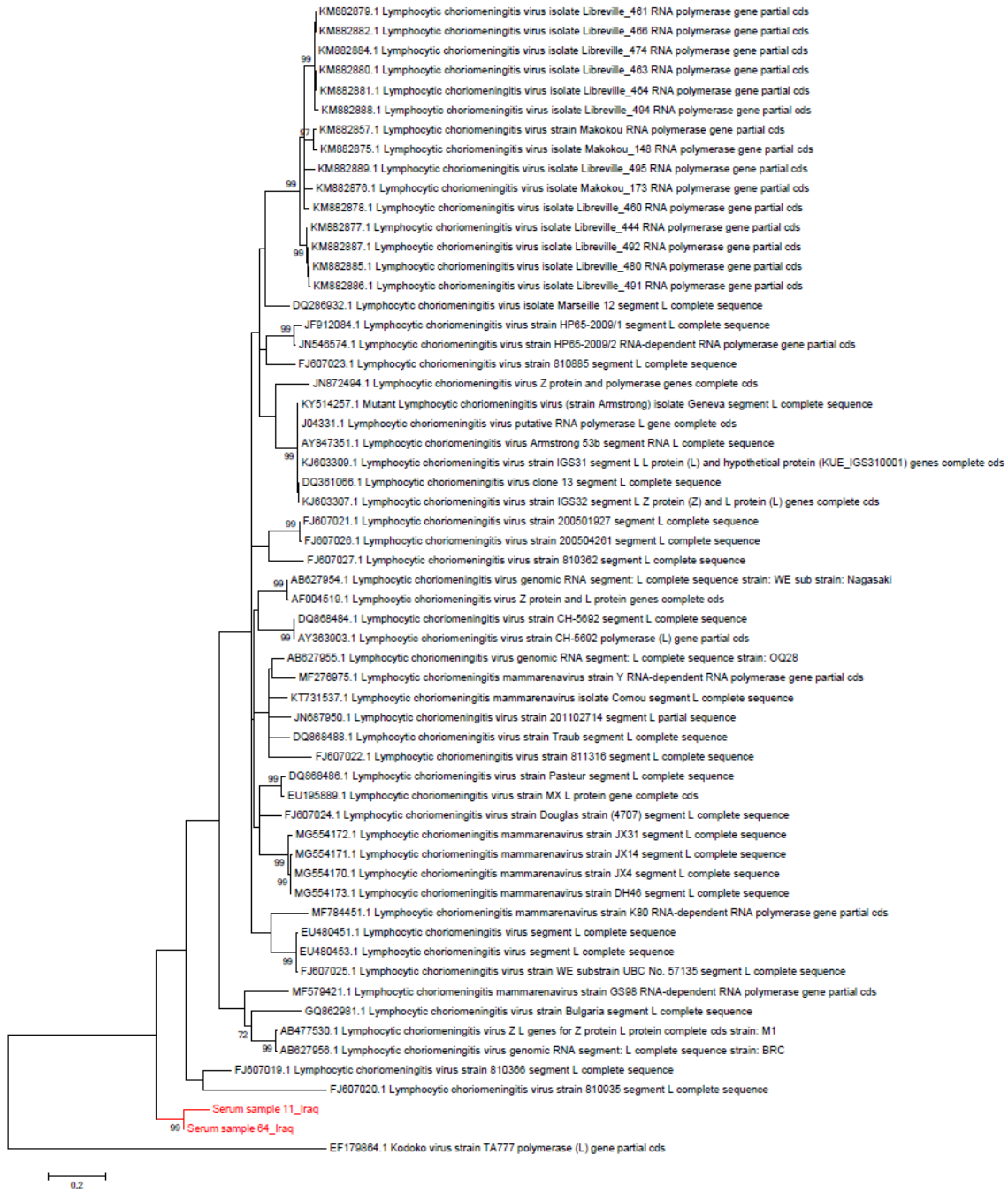

Figure 17: The phylogenetic tree was constructed with the MEGA6. Maximum likelihood tree constructed on the basis of complete L-segments of LCMV strains and partial L-segments of Kodoko virus and the LCMV strains isolated from Iraq. Bootstrap support values of >70 are shown at the nodes. The partial segments correspond to the sites 3706-4100 in the alignment of complete LCMV L-segments. The LCMV strains were compiled from the GenBank database. The newly sequenced LCMV strains from Iraq are in red. The scale bar represents 0.2 substitutions per amino acid positions.

5. Discussion

This study is the first comprehensive investigation of LCMV infection in the Southern Iraq.

LCMV is the prototypic *Mammarenavirus* that has been associated with the neurological disease LCM in rodents and humans (Bonthius, 2012). However, the prevalence and distribution of LCMV in the Southern Iraq were unknown. This study was performed on the basis of CSF and serum smaples collected from healthy and sick people from Nasiriyah city in the Southern Iraq to investigate the situation of LCMV. The present study is designed to give insights into the seroprevalence, genetic characterization, and evolution of LCMV in this region.

Serologically, LCMV antibodies were indicated in both healthy and sick people in this study using IFA. IgG is considered to be life-long, which suggests that most likely LCMV has circulated much longer in the area than sample collection time between 2012 and 2016. In the sense of recent exposure to the virus, evidence of IgM was observed in two serum samples (1.25%) only, whereas IgG was detected in 23 serum samples (8.6%). Old exposure to LCMV represented by 8.6% of IgG can be considered relatively high compared to seroprevalences range 1.7-4.6% reported previously in other countries (Lledo *et al*., 2003; Marrie & Saron, 1998; Park *et al*., 1997 ; Riera *et al*., 2005). There was a strong correlation between the PCR and IgM seropositive samples and patients presenting with neurological symptoms indicating that LCMV was the causative agent for these symptoms. However, additional studies should be conducted to rule out the other potential pathogenic agents.

