



**University of  
Nottingham**

UK | CHINA | MALAYSIA

# **Characterising the viral aetiology of undiagnosed acute febrile illness in Uganda**

**Timothy Byaruhanga, BSc, MSc.**

A thesis submitted to the University of Nottingham for the award of Doctor of Philosophy

## i. Abstract

Acute febrile illnesses (AFI) are a significant cause of morbidity in Uganda and sub-Saharan Africa (SSA). The epidemiology and burden of disease due to viral AFI are unknown in Uganda despite an active surveillance system at the Uganda Virus Research Institute (UVRI).

This work describes the epidemiology of arboviruses and other AFI from data and samples stored (from 2016 to 2019) at the UVRI, providing a candid picture of the spatiotemporal distribution and seasonality of arboviruses in Uganda. Furthermore, for the arbovirus-negative specimens stored at the UVRI, a virus discovery approach was applied using the widely affordable conventional Polymerase Chain Reaction (PCR) method and later augmented by high throughput sequencing (HTS). We identified the age group 5-14 (years) most at risk of arboviral infections. Generally, arboviruses were more prevalent during the first rainy season from March to May, with yellow fever virus (YFV) as the most prevalent arbovirus. A significant spatiotemporal cluster was observed in the central region, where most arbovirus activities occurred. Among AFI, conventional PCR detected enteroviruses (Coxsackie A4, B1 viruses and human echovirus E4), measles virus (MeV) genotype B3 and West Nile virus (WNV) lineage two. The MeV was phylogenetically analysed and determined to be different from previously circulating B3 strains in Uganda but identical to strains in neighbouring Kenya and Rwanda. Furthermore, the WNV, although similar to the ancestral 1937 strain, had acquired a motif (NYSTQ) in the envelope protein similar to that found in lineage one WNV strains. In addition, HTS discovered viruses undetected by PCR, such

as Hepatitis C virus, Pegiviruses, and sequences belonging to Orbiviruses. The ever-decreasing costs of sequencing will support the application of PCR augmented by regular HTS in the surveillance of AFI and, thus, capture emerging and re-emerging viruses.

## ii. Publications

Chappell JG, Byaruhanga T, Tsoleridis T, Ball JK, McClure CP, 2019. Identification of infectious agents in High throughput sequencing datasets is easily achievable using free cloud-based bioinformatics platforms. *Journal of Clinical Microbiology*.

Byaruhanga T, Kayiwa J.T, Nankya A.M, Ataliba I.J, McClure C. P, Ball J.K, and Lutwama J.J.(2023) Arbovirus circulation, epidemiology and spatiotemporal distribution in Uganda. *IJID Regions*

Timothy Byaruhanga, Jonathan Ball, John Kayiwa, Annet Nankya, Irene Ataliba, Joseph Chappell, Theocharis Tsoleridis, Julius Lutwama, Patrick McClure, 2022. West Nile Virus in a febrile patient, South-west Uganda, 2018 (Submitted to the *Emerging Infectious Disease Journal*)

### iii. Acknowledgements

I am very grateful to my supervisors, Professor Jonathan Ball and Dr Patrick McClure, for the incredible supervision and guidance, especially during the unprecedented SARS-CoV-2 pandemic and lockdown. Patrick, thank you for believing in me even at my lowest moments.

I thank Dr Joseph Chappell for the technical support during my early years in the Virology Research Group (VRG) laboratories. I thank the VRG team for showing interest in my research, my fellow PhD students, and the VRG post-doctoral members for your approachable personalities during my stay.

I thank Dr Julius Lutwama for his extraordinary vision, selflessness and support throughout my study. Access to the samples and laboratory database on which this thesis is built was granted due to his vision of capacity building. Without his support, commencing this PhD in October 2018 was impossible. Thank you for nurturing my career as a scientist since 2012 when I first professionally met you. I am indebted to your generosity.

Mum and dad, thank you for the relentless spiritual support. It is hard to believe how far this journey has come! You prophesied, I believed, and the Lord was faithful to accomplish!

My wife, Mary, you are a superstar and a treasure I hold so dearly. You are the engine of my success, a permanent tattoo in my life story. My children, Talia, Eben, Joshua and Blessed, you kept the fire burning in my heart.

To Mary, Talia, Eben, Joshua, Emma, Bella, Margie, Joram, Telios, and Shalom and those yet to come.

*"I returned, and saw under the sun, that the race is not to the swift, not the battle to the strong, neither yet bread to the wise, nor yet riches to men of understanding, nor yet favour to men of skill; but time and chance happeneth to them all."*

**Ecclesiastes 9:11**

#### iv. List of Abbreviations

|              |                                                    |
|--------------|----------------------------------------------------|
| AFI.....     | Acute Febrile Infections/ Acute Febrile infections |
| AIDS.....    | Acquired immunodeficiency syndrome                 |
| APMV.....    | Avian Paramyxovirus                                |
| cDNA.....    | complementary Deoxyribonucleic acid                |
| DNA/RNA..... | Deoxyribonucleic acid/ Ribonucleic acid            |
| DHF.....     | Dengue Haemorrhagic Fever                          |
| FUO.....     | Fever of unknown origin                            |
| HTS.....     | High Throughput Sequencing                         |
| JEV.....     | Japanese encephalitis virus                        |
| MCV.....     | Measles containing vaccine                         |
| NGS.....     | Next-Generation Sequencing                         |
| NPEV.....    | non-polio enteroviruses                            |
| NVD.....     | Newcastle virus disease                            |
| ORF.....     | Open Reading Frame                                 |
| PCR.....     | Polymerase Chain Reaction                          |
| RIN.....     | RNA Integrity Number                               |
| rRNA.....    | ribosomal Ribonucleic acid                         |

SIA.....Supplementary immunisation activities

SSA.....Sub- Saharan Africa

UVRI.....Uganda Virus Research Institute

VRG.....Virus Research Group

WNV.....West Nile Virus

YFV.....Yellow Fever Virus

ZIKV.....Zika virus

## v. Table of Contents

### Contents

|            |                                                                                         |             |
|------------|-----------------------------------------------------------------------------------------|-------------|
| i.         | Abstract.....                                                                           | i           |
| ii.        | Publications.....                                                                       | iii         |
| iii.       | Acknowledgements.....                                                                   | iv          |
| iv.        | List of Abbreviations.....                                                              | vi          |
| <b>v.</b>  | <b>Table of Contents.....</b>                                                           | <b>viii</b> |
| vi.        | List of Figures.....                                                                    | xii         |
| Vii.       | List of Tables.....                                                                     | xiv         |
| <b>1.</b>  | <b>INTRODUCTION .....</b>                                                               | <b>1</b>    |
| 1.1        | Acute Febrile Illnesses.....                                                            | 1           |
| 1.1.1      | Global burden and overview of Acute Febrile Illness.....                                | 2           |
| 1.1.2      | Acute Febrile Illness in Uganda.....                                                    | 5           |
| 1.2        | Viruses as aetiological agents of AFI .....                                             | 7           |
| 1.2.1      | Flaviviridae and its role in AFI.....                                                   | 7           |
| 1.2.2.     | Picornaviridae as causative agents of AFI .....                                         | 14          |
| 1.2.3      | Paramyxoviruses as causative agents of AFI .....                                        | 19          |
| 1.2.4      | Togaviridae and its role in AFI .....                                                   | 25          |
| 1.2.5      | The role of Orbiviruses in AFI .....                                                    | 26          |
| 1.2.6      | Polyomaviridae and its role in AFI.....                                                 | 28          |
| 1.3        | Other non-viral causes of AFI in Uganda .....                                           | 30          |
| 1.4        | Virus Discovery.....                                                                    | 30          |
| 1.4.1      | Diagnostic methods for AFI .....                                                        | 31          |
| 1.5        | Aim of the study.....                                                                   | 43          |
| 1.5.1      | Research and Significance.....                                                          | 43          |
| 1.5.2      | Objectives of the study.....                                                            | 44          |
| <b>2.0</b> | <b>METHODS AND MATERIALS.....</b>                                                       | <b>45</b>   |
| 2.1        | Investigating the seroprevalence and spatial distribution of Arboviruses in Uganda..... | 45          |
| 2.1.1      | Study population, sentinel sites and period .....                                       | 45          |
| 2.1.2      | Selection and exclusion criteria .....                                                  | 45          |
| 2.1.3      | Detection of Arboviruses in samples collected in Uganda .....                           | 46          |

|                                                                                                         |           |
|---------------------------------------------------------------------------------------------------------|-----------|
| 2.1.4 Data analysis for Arboviruses .....                                                               | 46        |
| 2.2 Screening for viral AFI in Arbovirus negative samples using RT-PCR .....                            | 49        |
| 2.2.1 Designing the positive control concatemer .....                                                   | 49        |
| 2.2.2 Primer selection for virus detection.....                                                         | 49        |
| 2.2.3 Sample selection criteria.....                                                                    | 56        |
| 2.2.4 Sample Pooling .....                                                                              | 57        |
| 2.2.5 Sample shipment .....                                                                             | 57        |
| 2.2.6 Nucleic acid extraction, cDNA synthesis and quality control .....                                 | 57        |
| 2.2.7 Polymerase Chain Reaction.....                                                                    | 58        |
| 2.2.8 Gel electrophoresis.....                                                                          | 60        |
| 2.2.9 Sanger sequencing and analysis .....                                                              | 61        |
| 2.3 Molecular detection of West Nile virus .....                                                        | 62        |
| 2.3.1 Primer design for West Nile virus .....                                                           | 62        |
| 2.3.2 PCR, Sanger sequencing and analysis of West Nile virus .....                                      | 62        |
| 2.3.3 West Nile virus genome analysis .....                                                             | 65        |
| 2.4 Screening samples for viruses using next-generation sequencing.....                                 | 67        |
| 2.4.1 Sample selection and RNA quantification.....                                                      | 67        |
| 2.4.2 Library preparation from RNA.....                                                                 | 67        |
| 2.4.3 Next-generation sequencing .....                                                                  | 68        |
| 2.4.4 Using IDseq to analyse NGS data for virus detection.....                                          | 69        |
| 2.4.5 Validation of metagenomics hits.....                                                              | 69        |
| <b>3.0 Investigating the prevalence and spatial-temporal distribution of Arboviruses in Uganda.....</b> | <b>74</b> |
| 3.1 Introduction.....                                                                                   | 74        |
| 3.2 Results .....                                                                                       | 76        |
| 3.2.1 Description of the study population characteristics .....                                         | 76        |
| 3.2.2 Seasonality of arboviruses .....                                                                  | 78        |
| 3.2.3 Risk factors for Arbovirus infections. ....                                                       | 81        |
| 3.2.4 Spatial and spatiotemporal analysis of Arboviruses.....                                           | 83        |
| 3.3 Discussion .....                                                                                    | 88        |
| 3.4 Conclusion.....                                                                                     | 91        |
| <b>4.0 Screening for viral AFI in IgM - negative using RT-PCR .....</b>                                 | <b>92</b> |
| 4.1 Introduction.....                                                                                   | 92        |

|                                                                                                                                                                       |            |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 4.2 Results .....                                                                                                                                                     | 94         |
| 4.2.1 Designing the positive control concatemer .....                                                                                                                 | 94         |
| 4.2.2 General analysis of viral and non-viral AFI .....                                                                                                               | 97         |
| 4.2.3 Detection of Enteroviruses in Serum and whole blood .....                                                                                                       | 98         |
| 4.2.4 Detection of Measles virus in whole blood.....                                                                                                                  | 102        |
| 4.2.5 Screening for Arenaviridae, Filoviridae, Hepeviridae, Nairoviridae,<br>Peribunyaviridae, Phenuiviridae, Pneumovirinae, Togaviridae, and Rhabdoviridae.<br>..... | 105        |
| 4.2.6 Screening for Flaviridae .....                                                                                                                                  | 106        |
| 4.2.7 Quality control of samples shipped from Uganda to the United Kingdom .....                                                                                      | 106        |
| 4.3 Discussion .....                                                                                                                                                  | 107        |
| 4.4 Conclusion.....                                                                                                                                                   | 111        |
| <b>5.0 Detecting West Nile Virus after 80 years of silent transmission in Uganda .....</b>                                                                            | <b>112</b> |
| 5.1 Introduction.....                                                                                                                                                 | 112        |
| 5.2 Results .....                                                                                                                                                     | 114        |
| 5.2.1 PCR, Gel electrophoresis and Sanger sequencing of Flaviviruses .....                                                                                            | 114        |
| 5.2.2 Analysis of West Nile virus primers.....                                                                                                                        | 115        |
| 5.2.3 Analysis of West Nile Virus genome recovery .....                                                                                                               | 116        |
| 5.2.4 Phylogenetic analysis of WNV .....                                                                                                                              | 117        |
| 5.2.5 Analysis of the genetic diversity of WNV .....                                                                                                                  | 119        |
| 5.3 Discussion .....                                                                                                                                                  | 124        |
| 5.4 Conclusion.....                                                                                                                                                   | 127        |
| <b>6.0 Screening samples for viruses using metagenomics Next-Generation Sequencing ..</b>                                                                             | <b>128</b> |
| 6.1 Introduction.....                                                                                                                                                 | 128        |
| 6.2 Results .....                                                                                                                                                     | 131        |
| 6.2.1 Sample selection.....                                                                                                                                           | 131        |
| 6.2.2 Quality control of the laboratory procedure .....                                                                                                               | 132        |
| 6.2.3 Analysis of the positive control sample (Polyomavirus).....                                                                                                     | 135        |
| 6.2.4 Sequence analysis using the IDseq platform .....                                                                                                                | 136        |
| 6.2.5 Analysis of Pegiviruses using the IDseq platform.....                                                                                                           | 140        |
| 6.2.6 Analysis of Measles virus using the IDseq platform.....                                                                                                         | 144        |
| 6.2.7 Analysis of the Polyomavirus .....                                                                                                                              | 144        |
| 6.2.8 Analysis of the non-redundant nucleotide reads.....                                                                                                             | 145        |

|                                                   |            |
|---------------------------------------------------|------------|
| 6.3 Discussion .....                              | 147        |
| 6.4 Conclusion.....                               | 153        |
| <b>7.0 Final discussion and conclusions .....</b> | <b>154</b> |
| <b>8.0 Future Work .....</b>                      | <b>162</b> |
| <b>9.0 References. ....</b>                       | <b>163</b> |

## vi. List of Figures

|                                                                                                                                                                                  |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1: Phylogenetic tree of family Flaviviridae.....                                                                                                                          | 9   |
| Figure 2 Cartoon structure of Picornavirus genome arrangement.....                                                                                                               | 15  |
| Figure 3: A Cartoon cross-section of the structure of Paramyxoviruses..                                                                                                          | 21  |
| Figure 4. Genomic organisation of Paramyxoviruses.....                                                                                                                           | 21  |
| Figure 5. Sample collection by region.....                                                                                                                                       | 77  |
| Figure 6. Regions with arboviruses confirmed by PRNT.....                                                                                                                        | 78  |
| Figure 7. Number of arbovirus positives by month in Uganda from March 2016 to November 2019. ....                                                                                | 79  |
| Figure 8. PRNT confirmed YFV (red) and other arboviruses (blue) cases observed in Uganda from 2016 to 2019..                                                                     | 80  |
| Figure 9. Spatio-temporal distribution of arboviruses in Uganda in clock-wise direction..                                                                                        | 87  |
| Figure 10: Order of the APC concatemer..                                                                                                                                         | 95  |
| Figure 11. Gel electrophoresis of APC amplicons of PCR assay after a two-fold serial dilution..                                                                                  | 97  |
| Figure 12. Gel electrophoresis image of enterovirus-positive samples using the Nix et al. assay and method.....                                                                  | 99  |
| Figure 13. Phylogenetic tree showing selected enteroviruses following the amplification of the VP1 region..                                                                      | 101 |
| Figure 14. Phylogenetic tree of Measles viruses in Uganda and select sequences from elsewhere, following the amplification of PCR product of circa 600bp in the N450 region..... | 103 |
| Figure 15 Visualisation of the H gene's PCR amplicons (approximately 600bp) after gel electrophoresis..                                                                          | 104 |
| Figure 16 Phylogenetic tree for MV B3 genotype after amplification of the H-gene..                                                                                               | 105 |
| Figure 17: Visualisation of flavivirus.....                                                                                                                                      | 114 |
| Figure 18. Cartoon of WNV genome. ....                                                                                                                                           | 117 |
| Figure 19: Molecular phylogenetic analysis of WNV using the maximum-likelihood method with GTR model..                                                                           | 119 |
| Figure 20. Sequence distance comparison between WNV Kasere versus WNV1937 group. ....                                                                                            | 121 |
| Figure 21. N- glycosylation of PrM and Envelope proteins. ....                                                                                                                   | 122 |
| Figure 22. N-linked glycosylation of the lineage 1 virus.....                                                                                                                    | 123 |
| Figure 23. The signs and symptoms associated with the patient samples selected for NGS. ....                                                                                     | 132 |
| Figure 24. Assembled PyV positive control recovered by IDseq pipeline following mapping to the reference genome LR216248.1.....                                                  | 136 |
| Figure 25. Analysis of pool 609 using the IDseq bioinformatics pipeline. ....                                                                                                    | 136 |

|                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 26. Phylogenetic tree showing 38 selected HCV from the East African region except accession number MN894517, from Egypt..... | 139 |
| Figure 27. Cartoon of Pegivirus C reads and contigs identified in pool 202.. ..                                                     | 140 |
| Figure 28. A single pegivirus A contig was identified in pool 202.. ..                                                              | 141 |
| Figure 29. Pegivirus A in pool 609.. ..                                                                                             | 141 |
| Figure 30. Phylogenetic tree showing 41 selected pegiviruses complete genomes .....                                                 | 143 |
| Figure 31. Cartoon of the MeV genome recovered using IDseq.....                                                                     | 144 |
| Figure 32. Cartoon of the PyV genome recovered using IDseq. ....                                                                    | 145 |

