

Diversity of Viruses in Hard Ticks (Ixodidae) from Select Areas of a Wildlife-livestock Interface Ecosystem at Mikumi National Park, Tanzania

Donath Damian^{1,3,*}, Modester Damas¹, Jonas Johansson Wensman², Mikael Berg³

¹Department of Molecular Biology and Biotechnology, University of Dar es Salaam, Dar es Salaam, Tanzania

²Section of Ruminant Medicine, Department of Clinical Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

³Section of Virology, Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, Uppsala, Sweden

Email address:

donath.damian@yahoo.com (D. Damian)

*Corresponding author

To cite this article:

Donath Damian, Modester Damas, Jonas Johansson Wensman, Mikael Berg. Diversity of Viruses in Hard Ticks (Ixodidae) from Select Areas of a Wildlife-livestock Interface Ecosystem at Mikumi National Park, Tanzania. *American Journal of BioScience*.

Vol. 8, No. 6, 2020, pp. 150-157. doi: 10.11648/j.ajbio.20200806.12

Received: December 3, 2020; **Accepted:** December 16, 2020; **Published:** December 28, 2020

Abstract: Many of the recent emerging infectious diseases have occurred due to the transmission of the viruses that have wildlife reservoirs. Arthropods, such as ticks, are known to be important vectors for spreading viruses and other pathogens from wildlife to domestic animals and humans. In the present study, we explored the diversity of viruses in hard ticks (Ixodidae) from select areas of a wildlife-livestock interface ecosystem at Mikumi National Park, Tanzania using a metagenomic approach. cDNA and DNA were amplified with random amplification and Illumina high-throughput sequencing was performed. The high-throughput sequenced data was imported to the CLC genomic workbench and trimmed based on quality (Q = 20) and length (≥ 50). The trimmed reads were assembled and annotated through Blastx using Diamond against the National Center for Biotechnology Information non-redundant database and its viral database. The MEGAN Community was used to analyze and to compare the taxonomy of the viral community. The obtained contigs and singletons were further subjected to alignment and mapping against reference sequences. The viral sequences identified were classified into bacteria, vertebrates, and invertebrates, plants, and protozoans viruses. Sequences related to known viral families; *Retroviridae*, *Flaviviridae*, *Rhabdoviridae*, *Chuviridae*, *Orthomyxoviridae*, *Phenuiviridae*, *Totiviridae*, *Rhabdoviridae*, *Parvoviridae*, *Caulimoviridae*, *Mimiviridae* and several Phages were reported. This result indicates that there are many viruses present in the study region, which we are not aware of and do not know the role they have or if they have the potential to spread to other species and cause diseases. Therefore, further studies are required to delineate the viral community present in the region over a large scale.

Keywords: Ticks, Viral Metagenomics, Virus, Mikumi National Park

1. Background

Ticks as obligate blood-sucking parasitic arthropod vector, harbor and transmit a wide range of pathogens of both veterinary and human medical importance [1]. In recent years, the reported number of tick-borne infections has increased worldwide [1]. Several emerging tick-borne viruses, bacteria, and protozoa which cause a threat to human, livestock, and wild animal's diseases have been

recognized [2, 3]. Viral diseases including tick-borne encephalitis virus, Crimean-Congo hemorrhagic fever virus, Kyasanur forest diseases virus, severe fever with thrombocytopenia syndrome virus, Heartland virus, African swine fever virus (ASFV), Nairobi sheep disease virus and Louping ill virus have been reported [1-4].

About 800 tick species are widely distributed all over the world, predominantly in tropical and subtropical countries [5]. Hard tick species (Ixodidae family) especially of the