Among all viral infections, there are much more non-symptomatic infections than symptomatic. The two IgM-positive samples indicate that there is active infection with LCMV in Southern Iraq currently, assuming that the patients were not infected whilst visiting other areas. These samples were checked twice to confirm the results. Deeper investigations into the seroprevalence of LCMV is needed. Concerning LCMV infection symptoms, most of LCMV infections are asymptomatic or may show mild symptoms and signs similar to the common cold and therefore most likely remain undiagnosed (Barton & Hyndman, 2000; Blasdell *et al.*, 2016). The similarity to common cold symptoms makes the diagnostic of LCMV infection challenging. For example, Baum *et al*. (1966) found that LCMV was the causative agent of a "common cold" in infected lab workers. However, in the current study, the virus was detected from patients suffering neurological symptoms, such as severe muscle weakness and others were combined with exhibited mild symptoms. Fever was noticed in all the patients and most of them were suffering from headache as shown in the Appendix where the symptoms and their distribution amongst the suspected patients is shown.

There was no significant statistical differences noticed with respect to the rural and urban areas. This is most likely due to the ecological similarities in rural and urban areas in the Southern Iraq which is characterized by drought. The house mouse *Mus musculus*

population is spreading significantly in Iraq and neighbouring countries (Suzuki *et al.*, 2013). This factor facilitates a high spreading of LCMV in this area among dwellings and may cause new outbreaks unless the vector distribution is controlled. Some previous studies have shown similar outcomes that the spread of the virus is relatively similar in rural and urban areas (Lledo, Gegundez, Saz, Bahamontes, & Beltran, 2003).

The study findings suggest that females are more likely to be infected with LCMV than males. This is indeed in agreement with previously published studies related to the identification of *Arenavirus*es (Marrie & Saron, 1998). Also, it is important to consider the high seropositivity in females, since LCMV has been mentioned as a teratogonic human pathogen (Barton, Peters, & Ksiazek, 1995). The association between the high seropositivity and females is, so far, not clear (Lledo *et al.*, 2003). A major contributor to this may be, that women are involved in cleaning the household in these communities, therefore they be more exposed to rodent excreta than men.

Regarding the age groups, it's known that LCMV is a potential agent to infect all human ages including fetuses (Bonthius, 2012). The study results showed that people between 12-60 years old, but not older than 60, were infected with LCMV. Our results suggested that people less than 40 years old have had more LCMV infections compared to other age groups. If all age groups were similarly exposed to the virus, one would expect a gradual increase in the IgG seroprevalence towards the older age cohorts. However, such a phenomenon was not observed in the current study. This may implicate that the virus has become more common. Another potential reason would be waning IgG levels over time. This could be further studied by measuring the IgG levels instead of merely the presence of IgGs. This could be investigated by making dilution series of the samples and determining the cut-off titer for antibodies. According to this hypothesis, young people are considered at high risk of LCMV infection in the presence of viral transmitting vectors.

Another consideration in this study is related to the viral genome. Many studies have targeted L-segment of *Arenavirus*, because it is correlated with pathogenicity and some studies described it as a virulence factor in mice (Riviere & Oldstone, 1986). In this study, the sequencing was set to target L-segment of LCMV, because polymerase regions are more conserved among the virus species. Sequencing outcomes suggest that short sequences were detected of *Mammarenaviruses* and the majority of sequences were closely related to LCMV. Subsequently, metagenomics sequencing directly from the sample (i.e. without sequence-specific amplification) was attempted from confirmed PCR positive samples, but no additional LCMV sequences were obtained. This was not unexpected, since viral sequences represent only a very small proportion of total RNA of the sample. In order to enrich the viral RNA, the preparation of the libraries included several steps that should concentrate the virus. Therefore, ApoH was worth using. The goal of using ApoH is to concentrate and protect it during the sample treatment.

Additionally, the samples underwent nuclease treatment in order to degrade free nucleic acids, such as those that are not protected by viral capsid. Also, I had used precise filters with pore size that permits the viral particles but not bacterial or eukaryotic cells to pass

through it. Further attempts should be conducted either by increasing the sequencing depth or by enriching the virus genome by PCR (Quick *et al.*, 2017) or virus-specific oligonucleotide capture based methods.

Through sequencing, I have noticed some unexpected sequences associated with bacteriophages, bacteria and viruses. Obviously, some of those sequences resulted due to the contaminations that are commonly found in nucleotide extraction and PCR kits, and some may represent 'normal microbiota', such as *anelloviruses* that are very frequent in blood samples of healthy people. All positive results from partial genome sequencing were derived from serum samples but not from CSF samples. Despite the fact that RNA was extracted from all the CSF samples and also they were tested by amplicon PCR I could not detect any positive samples. The potential reason is the low RNA concentration in the PCR reaction.

Importantly, Q score, which is the indicator that reflects the sequencing performance and the quality, revealed a significant value of the run 89.9%. According to Illumina instructions, Q 30 quality score is considered a high value of run performance. Also, it refers to reliable sequencing in which only one base call in 1000 is probably incorrect. Therefore, to reassure that the sequences are correct, the bases with quality score lower than 30 were removed before assembly.

Phylogenetic viral genome analysis suggested that the LCMV strains sequenced from Iraq form a distinct LCMV cluster. However, since the available sample set was limited to only a few years and to a restricted geographical region in Southern Iraq, further studies should be conducted to characterize the diversity of LCMV strains in Middle East region. In the partial sequencing, I detected 12 positive samples of LCMV, however, high quality sequences were obtained only for 2 novel strains shown in the phylogenetic tree. Furthermore, during RT-PCR screening products of right size was detected from 22 samples, whereas virus sequences were detected only from 12 of these. Most likely the other 10 samples were false positive. Sometimes the primers anneal to the host genome which it leads to false results. Therefore, improving and developing of primers and specific PCR for LCMV and Arenaviruses in general is needed.From the epidemiological viewpoint, the identification of LCMV infection was expected, but nevertheless, this is the first unambiguous detection of novel LCMV strains in Southern Iraq. The results from this study suggest that human infection with LCMV occurs widely throughout Southern Iraq, and probably result in LCM disease in the resident people.