## Vii. List of Tables

|                                                                                                                                                                                                            |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 1. Table of the Pros and Cons of the different methods of virus detection. ....                                                                                                                      | 41  |
| Table 2. Primers used to screen for viruses according to their Viral families or genera.<br>.....                                                                                                          | 50  |
| Table 3. In-house primers designed to detect the MeV Hemagglutinin gene and the<br>N450 region primers for the detection of the Nucleoprotein gene adopted from<br>Bankamp <i>et al.</i> *.....            | 55  |
| Table 4 shows Primer sets generated using Primal Scheme software showing their<br>fragment size and genomic position. ....                                                                                 | 63  |
| Table 5. Primers for the detection of HCV adopted from Adeboyejo <i>et al.</i> ....                                                                                                                        | 70  |
| Table 6. Pegivirus primers for the first and second PCR and the expected size of the<br>amplicon. ....                                                                                                     | 71  |
| Table 7. Primers by Palacios <i>et al.</i> OrbiVP1-F2494-1 targets Orungo virus, OrbiVP1-<br>F2494-2 and F2494-4 target the AHSV serotypes, OrbiVP1-F2494-3 targets Wallal<br>and Changuinola viruses..... | 72  |
| Table 8. In house primers designed from the IDseq translated nucleotide orbivirus<br>reads.....                                                                                                            | 72  |
| Table 9. The pan polyoma primer set was adopted from Johne <i>et al.</i> to detect<br>polyomaviruses.....                                                                                                  | 73  |
| Table 10. PRNT confirmed arboviruses and their distribution across age groups,<br>geographical regions and gender.....                                                                                     | 80  |
| Table 11. Epidemiological characteristics of arboviruses. Abbreviation: HCW-<br>Healthcare workers. ....                                                                                                   | 82  |
| Table 12. Table of pathogens used to design the APC in their synthetic order. The<br>PCR amplicon size (bp) differed from the expected size.....                                                           | 95  |
| Table 13. Clinical signs and symptoms of serotyped enteroviruses .....                                                                                                                                     | 100 |
| Table 14. Table showing concentration and corresponding RIN for whole blood<br>specimens.....                                                                                                              | 133 |
| Table 15. Library concentrations (ng/μL) before and after HTS library preparation<br>procedure.....                                                                                                        | 134 |
| Table 16. Viruses identified by NGS/HTS after bioinformatics analysis of IDseq data..<br>.....                                                                                                             | 137 |

# 1. INTRODUCTION

## 1.1 Acute Febrile Illnesses

Acute Febrile Illnesses or Infections (AFI) have been defined differently by several researchers, but as a consensus, it is a fever spanning up to 7 days without a known cause of infection [1]. Fever in AFI is a measured auxiliary temperature  $\geq 38^{\circ}\text{C}$  [2]; however, some individuals may take anti-pyretic drugs before visiting health centres. The signs and symptoms of AFI do not vary from other febrile illnesses of known causes. The clinical presentation of AFI may include but not be limited to; fever, malaise, headache, generalized myalgia and arthralgia, neck stiffness, a metallic taste, sore throat, abdominal pain sometimes with diarrhoea, running nose and maculopapular rash [3]. These symptoms vary with the disease agent and probably the individual's immune status. In another observation by Daumas *et al.*, the authors concluded that some physical signs were infrequent or presented with mild intensity[1]. In this study, the authors also included exanthema, pallor, lymphoid enlargement and eye congestion as part of the AFI clinical presentation.

Fevers of unknown origin (FUO) or Pyrexia of unknown origin (PUO) differ from AFI by nature of their prolonged illness (which spans for more than three weeks) [4] and thus fall out of the most agreeable definition of AFI. However, if a patient presents to the hospital during the early manifestation of the signs and symptoms of the illness, there would be no clear-cut distinction between the two; thus, a classical FUO can qualify as AFI.

Therefore given the above definition, the signs and symptoms; AFI can be of bacterial, fungal, parasitic or viral origin.

### 1.1.1 Global burden and overview of Acute Febrile Illness

The global burden of AFI encompasses bacterial [5] and fungal bloodstream infections, parasites such as malaria, and viral infections [6], [7], [8]. These infections have been investigated in different parts of the world, with investigations at times focusing on specific pathogens [9] whilst others broader [10], [11].

A high percentage of the known global burden of AFI is attributed to Malaria, Tuberculosis (TB) with and without the Human immunodeficiency virus (HIV), and bacterial infections [12]. With significant positive progress towards effective malaria control, viral AFI has become more noticeable globally. In most instances, viral AFI has been known to be self-limiting [13]; however, with the emergence of Ebola [14], Nipah [15] and more recently SARS-CoV- 2 viruses, among other emerging and re-emerging viral infections; the threat posed cannot be underestimated. For instance, alphaviruses such as the Mayaro [16] and Madariaga viruses [17] were detected in routine surveillance studies conducted among school children in Haiti. Other alphaviruses (Venezuelan equine encephalitis), flaviviruses (Dengue virus serotypes 1-4) and orthobunyaviruses (Oropouche virus, Group C viruses, and Guaroa virus) were detected in over 32% of participants in a study carried out in Bolivia, Ecuador, Paraguay, and Peru [18]. It is also important to note that a significant proportion of infections still go undiagnosed. For example, 26% of the 403 AFI patients had a bacterial infection in the Caribbean Islands, and only 1.7% was attributed to a virus. Over 42.7% of the patients were discharged without a diagnosis [19]. Another study in Thailand showed that bacterial and viral pathogens caused 38.7% of the

AFI, whilst 61.3% of the AFI also remained unknown [20]. Given the above data, it is evident that non- malarial infectious diseases with a bias to arboviral pathogens are an emerging human threat on the South American continent.

In Asia, an observational study carried out between 2007 to 2008 in South India showed in descending order; scrub typhus, malaria, enteric fever, dengue fever, leptospirosis, spotted fever, rickettsiosis and Hantavirus as the leading aetiologies of AFI [21]. The study, however, does not provide a comprehensive AFI burden since diagnosis was limited to enzyme-linked immunosorbent assay (ELISA) technique and hospitalised patients only. The general picture of AFI in Cambodia was more or less similar to that seen in south India but instead provided an in-depth investigation to include bacterial, viral, tick-borne and mosquito-borne parasites; as aetiologies of both blood-borne and respiratory illnesses [22]. In agreement with other investigations carried out on the continent (that is, Indonesia, Malaysia, Philippines, Sri Lanka, Thailand and Vietnam), and including countries previously mentioned above (Cambodia and India), the most common viral causes of AFI were Chikungunya (CHIKV) and Dengue (DENV) viruses [23], [24], [25]. A report from Indonesia showed that after excluding Dengue virus infection, the Japanese Encephalitis virus (JEV) was a cause of non-encephalitic AFI in 18% of the patients from Bali, Indonesia [26]. Importantly, these investigations did not aim to uncover other unknown causes of AFI outside their available laboratory testing capacities. In other words, all these investigations were tailored to identify specific aetiological agents, which is a massive disadvantage towards pathogen discovery, as a method of detecting previously unknown agents of microbial disease.

In sub-Saharan Africa (SSA), AFI has remained poorly diagnosed [10], yet they are a crucial threat to public health. This challenge may be attributed to several factors, including the diagnostic capacity of laboratories, quality of test specimens, and timely collection of specimens during the acute phase of the illness. In addition, there may be different reasons that influence a probable patient's decision to seek medical attention at a health centre, such as the availability of essential drugs and the distance to the health centre. Thus the ability to capture all viral AFI during the acute phase is likely under-reported. Nevertheless, viral AFI has been detected and shown to cause respiratory tract infections [27] and bloodstream infections [28]. According to studies done in Tanzania and Sierra Leone, the undiagnosed burden of AFI in investigated cases ranged from 7.5% to 70% [28], [29] and [30]. Further investigations in Sierra Leone showed that of the 70% Lassa virus-negative specimens, other alternative viral causes were identified in 25% of individuals [31].

In the East African region, most research in Tanzania and Malawi has focused on bacterial, fungal and parasitic pathogens as aetiological causes of AFI in bloodstream infections (BSI), [32], [33]; with more emphasis on bacterial causes. Hercik *et al.* stated that Plasmodia was the primary cause of AFI in Tanzania, followed by leptospirosis, and only 1% ascribed to the West Nile virus [34]. However, another report in Tanzania has shown that malaria may not necessarily cause fever in a patient, whereby in 60% of cases with clinical malaria, only 1.6% of the cases presented with a fever [35]. The report identified Chikungunya virus in 7.9% of the samples tested; however, no diagnosis was achieved in 40% of the samples.

### 1.1.2 Acute Febrile Illness in Uganda.

The diagnosed zoonotic blood-borne viral aetiological agents of febrile illness in Uganda include the Rift valley fever virus (RVFV) [36], the yellow fever virus (YFV) [37], Marburg virus [38] and Ebola virus. In addition, other notable viruses such as the Bwamba, Semliki forest, Bunyamwera, Ntaya, Uganda S, Mengo, West Nile (WNV), and Zika (ZIKV) viruses [39] were first identified in Uganda. Incidences of disease outbreaks in Uganda date back from the 1930s, and among the first recorded were bacillary dysentery, yaws, and venereal disease [40], [41]. Since then, several viral AFI outbreaks have been documented, such as the Ebola haemorrhagic fever outbreak in 2000 [42],[43], Marburg virus outbreak in 2017 [38], ten recorded outbreaks of RVFV [44],[45], influenza viruses [46], measles virus among others [47]. From August 2017 to August 2018 alone, eight different disease outbreaks were recorded and attributed to haemorrhagic fevers, viral infections, and other bacterial causes [48]. In response to the recurrent outbreaks, multi-sectoral organizations in animal welfare, zoonotic diseases and vectors, and the Ministry of Health (MoH) unified under the one health umbrella to mitigate the problem [49]. The one health approach concluded that the zoonotic diseases of high priority in Uganda are anthrax, zoonotic influenza viruses, viral haemorrhagic fevers (VHF), brucellosis, African trypanosomiasis, plague, and rabies [49]. Whilst some viral families such as Flaviviridae (YFV in particular), Phenuiviridae (for RVF virus), and Nairoviridae (for CCHF virus) have been implicated in outbreaks of VHF in Uganda, the above approach does not underscore the urgency to monitor and detect emerging and re-emerging viral diseases, that do not lead to VHF syndromes.

Given the imminent threat of bloodborne pathogens to the region, the role of systematic scientific surveillance of AFI has never been any more critical. For this reason, the Department of Arbovirology at the Uganda Virus Research Institute (UVRI), with support from the MoH, rose to the challenge. The surveillance activities of the department encompass both blood-borne and respiratory borne pathogens. Since 2006, the UVRI has carried out active surveillance of influenza viruses and occasional investigation of other viral causes of respiratory disease [50]. Previously detected respiratory viruses include adenoviruses, human rhinovirus, coronavirus strains OC43 and 229E, human metapneumoviruses and respiratory syncytial viruses [51]. Thus, these activities provided a strong foundation for the surveillance of the novel pandemic betacoronavirus severe acute respiratory syndrome (SARS-CoV 2) in 2019 [52]. In addition, the UVRI carries out routine serological surveillance of YFV, WNV, ZIKV, DENV and alphaviruses (CHIKV and ONNV).

New laboratory diagnostic techniques have highlighted the fact that malaria is probably clinically over- implicated as the cause of AFI in SSA [35], [53] and where confirmed, it may exist as a co-infection [54]. There is an urgency to understand the scope of other fever-related viral pathogens with non- haemorrhagic fever syndromes (nVHF) in Uganda. Bridging this knowledge gap with viral pathogens that are co-infecting with malaria [55], [56], [57] would improve patient management, provide accurate pathogen diagnosis and also encourage appropriate drug prescription and dispensing in the SSA region.

## 1.2 Viruses as aetiological agents of AFI

From the previously described history of outbreaks in Uganda, important nVHF have been described for viral families *Nairoviridae*, *Peribunyaviridae*, *Phenuiviridae*, *Flaviviridae*, *Paramyxoviridae* and *Picornaviridae*. For this thesis, the viral families of public health importance in Uganda, that is to say, *Flaviviridae*, *Picornaviridae*, *Paramyxoviridae*, and *Togaviridae*, will be discussed.

### 1.2.1 *Flaviviridae* and its role in AFI

#### *1.2.1.1 Taxonomy and Host range*

This family consists of four genera, namely, *Flavivirus* (53 species), *Hepacivirus* (14 species), *Pegivirus* (11 species) and *Pestivirus* (11 species).

Initially termed as “Group B arboviruses”, to differentiate them from “Group A arboviruses” (now known as *Alphaviruses*), *flaviviruses* emerged as a distinct genus within the *Flaviviridae* family [58]. The *flavivirus* genus derives its name from the Latin word *flavus* for ‘yellow’ due to its type species YFV. *Flaviviruses* were classified into eight complexes [59] according to antigen relatedness of members by cross-neutralisation tests. These separate antigenic complexes are; Tick-borne encephalitis comprising of 12 viruses, Rio Bravo (6 viruses), Japanese encephalitis (10 viruses), Tyuleny (3 viruses), Ntaya (with 5 viruses), Uganda S (4 viruses), Dengue (4 viruses), and Modoc (4 viruses), while several other viruses were left unclassified antigenically [60]. The International Committee on Taxonomy of Viruses (ICTV) has updated the virus species groups by defining them according to their known vectors and hosts, and thus groups them as; tick-borne with a

mammalian host (14 species), tick-borne with a seabird as a host (3 species), probable tick-borne (one species), mosquito-borne, Aroa virus group (4), mosquito-borne, Dengue virus group (4), mosquito-borne, Japanese encephalitis group (10), Mosquito-borne, Kokobera virus group (2), mosquito-borne, Ntaya virus group (7), mosquito-borne, yellow fever virus group (3), probable mosquito-borne, Kedougou and Edge Hill virus groups (8), unknown vector, Entebbe bat, Modoc and Rio Bravo virus groups (17 species). Members in this genus are primarily arthropod-borne viruses. Whilst DENV and YFV have a narrow host range from a few species of arthropods and vertebrates, WNV has been detected in over 49 species of mosquitoes and ticks, 225 species of birds, and over 30 animals [61], [62], [63].

Hepaciviruses have been classified into 14 species, Hepacivirus A – N, with the most extensively studied Hepacivirus as Hepatitis C virus grouped in Hepacivirus C species. The viral species in this genus are defined by their host range, that is, Hepacivirus A in non – primates (horses and dogs) [64], [65], Hepacivirus B in primates [66], Hepacivirus D in monkeys [67], Hepacivirus E- J in various rodent species [68], Hepacivirus K- M infecting bats [69], and Hepacivirus N infecting cattle [70], [71].

Pestiviruses was initially comprised of 4 species but later revised to be host-independent, and hence re-categorised as 11 Pestivirus species, A- K [72]. These viruses infect pigs and a range of ruminant animals [73]. Pegiviruses were initially classified within Hepaciviruses [74], but due to improvements in sequence datasets available for phylogenetic, phylogenetic relatedness and sequence distance were re-examined and thus, species Pegivirus A- K was created [75].