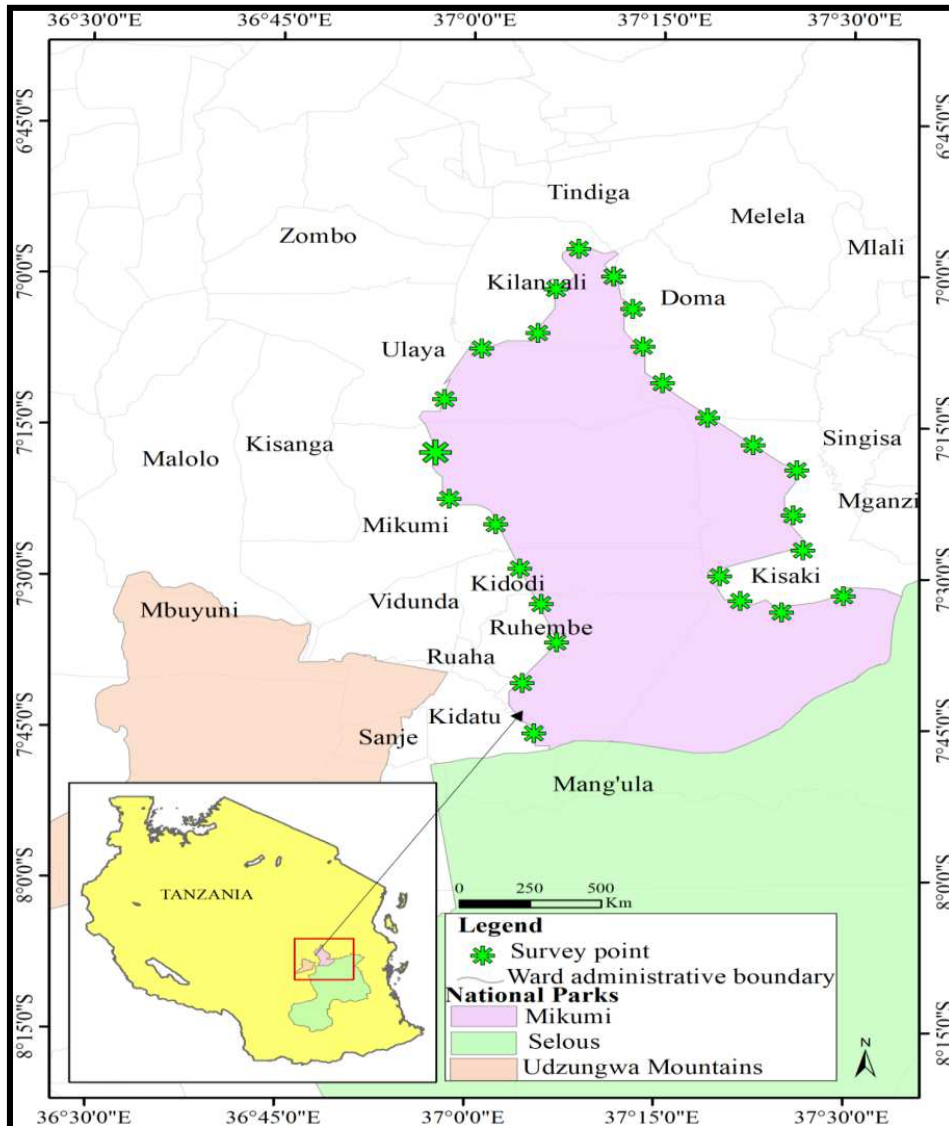
genera *Hyalomma*, *Rhipicephalus*, and *Ambryomma* are most widely distributed in Africa [5].

Nomadic and pastoralist lifestyles, especially those at the wildlife-livestock interface, can allow direct and indirect contact with wild animals that can facilitate exposure and sharing of previously isolated ticks and tick-borne pathogens [6]. Mikumi National Park is a protected conservation area and is home to a wide range of wildlife species. In areas that lie at the border of Mikumi National Park, people practice nomadic pastoralism, keeping a large number of indigenous cattle and goats. During dry seasons there is the movement of people and livestock to areas very close and sometimes entering beyond the National Park boundary, where water and pasture are abundant long after the rains have gone. Likewise, there is the migration of wild animals outside of the National Park boundaries. These migration patterns facilitate the potential transfer of ticks across great distances, presenting the opportunity for the exchange of diverse tick species between the domestic, wild animal, and

even human populations, facilitating risk of the exposure to tick-borne disease [6]. Therefore, the area is considered to be one of the epidemic foci of tick species and possibly tick-borne diseases, although the diversity of the tick viruses circulating the local area is still unknown.

Identification of tick viruses is essential to the control of tick-borne diseases. Since many viruses cannot be cultivated and lack common viral gene markers, the detection of viruses using conventional, immunological, and PCR methods is very difficult [7]. Furthermore, many as yet undiscovered viruses cannot be identified using these specific methods.

This study focused on viruses of ticks in the wildlife-livestock interface ecosystem of Mikumi National Park, Morogoro region, Tanzania based on metagenomic as a powerful tool to investigate emerging and re-emerging viruses [7]. To our knowledge, this is the first report from Tanzania in which ticks viruses are quantified using metagenomic techniques.



Data source; (National Bureau of Statistics GIS Data base; <https://www.nbs.go.tz/index.php/en/>)

Figure 1. Map showing wildlife-livestock-human interface ecosystem of Mikumi national park.

2. Methods

2.1. Sample Collection

Hard ticks were collected from the body of domestic animals (cattle and goats) and from the environment (free-living ticks) [7]. Ticks collection was done from April to July 2019 in the wards which lie at the border of Mikumi National park, Morogoro Region, Tanzania (figure 1). Adult ticks were collected by plucking using blunt forceps from domestic animals. Whereas, the questing (free-living) ticks were collected from purposefully chosen areas within logistical constraints, and include host resting areas and burrows, host routes and areas surrounding watering holes [7]. After collection, all ticks were transported to the laboratory at the Department of Parasitology and Entomology, Faculty of Veterinary Medicine, Sokoine University of Agriculture, Tanzania. All samples were stored at -80°C . All procedures, protocols, and methods performed in this study were approved based on guidelines and regulations of research by the Ethics Committee of the University of Dar es Salaam and Tanzania Commission for Science and Technology.