The hypothesis of this study mirrors the observation of known disease, due to the association between the LCMV and rodents that spread in the study site. Based on these results, the spreading of LCMV in Southern Iraq is confirmed. These results should be taken into considerations in the next studies related to LCMV. See Figure 18 where the updated distribution of LCMV is shown.

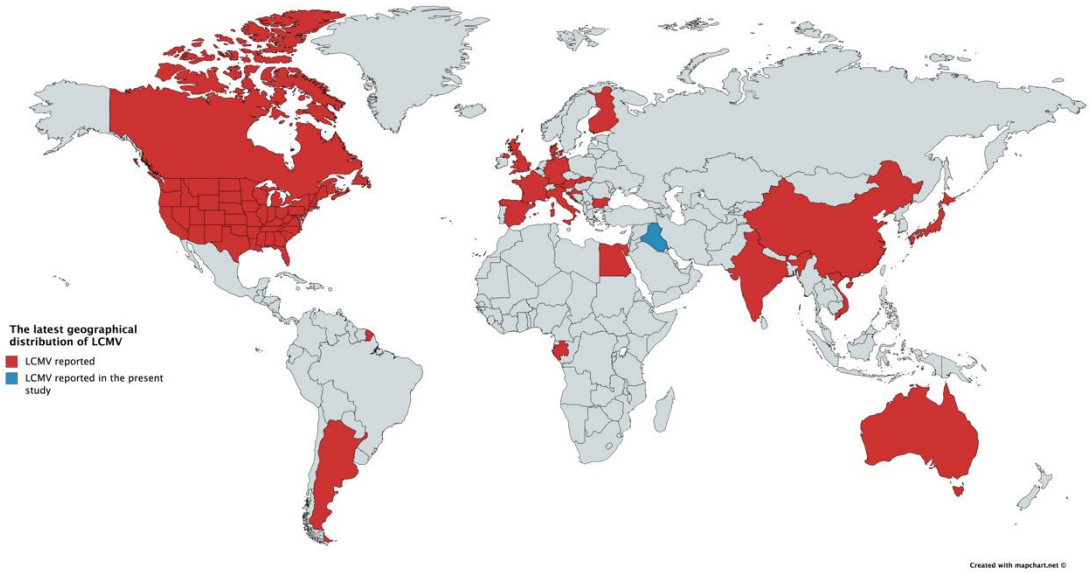

Figure 18: The latest geographical distribution of NW and OW *Mammarenavirus*es. The distribution of LCMV in Southern Iraq is mentioned in blue for the first time. The LCMV distribution rate was gained from (Pontremoli *et al.*, 2018). Map was created with mapchart.net.

6. Conclusion

In this work I focused on the seroprevalence and evolution of LCMV in the Southern Iraq, to extend the knowledge about LCMV. The seroprevalence was detected in sick and healthy individuals which is given me a significant signal of the presence of LCMV. As known, LCMV infects adults of different age groups in addition to the infants, but in this study, I have proved that most people less than 81 years old have probably been infected with LCMV. In addition, I characterized two novel LCMV strains that were circulating in Southern Iraq. These new strains were highly divergent from the previously known LCMV strains. There are surprisingly little sequence data on LCMV despite worldwide distribution. Therefore, more sequences are needed to gain information on the molecular epidemiology and evolution of LCMV. For example, phylogeographical studies could be performed to assess the dispersal of the virus from one geographic location to another. Therefore, the study should be extended to include other regions in Southern Iraq not only restricted in the Nasiriyah city. Also, whole genome sequencing of both L- and S-segments is quite necessary to confirm the evolution rate of LCMV. In addition, improving the tools using to study LCMV are very important. For example, designing of precise primers targeting both L- and S-segments will be useful to gain accurate results. Furthermore, isolating more LCMV strains in tissue culture would be useful e.g. for comparative studies on potential phenotypic differences between the strains. .

Regarding the transmission, the common house mouse *Mus musculus* is considered to be the main LCMV reservoir (Ackermann, Bloedhorn, Kupper, Winkens, & Scheid, 1964; Childs *et al.*, 1992; Morita *et al.*, 1996). However, the virus has been commonly detected also in other rodents. Therefore, targeting the precise reservoirs of LCMV in Southern Iraq is needed. In order to plan vector control and public health recommendations it is very important to understand the nature of LCMV reservoirs. That is could be done through examine rodent samples.

The interaction between viral proteins and host receptors influences the signs and symptoms manifestations (Takagi *et al.*, 2017). However, this kind of interactions was not investigated during this study. But understanding how chronic viruses interact with the host receptor probably it helps to develop a new antiviral therapeutic against LCMV.

Finally, samples from pregnant women and immuno-compromised individuals are necessary to evaluate the severity of disease in these target groups (Bonthius, 2012; Price, Fisher-Hoch, Craven, & McCormick, 1988).

7. ACKNOWLEDGEMENT

This study was done in the unit of viral zoonosis leading by Professor Olli Vapalahti.

I would also to thank the experts and main supervisors who were involved in the supervision of this thesis, Dr. Haider Al-Hello (International institute of health and welfare THL, health safety department) and Dr. Teemu Smura (Medicum-Department of virology of the Faculty of Medicine, University of Helsinki). The thesis work could not have been successfully conducted without their passionate participation and input.