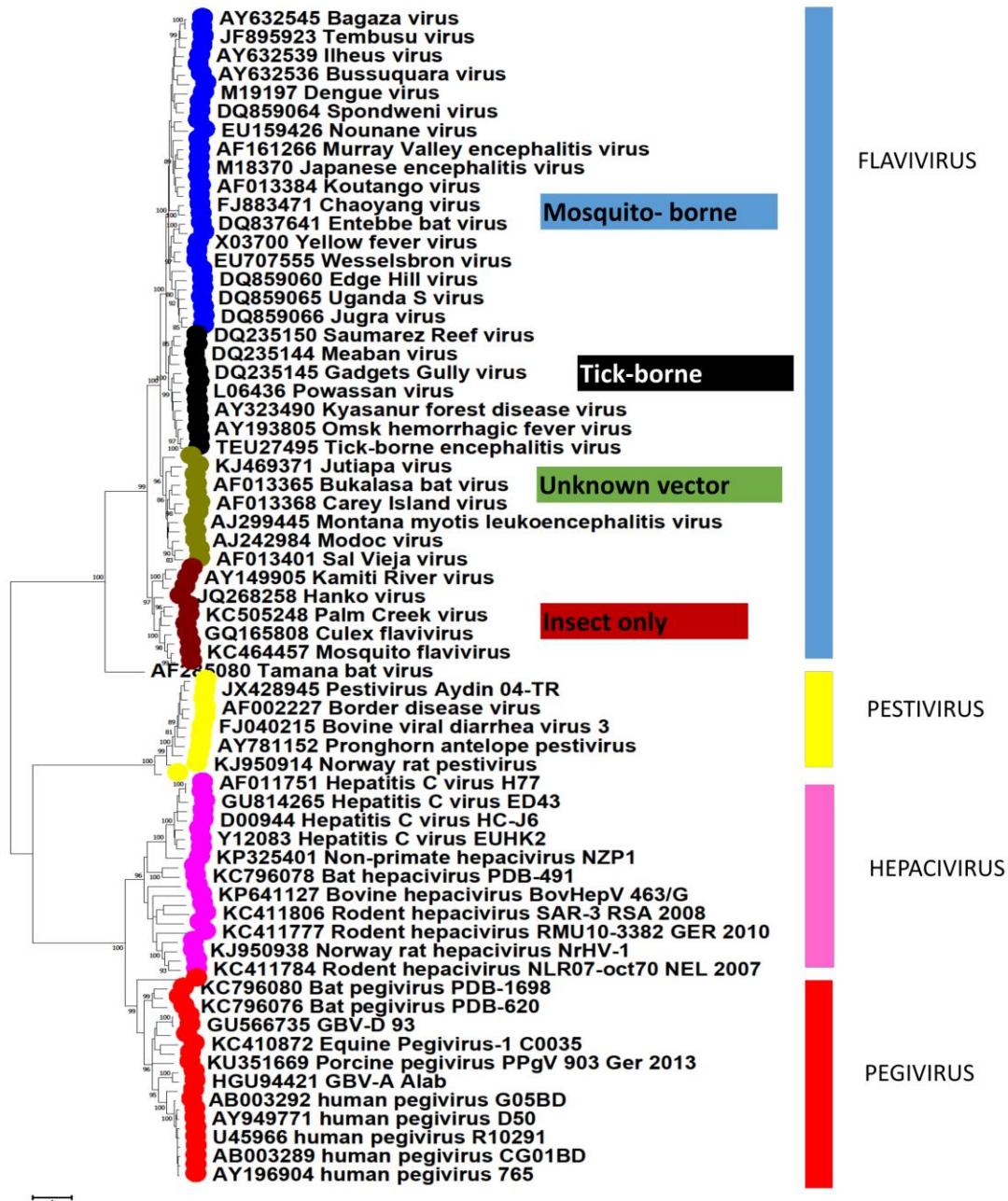


Figure 1: Phylogenetic tree of family Flaviviridae. The tree shows the 4 genera (Flavivirus- in turquoise, Pestivirus in yellow, Hepacivirus in pink and Pegivirus in red) within the family. Genus Flavivirus has been sub- divided to show viruses that are transmitted by mosquitoes, ticks, unknown vectors, and insects. The tree represents conserved regions of RNA dependent RNA Polymerase (RdRp) of members of Flaviviridae. Representative sequences were aligned with MEGA v.7 using the optimal substitution model (LG+G+I), for 100 replications. Accession numbers are provided in

*taxa labels and the values next to the branches are  $\geq 70\%$  as adapted from*  
[https://talk.ictvonline.org/ictv-reports/ictv\\_online\\_report/positive-sense-rna-viruses/w/flaviviridae](https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/flaviviridae).

### *1.2.1.2 Morphology and genome*

Members have a spherical, 40- 60nm diameter virion and a lipid envelope. Flaviviridae viruses are small enveloped viruses with positive-sense, single-stranded (ss) RNA genomes, ranging between 9.2 -11, 8.9 -10.5, 8.9 -11.3, 12.3 - 13kb for Flavivirus, Hepacivirus, Pegivirus, and Pestivirus, respectively [76]. Generally, the genome in the four genera encodes a variant sum of structural and non-structural proteins but in principle contains Capsid and Envelope proteins and NS1 (in Flaviviruses), NS2a, NS2b, NS3, NS4a, NS4b and NS5) respectively, flanked by non-coding regions (NCR) at the 5'- and 3'- terminal ends. Both NCR regions are involved in viral RNA translation, replication and packaging whilst viral NS2B-NS3 serine and cellular proteases process the viral polyprotein to form mature structural and non-structural proteins. The translation is initiated in the 5' NCR by a type IV internal ribosomal entry site (IRES) in Hepaciviruses, Pestiviruses, and Pegiviruses, although some members of Pegiviruses possess a type 1 IRES. In Flaviviruses, translation is initiated by a type I cap ( $m^7GpppAmp$ ) - dependent. Briefly, Type I cap-dependent translation involves the recognition of the  $m^7GpppN$  cap structure at the 5' end of mRNA by the eukaryotic translation initiation factor 4E (eIF4E). This interaction recruits other initiation factors, including eIF4G and eIF4A, to form the eIF4F complex. The eIF4F complex then binds to the mRNA and scans the 5' untranslated region (UTR) to identify the start codon. Once the start codon is recognized, the ribosome is recruited and translation begins. The cap structure helps to stabilize the mRNA and promote translation initiation by facilitating the recruitment of initiation factors

and protecting the mRNA from degradation. The envelope protein forms two glycoproteins, E1 and E2, in Hepaciviruses, Pegiviruses, and Pestiviruses.

#### *1.2.1.3 Host range and economic burden of Flaviviridae*

Perhaps the most studied genera of this family are Flavivirus and Hepacivirus (in particular, the hepatitis C virus) due to their economic and public health importance.

Among Flaviviruses, Dengue accounts for up to 200 million cases globally and an estimated 20,000 mortality annually [77]. In Southeast Asia alone, between 2001- 2010, there was an annual average of 2.9 million Dengue virus disease episodes with over 5,900 deaths and a yearly economic loss of US\$950m in 12 countries [78]. Globally 71 million people are infected with HCV, and over 1.75 million new infections every year [79]. A 2017 publication by Blach et al. estimated the global seroprevalence of HCV at 1.1% according to data acquired from 100 countries [80]. This report indicated a generally low prevalence of HCV in SSA. However, data showing higher seroprevalence of HCV (Rwanda - 5%, Democratic Republic of Congo - 2.3%, Kenya - 13% including HCV RNA positive individuals, and Tanzania - 50%) has been reported [81] , [82], [83], [84]. A model by Lee et al. predicted that if the Zika virus were to spread in the six high-risk states of the USA at an attack rate of 2%, it would result in an economic burden of more than \$2 billion [85].

#### *1.2.1.4 Disease and Transmission*

Viruses within this family can cause various clinical illnesses, from non-specific viral syndromes to severe and fatal diseases. Disease manifestation resulting from pathogens within this family can be congenital microcephaly syndrome, as seen in ZIKV [86].

Dengue virus (DENV), which consists of four antigenically distinct serotypes, causes an undifferentiated fever, which may progress to Dengue fever [87], dengue haemorrhagic fever (DHF) [88] or Dengue shock syndrome, subject to either primary or secondary infection [89]. Infections with JEV may progress to neurological symptoms such as encephalitis [90], whilst WNV disease may be associated with encephalitis and meningitis [91]. Hepatitis C virus (HCV), causes acute liver inflammation (hepatitis), which may lead to liver cirrhosis and or renal disease [92], [93]. Other non-specific disorders associated with Flaviviridae infections include joint and muscle pain, macular rash, and gastrointestinal pain.

Transmission can occur through a bite from vectors such as mosquitoes for mosquito-borne viruses and ticks – for tick-borne viruses. However, the main route of infection with HCV occurs through blood transfusion, high-risk drug usage and sexual exposure [94]. It is also worth noting that ZIKV, previously thought to be an exclusive mosquito-borne virus, was implicated in sexual transmission [95]. Other modes of transmission include organ transplant and blood transfusion [96] [97], breastfeeding, placental transmission, and occupational exposure for health workers.

#### *1.2.1.5 Treatment and Management*

Once a naïve individual is inoculated with an infectious dose of a pathogen, recovery from disease in the absence of vaccines and therapeutics is entirely up to the immune system. Once an infection has been established, disease prognosis varies among individuals (i.e. age, immune status) from asymptomatic, mild to severe or a fatal outcome.

The most effective method for controlling the spread of infectious pathogens is through vaccination. If there are no known therapeutics or vaccines, treatment is usually through the management of symptoms and supportive care. One of the breakthroughs in the treatment of flaviviral infections was achieved when an individual inoculated with a mouse-attenuated YFV coupled with serum from a YFV survivor was able to mount an immune response against wild-type YFV in the 1930s [98]. By the end of 1953, over 56 million Africans and other high-risk populations had been vaccinated against YFV [99]. However, this vaccine posed two limitations: its association with a high risk of neurotropism and the combination with immune sera was not practically sustainable. A chance mutation in an *in vitro* YFV culture led to the discovery of the current YFV 17D vaccine, which has improved lives globally [100], [101]. With success attained in controlling YFV, other vaccines such as the Tick-borne encephalitis virus (TBEV) vaccine in 1973 [102] and the JEV vaccine (1985) [103], [104] were introduced. WNV vaccines mounting an acceptable immune response in the equine species were developed and are currently used [105], [106]. However, this success story of WNV vaccines is eclipsed by the lack of a human vaccine, with the best candidates yet to proceed beyond phase II of the clinical trials [107].

A significant effort has been directed toward developing vaccines for the other Flaviviridae members. Currently, the most effective treatment for HCV is by administration of direct-acting antivirals (DAA) [108], [109]. A thorough discussion of DAA disadvantages, such as the cost of therapeutics, which significantly limits its use in developed countries and the emergence of DAA-resistant strains resulting from resistance-associated substitutions in circulating strains, have been discussed by Duncan et al. [110]. Earlier

attempts to design vaccines for HCV included the use of Hepatitis C virus-like particles to elicit an immune response in mice [111] and non-human primates [112]. Other more recent attempts have been promising, with some vaccines proceeding to early phases of clinical trials. These vaccines involved the use of HCV structural proteins, that is, envelope (E) 1 and 2 [113], while others are non-structural proteins inserted in a viral vector [114]. However, these did not proceed beyond phase II after failing to offer increased protection against chronic HCV compared to the placebo [115], [116]. Among Pestiviruses, vaccines are available for the classical swine fever virus among pigs [117], [118], whilst at the time of writing, no documented Pegivirus vaccine has been developed.

## 1.2.2. Picornaviridae as causative agents of AFI

### *1.2.2.1 Taxonomy and Morphology*

This family is thought to be growing as more viruses are being discovered whilst others remain unclassified by the ICTV. The name 'Picorna', earlier criticized by Andre Lwoff as translated to mean a unit of measure,  $10^{-12}$  (pico) RNA (rna), has stood the test of time to become an internationally recognised scientific virus family name [119]. This family consists of small, non-enveloped icosahedral viruses approximately 30 – 32nm in diameter. With few exceptions, such as Kobuviruses, they are nearly a feature-less sphere under an electron microscope and comprise 63 genera with 147 species. They possess a single-stranded, positive-sense RNA genome, ranging from 6.7 – 10.1 kb. This genome has a single open reading frame (ORF), except for Dicipiviruses, which possess two ORFs. The ORF encodes the capsid proteins whilst the latter encodes non-structural proteins [120]. The capsid, which consists of 60 protomers, if completely cleaved, makes

up the four capsid proteins, namely; 1A (Viral Protein 4, VP4), 1B (VP2), 1C (VP3) and 1D (VP1) or three capsid proteins if 1AB (VP0) remains un- cleaved [121]. Due to increased awareness of the variation in molecular weight of VP4 - VP1 within the family, describing the viral genome VP as a measure of size has become unreliable [122]. Thus it is generally recommended to adopt the 1A – 1D system that describes the arrangement of the proteins in the polyprotein. The general layout of the Picornavirus genome is shown in Figure 2.

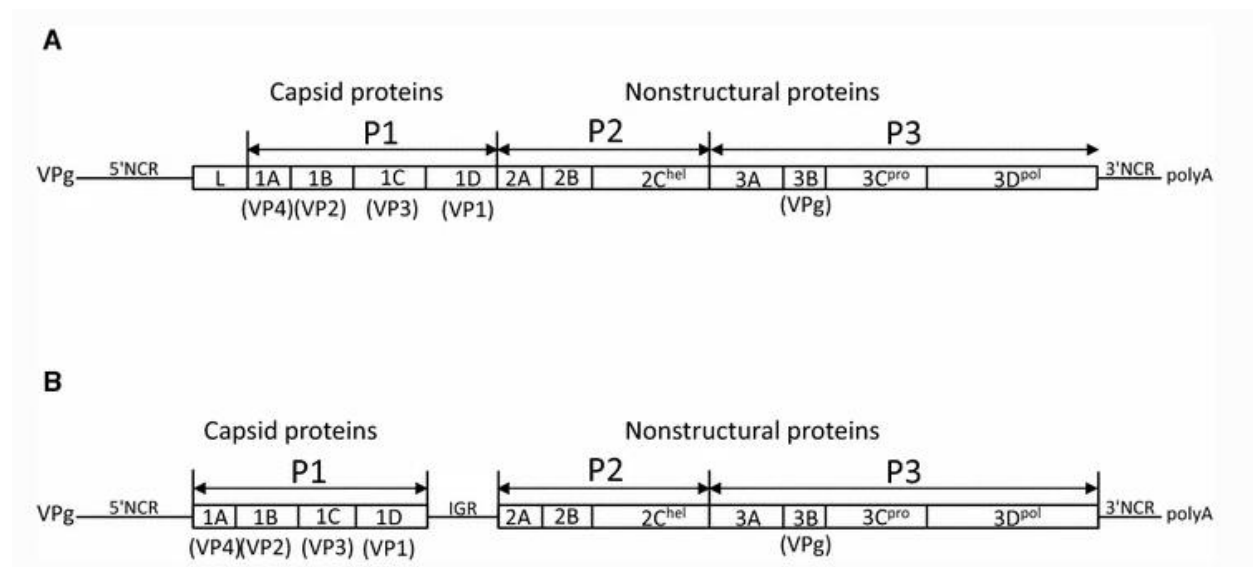


Figure 2. Cartoon structure of Picornavirus genome arrangement. Adopted from Roland Zell et al. [122]. Cartoon A shows the general genome organisation of picornaviruses flanked by a non-coding region (NCR) on the 5' and 3' end. The 5' end is linked to the tyrosine-3 residue of a small peptide (VPg) whilst the 3' end is polyadenylated. The ORF encodes for the capsid protein, which forms 1A-AD and the non-structural proteins P2 and P3, whilst B shows the genome arrangement of dicapiviruses in genus *Dicapivirus*, which encodes for 2 polyproteins, i.e. the capsid and non-structural proteins.

### *1.2.2.2 Transmission and Disease*

The transmission mode of most picornaviruses is via the faecal-oral route, through the consumption of contaminated food and water, or directly from person to person via contaminated hands [123]. Other forms of transmission such as saliva, respiratory droplets or discharge, such as rhinoviruses, conjunctival secretion, and skin lesions exudates, have also been reported [124]. Notably, these viruses cause disease in a broad spectrum of hosts such as humans [125], monkeys and non-human primates [126], and fish species [127], with birds [128], insects, and flies as potential vectors [129]. As with most viruses, picornaviruses are self-limiting and associated with asymptomatic to mild disease. However, where clinical manifestations occur, they can range from a mild, undifferentiated AFI to a life-threatening or fatal outcome. Apart from the Poliovirus, infections caused by non- Polio enteroviruses (NPEV) are not reported internationally. However, published data shows that NPEVs largely contribute to the global disease burden, with several clinical syndromes such as acute flaccid paralysis [130], respiratory symptoms, hand, foot and mouth disease, AFI [131], and acute haemorrhagic conjunctivitis [132]. In addition, enterovirus species and type 1 diabetes are correlated since these viruses were detected in significant abundance within the risk population compared to matched control groups [133], [134].