2.2. Sample Preparation

20 ticks per pool were homogenized with a mortar and pestle in DNase buffer [7]. The homogenates were centrifuged at $10,000\times g$ for 10 minutes at 4°C . Supernatants from tick pools were filtered using $0.4\ \mu\text{m}$ -pore-size filters to minimize the presence of cells, debris, and bacteria [7, 8]. To remove the naked DNA and RNA, $200\ \mu\text{l}$ of the resuspended pellet from each pooled samples were digested in a cocktail with 20U of Turbo DNase, 25U of benzonase, and $0.1\ \text{mg/ml}$ of RNase A at 37°C for 2 hours in $20\ \mu\text{l}$ of 10X DNase buffer [7, 8]. Then, the capsid-protected viral genomes were extracted with a QIAamp viral DNA and RNA mini kit (Qiagen) according to the manufacturer's instructions.

2.3. Nucleic acid Labeling and cDNA Synthesis

SuperScript III First-Strand System kit (Invitrogen, USA) was used to synthesis first-strand cDNA. An $8\ \mu\text{l}$ volume of purified viral nucleic acids from each tick pool was mixed with $1\ \mu\text{l}$ of $10\ \text{mM}$ deoxynucleoside triphosphate and $1\ \mu\text{l}$ ($50\text{ng}/\mu\text{l}$) of random Hexamer, after which the solution was denatured at 65°C for 5 min, and placed on ice for 1 min. A $10\ \mu\text{l}$ aliquot of the cDNA synthesis mix containing $2\ \mu\text{l}$ of $10\times$ RT buffer, $4\ \mu\text{l}$ of $25\ \text{mM}$ MgCl_2 , $2\ \mu\text{l}$ of $0.1\ \text{M}$ DTT, $1\ \mu\text{l}$ of RNase OUT, and $1\ \mu\text{l}$ of SuperScript III RT were added to each nucleic acid-primer mixture. The reaction mixture was incubated at 25°C for 10 min and 50°C for 50 min, followed by enzyme inactivation at 85°C for 5 min and chilling on ice for 1 min. For the second strand cDNA synthesis, $0.5\ \mu\text{l}$ ($20\ \text{pmol}$) of random primer, $2.5\ \mu\text{l}$ of $10\times$ Klenow fragment buffer, and $2\ \mu\text{l}$ of Klenow fragment were added. The reaction mixture was incubated at 25°C for 10 min and 37°C for 60 min, followed by enzyme inactivation at

75°C for 10 min [8, 9]. The DNA was also labeled using the same primer and Klenow fragment (NEB).

2.4. Random Amplification

PCR amplification was performed on both DNA and cDNA using the primer FR-26RV (GCC GGA GCT CTG CAG ATA TC) according to the following procedure: 1 PCR buffer, $2.5\ \text{mM}$ MgCl_2 , $2.5\ \text{mM}$ deoxynucleoside triphosphates (dNTP), $0.4\ \text{mM}$ F primer, $0.4\ \text{mM}$ R primer, and $1.25\ \text{U}$ AmpliTaq Gold DNA polymerase (Applied Biosystems). The amplification was initiated with a 10-min heating step at 95°C , followed by 40 cycles of 30 s at 95°C , 30 s at 58°C , and 90 s at 72°C . The reaction was ended with an extra elongation step at 72°C for 10 min [8, 9]. The product was purified using a PCR purification kit (Qiagen) according to the manufacturer's instructions before being sent for large scale sequencing. All the procedures were conducted at the Center for Infectious Diseases (CID), Tanzania Veterinary Laboratory Agency (TVLA), Temeke.

2.5. Large-scale Sequencing

The sample pools were pooled together and sequenced at Macrogen Europe (Macrogen Inc.) by Illumina platform using the TrueSeq Nano DNA library preparation method. The sequencing library was prepared by random fragmentation of the DNA and cDNA pooled samples, followed by 5' and 3' adapter ligation. Adapter-ligated fragments were then PCR amplified and gel purified. For cluster generation, the library was loaded into a flow cell where fragments were captured on a lawn of surface-bound oligos complementary to the library adapters. Each fragment was then amplified into distinct, clonal clusters through bridge amplification.

2.6. Bioinformatics

The High Throughput Sequenced (HTS) data was imported to the CLC genomic workbench (version 11). Filtering and trimming to remove adaptors and bad sequences were based on quality score of 20 ($Q = 20$) and length of greater or equal to 50 (≥ 50) [10]. The trimmed reads were assembled using the same CLC genomic workbench and annotated through Blastx ($E\text{-value} \leq 0.0001$) using Diamond (version 0.9.10) against the National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>) non-redundant database (nr) and its viral database [11]. The MEGAN Community Edition (Version 6.19.1) analysis toolset (<http://ab.inf.uni-tuebingen.de/software/megan/>) was used to analyze and to compare the taxonomy of the viral community [12, 13]. The obtained contiguous sequences (contigs) together with singletons were further subject to BLASTx analysis before alignment and mapping against the reference protein genome to find the consensus sequences.

2.7. Confirmation and Retrieval of Viral Genome Sequences

Viral contigs and singletons were used to design primers for specific selected viral sequences to confirm the presence of the viruses as well as to amplify longer genome regions of the selected viruses from the original material. Thermal cycling was initiated with a denaturation step at 95°C for 10min; followed by 35 cycles of 95°C for 30 sec, 58–60°C 30 sec and 72°C for 1 min; and a final extension at 72°C for 7 min [9].