Also, I would like to express my gratitude to my minor supervisor Professor Sarah Butcher (The Faculty of Biological and Environmental Sciences at the University of Helsinki). The door to her office was always open whenever I needed consultation and she was always ready to answer my questions. She consistently allowed me to work independently, steering me in the right direction whenever she thought I needed it.

I would also like to acknowledge Professor Olli Vapalahti, I am gratefully indebted to him for his support and encouragement during the thesis work.

Finally, I must express my very profound gratitude to all members of my family for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author

Hussein Alburkat

8. References

Ackermann, R., Bloedhorn, H., Kupper, B., Winkens, I., & Scheid, W. (1964). [Spread of the lymphocytic choriomeningitis virus among West German mice. I. Investigations mostly on domestic mice (Mus musculus)]. *Zentralbl Bakteriol Orig, 194*(4), 407-430.

Adlhoch, C., Kaiser, M., Hoehne, M., Mas Marques, A., Stefas, I., Veas, F., & Ellerbrok, H. (2011). Highly sensitive detection of the group A Rotavirus using Apolipoprotein H-coated ELISA plates compared to quantitative real-time PCR. *Virol J, 8*, 63. doi:10.1186/1743-422X-8-63

Albarino, C. G., Palacios, G., Khristova, M. L., Erickson, B. R., Carroll, S. A., Comer, J. A., Hui, J., Briese, T., St George, K., Ksiazek, T. G., Lipkin, W. I., & Nichol, S. T. (2010). High diversity and ancient common ancestry of lymphocytic choriomeningitis virus. *Emerg Infect Dis, 16*(7), 1093-1100. doi:10.3201/eid1607.091902

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol, 215*(3), 403-410. doi:10.1016/s0022-2836(05)80360-2

Baker, M. (2012). De novo genome assembly: what every biologist should know. *Nature Methods, 9*, 333. doi:10.1038/nmeth.1935

Barton, L. L., & Hyndman, N. J. (2000). Lymphocytic choriomeningitis virus: reemerging central nervous system pathogen. *Pediatrics, 105*(3), E35.

Barton, L. L., Peters, C. J., & Ksiazek, T. G. (1995). Lymphocytic choriomeningitis virus: an unrecognized teratogenic pathogen. *Emerging Infectious Diseases, 1*(4), 152-153. doi:10.3201/eid0104.950410

Bausch, D. G., Rollin, P. E., Demby, A. H., Coulibaly, M., Kanu, J., Conteh, A. S., Wagoner, K. D., McMullan, L. K., Bowen, M. D.,

Peters, C. J., & Ksiazek, T. G. (2000). Diagnosis and clinical virology of Lassa fever as evaluated by enzyme-linked immunosorbent assay, indirect fluorescent-antibody test, and virus isolation. *J Clin Microbiol, 38*(7), 2670-2677.

Beyer, W. R., Popplau, D., Garten, W., von Laer, D., & Lenz, O. (2003). Endoproteolytic processing of the lymphocytic choriomeningitis virus glycoprotein by the subtilase SKI-1/S1P. *J Virol, 77*(5), 2866-2872.

Blasdell, K. R., Duong, V., Eloit, M., Chretien, F., Ly, S., Hul, V., Deubel, V., Morand, S., & Buchy, P. (2016). Evidence of human infection by a new mammarenavirus endemic to Southeastern Asia. *Elife, 5*. doi:10.7554/eLife.13135

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120. doi:10.1093/bioinformatics/btu170

Bonthius, D. J. (2009). Lymphocytic choriomeningitis virus: a prenatal and postnatal threat. *Adv Pediatr, 56*, 75-86. doi:10.1016/j.yapd.2009.08.007

Bonthius, D. J. (2012). Lymphocytic choriomeningitis virus: an underrecognized cause of neurologic disease in the fetus, child, and adult. *Semin Pediatr Neurol, 19*(3), 89-95. doi:10.1016/j.spen.2012.02.002

Bonthius, D. J., & Karacay, B. (2002). Meningitis and encephalitis in children. An update. *Neurol Clin, 20*(4), 1013-1038, vi-vii.

Bonthius, D. J., Wright, R., Tseng, B., Barton, L., Marco, E., Karacay, B., & Larsen, P. D. (2007). Congenital lymphocytic choriomeningitis virus infection: spectrum of disease. *Ann Neurol, 62*(4), 347-355. doi:10.1002/ana.21161

Buchmeier, M. J., Lewicki, H. A., Tomori, O., & Oldstone, M. B. (1981). Monoclonal antibodies to lymphocytic choriomeningitis and pichinde viruses: generation, characterization, and cross-reactivity with other arenaviruses. *Virology, 113*(1), 73-85.

Buckley, S. M., & Casals, J. (1970). Lassa fever, a new virus disease of man from West Africa. 3. Isolation and characterization of the virus. *Am J Trop Med Hyg, 19*(4), 680-691.

Calisher, C. H., Tzianabos, T., Lord, R. D., & Coleman, P. H. (1970). Tamiami virus, a new member of the TaCaribe group. *Am J Trop Med Hyg, 19*(3), 520-526.

Charrel, R. N., de Lamballerie, X., & Emonet, S. (2008). Phylogeny of the genus Arenavirus. *Curr Opin Microbiol, 11*(4), 362-368. doi:10.1016/j.mib.2008.06.001

Childs, J. E., Glass, G. E., Korch, G. W., Ksiazek, T. G., & Leduc, J. W. (1992). Lymphocytic choriomeningitis virus infection and house mouse (Mus musculus) distribution in urban Baltimore. *Am J Trop Med Hyg, 47*(1), 27-34.