Perhaps wealthier countries acknowledge the impact of picornaviruses through economic losses in the agriculture and food processing industry. For example, species from the genus *Limnivirus* were isolated and detected as the aetiological agent of large-scale mortality in ornamental fish, leading to significant losses in a multimillion-dollar industry in the USA [127]. In the UK, the foot and mouth disease virus led to the culling of an

estimated 6.5 million animals [135], with over US\$ 4 billion lost in tourism control measures in 2001, not to forget the annual loss of US\$ 6 – 21 billion in endemic regions [136].

#### *1.2.2.4 Treatment and Management*

Perhaps one of the most iconic vaccines in this family was the introduction of the inactivated [137] and later the live-attenuated Poliomyelitis virus vaccines [138]. Other important vaccines targeting pathogens within Picornaviridae for human and animal diseases have been manufactured. Among the human vaccines, the Enterovirus – A71 vaccines, which immunise against hand, foot and mouth disease, mounting over 90% neutralising antibodies (nAbs), have completed phase III of human clinical trials [139], [140]. These are currently approved for use in China [141]. Enterovirus D68 of species D has attracted attention due to its Polio-like neurologic symptoms [142], [143]; thus, approaches to vaccine design using suitable models are underway [144], [145]. Attempts to design rhinovirus vaccines have proved problematic due to numerous viral serotypes and the lack of cross-serotype protection [146]. However, Lee *et al.* proposed that using an inactivated polyvalent Rhinovirus vaccine would elicit broadly nAbs in the presence of an adjuvant [147]. Immunisation against Hepatitis A virus, one of many viral aetiologies of liver disease in all ages, has been available since 1995 [148]. Among domesticated animals and fowls, vaccines are available against Foot and Mouth Disease virus (FMDV) in cloven- hoofed animals [149] and avian encephalomyelitis virus in birds [150].

#### *1.2.2.5 Enteroviruses as aetiological agents of disease in Uganda*

Enteroviruses are a large genus within the family Picornaviridae and comprise 15 species, including 12 Enterovirus A – L and 3 species of rhinovirus A-C. The burden of disease associated with enteroviruses has not been established in Uganda. This may be because most enterovirus infections resolve without any clinical manifestation. In domesticated animals, specifically pigs, untreated manure has been indicated as a potential source of porcine enterovirus transmission [151]. A previous report on a lethal respiratory infection caused by human rhinovirus C in wild chimpanzees underscores how disastrous virus cross-over to new hosts can be [152]. Previous research detected enteroviruses in sewage water in Kampala, the capital of Uganda [153]. There has been limited information on the public health burden of enteroviruses in Uganda. In 1977, enterovirus species were considered an insignificant cause of respiratory illness among children in Uganda [154]. Epidemiological data appeared in a French Guiana publication, suggesting an outbreak of acute haemorrhagic conjunctivitis caused by Coxsackievirus A24 in a Ugandan prison in 2016 [132]. Not until recently in 2021, Tushabe *et al.* showed that Enterovirus A and B species isolates from stool were associated with acute flaccid paralysis (AFP), a debilitating syndrome in humans following a decade of surveillance of non-polio enteroviruses (NPEV) [155]. The investigation further showed that Enterovirus B species contributed to over 90% of the total isolates, indicating a unique dominance by the species for 10 years. Other research focusing on paediatric febrile patients detected several species of picornaviruses in serum using the NGS platform [156].

### 1.2.3 Paramyxoviruses as causative agents of AFI

Paramyxoviruses are RNA viruses, with some presenting highly contagious and pathologic phenotypes. Paramyxoviridae is divided into a total of four subfamilies, each further divided into genera and a total of 78 species, as described by Rima et al. 2019 [157].

#### *1.2.3.1 Taxonomy*

The subfamilies within Paramyxoviridae are Avulavirinae, Metaparamyxovirinae, Orthoparamyxovirinae and Rubulavirinae. Firstly, the sub-family Avulavirinae consists of 3 genera (Metaavulavirus, Orthoavulavirus and Paraavulavirus) and 22 species exclusive to birds. Metaavulavirus comprises avian paramyxoviruses (APMV) 2, 5 - 8, 10, 11, 14, 15 and 20. Orthoavulaviruses consist of 10 species, with Newcastle virus disease (NVD), also known as APMV-1 as the most important member. Members of this genus have been isolated from domestic and wild birds such as penguins [158], [159], whilst Paraavulavirus consists of two species; APMV 3 and 4. Secondly, Subfamily Metaparamyxovirinae consists of only one genus, Synodonvirus, first detected through metagenomics analysis of the triple cross lizardfish RNA [160]. The third subfamily Orthoparamyxovirinae is composed of 8 genera and 32 species. Several species within the eight genera in this subfamily are of medical and public health importance. This includes species in the genera Henipavirus, Morbillivirus, and Respirovirus, such as Nipah (NiV), Measles (MeV) and Human parainfluenza viruses 1 (PIV- 1) and 3 (PIV- 3), respectively. Finally, the fourth subfamily, Rubulavirinae, comprises two genera and 18

species. The viruses belonging to the two genera in this subfamily have been isolated from bats, pigs [161], [162], and humans [163], [164].

#### *1.2.3.2 Morphology and genome*

Morphologically, Paramyxoviruses are large enveloped RNA viruses with a genome size of 14.2 – 20.1kb. The virion size is also equally large, ranging from 150nm to 500nm in diameter. The virion is composed of a lipid envelope and, depending on the genus, may contain 2- 4 transmembrane glycoproteins. In all Paramyxoviridae, the fusion (F) and receptor-binding (RNP) transmembrane proteins are common. In some viruses, the genome can encode several proteins, such as the non-structural protein (C), a cysteine-rich protein that binds  $Zn^{2++}$  (V), and a small integral protein (SH). Towards the inner face of the envelope protein lies the matrix protein (M), surrounding the nucleocapsid protein (N). The N protein, together with the Phosphoprotein (P), the large protein (L) – which contains the Polymerase, capping and methylation activities- make up the ribonucleoprotein complex of the virion (Figure 3).

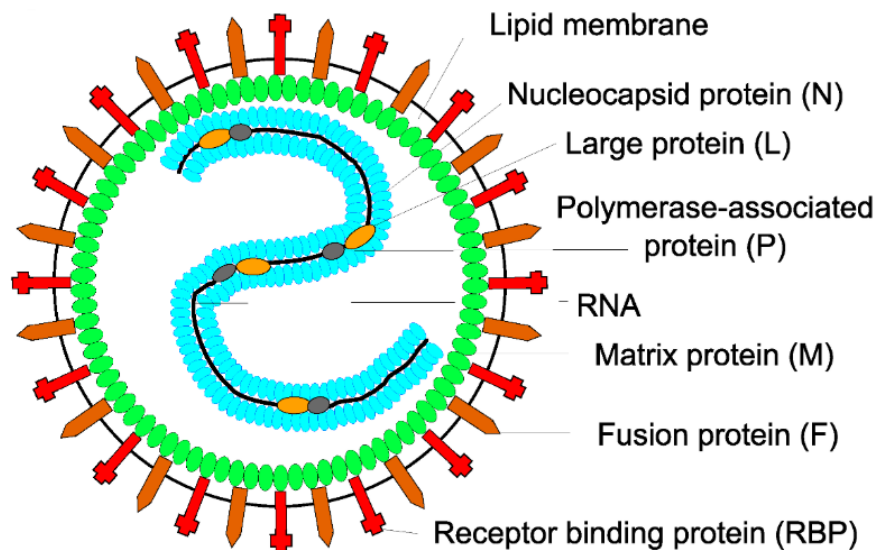


Figure 3: A Cartoon cross-section of the structure of Paramyxoviruses. *Adopted from Rima et al. 2019 [157].*



Figure 4. Genomic organisation of Paramyxoviruses. *The coding regions of the genome are flanked by non-coding regions (NCR) shown as uncoloured areas. Adopted from Rima et al. 2019 [157].*

### 1.2.3.3 Subfamily, Genus and Species demarcation

Like all negative-sense RNA viruses, Paramyxoviruses are thought to have evolved from double-stranded (ds) RNA viruses [165]. Inevitably, due to improving diagnostics and technology, more new paramyxoviruses have been discovered through viral metagenomics [166]. This has led to a shift in species demarcation from biological characteristics to a universal gene common to all RNA viruses, the Large (L) gene, encoding the RNA-dependent RNA polymerase (RdRp) [157], [165]. Thus the ICTV

recommends determining virus relatedness within species, genera and the family by measuring genetic distances on phylogenetic trees in branch length from the nodes.

#### *1.2.3.4 Transmission and Disease*

Transmission of Paramyxoviruses is through respiratory droplets, direct contact with an infected individual(s), and fomites [167]. Due to their mode of transmission, it is evident that these viruses are highly contagious, a fact greatly synergised by the interactive social nature between the virus hosts. Disease-associated with most paramyxoviruses is considered self-limiting, with most infections asymptomatic. However, where clinical disease develops, it is characterised by fever, rash and the classical measles virus-related syndrome, i.e. cough, coryza and conjunctivitis [168]. The ancient known paramyxoviruses associated with diseases have been the measles virus (MeV), Mumps, parainfluenza viruses, respiratory syncytial virus (RSV), and Newcastle virus disease (NVD). Other emerging paramyxovirus outbreaks such as Nipah (NiV) [169], [170] and Hendra viruses [171] have been reported in humans, pigs and horses. Although most paramyxoviruses are considered self-limiting, It is also important to note that they can still be associated with massive mortality and morbidity. For example, a NiV was implicated in over 100 human deaths and the near-collapse of the billion-dollar pig-farming industry in Malaysia [172]. Moreover, these viruses, perhaps similar to most emerging viruses, are characterised by a sequel of unorthodox clinical manifestations within their respective family/ genera or species, such as encephalitis [173], [174].

### *1.2.3.5 Treatment and Management*

Given that most viral infections are self-limiting, no treatment is given for mild infections. However, a clinical approach is required, involving treatment and management by supportive care for patients admitted with severe disease. This was the case for the treatment and management of NiV in Bangladesh, where patients were managed with anticonvulsants, mechanical ventilation, rehabilitation or treatment of secondary infections [175]. Experimental animal testing indicated that the human monoclonal antibody m102.4 has an exceptional potency against NiV [176] and HeV [177] and could act as a potential post-exposure future therapy for these viruses. In addition, broad-spectrum antivirals such as ribavirin, which can cross the blood-brain barrier [178], have also been administered to patients, showing a reduction in mortality due to RSV and PIV among immunosuppressed individuals receiving transplants [179].

Perhaps the greatest success in treating paramyxoviruses has been the near eradication of wild-type MeV using the safe and cost-effective measles vaccines since 1963 [180]. However, maintaining a MeV –free status comes with challenges globally, as some countries have been reported to lose this status [181], [182]. Some of these hurdles include; migration, vaccine hesitancy, and inadequate vaccine coverage [183].

### *1.2.3.6 Measles virus as an aetiological agent of disease in Uganda*

A new MeV genotype identified as D10 circulated during the 2000 – 2002 outbreak in Uganda as documented in 2005 [184]. Around this period, vaccination coverage was estimated at 74%. Another MeV genotype, B3, was detected in Uganda, Kenya, Burundi, and Tanzania in 2005 [185]. An inevitable race against time was begun, which led to

devising new ways of controlling the spread of MeV. With an increase in vaccination coverage of 83 -100% from 2003 – 2009, the spread of MeV was probably interrupted in Uganda [186]. However, the interruption of MeV spread in Uganda did not last long. In 2015 a new outbreak was reported in the western region [187]. Several clusters followed, such as Buvuma island [188] and Lyantonde district [189] in 2017. Unpublished data showed that a larger epidemic in 2019 led to the mass vaccination of approximately 18 million children aged between 6 months – 15 years. Uganda's MeV eradication plan involves routine vaccination of 9 months old infants with the measles-containing vaccine (MCV). This is usually followed by the supplementary immunisation activities (SIAs) carried out every three years across the country. However, this has not sufficed to break the MeV transmission chain due to the following reasons. Firstly, it is probable that there was laxity in the control measures after the 83-100% successful immunisation coverage campaign. Secondly, despite the unparalleled benefits of SIAs in the elimination of MeV, low-income countries still have incidences of prioritising government expenditures [190]. Thirdly, the global ease of movement may involve transmitting highly infectious genotypes to a naïve population by asymptomatic individuals. For example, MeV genotype D4, thought to have originated from a British tourist, was identified in a cluster of infants in central Uganda [51]. Perhaps the most detrimental challenges are the cessation of vaccination, continuous monitoring and surveillance of MeV. Therefore, these coupled with laboratory confirmation of probable clinical cases are key to eradicating the virus in Uganda. Without these control measures, MeV will inevitably remain a global challenge for years to come.

### 1.2.4 Togaviridae and its role in AFI

As of April 2019, the ICTV re-classified Togaviridae as a family of one genus, Alphavirus consisting of 32 viral species. These are spherical and enveloped virions with a single-stranded, positive-sense RNA genome of 10 – 12 kb. Before the advent of phylogenetic analysis, Alphaviruses were taxonomically structured by their antigenic relatedness, giving rise to antigenic sero-complexes [191]. Currently, phylogenetic analysis has widened our knowledge of Alphavirus classification, improving the previous grouping by identifying a recombination event that gave rise to the Western equine encephalitis virus (WEEV) ancestor. This occurred through recombining the ancestral Sindbis virus and Eastern equine encephalitis virus (EEEV) [192]. For medical purposes, these viruses can be grouped into encephalitic Alphaviruses (EA), that is, viruses that are associated with neurological disease [193] and Arthritogenic Alphaviruses (AA), which are characterised by human rheumatic disease [194].

The geographical distribution of EA is restricted to the continents of North and South America. These viruses include the WEEV [195], the EEEV, and the Venezuelan equine encephalitis virus (VEEV). The EA viruses are maintained in the population by passerine birds and *Culex tarsai* mosquito species; however, other mosquito species can act as vectors [193]. In addition, virus transmission has been documented in horses, humans, dogs and rodents. Where clinical disease manifests, it is characterised by febrile illness, respiratory symptoms, and severe disease that can lead to coma and or death. The mortality rate ranges from 1% for VEEV, 3% (WEEV) to 75% (EEEV) in humans, whilst equids show the highest mortality ranging from 3% (WEEV), 80% (VEEV), to 90% (EEEV)

[193]. Management is by vaccination with a formalin-inactivated virus, given in two doses to mount immunity.

Arthritogenic Alphaviruses associated with the arthritic disease include Chikungunya virus (CHIKV), O'nyong'nyong virus (ONNV), Ross River virus, Sindbis virus, Barmah Forest virus (BFV) and Mayaro virus. The most notable member of AA viruses is the CHIKV which caused an epidemic that started in Kenya in 2004 and later spread across the Indian Ocean, reaching India in 2005 [196] and lasting till 2011 [197], [194]. It caused an average of 4 million cases spanning over 40 countries [194]. Other outbreaks that AA viruses have caused include the ONNV outbreak in West Africa [198], Pogosta virus related to the Sindbis virus, which caused an outbreak in Finland [199], and BFV outbreak in Australia [200]. One remarkable feature of AA viruses is their ability to persist in host cells for a long time despite eliciting robust immune responses [201], [202], [203]. The disease symptoms are generally similar to AFI infections but more biased to fever, rash, myalgia and arthritis. Bleeding of gums has also been reported for CHIKV, ONNV and Mayaro virus [194].

Diagnosis of Alphaviruses can be achieved through virus isolation by cell culture, IgM antibody detection during the acute phase of infection, detection of viral antigens by immunofluorescence, PCR [204], or, more recently, by high throughput sequencing (HTS) [205].