Table 1. Sample sequencing, filtering, and trimming information.

Number of reads	Avg. length	Number of reads after trim	Percentage trimmed (%)	Avg. length after trim
35,339,795	151	35,339,742	100	143

3.2. The Percentage Distribution of the Metagenomic Reads

To perform taxonomic profiling of the tick viral metagenome, binning of the metagenomic reads to their respective taxonomic groups based on the most significant similarities was performed. The majority of the reads (99%) were classified as cellular organism genomes, mainly Eukaryota and bacteria. 0.535 and 0.028% of the reads were identified as unclassified sequences and other sequences respectively and 0.437% of the reads were classified as viral sequences (Table 2).

Table 2. The percentage distribution of the sequences after binning.

Group	Cellular organisms	Unclassified sequences	Viruses	Other sequences
Percentage	99	0.535	0.437	0.028

3.3. Viral Groups Related Sequences

The viral sequences identified were classified into bacteria, plants, vertebrates, invertebrates, and protozoans' viruses. The viral sequences related to Retroviridae, Rhabdoviridae, Chuviridae, Orthomyxoviridae, Phenuiviridae, Totiviridae, Parvoviridae, Caulimoviridae, Mimiviridae, Siphoviridae, Podoviridae, Myoviridae, Herelleviridae as well as unclassified +ssRNA and unclassified Riboviria were recorded in different proportions (Table 3).

Table 4. The summary of BLASTx analysis and mapping of the identified viral sequences to the reference genes.

Species	Identity	Ref. Gene	Accession	Family
Taishuni tick virus	70-88	RdRp	QBQ65007	Rhabdoviridae
Taishuni tick virus	72-78	Nucleoprotein	QBQ65008	Rhabdoviridae
Yongjia tick virus	78-94	Nucleoprotein	AXH64984	Rhabdoviridae
Piry virus	63-72	Nucleoprotein	YP009505532	Rhabdoviridae
Hubei toti-like virus	40-50	Hypothetical protein	YP009336907	RNAvirus
Bole tick virus	81-89	Polyprotein	QFR36180	ssRNA virus
Tjuloc virus	65-73	Hemagglutinin	AFN73046	Orthomyxoviridae
Lihan tick virus	57-70	Nucleoprotein	QDW81043	Phenuiviridae
Xinjiang tick virus	55-65	Polymerase	QBQ65054	Totiviridae
Lonestar tick virus	76-86	Polymerase	AUX13136	Totiviridae
Lonestar tick virus	76-86	ORF1	AUX13135	Totiviridae
Mivirus	57-72	Nucleoprotein	QDW81056	Chuviridae
Iteravirus	56-69	Nucleoprotein	AWA28232	Parvoviridae
Mimiviridae sp.	52-73	Ubiquitin enzyme	QDY51857	Mimiviridae
Tungrovirus	98-99	Hypothetical protein	FAA00009	Caulimoviridae
Bovine retrovirus	76-84	DNA polymerase	YP009243641	Retroviridae

3. Results

3.1. High Throughput (HT) Sequencing Results

The sequencing resulted in 35,339,795 reads with an average length of 151 base pairs. After filtering and trimming, 53 reads were removed and the remained 35,339,742 (99.9%) with an average length of 143 base pairs were classified as the good quality reads (Table 1).

Table 3. Proportions of the viral group related sequences.

SN	Viral family	Percentage of reads
1	Rhabdoviridae	5.4
2	Totiviridae	2.2
3	Unclassified +ssRNA	3.8
4	Retroviridae	60
5	Caulimoviridae	5.9
6	Unclassified Riboviria	1.6
7	Mimiviridae	2.7
8	Iridoviridae	2
9	Orthomyxoviridae	1.2
10	Flaviviridae	1
11	Chuviridae	1.2
12	Phenuiviridae	1
13	Podoviridae	1.2
14	Herelleviridae	4.8
15	Myoviridae	1
16	Siphoviridae	2.6
17	Parvoviridae	1.2

3.4. Further BLASTx Analysis and Mapping of the Viral Sequences

In subsequent BLASTx analysis and mapping of the viral singletons and contigs, several viral related proteins genomes of the classified and nonclassified viruses were identified. The analysis of viruses' proteins genomes based on the BLASTx search and mapping for each tick group is described in table 4.

Species	Identity	Ref. Gene	Accession	Family
Bacillus phage	94-98	Hypothetical protein	YP007678111	Siphoviridae
Staphylococcus phage	82-92	Hypothetical protein	YP009283990	Podoviridae
Escherichia phage	98-100	Putative protein	QBQ78284	Myoviridae phage
Salmonera phage	92-98	Putative protein	AUO79628	Herelleviridae

3.5. *Rhabdoviridae* and *Totiviridae*

The Rhabdoviridae related sequences generated in the present study recorded to have the amino acids identity of 63-88% to the RNA dependent RNA polymerase (RdRp) gene of the Taishun tick virus (Accession; QBQ65007) which was further confirmed by PCR, and 63-94% to the Nucleoprotein (NP) of the Taishun tick virus (Accession; QBQ65008) (Table 4).