Downs, W. G., Anderson, C. R., Spence, L., Aitken, T. H., & Greenhall, A. H. (1963). Tacaribe virus, a new agent isolated from Artibeus bats and mosquitoes in Trinidad, West Indies. *Am J Trop Med Hyg, 12*, 640-646.

Duffy, S., Shackelton, L. A., & Holmes, E. C. (2008). Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet, 9*(4), 267-276. doi:10.1038/nrg2323

Emonet, S. F., de la Torre, J. C., Domingo, E., & Sevilla, N. (2009). Arenavirus genetic diversity and its biological implications. *Infect Genet Evol, 9*(4), 417-429. doi:10.1016/j.meegid.2009.03.005

Fenner, F. (1976). Classification and nomenclature of viruses. Second report of the International Committee on Taxonomy of Viruses. *Intervirology, 7*(1-2), 1-115. doi:10.1159/000149938

Fischer, S. A., Graham, M. B., Kuehnert, M. J., Kotton, C. N., Srinivasan, A., Marty, F. M., Comer, J. A., Guarner, J., Paddock, C. D., DeMeo, D. L., Shieh, W. J., Erickson, B. R., Bandy, U., DeMaria, A., Jr., Davis, J. P., Delmonico, F. L., Pavlin, B., Likos, A., Vincent, M. J., Sealy, T. K., Goldsmith, C. S., Jernigan, D. B.,

Rollin, P. E., Packard, M. M., Patel, M., Rowland, C., Helfand, R. F., Nichol, S. T., Fishman, J. A., Ksiazek, T., & Zaki, S. R. (2006). Transmission of lymphocytic choriomeningitis virus by organ transplantation. *N Engl J Med, 354*(21), 2235-2249. doi:10.1056/NEJMoa053240

Foscaldi, S., D'Antuono, A., Noval, M. G., de Prat Gay, G., Scolaro, L., & Lopez, N. (2017). Regulation of Tacaribe Mammarenavirus Translation: Positive 5' and Negative 3' Elements and Role of Key Cellular Factors. *J Virol, 91*(14). doi:10.1128/JVI.00084-17

Frame, J. D., Baldwin, J. M., Jr., Gocke, D. J., & Troup, J. M. (1970). Lassa fever, a new virus disease of man from West Africa. I. Clinical description and pathological findings. *Am J Trop Med Hyg, 19*(4), 670-676.

Hepojoki, J., Hepojoki, S., Smura, T., Szirovicza, L., Dervas, E., Prähauser, B., Nufer, L., Schraner, E. M., Vapalahti, O., Kipar, A., & Hetzel, U. (2018). Characterization of Haartman Institute snake virus-1 (HISV-1) and HISV-like viruses—The representatives of genus Hartmanivirus, family Arenaviridae. *PLOS Pathogens, 14*(11), e1007415. doi:10.1371/journal.ppat.1007415

Hepojoki, J., Salmenpera, P., Sironen, T., Hetzel, U., Korzyukov, Y., Kipar, A., & Vapalahti, O. (2015). Arenavirus Coinfections Are Common in Snakes with Boid Inclusion Body Disease. *J Virol, 89*(16), 8657-8660. doi:10.1128/JVI.01112-15

Johnson, K. M., Wiebenga, N. H., Mackenzie, R. B., Kuns, M. L., Tauraso, N. M., Shelokov, A., Webb, P. A., Justines, G., & Beye, H. K. (1965). Virus Isolations from Human Cases of Hemorrhagic Fever in Bolivia. *Proc Soc Exp Biol Med, 118*, 113-118.

Kallio-Kokko, H., Laakkonen, J., Rizzoli, A., Tagliapietra, V., Cattadori, I., Perkins, S. E., Hudson, P. J., Cristofolini, A., Versini, W., Vapalahti, O., Vaheri, A., & Henttonen, H. (2006).

Hantavirus and arenavirus antibody prevalence in rodents and humans in Trentino, Northern Italy. *Epidemiol Infect, 134*(4), 830-836. doi:10.1017/S0950268805005431

Knust, B., Macneil, A., Wong, S. J., Backenson, P. B., Gibbons, A., Rollin, P. E., & Nichol, S. T. (2011). Exposure to lymphocytic choriomeningitis virus, New York, USA. *Emerg Infect Dis, 17*(7), 1324-1325. doi:10.3201/eid1707.101349

Knust, B., Stroher, U., Edison, L., Albarino, C. G., Lovejoy, J., Armeanu, E., House, J., Cory, D., Horton, C., Fowler, K. L., Austin, J., Poe, J., Humbaugh, K. E., Guerrero, L., Campbell, S., Gibbons, A., Reed, Z., Cannon, D., Manning, C., Petersen, B., Metcalf, D., Marsh, B., Nichol, S. T., & Rollin, P. E. (2014). Lymphocytic choriomeningitis virus in employees and mice at multipremises feeder-rodent operation, United States, 2012. *Emerg Infect Dis, 20*(2), 240-247. doi:10.3201/eid2002.130860

Ksiazek, T. G., Rollin, P. E., Williams, A. J., Bressler, D. S., Martin, M. L., Swanepoel, R., Burt, F. J., Leman, P. A., Khan, A. S., Rowe, A. K., Mukunu, R., Sanchez, A., & Peters, C. J. (1999). Clinical virology of Ebola hemorrhagic fever (EHF): virus, virus antigen, and IgG and IgM antibody findings among EHF patients in Kikwit, Democratic Republic of the Congo, 1995. *J Infect Dis, 179 Suppl 1*, S177-187. doi:10.1086/514321

Kunz, S., Campbell, K. P., & Oldstone, M. B. (2003). Alpha-dystroglycan can mediate arenavirus infection in the absence of beta-dystroglycan. *Virology, 316*(2), 213-220.