### 1.2.5 The role of Orbiviruses in AFI

Orbiviruses are a genus within the family Reoviridae. Several members have been documented within this genus, but of economic importance, the Blue Tongue virus (BTV)

in ungulates, Epizootic haemorrhagic disease virus (EHDV) in deer, and the African horse sickness virus (AHSV) in equids are more prominent in the genus [206]. These viruses are transmitted by mosquitoes, hematophagous culicoides insects, black flies, sand flies and ticks [207]. Except for Antarctica, Orbiviruses show a global geographical distribution, for example, the Middle East as early as the year 1327 in Yemen [208], North America [209], South America [210], Europe [211], Africa [208], Asia [212] and Australia [210]. The viral genome consists of 10 dsRNA segments (VP1 – VP10) which encode for 7 structural proteins (VP1 –VP7) and 4 non-structural proteins (NS1 –NS3 and NS3a). The structural proteins are organised to form two structural shells, with VP2 and VP5 forming the virus's outer shell whilst VP1, VP3, VP4, VP6 and VP7 forming the inner core. This inner core surrounds the dsRNA genome [213]. There are currently 29 serotypes of BTV [214], 9 serotypes of AHSV [206] and 8 serotypes of EHDV [215].

Clinical disease manifestations of these viruses include BTV- induced vascular injury leading to haemorrhage; coronitis and laminitis; facial, intermuscular and pulmonary oedema; impaired respiratory and vascular functions; and pyrexia [206], [216]. MacLachlan *et al.* reported that more than 90% of the animals infected with AHSV succumb to the infection, whilst ungulates that recover from BTV show lifelong immunity to the BTV serotype [206]. A study carried out between Uganda and Nigeria reported serological evidence of the Orungo virus, an Orbivirus, in man [217]. More recently, another study reported the discovery and characterisation of the Bukakata virus from the Egyptian fruit bat in Uganda [218]. Two types of BTV vaccines are currently in use: the modified live virus vaccines containing the weakened virus and the attenuated vaccines

consisting of killed BTV [219]. Live attenuated vaccines are used for the AHSV as reviewed by Dennis *et al.* [220].

### 1.2.6 Polyomaviridae and its role in AFI

Polyomaviruses (PyV) belong to the family Polyomaviridae. They are small non-enveloped viruses with dsDNA genomes of approximately 5,000bp. Phylogenetically, polyomaviruses are classified into four genera; i.e. Alpha, Beta, Gamma and Delta polyomaviruses. Altogether, these genera comprise 102 species infecting mammals [221] and birds [222]. Other PyV species such as *Centropristis striata polyomavirus 1*, *Rhynchobatus djiddensis polyomavirus 1* and *Trematomus pennellii polyomavirus 1* have been detected in fish [223], [224].

The size of the virion varies in diameter from 40- 45 nm. The viral genome usually encodes three capsid proteins, namely VP1 to VP3 and two regulatory proteins, the large tumour antigen (LTA) and the small tumour antigen (STA). The three capsid proteins are expressed during DNA replication, whilst LTA and STA are expressed earlier during infection. Other proteins expressed by different species of polyomaviruses include VP4, alternative tumour antigen, and the middle tumour antigen.

Among mammals, the Alpha, Beta and Delta- polyomavirus genera contain viruses known to infect humans. According to previous serological studies, over 90% of the human population is presumed to have been exposed to PyV [225]. Several studies have also documented the occurrence of PyV in healthy and immunocompetent individuals, with similar seroprevalence in both the young and old populations [225], [226].

Besides their existence in healthy and immunocompetent populations, PyV has also caused clinical disease in immunocompromised individuals. For example, BK polyomavirus (BKPyV) was detected as an aetiological agent of nephropathy in renal transplantation [227], [228]. Multifocal leukoencephalopathy in an AIDS patient was believed to be caused by JCV after isolation from brain tissue [229]. *Trichodysplasia spinulosa* polyomavirus (TSPyV) was detected in an immunocompromised 4-year-old with signs of respiratory disease [230]. Therefore, it is rather this ambiguous relationship of the viruses that somewhat eludes their importance as potential pathogens in humans.

#### *1.2.6.1 Polyomaviruses in Uganda*

In Uganda, studies to ascertain the presence of PyV in both animals [231] and humans have been carried out. Stoner *et al.* 1998, associated PyV (JC virus in particular) as the aetiological agent of fatal progressive multifocal encephalopathy (PML) in a Ugandan patient [232]. This was in agreement with other previous studies carried out in the USA and Europe that associated PML to JCPyV in AIDS patients [233], [234], [235]. Another serological study was done to ascertain the relationship between several cancer forms and BKPyV but with inconclusive results due to the small sample size [236].

Despite evidence of the association of PyV with severe AIDS in HIV-positive patients, there is no documented prevalence of PyV in Uganda. The lack of such vital information limits our knowledge of the extent of morbidity and AIDS-related causes of death among HIV patients in the SSA region.

### 1.3 Other non-viral causes of AFI in Uganda

Among the zoonotic diseases of non-viral origin with high priority in Uganda are anthrax, brucellosis, African trypanosomiasis, and plague [49]. Not included on this list is *Plasmodia* spp causing malaria disease, and pathogenic *Leptospira* spp causing leptospirosis. In 2019, up to 229 million cases and over 400,000 deaths due to malaria were reported in 87 endemic countries [237]. Uganda contributed up to 5% of the total disease burden in 29 endemic countries, with over 5,000 reported deaths in 2019. Reduction in malaria cases and deaths in Uganda has been attributed to political engagement and support, multi-sectorial and community mobilisation. However, despite the reduction in malaria cases in Uganda, some cases still go undetected using the available rapid diagnostic tests [238], thus contributing to the surge of undiagnosed AFI in Uganda.

### 1.4 Virus Discovery

The capability to develop assays that can rapidly and reliably detect novel and already known viruses causing AFI has never been as important as today. The recent unprecedented SARS-COV 2 (2019) global pandemic [239], [52] and previous outbreaks such as the Zika virus (ZIKV) on the Yap Island in 2007 [240] and later in Brazil (2013 – 2015) [241], and Ebola virus in West Africa (2013 -2016) [242], emphasise the need for adequate surveillance and rapid diagnostics.

Whilst some virus discoveries have been out of curiosity [243] or serendipitous [244], [245], other discoveries are the result of a quest to provide answers in the face of public

health concerns [246], [247], [248]. In the agricultural sector, viruses of equally significant importance have been detected in domesticated animals with the potential to disrupt the food supply chain [249]. Wild animals have not been spared, as disease-causing viruses have been identified in turtles [250], bats [251] and other animals. The fluidity of the situation is that viruses initially thought to be zootic can mutate or undergo genetic reassortment and jump previously existing barriers to cause zoonotic diseases. Thus, the need to detect these viruses at the earliest opportunity cannot be overemphasized.

### 1.4.1 Diagnostic methods for AFI

The lack of diagnostic techniques to effectively detect the continuous threat of AFI has led to several challenges in developing countries. These include an increase in misdiagnosis [252], self-medication [253] and over-prescription of antimalarial drugs[254]. Below are some of the available diagnostic techniques and in context to their application to the SSA region.

To date, AFI diagnosis in patients can simply be based on the observational and clinical experience of the attending clinician [34]. For this reason, most AFI has been attributed to malaria in malaria-endemic regions. However, besides observational and clinical experience, reliable laboratory techniques such as virus isolation by culture, nucleic acid detection, antigen detection, microscopy and serological techniques have been developed to diagnose AFI globally [255].

The simple use of a light microscope to detect the presence of malaria is widely applied in SSA and malaria-endemic regions. However, besides South Africa, SSA countries have not adopted electron microscopy technology as a microscopy technique. This is

largely due to the associated purchase, operational and maintenance costs, researcher training and image artefacts resulting from specimen preparation. Another diagnostic technique is virus isolation which can be done by tissue culture [256], [257, 258], embryonated chicken eggs or inoculation in baby mice [259]. By using this technique, an infectious or a live virus attenuated by several passages can be recovered. Also, virus isolation by cell culture can be achieved at low operational costs. The downside of this approach is that the process is laborious and cannot be used for fastidious viruses or, let alone category 4 (four) pathogens.

Another tedious and equally labour-intensive method involves the use of serological assays such as the enzyme-linked immunosorbent assay (ELISA). ELISAs are used to detect the presence of particular biomarkers of disease such as antibodies (Immunoglobulin A, G, and M) which provide an excellent prediction of infectious diseases such as SARS-CoV-2 [260] and arboviruses [261] within a host. In addition to the tedium of the assay, ELISAs suffer from non-specificity as a result of cross-reacting antibodies within pathogen families [262]. The cross-reactivity challenge is mitigated by the ELISA test result confirmation using the plaque reduction neutralisation assay (PRNT). This, however, increases expenses and the length of results turnaround time and therefore, is not suitable for patient management, making it less relevant in clinical decision-making. Much-advanced methods for diagnosis such as microarray [263], PCR [264], [265], large-scale sequencing using cDNA-amplified restriction length polymorphism (cDNA-AFLP) [266], [267] and NGS [268], [269] have been developed. All these advanced techniques have pros and cons, as expounded by Storch *et al.* [255]. On the advantageous side, firstly, there is a reduced risk of technician self-inoculation as diagnosis involves

processed non-infectious specimens. Secondly, advanced PCR methods are increasingly affordable even in developing countries compared to earlier days with shorter turnaround time. Lastly, microarray technology and NGS platforms are high throughput methods, thus providing large volumes of data to aid pathogen detection and discovery opportunities. However, these advanced technologies also come with shortcomings. For example, PCR and microarray technology rely on already known genomic sequences [270], especially a conserved target region such as the polymerase gene [265], [271] for diagnosis. Microarray technology still has to overcome challenges of high background due to cross-hybridization and a limited dynamic range of detection [272], among other factors [273]. NGS overcomes these issues as no prior knowledge of the pathogen's genetic sequence is needed. Microarray and NGS technologies are still expensive and not easily sustainable, especially in SSA.

Given the above advances in diagnostics, the ability to detect AFI in developing countries may not be limited to inadequate diagnostic capacity alone. There is a need for a diagnostic method that can capture the emergence of novel pathogens and detect significant mutations in the genetic structure of already known pathogens [268] at affordable and realistic costs. Therefore, a combination of techniques may be necessary to overcome these challenges, such as the continuous use of readily available PCR supplemented by the unbiased capacity of NGS to periodically screen for pathogens. This will also provide the benchmark for pathogen discovery at an earlier stage and thus reduce the risk of economically costly and global health emergencies of infectious diseases.

#### *1.4.1.1 Using degenerate RT- PCR for virus detection and discovery.*

Kary Mullis is credited with having come up with the idea of synthesising millions of identical DNA copies from a single DNA molecule in 1983 [274]. This was the foundation of the polymerase chain reaction (PCR) technique. It was developed to generate the DNA sequence between two oligonucleotides (i.e. forward and reverse primers); each designed to complementary anneal to either end of the target region (5' and 3'). This PCR (also known as conventional PCR) is a qualitative method that requires post-PCR visualisation of DNA by an intercalating dye (e.g. ethidium bromide) under UV light. The PCR technique has been modified to improve performance, specificity, and use RNA as a template. Some of these modifications include i) A multiplex PCR used to simultaneously amplify three or more DNA sequences; ii) Nested PCR which uses the first PCR DNA product as a template for the second PCR; iii) Real-Time reverse transcriptase PCR which accomplishes the reverse transcription of RNA to cDNA together with relative quantification of nucleic acid copies produced. Several other PCR modifications are documented. PCR has extended its versatility by answering questions in criminal forensic investigations, food science, ecological field studies and diagnostic medicine [275]. The use of conventional PCR requires confirmation of the visualised DNA product by applying the Sanger sequencing method. Briefly, Sanger sequencing applies oligonucleotides (primers) annealing at known regions of the template. The reaction is extended by DNA polymerase in the presence of the four deoxynucleotide triphosphates (dNTPs: A, C, G and T). The extension is terminated when the DNA polymerase randomly adds a distinct dideoxynucleotide (ddNTP) labelled with fluorescently labelled dye. The

amplicons are analysed by electrophoresis, and the raw data is converted into sequences by Basecaller software [276].

The use of PCR to detect the presence of pathogen nucleic acids (Target –specific PCR) has become an efficient and reliable diagnostic method. Viruses such as the transfusion-transmitted virus (TTV) [277] and the TTV mini-like virus (TLMV) [278] were first detected by PCR. Across SSA, this method has been dependably used in research and surveillance activities of viruses such as influenza [279], HIV/AIDS [280] and the diagnosis of malaria [252], among other known infectious diseases. However, it has also brought to light the existence of many fevers of unknown origin (FUO) and, most importantly, undiagnosed AFI.

#### *1.4.1.2 Using NGS technology for virus detection and discovery*

Next-Generation Sequencing (NGS), also known as high-throughput sequencing (HTS) or massively parallel sequencing, is the most contemporary technology used for unbiased virus/ pathogen discovery. Billions of DNA fragments are simultaneously and independently sequenced [281]. An in-depth description of available NGS sequencing platforms and their underlying principles are described by Gu *et al.* [281]. Next, I will briefly discuss the most relevant NGS technologies and their application to pathogen detection and discovery.

- I. **Single-molecule real-time sequencing.** Pacific Biosciences (PacBio) developed the single-molecule real-time (SMRT) sequencing method to generate long sequence reads in a real-time manner. In principle, a circular dsDNA template, also known as DNA SMRTbell, is created by the ligation of hairpin adaptors to both ends of the template. This

template is then loaded onto a chip where it diffuses into a sequencing cell/hole called a Zero-mode Waveguide (ZMW), onto which a DNA polymerase is attached. The principle of operation of a ZMW is similar to that of a microwave. It consists of holes of ~30nm in diameter, which prevent laser light of ~600nm from passing through but emits fluorescent signals of 30nm. Thus, when the fluorophore-labelled nucleotides flood the ZMW, only the nucleotide detected by the DNA polymerase will be added to the growing sequence. Upon attachment onto the growing strand, the fluorophore is cleaved, leading to light emission whilst the rest of the nucleotides float away without emitting any fluorescence [282], [283]. This light is detected as a flash whose colour is identical to a specifically labelled nucleotide that is: A, C, T or G. In this way, sequencing continues without the requirement of halting the process or ZMW scanning and washing as experienced in Illumina NGS. For this reason, SMRT is also referred to as third-generation sequencing (TGS). The absence of 'scanning and washing' also shortens the sequencing time, producing long reads in hours as compared to the days required by Illumina. The real-time monitoring of nucleotide addition by DNA polymerase can also allow the analysis of the kinetics and thus determine methylation points in the synthesised sequence. In addition, the ability to generate long reads provides an easier route to *de novo* assembly and detection of repetitive regions. The challenges of SMRT are its high error rates (5 – 11%) which are reported as insertions and deletions (INDELs). Because SMRT technology is designed to generate long reads, it may not be the method of choice for degraded or long term archived specimens.

- II. **Ion Torrent.** Ion Torrent uses adapter-ligated DNA libraries that are clonally amplified on beads. The amplified DNA on beads is then channelled into the loading port of the chip,

together with the bound sequencing primers and DNA polymerase. Upon incorporating dNTP to the nascent DNA polymer, H<sup>+</sup> ion is released by DNA polymerase, which leads to a change in pH. The change in pH is detected by sensors incorporated at the bottom of the wells in the chip, converting it into a digitalised voltage read by external electronics. This provides a non-optical sequencing option and eliminates the need for cameras and fluorescently labelled nucleotides [284].

- III. **Oxford Nanopore.** Among all HTS technologies, Oxford Nanopore developed the first portable sequencing device, the MinION weighing 90g [285]. The device is easily inserted into a computer USB port for sequencing to commence. The mode of operation also involves channelling adapter-ligated cDNA or DNA into the nanopores of the MinION flowcell. Two types of adapters are used, i.e. the lead adapter from which sequencing initiates and the hairpin adapter for the complementary strand sequencing [286]. A motor protein unzips the dsDNA allowing the first template to channel into the nanopore, and once the hairpin adapter is reached, the hairpin protein allows the second strand to pass through the nanopore. Depending on the method used for base-calling, i.e. 1D (if bases are called for only the first strand) or 2D (if bases are called for both the first and second strand of the dsDNA), the base quality can be determined. Once sequencing is initiated in the nanopore, a change in current is observed and measured by a sensor that streams the data to a microchip.
- IV. **Illumina.** Illumina manufactures a variety of sequencing platforms such as the MiSeq, HiSeq series, and NextSeq 500 platforms. The data throughput and turnaround time vary among these Illumina platforms ranging from 4 hours using a Miseq to 6 days using the HiSeq 2500, generating 1 TB of data [287]. The principle behind Illumina sequencing

technologies involves the attachment of DNA (ligated with unique adapters per sample) fragments onto a glass slide (flow cell). The DNA fragments are then flooded with reagents containing fluorescently labelled nucleotides, and incorporated onto the growing DNA strand, followed by termination of the reaction, washing off excess reagents from the flow cell. The flow cell is scanned to identify the incorporated base using couple-charged device cameras, and the fluorescent dye is cleaved off the growing strand to allow the continuation of the polymerase reaction. Thus base incorporation, reaction termination, washing, and scanning process are repeated for several cycles, which can last days. Because Illumina technologies generate shorter fragment sizes (150 - 250bp), the error rate across all models is below 1% [288].