Some of the viral sequences related to the *Totiviridae* family generated in the present study had 65-86% amino acid identity to the polymerase gene sequences of the Xinjiang tick totivirus (Accession; QBQ65054) and some to the polymerase (Accession; AUX13136), and the ORF1 sequences of the Lonestar tick totivirus (Accession; AUX13135) (Table 4) which was further confirmed by PCR.

3.6. *Phenuiviridae* and *Chuviridae*

For the *Phenuiviridae*, sequences related to the small (S) segment which encodes a nucleoprotein (NP) were recorded in the present study. Based on Blastx analysis, we identified that the sequences shared 57-70% amino acid identity with a nucleoprotein (NP) of the Lihan tick virus described in *Rhipicephalus microplus* collected in China (Accession; QDW81043) (Table 4).

The *Chuviridae* sequences related to the segment which encodes a nucleoprotein (NP) were recorded in the present study. Based on Blastx analysis, we identified that the sequences shared 57 to 71% amino acid identity with a nucleoprotein (NP) gene of the Mivirus sp., described previously elsewhere (Accession; QDW81056) (Table 4).

3.7. *Orthomyxoviridae* and *Parvoviridae*

The *Orthomyxoviridae*, sequences related to the segment which encodes a hemagglutinin protein were recorded in the present study. Based on Blastx analysis, we identified that the sequences shared 65 to 73% amino acid identity with the hemagglutinin protein of the Tjuloc tick virus described elsewhere (Accession; AFN73046) (Table 4).

For the *Parvoviridae*, sequence related to the segment which encodes a nucleocapsid was recorded in the present study. Based on Blastx analysis, we identified that the sequences shared 56 to 69% amino acid identity with a nucleocapsid protein of the Iteravirus described previously (Accession; AWA28232) (Table 4).

3.8. *Mimiviridae* and *Caulimoviridae*

The viral sequences related to the *Mimiviridae* family generated in the present study had 53-73% amino acid similarities to the ubiquitin enzymes of the *Mimiviridae* sp (Accession; QDY51857) (Table 4).

The viral sequences related to the *Caulimoviridae* family generated in the present study had high identity of 98-99% amino acid to the hypothetical protein of the Tungrovirus (Accession; FAA00009) (Table 4).

3.9. *Retroviridae* and *Phages*

In the present study, several sequences of the retroviruses were recorded, but the sequences related to the bovine retrovirus were recorded in large number. For the bovine *retroviridae*, the sequences related to the segment which encodes a polymerase protein were recorded. Based on Blastx analysis, we identified that the sequences shared 76-84% amino acid identity with the polymerase protein of the bovine *retroviridae* (Accession; YP009243641) (Table 4).

The tick metagenomic viromes generated in the present study contain the large diversity of phages, including members from the *Myoviridae*, *Podoviridae*, *Herelleviridae*, *Siphoviridae* as well as the unclassified *Caudovirales* phages. The phage sequences found in the ticks shared amino acid identity with the known phages species. Most of the phage sequences reported in the tick's sequences generated in the present study shared amino acid identity scores of 82-100% with the known phages proteins genomes mostly from the *Staphylococcus*, *Escherichia*, *Salmonella*, and *Bacillus* phages (Table 4).

3.10. *Unclassified +ssRNA* and *Unclassified RNA virus*

For the unclassified +ssRNA virus, sequences related to the segment which encodes the polyprotein were recorded in the present study. Based on Blastx analysis, we identified that the sequences shared 81-89% amino acid identity with a polyprotein of the Bole tick virus described previously elsewhere (Accession; QFR36180) (Table 4).

For the unclassified RNA virus, sequences related to the segment which encodes the hypothetical protein were recorded in the present study. Based on Blastx analysis, we identified that the sequences shared only 40% amino acid identity with a hypothetical protein of the Hubei toti-like tick virus described previously elsewhere (Accession; YP009336907) (Table 4).

4. Discussion

Many of the recent emerging infectious diseases have occurred due to the transmission of the viruses that have the wildlife reservoir [14]. Arthropods, such as ticks, are known to be important vectors for spreading viruses and other pathogens from wildlife to domestic animals and humans [14]. The present study is the first where metagenomic was used to explore the viral diversity that exists in a community of hard ticks in Tanzania. High-throughput sequencing was

performed by Illumina technology. Through BLASTx analyses and mapping, the sequences corresponded to diverse viral groups were recorded.

The viral sequences identified were classified into bacteria, fungi, protozoans, plants, invertebrates, and vertebrates' viruses. The sequences related to several viral families have been recorded in the present study namely; *Siphoviridae*, *Podoviridae*, *Myoviridae*, *Herelleviridae*, *Caudoviridae*, *Caulimoviridae*, *Retroviridae*, *Rhabdoviridae*, *Phenuiviridae*, *Chuviridae*, *Orthomyxoviridae*, *Totiviridae*, *Parvoviridae*, *Mimiviridae*, and other unclassified viruses. Such a large number of viral families reflect the high diversity of the viruses present in the study region.