Leibler, J. H., Zakhour, C. M., Gadhoke, P., & Gaeta, J. M. (2016). Zoonotic and Vector-Borne Infections Among Urban Homeless and Marginalized People in the United States and Europe, 1990-2014. *Vector Borne Zoonotic Dis, 16*(7), 435-444. doi:10.1089/vbz.2015.1863

Lenz, O., ter Meulen, J., Klenk, H. D., Seidah, N. G., & Garten, W. (2001). The Lassa virus glycoprotein precursor GP-C is

proteolytically processed by subtilase SKI-1/S1P. *Proc Natl Acad Sci U S A, 98*(22), 12701-12705. doi:10.1073/pnas.221447598

Lischer, H. E. L., & Shimizu, K. K. (2017). Reference-guided de novo assembly approach improves genome reconstruction for related species. *BMC Bioinformatics, 18*(1), 474. doi:10.1186/s12859-017-1911-6

Lledo, L., Gegundez, M. I., Saz, J. V., Bahamontes, N., & Beltran, M. (2003). Lymphocytic choriomeningitis virus infection in a province of Spain: analysis of sera from the general population and wild rodents. *J Med Virol, 70*(2), 273-275. doi:10.1002/jmv.10389

Maes, P., Alkhovsky, S. V., Bào, Y., Beer, M., Birkhead, M., Briese, T., Buchmeier, M. J., Calisher, C. H., Charrel, R. N., Choi, I. R., Clegg, C. S., de la Torre, J. C., Delwart, E., DeRisi, J. L., Di Bello, P. L., Di Serio, F., Digiaro, M., Dolja, V. V., Drosten, C., Druciarek, T. Z., Du, J., Ebihara, H., Elbeaino, T., Gergerich, R. C., Gillis, A. N., Gonzalez, J.-P. J., Haenni, A.-L., Hepojoki, J., Hetzel, U., Hồ, T., Hóng, N., Jain, R. K., Jansen van Vuren, P., Jin, Q., Jonson, M. G., Junglen, S., Keller, K. E., Kemp, A., Kipar, A., Kondov, N. O., Koonin, E. V., Kormelink, R., Korzyukov, Y., Krupovic, M., Lambert, A. J., Laney, A. G., LeBreton, M., Lukashevich, I. S., Marklewitz, M., Markotter, W., Martelli, G. P., Martin, R. R., Mielke-Ehret, N., Mühlbach, H.-P., Navarro, B., Ng, T. F. F., Nunes, M. R. T., Palacios, G., Pawęska, J. T., Peters, C. J., Plyusnin, A., Radoshitzky, S. R., Romanowski, V., Salmenperä, P., Salvato, M. S., Sanfaçon, H., Sasaya, T., Schmaljohn, C., Schneider, B. S., Shirako, Y., Siddell, S., Sironen, T. A., Stenglein, M. D., Storm, N., Sudini, H., Tesh, R. B., Tzanetakis, I. E., Uppala, M., Vapalahti, O., Vasilakis, N., Walker, P. J., Wáng, G., Wáng, L., Wáng, Y., Wèi, T., Wiley, M. R., Wolf, Y. I., Wolfe, N. D., Wú, Z., Xú, W., Yang, L., Yāng, Z., Yeh, S.-D., Zhāng, Y.-Z., Zhèng, Y., Zhou, X., Zhū, C., Zirkel, F.,

& Kuhn, J. H. J. A. o. V. (2018). Taxonomy of the family Arenaviridae and the order Bunyavirales: update 2018. *163*(8), 2295-2310. doi:10.1007/s00705-018-3843-5

Maes, P., Alkhovsky, S. V., Bao, Y., Beer, M., Birkhead, M., Briese, T., Buchmeier, M. J., Calisher, C. H., Charrel, R. N., Choi, I. R., Clegg, C. S., de la Torre, J. C., Delwart, E., DeRisi, J. L., Di Bello, P. L., Di Serio, F., Digiaro, M., Dolja, V. V., Drosten, C., Druciarek, T. Z., Du, J., Ebihara, H., Elbeaino, T., Gergerich, R. C., Gillis, A. N., Gonzalez, J. J., Haenni, A. L., Hepojoki, J., Hetzel, U., Ho, T., Hong, N., Jain, R. K., Jansen van Vuren, P., Jin, Q., Jonson, M. G., Junglen, S., Keller, K. E., Kemp, A., Kipar, A., Kondov, N. O., Koonin, E. V., Kormelink, R., Korzyukov, Y., Krupovic, M., Lambert, A. J., Laney, A. G., LeBreton, M., Lukashevich, I. S., Marklewitz, M., Markotter, W., Martelli, G. P., Martin, R. R., Mielke-Ehret, N., Muhlbach, H. P., Navarro, B., Ng, T. F. F., Nunes, M. R. T., Palacios, G., Paweska, J. T., Peters, C. J., Plyusnin, A., Radoshitzky, S. R., Romanowski, V., Salmenpera, P., Salvato, M. S., Sanfacon, H., Sasaya, T., Schmaljohn, C., Schneider, B. S., Shirako, Y., Siddell, S., Sironen, T. A., Stenglein, M. D., Storm, N., Sudini, H., Tesh, R. B., Tzanetakis, I. E., Uppala, M., Vapalahti, O., Vasilakis, N., Walker, P. J., Wang, G., Wang, L., Wang, Y., Wei, T., Wiley, M. R., Wolf, Y. I., Wolfe, N. D., Wu, Z., Xu, W., Yang, L., Yang, Z., Yeh, S. D., Zhang, Y. Z., Zheng, Y., Zhou, X., Zhu, C., Zirkel, F., & Kuhn, J. H. (2018). Taxonomy of the family Arenaviridae and the order Bunyavirales: update 2018. *Arch Virol, 163*(8), 2295-2310. doi:10.1007/s00705-018-3843-5

Marrie, T. J., & Saron, M. F. (1998). Seroprevalence of lymphocytic choriomeningitis virus in Nova Scotia. *Am J Trop Med Hyg, 58*(1), 47-49.