The advantages and limitations, and in most aspects, the workflow of wet and dry (bioinformatics) laboratory procedures is well described in Gu et al.'s review [281]. Briefly, one of the advantages of NGS is that there is always an unbiased diagnosis of illnesses, whether caused by known or unknown pathogens. The major drawback is the high amount of host nucleic acid sequenced together with pathogen nucleic acids. Available measures in the NGS pipeline of wet laboratory workflow and bioinformatics for host nucleic acid depletion overcome this challenge. The high risk of contaminants in the workflow also emphasises the need for stringent adherence to methods, reagents and a sterile working area. The cost of NGS is still high and may not be easily affordable by populations in SSA. However, the continuous improvement in the technology leads to a steady reduction in the cost per run. The importance of NGS cannot be overstated in virus discovery [289] and its role in reducing uncertainty in the diagnosis of AFI in high-risk areas [290].

#### 1.4.1.2.1 Next Generation Sequencing: Library preparation approaches

Several methods are currently available for the generation of NGS libraries. These approaches depend on several factors such as the sample characteristics, quality, research objectives and resource availability. Briefly, library preparation approaches are described below including their advantages and disadvantages.

Whole genome Sequencing (WGS) involves sequencing the entire genome population including protein-coding genes, non-coding regions, repetitive sequences, and intergenic regions, in a single experiment and thus provides a comprehensive description of the individual's genetic variation [291]. The applications of WGS can be observed in the identification of rare diseases and the development of personalised medicine in clinical care and research [292]. On the other hand, the vast amount of knowledge associated with this approach may affect a patient's life decisions or strategies for coping. Furthermore, information on genetic predispositions such as the increased risk for a sudden cardiac arrest, psychiatric disorders or behavioural traits may come as unpleasant surprises to patients [293].

Targeted Sequencing is another approach that involves sequencing specific regions of the genome. Unlike WGS, this selective method of genome recovery has lowered sample processing costs and efforts. Some of the methods used to achieve targeted sequencing include; hybrid capture of targeted nucleic acids using complementary DNA fragments; selective circularisation in which single-stranded DNA circles that include targeted regions are formed in a highly selective manner, creating DNA structures that are then used for selective amplification of the targeted regions; and lastly, PCR amplification where

polymerase reaction is carried out toward the targeted region [294]. On the flipside, this approach requires large amounts of input DNA and PCR conditions may influence the effectiveness of the procedure.

RNA sequencing is another indispensable approach for transcriptome-wide analysis of differential gene expression and splicing of mRNAs. Its applications have diversified including pathogen discovery in clinical and non-clinical specimens. The standard workflow in this approach involves the extraction of RNA, followed by enrichment of mRNAs, ribosomal RNA depletion, cDNA synthesis and preparation of an adaptor-ligated library ready for sequencing [295]. The benefits of this approach in but are not limited to pathogen/genome discovery, revealing the precise location of transcription boundaries, analysis of sequence variation and its application to low input abundance and degraded specimens [296]. Some of the shortcomings of RNA sequencing include; being very expensive, taking days to complete a library and challenges with RNA stability. While working with mRNAs, it is common to process long fragments of RNA in well-preserved samples with nearly intact RNA. Depending on the sequencing platform, such as Illumina which uses short reads, RNA fragmentation may be necessary. Different methods of RNA fragmentation such as chemical fragmentation (e.g., heat, alkaline hydrolysis), enzymatic fragmentation (e.g., RNase III, fragmentation enzymes), and mechanical fragmentation (e.g., sonication) can be applied during library preparation. However, it is important to know the advantages and disadvantages of RNA fragmentation before its applications. The pros include; the simple application of chemical fragmentation without the need for specialized equipment, the high specificity and the ease of control of enzymatic fragmentation. Lastly, mechanical fragmentation allows for fine-tuning of fragment size by

adjusting sonication parameters. The disadvantages of RNA fragmentation methods include the risk of the introduction of sequence-dependent bias and non-uniform fragment sizes through chemical fragmentation, limited reproducibility and probable interference by RNA secondary structure in enzymatic fragmentation, and the need for specialised equipment for mechanical fragmentation.

Table 1. Table of the Pros and Cons of the different methods of virus detection.

| <b>METHOD</b>                      | <b>PROs:</b>                         | <b>CONS:</b>                                                                                                           |
|------------------------------------|--------------------------------------|------------------------------------------------------------------------------------------------------------------------|
| CELL CULTURE                       | low cost and affordable              | Lengthy and tedious<br>It cannot be used to isolate fastidious viruses                                                 |
| Microscopy<br>a). Light Microscope | Low-cost and easily affordable       | Lower magnification, resolving, and resolution power (1500X, 1nm and 0.2µm respectively) compared to other microscopes |
| b). Electron Microscope            | Up to 1,000,000X magnification power | Higher operational costs<br>Sensitive to external magnetic fields and vibrations                                       |
| DNA Microarrays                    | High sensitivity                     | Low accuracy, precision, and specificity<br>Long results turnaround time                                               |
| PCR                                | low cost and affordable              | Biased to targeted amplification                                                                                       |
| NGS                                | Less than 1% error rate              |                                                                                                                        |

|                             |                                                                           |                                                                                   |
|-----------------------------|---------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| a). Illumina                | 120 GB – 1 TB of data in 30hr – 6 days<br><br>2- 8 hours turn-around time | Not suitable for use in <i>De Novo</i> assembly                                   |
| b) Ion Torrent              | Generates up to 1 GB of data in 4 hr<br>Suitable for De Novo assembly     | High error rate, prone to insertions and deletions                                |
| c) Pacific Bio              | Minimal library preparation time<br><br>Long read lengths (6 - 60kb)      | Higher error rate (11%)<br>Lower throughput data<br>Higher costs per base calling |
| d) Oxford Nanopore (MinION) | About 18 hrs turn-around time<br>Portable devices available               | High error rates and low throughput                                               |

## 1.5 Aim of the study

### 1.5.1 Research and Significance

The fact that AFI of different aetiologies presents a similar clinical picture complicates patient treatment and management [297]. This situation is worsened by the inadequate facilitation of clinical and diagnostic laboratories in developing countries, resulting in high complications and case fatalities. Although efforts are in place to monitor respiratory tract infections by the National Influenza Centre at the Uganda Virus Research Institute [298],[299], other infections such as bloodstream infections may be left undiagnosed. Serological surveillance of arboviruses [300] is not sufficient to understand the burden of bloodstream infections within the population. As much as NGS technology is increasingly becoming more available and cheaper, it is still less affordable by the authorities in the SSA region; more so, its use in routine diagnosis of AFI. This leaves a gap for a more affordable yet readily available diagnostic method that can efficiently survey and monitor trends in AFI on a routine basis in patients visiting hospitals. Using degenerate PCR as a bridge between NGS and routine diagnostics might provide a timely and dependable solution that will help in the control, management, and prevention of AFI in SSA.

The utmost importance of understanding the aetiology of AFI is to allow for the right prescription of antimicrobials to the right patients, early detection of emerging and re-emerging diseases and identifying vaccination gaps [290]. Furthermore, by carrying out this research, information much-needed on other causes of fever in patients will be provided. Pathogens that have always escaped routine diagnosis by target-specific PCR will be identified using metagenomics NGS.

## 1.5.2 Objectives of the study

This research aims to identify other pathogens that account for disease in clinically morbid individuals with undiagnosed fever presenting at the health centres. Another key question is whether degenerate PCR is feasible with routine diagnostics of AFI in developing countries.

The main objective of the research is to determine causative agents of undiagnosed acute febrile illness in patients presenting at health centres in Uganda. This will be done by utilising a combination of conventional PCRs and next-generation sequencing (NGS) to screen for pathogens in sera and whole blood samples. The conventional PCRs will target conserved regions within the genomes of pan-family or genera of the most likely viruses. The likely viral families will be determined by literature search of previously detected viruses in the SSA while NGS will apply the metagenomics approach.

The secondary objective is to describe the epidemiological characteristics of arbovirus-positive cases in Uganda. This was done by analysing IgM positive cases available from data collected from 2016 – 2019 by the Arbovirology Department, UVRI. Furthermore, this study will describe the genetic characteristics of rarely isolated viruses such as the West Nile virus and compare them to previously isolated strains.

## **2.0 METHODS AND MATERIALS**

### **2.1 Investigating the seroprevalence and spatial distribution of Arboviruses in Uganda**

#### **2.1.1 Study population, sentinel sites and period**

Demographic and laboratory data were selected from a cohort of patients enrolled at the hospital-based sentinel sites. The sentinel sites were established for arboviruses surveillance across Uganda [301]. The sentinel sites cover the Central, Eastern, Northern and Western regions and are distributed in over 65 districts. In total, 1900 samples with demographic and laboratory results, and collected between January 2016 to September 2019, were selected for analysis. In addition, data on the national population census and yearly projections per district were obtained as available on the Uganda Bureau of Statistics (UBoS) website [302].

#### **2.1.2 Selection and exclusion criteria**

The data (provided with written consent or assent from a caretaker for minors) were from patients aged two months and older who had a recorded temperature of  $\geq 37.5^{\circ}\text{C}$  or a history of fever not spanning more than seven days at the time of presentation. A temperature of  $37.5^{\circ}\text{C}$  was acceptable considering some patients tend to self-medicate with antipyretics before seeking medical attention. Samples with an initial volume of less than  $250\mu\text{L}$ , visibly haemolysed blood specimens and samples referred to the Uganda

Virus Research Institute (UVRI) from other countries (South Sudan, Democratic Republic of Congo, and Somalia) were excluded. In total, 1,441 samples with complete epidemiological data were analysed.

### 2.1.3 Detection of Arboviruses in samples collected in Uganda

The data used in this study were acquired from diagnostic results of serological tests carried out at the UVRI. We applied these results to determine the seroprevalence of arboviruses in Uganda. Immunoglobulin M (IgM) antibodies for YFV and ZIKV were detected using the CDC IgM antibody capture- enzyme-linked immunosorbent assay (CDC-MAC-ELISA) [303]. Diagnosis of WNV, DENV, and CHIKV virus was carried out using the InBios West Nile virus Detect IgM Capture ELISA, DENV Detect IgM Capture ELISA, and CHIKV Detect IgM Capture, respectively (InBios international, inc., Seattle, WA, USA). In addition, virus-specific plaque reduction neutralisation tests (PRNT) were carried out for arbovirus confirmation, where possible. However, because the results data on the PRNT were incomplete, all analyses and conclusions were made based on the IgM test results.

### 2.1.4 Data analysis for Arboviruses

Interpretation of results was performed as described by Kayiwa *et al.* [300]. Briefly, a sample was considered “positive” if the positive/ negative (P/N) ratio at mean OD<sub>450</sub> of the patient serum on viral antigen divided by the mean OD<sub>450</sub> of the negative control on the viral antigen was greater than two on serological testing. These were then reported and

recorded in the laboratory Microsoft excel sheet as "IgM positives", "negative", or "inconclusive". SPSS version 26 software (linear regression, odds ratios, and p values) was used to analyse the IgM positive results on variables such as age group, region, occupation, and gender. The epidemiologic and demographic data at this stage were filtered to remove samples lacking information on the region of sample collection, age, occupation, or gender of the patient. Univariate linear regression was carried out to determine arboviruses' impact on each parameter (age group, occupation, region of sample collection, and gender of the patient). The association between variables was analysed using the Chi-square test (Odds ratio and p-value). SPSS v26 was also used to determine the yearly arbovirus transmission pattern using the bar graph function.

#### *2.1.4.1 Spatial and spatiotemporal analysis*

We sought to identify geographical areas in Uganda with an increased risk of arboviral transmission/ activity. Thus, using the SaTScan software (version 10.0.1, Boston, MA) [304], we applied the space-time and spatial-only models to statistically detect clusters of YFV and alphaviruses (CHIKV and ONNV) with greater than expected numbers in a geographical area [305], [306]. The expected cases were determined based on the assumption of a Poisson distribution of cases in the population. Since a single laboratory or clinically confirmed case of YFV is reportable for investigation, we set the SaTScan parameters to include a maximum radial distance of 50km, a temporal window of 3 months, and a cluster number of 3 cases. The parameters for alphaviruses were similar except for the cluster adjusted to at least five (5) cases since no public health action is required upon laboratory confirmation. For all arboviral cases, geo-coordinates (latitudes

and longitudes) were acquired based on google maps and uploaded into the SaTScan software for the space-time and spatial-only models. Data output were visualised using HTML files for google maps. Incidence rates were applied as a measure of the likelihood of an outbreak or cluster of an arbovirus infection ( An observed incidence value greater than the expected incidence value suggested that there is a higher than expected number of cases and thus a cluster or outbreak while if the observed incidence value lower than the expected incidence value suggested a lower than expected number of cases).

#### *2.1.5 Ethical approval*

Ethical approval for this study was sought from and waived by the Uganda Virus Research Institute's review board.

## 2.2 Screening for viral AFI in Arbovirus negative samples using RT-PCR

### 2.2.1 Designing the positive control concatemer

To ensure that PCR assays were valid, both negative and positive controls were included in PCR reactions. In circumstances where there was no initial positive control material, a positive control concatemer was designed. *In silico*, construction of the concatemer was done targeting 15 conserved genes in viruses known to cause AFI [307]. The positive control concatemer consisted of RdRp, UTRs and non-structural genomic sequences from 15 viral spp, namely; the Crimean Congo haemorrhagic fever virus (CCHFV), Lassa fever virus (LASV), YFV, the enterovirus 5' untranslated region (UTR), Henipah/ Morbillo/ respirovirus, Avula/ Rubulavirus, metapneumovirus, Bunyamwera virus, CHIKV, Bwamba virus, RVFV, Hepatitis E virus, Filovirus (for Ebola and Marburg viruses) and Lyssavirus. In addition, to rule out malaria infection, brucellosis, and leptospirosis, conserved genomic sequences of plasmodia spp., Brucella spp., and Leptospira spp were added to the concatemer (Table 1). Where a positive sample was identified and confirmed, RNA was transcribed to cDNA and used in subsequent PCR assays.

### 2.2.2 Primer selection for virus detection

Previously published primers were used to detect pan-family viruses in Arenaviridae, Paramyxoviridae, Picornaviridae (genus Enterovirus), Flaviviridae, Togaviridae, and genus Orthohepevirus (Table 2).