The *Rhabdoviridae* is a family of viruses in the order Mononegavirales, vertebrates and invertebrates serve as natural hosts [15, 16]. Rhabdoviruses are characterized by a non-segmented, negative-sense RNA, with 11–15 kbp genomes coding for not less than five transcription units: nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and the RNA-dependent RNA polymerase (RdRp) [16]. Rhabdoviruses are mainly transmitted by arthropod vectors, and a few are Tick-borne viruses (TBV) [16]. We Identified sequences related to the gene encoding for the RNA-dependent RNA polymerase (RdRp) which was further confirmed by PCR and the gene encoding for the nucleoprotein (NP). Similar to the present study, metagenomic profile analysis of the viral populations carried in *Rhipicephalus spp.* ticks in China identified some sequences related to *Rhabdoviridae* that had not been described previously in the named study region [17]. Currently, there are more than 10 reorganized genera present in the family and many rhabdoviruses await taxonomic classification [16]. Large numbers of these viruses have been placed in unassigned group, this is due to the high genetic diversity of the rhabdoviruses from different hosts [16].

For the *Phenuiviridae* viral family, a nucleoprotein (NP) of Lihan tick virus was recorded to have amino acids similarity to sequences generated in the present study. Similar to the present study results, a study on viral diversity of *Rhipicephalus microplus* parasitizing cattle in Southern Brazil, authors revealed the presence of sequences of *Phenuiviridae* viruses related to Lihan and Wuhan tick viruses which had not reported in the study region before [18]. *Phenuiviridae* is a virus family belonging to the order Bunyavirus [19]. Ruminants, camels, and arthropods are the known hosts of members of this negative-sense single-stranded RNA virus (-ssRNA) [19]. The genus *Phlebovirus* is the only one in which some members are associated with ticks, which were usually named tick-borne phleboviruses (TBPVs) [19]. In recent years, newly emerging TBPVs able to induce severe diseases in humans have attracted increasing attention. Severe fever with thrombocytopenia syndrome virus (SFTSV) and Heartland virus (HRTV) are the novel pathogenic TBPVs [20].

The viral sequences related to the *Chuviridae* family recorded in the present study had amino acids identity to the nucleoproteins (NP) of the Mivirus sp. Similar to the present

study, Sameroff et al. 2019 in their study of viral diversity of tick species parasitizing cattle and dogs in Trinidad and Tobago identified the sequences related to the *Chuviridae* viral family representing two species of mivirus with amino acid similarity to Wuhan mivirus and Changping mivirus [21]. Likewise, the study on viral diversity of *Rhipicephalus microplus* parasitizing cattle in Southern Brazil [18], authors identified two sequences similar to Wuhan tick virus 2 reported previously in China [17], a virus that belongs to *Chuviridae*. *Chuviridae* is recently discovered RNA viral family of negative-sense single-stranded viruses (-ssRNA) and have variable genomic structure; unsegmented, bisegment, and circular. The viruses from the *Chuviridae* family infect several arthropod species including mosquito and tick species [17, 18].

As shown in the results, the viral sequences related to the *Orthomyxoviridae* family recorded in the present study had amino acids similarity to hemagglutinin of the Tjuloc virus, a member of the genus quaranjavirus. The present study results are in agreement with Cholet et al. 2018 in their metagenomic study on *Rhipicephalus* ticks in Mozambique in which the authors reported the sequences with closest identity to members of quaranjavirus; Johnston Atoll quaranjavirus, Tjuloc viruses, Wellfleet Bay viruses, and unclassified quaranjaviruses [22]. The family *Orthomyxoviridae* comprises viruses characterized with six to eight segments of linear, negative-sense RNA genomes [22]. One most notorious representative in this family *Orthomyxoviridae* is influenza virus which attracted significant medical and veterinary attentions. In this family, two recently proposed genera, *Quarjavirus* and *Thogotovirus*, contain members of TBVs. *Thogotovirus* infection has been reported to be pathogenic for sheep and has been associated with high level of abortion [22].

As recorded in the results, *Totiviridae* related sequences generated from the present study, have amino acids identity to a polymerase and the ORF1 of the Xinjiang tick toti virus, and the Lonestar tick virus which have been confirmed by PCR. *Totiviridae* is a family of viruses whereby *Giardia lamblia*, *Leishmania*, *Trichomonas*, and fungi harbored by ticks together with other organisms serve as natural hosts [22, 23]. Viruses in *Totiviridae* are non-enveloped, double-stranded RNA viruses (dsRNA) with icosahedral geometries [23]. It is possible that tick acquires totivirus during blood-feeding or it is possible totivirus originating from the *Giardia lamblia*, *Leishmania*, *Trichomonas*, and fungi harbored by ticks. Thus, so far, only a few studies reported the genome of the totiviruses, further studies are required to ultimately classify this virus.