Martin, V., Abia, D., Domingo, E., & Grande-Perez, A. (2010). An interfering activity against lymphocytic choriomeningitis

virus replication associated with enhanced mutagenesis. *J Gen Virol, 91*(Pt 4), 990-1003. doi:10.1099/vir.0.017053-0

Martinez-Sobrido, L., & de la Torre, J. C. (2016). Reporter-Expressing, Replicating-Competent Recombinant Arenaviruses. *Viruses, 8*(7). doi:10.3390/v8070197

Mets, M. B., Barton, L. L., Khan, A. S., & Ksiazek, T. G. (2000). Lymphocytic choriomeningitis virus: an underdiagnosed cause of congenital chorioretinitis. *Am J Ophthalmol, 130*(2), 209-215.

Mettler, N. E., Gianantonio, C. A., & Parodi, A. S. (1963). [Isolation of the causal agent of the hemolytic syndrome]. *Medicina (B Aires), 23*, 139-142.

Michele, D. E., & Campbell, K. P. (2003). Dystrophin-glycoprotein complex: post-translational processing and dystroglycan function. *J Biol Chem, 278*(18), 15457-15460. doi:10.1074/jbc.R200031200

Moreno, H., Gallego, I., Sevilla, N., de la Torre, J. C., Domingo, E., & Martin, V. (2011). Ribavirin can be mutagenic for arenaviruses. *J Virol, 85*(14), 7246-7255. doi:10.1128/JVI.00614-11

Morita, C., Tsuchiya, K., Ueno, H., Muramatsu, Y., Kojimahara, A., Suzuki, H., Miyashita, N., Moriwaki, K., Jin, M. L., Wu, X. L., & Wang, F. S. (1996). Seroepidemiological survey of lymphocytic choriomeningitis virus in wild house mice in China with particular reference to their subspecies. *Microbiol Immunol, 40*(4), 313-315.

Mueller, T. C., Siv, S., Khim, N., Kim, S., Fleischmann, E., Ariey, F., Buchy, P., Guillard, B., Gonzalez, I. J., Christophel, E. M., Abdur, R., von Sonnenburg, F., Bell, D., & Menard, D. (2014). Acute undifferentiated febrile illness in rural Cambodia: a 3-year prospective observational study. *PLoS One, 9*(4), e95868. doi:10.1371/journal.pone.0095868

Murphy, F. A., Webb, P. A., Johnson, K. M., Whitfield, S. G., & Chappell, W. A. (1970). Arenoviruses in Vero cells: ultrastructural studies. *J Virol, 6*(4), 507-518.

N, N. D., Berthet, N., Rougeron, V., Mangombi, J. B., Durand, P., Maganga, G. D., Bouchier, C., Schneider, B. S., Fair, J., Renaud, F., & Leroy, E. M. (2015). Evidence of lymphocytic choriomeningitis virus (LCMV) in domestic mice in Gabon: risk of emergence of LCMV encephalitis in Central Africa. *J Virol, 89*(2), 1456-1460. doi:10.1128/jvi.01009-14

Ortiz-Riano, E., Cheng, B. Y., Carlos de la Torre, J., & Martinez-Sobrido, L. (2013). Arenavirus reverse genetics for vaccine development. *J Gen Virol, 94*(Pt 6), 1175-1188. doi:10.1099/vir.0.051102-0

Park, J. Y., Peters, C. J., Rollin, P. E., Ksiazek, T. G., Katholi, C. R., Waites, K. B., Gray, B., Maetz, H. M., & Stephensen, C. B. (1997). Age distribution of lymphocytic choriomeningitis virus serum antibody in Birmingham, Alabama: evidence of a decreased risk of infection. *Am J Trop Med Hyg, 57*(1), 37-41.

Parodi, A. S., Greenway, D. J., Rugiero, H. R., Frigerio, M., De La Barrera, J. M., Mettler, N., Garzon, F., Boxaca, M., Guerrero, L., & Nota, N. (1958). [Concerning the epidemic outbreak in Junin]. *Dia Med, 30*(62), 2300-2301.

Perez, M., & de la Torre, J. C. (2003). Characterization of the genomic promoter of the prototypic arenavirus lymphocytic choriomeningitis virus. *J Virol, 77*(2), 1184-1194.

Pontremoli, C., Forni, D., & Sironi, M. (2018). Arenavirus genomics: novel insights into viral diversity, origin, and evolution. *Curr Opin Virol, 34*, 18-28. doi:10.1016/j.coviro.2018.11.001

Price, M. E., Fisher-Hoch, S. P., Craven, R. B., & McCormick, J. B. (1988). A prospective study of maternal and fetal outcome in acute Lassa fever infection during pregnancy. *BMJ, 297*(6648), 584-587.

Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T. F., Beutler, N. A., Burton, D. R., Lewis-Ximenez, L. L., de Jesus, J. G., Giovanetti, M., Hill, S. C., Black, A., Bedford, T., Carroll, M. W., Nunes, M., Alcantara Jr, L. C., Sabino, E. C., Baylis, S. A., Faria, N. R., Loose, M., Simpson, J. T., Pybus, O. G., Andersen, K. G., & Loman, N. J. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nature Protocols, 12*, 1261. doi:10.1038/nprot.2017.066

https://www.nature.com/articles/nprot.2017.066#supplementary-information

Radoshitzky, S. R., Bao, Y., Buchmeier, M. J., Charrel, R. N., Clawson, A. N., Clegg, C. S., DeRisi, J. L., Emonet, S., Gonzalez, J. P., Kuhn, J. H., Lukashevich, I. S., Peters, C. J., Romanowski, V., Salvato, M. S., Stenglein, M. D., & de la Torre, J. C. (2015). Past, present, and future of arenavirus taxonomy. *Arch Virol, 160*(7), 1851-1874. doi:10.1007/s00705-015-2418-y

Riera, L., Castillo, E., Del Carmen Saavedra, M., Priotto, J., Sottosanti, J., Polop, J., & Ambrosio, A. M. (2005). Serological study of the lymphochoriomeningitis virus (LCMV) in an inner city of Argentina. *J Med Virol, 76*(2), 285-289. doi:10.1002/jmv.20357

Rivers, T. M., & McNair Scott, T. F. (1935). Meningitis in Man Caused by a Filterable Virus. *Science, 81*(2105), 439-440. doi:10.1126/science.81.2105.439-a

Rivers, T. M., & Scott, T. F. (1936). Meningitis in Man Caused by a Filterable Virus : Ii. Identification of the Etiological Agent. *J Exp Med, 63*(3), 415-432.

Riviere, Y., & Oldstone, M. B. (1986). Genetic reassortants of lymphocytic choriomeningitis virus: unexpected disease and mechanism of pathogenesis. *J Virol, 59*(2), 363-368.

Salvato, M., Borrow, P., Shimomaye, E., & Oldstone, M. B. (1991). Molecular basis of viral persistence: a single amino acid change in the glycoprotein of lymphocytic choriomeningitis virus is associated with suppression of the antiviral cytotoxic T-lymphocyte response and establishment of persistence. *J Virol, 65*(4), 1863-1869.

Scott, T. F., & Rivers, T. M. (1936). Meningitis in Man Caused by a Filterable Virus : I. Two Cases and the Method of Obtaining a Virus from Their Spinal Fluids. *J Exp Med, 63*(3), 397-414.

Shao, J., Liang, Y., & Ly, H. (2015). Human hemorrhagic Fever causing arenaviruses: molecular mechanisms contributing to virus virulence and disease pathogenesis. *Pathogens, 4*(2), 283-306. doi:10.3390/pathogens4020283

Spiropoulou, C. F., Kunz, S., Rollin, P. E., Campbell, K. P., & Oldstone, M. B. (2002). New World arenavirus clade C, but not clade A and B viruses, utilizes alpha-dystroglycan as its major receptor. *J Virol, 76*(10), 5140-5146.

Stefas, I., Rucheton, M., D'Angeac, A. D., Morel-Baccard, C., Seigneurin, J. M., Zarski, J. P., Martin, M., Cerutti, M., Bossy, J. P., Misse, D., Graafland, H., & Veas, F. (2001). Hepatitis B virus Dane particles bind to human plasma apolipoprotein H. *Hepatology, 33*(1), 207-217. doi:10.1053/jhep.2001.20531

Suzuki, H., Nunome, M., Kinoshita, G., Aplin, K. P., Vogel, P., Kryukov, A. P., Jin, M. L., Han, S. H., Maryanto, I., Tsuchiya, K., Ikeda, H., Shiroishi, T., Yonekawa, H., & Moriwaki, K. (2013). Evolutionary and dispersal history of Eurasian house mice Mus musculus clarified by more extensive geographic sampling of mitochondrial DNA. *Heredity (Edinb), 111*(5), 375-390. doi:10.1038/hdy.2013.60

Takagi, T., Ohsawa, M., Morita, C., Sato, H., & Ohsawa, K. (2012). Genomic analysis and pathogenic characteristics of lymphocytic choriomeningitis virus strains isolated in Japan. *Comp Med, 62*(3), 185-192.

Takagi, T., Ohsawa, M., Yamanaka, H., Matsuda, N., Sato, H., & Ohsawa, K. (2017). Difference of two new LCMV strains in lethality and viral genome load in tissues. *Exp Anim, 66*(3), 199-208. doi:10.1538/expanim.16-0097

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol Biol Evol, 9*(4), 678-687. doi:10.1093/oxfordjournals.molbev.a040752

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution, 30*(12), 2725-2729. doi:10.1093/molbev/mst197

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res, 22*(22), 4673-4680.

Trapido, H., & Sanmartin, C. (1971). Pichinde virus, a new virus of the Tacaribe group from Colombia. *Am J Trop Med Hyg, 20*(4), 631-641.

Traub, E. (1935). A Filterable Virus Recovered from White Mice. *Science, 81*(2099), 298-299. doi:10.1126/science.81.2099.298

Urata, S., & de la Torre, J. C. (2011). Arenavirus budding. *Adv Virol, 2011*, 180326. doi:10.1155/2011/180326

Zerbino, D. R., & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res, 18*(5), 821-829. doi:10.1101/gr.074492.107

Zhang, L., Li, S., Huang, S. J., Wang, Z. D., Wei, F., Feng, X. M., Jiang, D. X., & Liu, Q. (2018). Isolation and genomic characterization of lymphocytic choriomeningitis virus in ticks from northeastern China. *Transbound Emerg Dis*. doi:10.1111/tbed.12946

9. APPENDIX

Table shows the symptoms' percentage's among the positive suspected samples

Symptoms	Percentage
fevere	100 %
headache	90 %
joint pain	68 %
vertigo	61 %
severe malaise	48 %
chills	46 %
cough	46 %
abdominal pain	34 %
drowsiness	30 %
anorexia	28 %
stiff neck	28 %
nausea	21 %
retro-orbital pain	19 %
diarrhea	18 %
vomiting	10 %
confusion	8 %
severe muscle weakness	6 %
conjunctivitis	3 %
Lymphadenopathy	3 %
rash	2 %
ataxia	1 %
shortness of breath	1 %