Table 2. Primers used to screen for viruses according to their Viral families or genera.

| Family/ Genus ( Assay reference)                    | Primer sequence                                                                       | PCR extension period (minutes) |
|-----------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------|
| Arenaviridae<br>(Mammarenavirus old world)<br>[271] | <b>Forward sequence (5'→ 3')</b><br><br>AGAATYAGTGAAAGGGARAG<br><br>YAAYTC            | 10                             |
|                                                     | <b>Reverse sequence (5'→ 3')</b><br><br>CACATCATTGGTCCCCATTTA<br><br>CTRTGRTC         |                                |
| (Paramyxoviridae)                                   |                                                                                       |                                |
| Avulavirus/ Rubulavirus [308]                       | <b>Forward sequence (5'→ 3')</b><br><br>GGT TAT CCT CAT TTI TTY GAR<br><br>TGG ATH CA | 1                              |
|                                                     | <b>Reverse sequence (5'→ 3')</b><br><br>GCA ATT GCT TGA TTI TCI CCY<br><br>TGN AC     |                                |
| Henipavirus/<br>Morbilivirus/ Respirovirus<br>[308] | <b>Forward sequence (5'→ 3')</b><br><br>GTT GCT TCA ATG GTT CAR<br><br>GGN GAY AA     | 7                              |
|                                                     | <b>Reverse sequence (5'→ 3')</b>                                                      |                                |

|                                                          |                                                                            |    |
|----------------------------------------------------------|----------------------------------------------------------------------------|----|
|                                                          | GCT GAA GTT ACI GGI TCI<br>CCD ATR TTN C                                   |    |
| <i>Pneumovirinae subfamily</i><br>/Metapneumovirus [308] | <b>Forward sequence (5'→ 3')</b><br>ACT GAT CTI AGY AAR TTY<br>AAY CAR GC  | 1  |
|                                                          | <b>Reverse sequence (5'→ 3')</b><br>GTC CCA CAA ITT TTG RCA<br>CCA NCC YTC |    |
| Picornaviridae<br>(Genus Enterovirus) [309]              | <b>Forward sequence (5'→ 3')</b><br>ACATGGTGTGAAGAGTCTATT<br>GAGCT         | 10 |
|                                                          | <b>Reverse sequence (5'→ 3')</b><br>CCAAAGTAGTCGGTTCCGC                    |    |
| Flaviviridae (Flavivirus) [310]                          | <b>Forward sequence (5'→ 3')</b><br>TGYRTBTAYAACATGATGGG                   | 10 |
|                                                          | <b>Reverse sequence (5'→ 3')</b><br>GTGTCCCAICCNCGNTRTC                    |    |
| Hepeviridae (Orthohepevirus)<br>[311]                    | <b>Forward sequence (5'→ 3')</b><br>ACYTTYTGTGCYYTITTTGGTC<br>CITGGTT      | 10 |
|                                                          | <b>Reverse sequence (5'→ 3')</b>                                           |    |

|                                                 |                                                                                                                                              |    |
|-------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|----|
|                                                 | CCGGGTTCRCCIGAGTGTTTC<br>TTCCA                                                                                                               |    |
| Togaviridae [312]                               | <b>Forward sequence (5'→ 3')</b><br><br>GGGGCnATGATGAArTCnGG                                                                                 | 10 |
|                                                 | <b>Reverse sequence (5'→ 3')</b><br><br>ATCTTCACTTCCATRRTTNANC<br>CA                                                                         |    |
| Bunyamwera virus, primers<br>adapted from [313] | <b>Forward sequence (5'→ 3')</b><br><br>TTACGCTTAACCTTGGGGGC<br>TGGA<br><br><b>Reverse sequence (5'→ 3')</b><br><br>CTAACATTGGTTTTCTTCC<br>A | 5  |
| Bwamba virus<br><br>Inhouse primers             | <b>Forward sequence (5'→ 3')</b><br><br>TTCTTCCTCAATGCCGCAAAG<br>GCC<br><br><b>Reverse sequence (5'→ 3')</b><br><br>CCATTTGATTGCTATTG<br>G   | 5  |
| Phleboviruses (RVFV)                            | <b>Forward sequence (5'→ 3')</b>                                                                                                             | 5  |

|                                                   |                                                                                                                                 |    |
|---------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|----|
| Adapted from Lambert <i>et al.</i><br>2009. [314] | GCYTATCAAGGATTTGATSC<br><b>Reverse sequence (5'→ 3')</b><br>CTK RCC CTY KTY TGA CAT<br>YTT                                      |    |
| Filoviridae (EBOV) [315]                          | <b>Forward sequence (5'→ 3')</b><br>GCMTTYCCNAGYAAYATGATGG<br>T<br><b>Reverse sequence (5'→ 3')</b><br>GCACANGANATRCWNGT<br>CCA | 10 |
| Rhabdoviridae (Lyssa virus)<br>[316]              | <b>Forward sequence (5'→ 3')</b><br>GGKMGRTTYTTYKCHYTDATG<br><b>Reverse sequence (5'→ 3')</b><br>CARAARGGNTGGASYNTH<br>BT       | 5  |
| Nairoviruses (CCHFV)<br>In house primers          | <b>Forward sequence (5'→ 3')</b><br>CWGARTAYAAAGTSCCTGG<br><b>Reverse sequence (5'→ 3')</b><br>GTGTCDATYTGWGCHCCYTG             | 5  |

### *2.2.2.1 Primers for Enterovirus serotyping*

Initial screening for enteroviruses was performed using primers targeting the conserved 5' UTR of enteroviruses (Dierssen *et al.*) [309], producing a 142bp fragment. Then, primers designed by Nix *et al.* [317] were applied in a semi-nested PCR to serotype the detected enteroviruses further. The first primer set, p224 (GCIATGYTIGGIACICAYRT) /p222 (CICCIGGIGGIAYRWACAT), targets a conserved region between the VP3 (region 1977 - 1996) and VP1 (2969 – 2951) genes of enteroviruses whilst the second primer set A89 (CCAGCACTGACAGCAGYNGARAYNGG) position 2602 – 2627 and A88 (TACTGGACCACCTGGNGGNAYRWACAT) position 2977 - 2951 targets the VP1 gene for a final PCR amplicon of approximately 375bp.

### *2.2.2.2 Primer design for Measles virus strain characterisation*

Measles virus (MeV) primers were designed based on the strain with the highest percentage of nucleotide identity score to the sequence obtained via Sanger sequencing the Paramyxoviridae PCR product. To phylogenetically characterise the detected MeV, first, previously available primers targeting the Neuraminidase (N) N450 region, designed for a nested PCR described by Bankamp *et al.* [318] were applied. And because the N gene cannot be used to infer the introduction of a new MeV strain in a given area, the coding sequence (CDS) of MeV genotype B3 accession number KT732214.1 was downloaded from GenBank and applied to design overlapping primers targeting the Hemagglutinin (H) gene using the Primal Scheme tool [319]. The primer sequences, target region and expected fragment size as shown in Table 3.

Table 3. In-house primers designed to detect the MeV Hemagglutinin gene and the N450 region primers for the detection of the Nucleoprotein gene adopted from Bankamp *et al.* \* The primer name shows the virus abbreviation (MV) while the number denotes the starting (for the forward primer) and the ending (reverse primer) position in accordance to the reference strain KT732214.

| HEMAGGLUTININ | Primer/position* | Sequence (5' to 3')              | Size (bp) |
|---------------|------------------|----------------------------------|-----------|
|               | MV_6871F         | AGG ATG CCA AGG AAT TGT<br>TGG A | 616       |
|               | MV_7487R         | TTA TGG ATC TCC GCG GTG<br>TAG A |           |
|               | MV_7379F         | ATG TTC TGC TGG CTG TTC<br>TGT T | 598       |
|               | MV_7977R         | TTC AAC TAG GTA GGT TCC<br>CCC A |           |
|               | MV_7869F         | AGA CAT GTC GCT GTC CTT<br>GTT G | 615       |
|               | MV_8484R         | ATC CTT CAA TGG TAC CCA<br>CTC G |           |
|               | MV_8387F         | CAG ATG ACA AGT TGC GRA<br>TGG A | 607       |
|               | MV_8994R         | GTC CCA TGT GAA GCA TTC<br>CAC T |           |

|               |             |                                    |                    |
|---------------|-------------|------------------------------------|--------------------|
|               | MV_8904F    | CCC AAG CCG CTC ATT TTC<br>TTA C   | 589                |
|               | MV_9493R    | TGG AAT ATG AGA GTG GGT<br>CGG A   |                    |
| NUCLEOPROTEIN |             |                                    |                    |
|               | MeV216_(1)F | TGG AGC TAT GCC ATG GGA<br>GT      | 633, ref.<br>[318] |
|               | MeV214_(1)R | TAA CAA TGA TGG AGG GTA<br>GG      |                    |
|               | MeV210_(2)F | GCT ATG CCA TGG GAG TRG<br>GAG TGG | 626                |
|               | MeV217_(2)R | CAA TGA TGG AGG GTA GG             |                    |

### 2.2.3 Sample selection criteria

Serum and whole blood samples with negative IgM antibody ELISA test results and complete epidemiologic and demographic data were selected. In addition, these samples fulfilled the AFI case definition, with no visible haemolysis and a sample volume of more than 250µL.

## 2.2.4 Sample Pooling

The samples were prepared by mixing five (5) samples to form a single pool according to their type; whole blood or serum. From each sample, 100µL was collected to form a final pool of 500 µL. Two hundred and eighteen pools were made from 1090 serum and whole blood samples and subsequently taken for further analysis.

## 2.2.5 Sample shipment

Sample pools that tested positive for any viral targets were selected and then added in a ratio of 1:1 to a DNA/RNA shield (Zymo Research, Irvine, USA); a preservative used for the inactivation, stabilisation, storage, and transport of biological materials [320].

These samples were then shipped at ambient temperature from the UVRI, Uganda, to the University of Nottingham (UoN), School of Life Sciences, Virology Research Group (VRG), where the laboratory assays included in 2.2.6 to 2.4 were carried out.

## 2.2.6 Nucleic acid extraction, cDNA synthesis and quality control

Total nucleic acids were extracted from serum samples using the QIAamp viral RNA kit (Qiagen, UK) according to the manufacturer's protocol. However, Carrier RNA was not added during extraction as this would interfere with the downstream NGS/ HTS process.

Total nucleic acids were extracted using the Monarch<sup>®</sup> Total RNA Miniprep Kit according to the manufacturer's protocol (New England Biolabs Inc.) for whole blood samples. All

extraction procedures were carried out according to the manufacturer's instructions. After nucleic acid extraction, 20µL of extracted RNA was used for cDNA synthesis using the RNA to cDNA EcoDry™ Premix kit (Takara, Japan) according to the manufacturer's instructions and stored at -20°C until further use. The sample quality was assessed at the University of Nottingham by measuring the RNA amount using the Nanodrop One spectrophotometer (Thermo Scientific).

## 2.2.7 Polymerase Chain Reaction

### *2.2.7.1 Components and conditions of PCR*

Initial screening for viral pathogens in the synthesised cDNA was done using previously published pan-family or genus primer sets. The degenerate PCR primers used for this study were as described in section 2.2.2. For each PCR reaction, a master mix was prepared using 1.25µL of a 10X reaction buffer (Manufactured by Qiagen), 0.5µL of 2.5mM dNTPs (Sigma, Aldrich), 0.48µL of forward and reverse primer (10 pM/µL each), 0.06µL of HotStarTaq DNA Polymerase enzyme (Qiagen, Hilden, Germany) and 1µL of cDNA and nuclease-free sterile deionised water to yield a final reaction volume of 12.5µL.

PCR assays were done using optimised cycles for each viral family. Briefly, the conditions of each PCR included an initial heat activation of HotStarTaq DNA Polymerase enzyme at 95°C for 15 minutes, then followed by 55 cycles of denaturation at 94°C for 15-30 seconds, annealing temperature between 54 -60°C for 20-30 seconds and an elongation temperature of 72°C for 20 seconds - 1 minute (individual elongation periods shown in Table 3).

### *2.2.7.2 PCR for detection of Plasmodia, Brucella and Leptospira spp*

Due to the endemicity of *Plasmodia*, *Brucella* and *Leptospira* spp. in the region, PCRs were carried out to exclude non- viral AFI from the cohort. The pan –*Plasmodia* spp. and the *Leptospira* spp. primers targeting the 18s and the 16s rRNA genes were adopted from Waggoner *et al.* [321]. Primers for the detection of *Brucella* spp. targeted the *bcs31* gene as described by Probert *et al.* [322]. The PCR conditions for *Plasmodia*, *Leptospira*, and *Brucella* spp. were carried out as described by Waggoner *et al.* and Probert *et al.*

### *2.2.7.3 PCR for Enterovirus genotype identification*

This PCR was carried out according to the assay described by Nix *et al.* [317], targeting the VP1 region. During PCR, the initial heat activation was done at 95°C for 15 minutes, followed by 55 cycles of amplification at 95°C for 30 seconds, 42°C for 30 seconds, and 60°C for 45 seconds. Then, a final extension at 72°C for 10 minutes was done. From this PCR product, 1 µL was used to carry out the second PCR using the second primer set A89/A88 (described in 2.2.2.1). Similar PCR amplification was carried out at 95°C for 30 seconds, 60°C for 20 seconds, and 72°C for 15 seconds.

### *2.2.7.4 PCR for the detection of Measles virus Hemagglutinin*

All cDNA synthesis was carried out as previously described in 2.2.5. The H gene was amplified by first carrying out an initial heat activation at 95°C for 15 minutes. The

subsequent PCR conditions were 55 cycles of amplification at 95°C for 20 seconds, 58°C for 20 seconds, and 72°C for 45 seconds with a final extension at 72°C for 1 minute.

#### *2.2.7.5 PCR for detection of Measles virus Nucleoprotein*

The PCR conditions were adapted to include an initial heat activation at 95°C for 15 minutes to amplify the N gene. After heat activation, the initial PCR included 55 cycles of amplification at 94°C for 15s, 55°C for 30 seconds, 72°C for 30 seconds, and a final extension at 72°C for 7 minutes. The second PCR required 2 µL of the first PCR product and 10pmol each of primers MeV210\_(2)F and MeV217\_(2)R (adapted from Bankamp *et al.*). The PCR conditions comprised an amplification at 94°C for 15 seconds, 50°C for 30 seconds, 68°C for 1 minute and a final extension at 72°C for 5 minutes.

### 2.2.8 Gel electrophoresis

The PCR product was mixed with 2.5 µL of loading dye and 6µL of the mix were individually loaded together with the appropriate size markers on a 2% gel (agarose and TAE buffer) for electrophoresis. The gel was run for 36 minutes at 90 volts, and the expected bands with a size corresponding to the positive control as viewed under a UV light illuminator were sent for sequencing to verify the putative virus detected.

## 2.2.9 Sanger sequencing and analysis

Putative positive PCR products and their appropriate forward primers were diluted in distilled water to final concentrations of 10ng/μL and 3.2pmol/ μL, respectively and sent for Sanger sequencing to SourceBioscience (Nottingham, UK). The nucleotide sequence chromatogram from SourceBioscience was analysed in FinchTV (Geospiza, Inc), blasted and compared to sequence data on GenBank using the basic local alignment search tool (BLAST) and BLASTn function. In addition, translated nucleotides were analysed using the BLASTx function.

## 2.3 Molecular detection of West Nile virus

### 2.3.1 Primer design for West Nile virus

The genome of the most identical WNV strain (KY523178.1) was downloaded from GenBank. This genome was used to design overlapping primers using Primal Scheme as described by Quick *et al.* [319]. The expected average size of the PCR product was 800bp. Where no PCR product was observed under UV- light, primers producing shorter amplicons of approximately 400bp or even 200bp were designed. The list of primers and the product size of amplicons are shown in Table 3.

### 2.3.2 PCR, Sanger sequencing and analysis of West Nile virus

The PCR components are described in section 2.2.6.1. The PCR conditions for the 800bp amplicon were optimised at 95°C for 15 minutes for the initial heat activation of HotStar Taq DNA polymerase. This was followed by 55 cycles of amplification with denaturation at 94°C for 20 seconds, annealing at 58°C for 20 seconds, and elongation at 72°C for 2 minutes. Similar conditions were set for the 400bp and 200bp amplicons except for 20 seconds of elongation time. Gel electrophoresis, Sanger sequencing, and sequence analysis were done as described in 2.2.7 and 2.2.8, respectively.