The *Parvoviridae* sequences identified in the current study had the closest similarity to the viral related capsid protein of the Iterovirus. Similar results were reported previously elsewhere by Cholet et al in their viral metagenomics study of *Rhipicephalus* ticks genera in Mozambique [22]. *Parvoviridae* is a family of small, rugged genetically compact ssDNA viruses, known collectively as Parvoviruses [23]. Members of this family infect a wide array of animal

hosts and have been divided into two subfamilies, which infect either vertebrates (the *Parvovirinae*) or invertebrates (*Densovirinae*). Parvoviruses are associated with a variety of chronic diseases in humans and animals. However, these viral sequences have the ability to integrate into the chromosomal DNA of a wide range of hosts and may transmit to offspring [24]. Identification of the sequences related to this family of the virus indicates the presence of the parvoviruses and possibly the diseases relating to this family of the virus in the study region.

As shown in the results, the majority of the viral sequences were those related to the *Retroviridae* family, especially bovine retrovirus. The bovine retrovirus infects mostly cattle, buffalo, and they are zoonotic. Viruses belonging to these families are usually identified in blood, and ticks may obtain viruses from viremic hosts during blood feeding. Identification of the sequences related to these viruses in large proportion, reflects a high abundance of these viruses circulating in the wildlife-livestock interface ecosystem of Mikumi National Park Tanzania.

Additionally, the tick pool viromes in the present study contain some sequences related to plant viruses of the *Caulimoviridae* family. The presence of sequences related to plants viruses has also been reported in a number of tick viral metagenomics studies. For example in the tick metagenomic study conducted in China [17] and Brazil [18], authors reported the sequences related the known proteins of plants viruses.

As indicated in the results, the tick viromes contained a large diversity of phages, including members from *Myoviridae*, *Podoviridae*, *Siphoviridae*, *Herelleviridae* as well as unclassified Caudovirales phages. However, most of the phages sequences found in the ticks shared amino acid identity with the *Staphylococcus* phages, *Salmonella* phages, *Bacillus* phages, and *Escherichia* phages. The *Staphylococcus* phages, *Salmonella* phages, *Bacillus*, and many other groups of phages have been reported in previous ticks' viral metagenomics studies [17, 18, 21, 22]. It is possible that the ticks acquire phages during blood feeding, or is possible phages originates from the bacteria harbored by the ticks.

The most interesting findings were the presence of viral sequences from the unclassified +ssRNA and unclassified RNA viruses in the tick viromes generated in the present study. The unclassified +ssRNA viral sequences recorded to have identity to the polyprotein of the unclassified virus named Bole tick virus, whereas for the unclassified viral sequences related to the RNA virus generated in the present study recorded to have identity to the hypothetical protein of the unclassified virus named Hubei toti-like virus. These sequences were highly divergent to their reference viral sequences from the GenBank indicating that perhaps these are new viruses which are not yet documented in the Genbank database. Therefore, further research work is needed to characterize in details these unknown viral sequences.

5. Conclusion

Using viral metagenomic, ticks were investigated and the

sequences from diverse viral families were recorded. Both arthropods specific and non-arthropods specific viruses were found. This result indicates that there are many viruses present in the study region, which we are not aware of and do not know the role they have or if they have the potential to spread to other species and cause diseases. Therefore, further studies are required to delineate the viruses present in the region over a large scale.

Funding

This work was supported by the Swedish Government through the Swedish International Development Agency (SIDA) in collaboration with the University of Dar es Salaam and the Swedish University of Agricultural Sciences.

Author's Contributions

DD designed the investigation, performed the experiments, conducted the investigation, and drafted the manuscript. MD designed the investigation, conceived the study, and revised the manuscript. JJW and MB conceived the study, designed the investigation, and revised the manuscript. All authors read and approved the final manuscript.

Ethics Approval and Consent to Participate

No specific permits were required for this study. The study did not involve endangered or protected species. Therefore, the local ethics committee deemed that approval was unnecessary.

Competing Interests

The authors declare that they have no competing interests

Publishers Note

The *American journal of BioScience* remains neutral with regards to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

The authors thank the Swedish Government and Swedish International Development Cooperation Agency (SIDA) for contributions and financial support. Special thanks is also extended to the University of Dar es Salaam, Tanzania, and Swedish University of Agricultural Sciences, Uppsala, Sweden, for supporting this work, as well as colleagues and researchers at the Swedish University of Agricultural Sciences Global Bioinformatics Centre (SGBC) who provided valuable knowledge and skills concerning bioinformatics. We thank the livestock keepers, village and ward leaders from each location for their help with tick collection, as well as entomologists from Sokoine University of Agriculture who assisted in tick

morphological identification.

References

- [1] Weaver, S. C.; Reisen, W. K. Present and future arboviral threats. *Antivir. Res.* 2010, 85, 328–345.
- [2] Olival, K. J.; Hosseini, P. R.; Zambrana-Torrel, C.; Ross, N.; Bogich, T. L.; Daszak, P. Host and viral traits predict zoonotic spillover from mammals. *Nature* 2017, 546, 646–650.
- [3] Valarcher, J. F.; Hagglund, S.; Juremalm, M.; Blomqvist, G.; Renstrom, L.; Zohari, S.; Leijon, M.; Chirico, J. Tick-borne encephalitis. *Rev. Sci. Tech.* 2015, 34, 453–466.
- [4] Guglielme, A. A.; Venzal, J. M.; Battesti, D.; Onofrio, V. C.; Trajano, E.; Firminho, J. V. L. Ticks (Acari: Ixodida) of the Neotropical Zoogeographic Region; Consortium on Ticks and Tick-borne Diseases (ICTTD-2): Houten, The Netherlands, 2003; p. 173.
- [5] Estrada-Pena A and de la Fuente J. The ecology of ticks and epidemiology of tick-borne viral diseases. *Antiviral Res.* 2014; 108: 104–128.
- [6] Labuda M, Nuttall PA. Tick-borne viruses. *Parasitology.* 2004; 129 (Suppl): S221–S45.
- [7] Damian D, Maghembe R, Damas M, Wensman JJ, Berg M. Application of viral metagenomics for study of emerging and reemerging tick-borne viruses. *Vector-borne and zoonotic diseases.* 2020; 20 (8), 557-565.
- [8] Blomstrom A, Ye X, Fossum C and Berg M. Characterisation of the Virome of Tonsils from Conventional Pigs and from Specific Pathogen-Free Pigs. *Viruses* 2018; 10: 38.
- [9] Blomstrom AL, Fossum C, Wallgren P, Berg, M. Viral Metagenomics Analysis Displays the Co- Infection Situation in Healthy and PMWS Affected Pigs. *PLoS ONE* 2016; 11: 11.
- [10] Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics.* 2011 Mar 15; 27 (6): 863–864.
- [11] Buchfink, B.; Xie, C.; Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 2015, 12, 59–60.
- [12] Huson, D. H.; Beier, S.; Flade, I.; Gorska, A.; El-Hadidi, M.; Mitra, S.; Ruscheweyh, H. J.; Tappu, R. MEGAN Community Edition—Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput. Biol.* 2016, 12, e1004957.
- [13] Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 2016, 33, 1870–1874.
- [14] Vandegrift, K. J.; Kapoor, A. The Ecology of New Constituents of the Tick Virome and Their Relevance to Public Health. *Viruses* 2019, 11, 529.
- [15] Shi, M.; Lin, X. D.; Tian, J. H.; Chen, L. J.; Chen, X.; Li, C. X.; Qin, X. C.; Li, J.; Cao, J. P.; Eden, J. S.; et al. Redefining the invertebrate RNA virosphere. *Nature* 2016, 540, 539–543.
- [16] Dietzgen, R. G.; Kondo, H.; Goodin, M. M.; Kurath, G.; Vasilakis, N. The family Rhabdoviridae: Mono- and bipartite negative-sense RNA viruses with diverse genome organization and common evolutionary origins. *Virus Res.* 2017, 227, 158–170.
- [17] Xia, H.; Hu, C.; Zhang, D.; Tang, S.; Zhang, Z.; Kou, Z.; Fan, Z.; Bente, D.; Zeng, C.; Li, T. Metagenomic profile of the viral communities in *Rhipicephalus* spp. ticks from Yunnan, China. *PLoS ONE* 2015, 10, e0121609.
- [18] Souza, W. M.; Fumagalli, M. J.; Torres Carrasco, A. O.; Romeiro, M. F.; Modha, S.; Seki, M. C.; Gheller, J. M.; Daffre, S.; Nunes, M. R. T.; Murcia, P. R.; et al. Viral diversity of *Rhipicephalus microplus* parasitizing cattle in southern Brazil. *Sci. Rep.* 2018, 8, 16315.
- [19] Sakamoto JM, Ng TFF, Suzuki Y, et al. Bunyaviruses are common in male and female *Ixodes scapularis* ticks in central Pennsylvania. *PeerJ.* 2016; 4: e2324.
- [20] McMullan LK, Folk SM, Kelly AJ, MacNeil A et al. A new phlebovirus associated with severe febrile illness in Missouri. *N. Engl. J. Med.* 2012; 367, 834–841.
- [21] Sameroff S, Tokarz R, Charles R, Jain K, Oleynik A, Che X, Georges K, Carrington C. Viral Diversity of tick species parasitizing Cattle and Dogs in Trinidad and Tobago. *Scientific reports.* 2019; 9: 10421.
- [22] Cholleti H, Hayer J, Berg M, Blomström AL et al. Viral metagenomics reveals the presence of highly divergent quaranjavirus in *Rhipicephalus* ticks from Mozambique. *Infection ecology & epidemiology* 2018; 8: 1.
- [23] Francois S, Filloux D, Roumagnac P, et al. Discovery of parvovirus-related sequences in an unexpected broad range of animals. *Sci Rep.* 2016 Sep 07; (6): 30880.
- [24] Liu H, Fu Y, Xie J, et al. Widespread endogenization of densoviruses and parvoviruses in animal and human genomes. *J Virol.* 2011 Oct; 85 (19): 9863–9876.