Table 4 shows Primer sets generated using Primal Scheme software showing their fragment size and genomic position.

| <b>Primer</b> | <b>Sequence 5' – 3'</b>       | <b>Genomic Position</b> | <b>Size (bp)</b> |
|---------------|-------------------------------|-------------------------|------------------|
| WNV_8 F       | GCCTGTGTGAGCTGACAACTT         | 8 - 813                 | 803              |
| WNV_813 R     | TTTTGTGCTGTCCAACCAAGCT        |                         |                  |
| WNV_1252 F    | TTGTGGACAGAGGATGGGGAAA        | 1252 - 2033             | 780              |
| WNV_2033 R    | ACCAGTCTTCCAACAGGTGTGA        |                         |                  |
| WNV_1888 F    | CGTTCAAATTCGCTGGGACTCC        | 1888 - 2672             | 783              |
| WNV_2672 R    | CACATTTGGTGCTCGAGTCTGG        |                         |                  |
| WNV_2509 F    | GGTGCGGAAGTGGAGTGTTTAT        | 2509 - 3275             | 766              |
| WNV_3275 R    | TCCCATGGACCTTGGTTCTGAG        |                         |                  |
| WNV_3130 F    | AATCATGCACCTGGCCAGAAAC        | 3130 - 3926             | 795              |
| WNV_3926 R    | TTCAAAACGTCAGGCACTTCCC        |                         |                  |
| WNV_3763 F    | TGGCTACATTCAAGATTCAACCAGT     | 3763 - 4515             | 604              |
| WNV_4515 R    | CATCCAAATTTTCCATGGCGCC        |                         |                  |
| WNV_4368 F    | GACATGTGGATTGAGAGGACGG        | 4368 - 5136             | 768              |
| WNV_5136 R    | TGCCGGTTCTTCCATTCTCTCT        |                         |                  |
| WNV_4953F     | GATTGGACTGTATGGRAACGGC        | 4953 - 5359             | 406              |
| WNV_5359R     | ACA RGT TGT AGT TGG GGA CTC T |                         |                  |
| WNV_5591F     | CCGTTTGGTTTGTTCOAAGTGT        | 5591 - 5987             | 396              |
| WNV_5987R     | TCTGTCCAGTGAGCAAAGTTGG        |                         |                  |

|            |                                |              |     |
|------------|--------------------------------|--------------|-----|
| WNV_6215F  | GACCTCGAACCAACACGATTCT         | 6215 - 6629  | 414 |
| WNV_6629R  | ATT ACT CCT CCC AAG CCA ATC TT |              |     |
| WNV_6521F  | CCCTTCAGACAATAGCTTTGATTGC      | 6521 - 6936  | 415 |
| WNV_6936R  | CGGCCGCAGATCAAGTAAGAAG         |              |     |
| WNV_7408F  | AGAGAGGCTGGAATTCTGACCA         | 7408 - 7826  | 418 |
| WNV_7826R  | AGATCCACAACCTTTCCCACC          |              |     |
| WNV_7721F  | GAGAGGGAAACATCACTGGAGG         | 7721 - 8131  | 410 |
| WNV_8131R  | TGATGCAGAATTCCTTCGGTCC         |              |     |
| WNV_9006F  | CAGAGCCATCTGGTTCATGTGG         | 9006 - 9391  | 385 |
| WNV_9391R  | TCTGGTCCTCTCTGGAGATGAC         |              |     |
| WNV_9297F  | CATCGAACTCACATACCGACACA        | 9297 - 9696  | 343 |
| WNV_9696R  | CGTCGAGGGTTTCCATTCCT           |              |     |
| WNV_9620F  | GCTTCGCCACATCACTACACT          | 9620 - 10008 | 388 |
| WNV_10008R | TTCTCCTTTTGCGTGGATCGAC         |              |     |
| WNV_9886F  | TGGCTGCTGCTTTATTTCCACC         | 9886 - 10295 | 409 |
| WNV_10295R | GTGTCCTCCACTACAATGGTGT         |              |     |
| WNV_5580F  | GCAGCCAGAGGATACATAGCAA         | 5763 - 5580  | 183 |
| WNV_5763R  | TCTCATTTCCCATTTTCACACTTGG      |              |     |

|           |                            |             |     |
|-----------|----------------------------|-------------|-----|
| WNV_5731F | TGGAACACTGGATATGARTGGATAAC | 5914 - 5731 |     |
| WNV_5914R | GCTYGCTTTRAAGTTRGCTCCCA    |             |     |
| WNV_9036F | AACAGAACCAATGGAAGAACGC     | 9245 - 9036 | 209 |
| WNV_9245R | CCAGCCAGTGGTCTTCATTGAG     |             |     |

## 2.3.3 West Nile virus genome analysis

### 2.3.3.1 *Phylogenetic analysis of WNV*

The coding sequences (CDS) of WNV were downloaded from GenBank and aligned together with the detected WNV in this study in MEGA7 software [323] using the MUSCLE function (Because of its superiority to ClustalW in terms of re-optimisation and accuracy). Phylogenetic analysis was carried out by constructing a maximum-likelihood phylogenetic tree based on the best-fit substitution model.

### 2.3.3.2 *Analysis of the genetic diversity of WNV*

The genetic diversity of the recovered WNV was analysed using the Simple Sequence Editor (SSE) version 1.4 [324] by carrying out a sequence distance analysis. Nucleotide sequences of strains clustering with the recovered WNV were compared to strains clustering with the WNV first detected in 1937 in Uganda. The output data were in the form of identity plots, showing genomic loci of increased genetic divergence or evolutionary distance between the groups of phylogenetically clustered WNV strains.

### *2.3.3.3 Analysis of the WNV Pre-membrane and Envelope nucleotide coding sequences*

The pre-membrane (Pre-M) and Envelope (E) sequences of the recovered WNV and the 1937 WNV (AY532665.1) were analysed for N-glycosylation using a web-based tool. This tool developed by Gupta and Brunak [325] predicts the N-glycosylation in proteins by examining the Asn-Xaa- Ser/Thr sequon. Output data was presented in the form of a graph, and  $\geq 0.5$  threshold level was associated with areas of potential glycosylation.

## **2.4 Screening samples for viruses using next-generation sequencing**

### **2.4.1 Sample selection and RNA quantification**

RNA samples were selected according to the quality of RNA as measured by the Agilent 4200 TapeStation (Agilent). For further processing, samples with less degradation on the electropherogram results (High sensitivity RNA screening on the TapeStation) were considered. The RNA Integrity Number (RIN) influenced the downstream processing of selected samples. Samples with RIN above 7 were considered to contain intact RNA and therefore required RNA fragmentation, whilst a RIN of  $\leq 1$  was of highly degraded samples, and no fragmentation was done.

### **2.4.2 Library preparation from RNA**

RNA libraries were prepared first by carrying out ribosomal RNA (rRNA) removal using the NEBNext rRNA depletion kit (human/mouse/rat) manufactured by New England Biolabs, USA. Then, RNA fragmentation (if required), first and second-strand cDNA synthesis, end repair, adapter ligation and PCR enrichment of libraries were carried out using the NEBNext Ultra II RNA library preparation kit for Illumina, following the manufacturer's protocol.

#### *2.4.2.1 Library quality control*

RNA from a known formalin-fixed paraffin-embedded (FFPE) tissue positive for human Polyomavirus was processed simultaneously with the pools to ensure successful library preparation and sequencing. In addition, given that the samples used in this study were shipped from Uganda to the UK on a DNA/RNA shield preservative, freshly collected whole blood was used as a control to assess the library preparation procedure.

#### *2.4.2.2 Library quantification*

The library quality was assessed by quantifying the amount of DNA using the High sensitivity D1000 ScreenTape on the Agilent 4200 TapeStation.

### **2.4.3 Next-generation sequencing**

The prepared libraries were sent to the Genewiz Company (Germany), where sequencing was done using the Illumina Hiseq 2500 platform. Fastq files were analysed as expounded in section 2.4.2.4.

## 2.4.4 Using IDseq to analyse NGS data for virus detection

IDseq, a cloud-based bioinformatics platform, was used to detect viruses' presence and relative abundance. Raw NGS data (FASTQ files) from Genewiz (Germany) was uploaded onto the IDseq web interface. Upon uploading, the pipeline is automatically designed to perform host filtration and quality control steps, execute an assembly-based alignment where reads and contigs are assigned and provide an output in the form of a taxonomic categorisation of pathogens [326]. Thus, the taxonomic relative abundances of pathogens were reported using the IDseq easy-to-use visual web-based application. The NGS reads that were concomitant to a pathogen were blasted and compared to sequence data on GenBank using the basic local alignment search tool n (BLASTn) function. Translated nucleotides were analysed using the BLASTx function of the basic local alignment search tool (BLAST). Contigs specific to the FFPE Polyomavirus and its relative abundance were analysed using the IDseq web-based interface.

## 2.4.5 Validation of metagenomics hits

### *2.4.5.1 Detection and confirmation of Hepatitis C virus*

Contigs specific to the Hepatitis C virus (HCV) and its relative pathogen abundance were analysed using the IDseq web-based interface.

#### 2.4.5.1.2 Primers and PCR confirmation of HCV

Primers for HCV confirmation were adopted from Adeboyejo *et al.* [327], as shown in Table 5. Pan-genotype detection was achieved using the NS5B primers targeting the RNA dependent RNA polymerase gene and where possible, genotype subtyping was

accomplished using the NS5a primers. The RNA extraction and cDNA synthesis were done as previously discussed in 2.2.5. The PCR reaction was set up with initial heat activation of HotStarTaq DNA Polymerase at 95°C for 15 minutes, followed by 55 cycles of amplification at 94°C for 20 sec, annealing at 56°C for 20 sec, and 72°C for 30 sec and a final extension at 72°C for 7 minutes.

Gel electrophoresis and Sanger sequencing were done as described in 2.2.7 and 2.28, respectively.

Table 5. Primers for the detection of HCV adopted from Adeboyejo *et al.* [327]

| <b>Name</b>  | <b>Nucleotide position</b> | <b>Primer sequence (5' to 3')</b> | <b>Size (bp)</b> |
|--------------|----------------------------|-----------------------------------|------------------|
| panHCV_NS5bF | 8256 - 8278                | TATGAYACCCGCTGYTTTGACTC           | 366              |
| panHCV_NS5bR | 8644 - 8622                | GCNGARTAYCTVGTTCATAGCCTC          |                  |
| HCV_NS5aFi   | 6081 - 6103                | CAGTGGATGAACMGICTIATIGC           | 759              |
| HCV_NS5aRi   | 6840 - 6823                | CCGGTTCIGGITCRCAIG                |                  |

#### 2.4.5.2.2 HCV prevalence

HCV prevalence was determined as the percentage ratio of the number of samples positive for HCV on PCR to the total number of samples (1090) in this cohort.

### 2.4.5.3 Detection and confirmation of Pegivirus

Contigs specific to Pegivirus and its relative pathogen abundance were analysed using the IDseq web-based interface.

#### 2.4.5.3.1 Primers and PCR confirmation of Pegiviruses

As described by Bonsall *et al.* [328], we used primers (Table 6) targeting a conserved region between 3931 and 4546 on the pegivirus genome. The PCR was carried out using

2 µL of template cDNA as indicated by Bonsall *et al.* The PCR conditions were an initial heat activation at 95°C for 15 minutes, 55 cycles of denaturation for 18 seconds at 94°C, annealing for 21 seconds 50°C, 60 seconds at 72°C and a final extension for 10 minutes at 72°C. Two microliters of the first PCR product were used for the semi-nested PCR under similar reaction conditions, giving a 265bp product visible under UV- light as described in 2.2.7. The second PCR amplicon was sent for Sanger sequencing as described in 2.2.8.

Table 6. Pegivirus primers for the first and second PCR and the expected size of the amplicon.

|   | Primer name          | Sequence                   | Size (bp) |
|---|----------------------|----------------------------|-----------|
| 1 | Pegivirus (1)_ 3931F | GSGCNATGGGNCCNTAYATGGA     | 615       |
| 2 | Pegivirus (2)_4546R  | GTNACYTCVACNACCTCCTCYACCA  |           |
|   | 2 <sup>ND</sup> PCR  |                            |           |
| 3 | Pegivirus (3)_4092F  | GTGGTNATHHTGYGAYGAGTGYCA   | 265       |
| 4 | Pegivirus (4)_ 4357R | TCRCACTCMRCCTTKGARTGRCARAA |           |

#### 2.4.5.3.2 Analysis of the seroprevalence of Pegiviruses

Using the confirmed pegivirus from the NGS sample as a positive control, all the samples in this cohort were screened for pegiviruses under similar PCR conditions. Pegivirus seroprevalence was determined as the percentage of samples positive for Pegivirus on PCR to the total number of samples (1090) in this cohort.

#### 2.4.5.4 Detection of Orbiviruses

Reads specific to Orbivirus and its relative pathogen abundance were analysed using the IDseq web-based interface.

##### 2.4.5.4.1 Primers and PCR confirmation of Orbivirus

Primers targeting the consensus VP1 region of Orbivirus were acquired from previous research by Palacios *et al.* [329] (Table 7). Primers were designed from the translated nucleotide Orbivirus reads from the IDseq for direct PCR (Table 8).

Table 7. Primers by Palacios *et al.* OrbiVP1-F2494-1 targets Orungo virus, OrbiVP1-F2494-2 and F2494-4 target the AHSV serotypes, OrbiVP1-F2494-3 targets Wallal and Changuinola viruses.

| Primer          | Sequence (5' – 3')            | Size (bp) |
|-----------------|-------------------------------|-----------|
| Orbivirus_2663R | CCYTGYYTNGCRTCNGTYTGNGTYTTYTC |           |
| OrbiVP1-F2494-1 | TCTGAGATGTAYGTYGGAGATGATA     | 175       |
| OrbiVP1-F2494-2 | TCTGAGATGTAYGTYGGTGATGACA     | 177       |
| OrbiVP1-F2494-3 | TCGGAACARTAYGTVGGNGAYGATA     | 177       |
| OrbiVP1-F2494-4 | TCNGARCARTAYGTKGGNGAYGACA     | 163       |

Table 8. In house primers designed from the IDseq translated nucleotide orbivirus reads.

| Primer   | Sequence (5' – 3')   | Size (bp) |
|----------|----------------------|-----------|
| Orbi_13F | GCATCTTTCAGGGGAAACTC | 137       |

|          |                       |  |
|----------|-----------------------|--|
| Orbi_13R | TCATCTCCGACGTACTGTTCA |  |
|----------|-----------------------|--|

#### 2.4.5.5 Detection of Polyomaviruses

Contigs and reads specific to Polyomaviruses and their relative pathogen abundances were analysed using the IDseq web-based interface.

##### 2.4.5.5.1 Primers and PCR confirmation of Polyomaviruses

The Pan-polyoma assay, developed by Johne et al., 2005 [330], was applied to detect polyomaviruses in the samples. The primers used in the assay are shown below (Table 9).

Table 9. The pan polyoma primer set was adopted from Johne *et al.* to detect polyomaviruses.

| Assay       | Primer | Sequence 5' to 3.'             | Size (bp) |
|-------------|--------|--------------------------------|-----------|
| Pan-polyoma | VP1–2f | ATGAAAATGGGGTTGGCCCNCTNTGYAARG | 274       |
|             | VP1–2r | CCCTCATAAACCCGAACYTCYTCHACYTG  |           |

## **3.0 Investigating the prevalence and spatial-temporal distribution of Arboviruses in Uganda**

### **3.1 Introduction**

Arboviruses are arthropod-borne viruses transmitted by mosquitoes, biting midges, and ticks; grouped within viral families such as; Flaviviridae, Rabdoviridae, Reoviridae, Togaviridae, Nairoviridae, Peribunyaviridae, and Phenuiviridae.

The global burden of arboviruses is significant, with up to 700,000 deaths annually attributed to arbovirus-related diseases [331]. These viruses also cause high morbidity. For example, in 2005 alone, yellow fever (YFV), chikungunya (CHIKV), rift valley fever (RVFV), and Japanese encephalitis (JEV) viruses were responsible for an estimated 5 million disability-adjusted life years (DALY) in the human population [332]. A more recent survey estimated an average loss of 106,000 and 44,000 DALY among CHIKV and ZIKV patients, respectively, between 2010 and 2019 [333]. Among mosquito-borne arboviruses, Dengue virus (DENV) has a massive impact on public health, with over 390 million cases reported annually, mainly affecting Asia and Latin America [334]. YFV, which is mainly a problem in Africa [335],[336] and South America [337],[338], causes over 200,000 global cases and 30,000 deaths annually [339]. In Uganda, outbreaks of YFV date back to the late 1930s [340],[341],[342]. A recent outbreak in 2016 in northern Uganda had a case fatality rate of over 24%, with a 7.5% seroprevalence amongst the population tested [343]. Earlier serological studies by Henderson *et al.* reported a 16.8% prevalence for WNV, YFV, and ZIKV [344]. The immunoglobulin G (IgG) seroprevalence

of CHIKV and ONNV (which are not easily distinguished serologically) is estimated at 31 - 38.6% in Uganda [344],[345]. Among alphaviruses, the first detection of CHIKV was in the central region [346] and was later continuously reported for 11 months around the Zika forest area [347] from May 1961 to March 1962.

The laboratory testing algorithm for arboviruses at the Uganda Virus Research Institute (UVRI) involves serological detection of IgM antibodies to an arbovirus (YFV, WNV, DENV 1-4, ZIKV, CHIKV), followed by plaque reduction neutralisation test (PRNT) confirmation and where possible RT-PCR for the detection of viral nucleic acids [261]. Although IgM serology provides evidence of recent transmission and detection during the acute phase of illness, the method is plagued by cross-reactivity of flavivirus antibodies with envelope proteins of related flaviviruses [348] in arbovirus endemic regions. Cross-neutralisation is common in regions where more than one arbovirus circulates or results from vaccination [349]. In addition, despite continuous surveillance, there is little information about the epidemiology, spatiotemporal distribution, and seasonality of mosquito-borne arboviruses in Uganda. The *Aedes* species is known as a competent virus transmitter of YFV, ZIKV, and CHIKV from a single bite [350],[351]. Therefore, given the abundance and distribution of *Aedes* species within Uganda, a directly proportional distribution pattern of YFV, ZIKV, and CHIKV would be expected. However, the distribution of these viruses within the country is not known. Therefore, in addition to defining the epidemiology, spatiotemporal distribution and seasonality, this study aims to provide evidence supporting the co-circulation of YFV, ZIKV, and CHIKV during outbreaks.

## 3.2 Results

### 3.2.1 Description of the study population characteristics

Samples were collected from the Northern, Central, Western, and Eastern regions of Uganda (Figure 5). Demographic and laboratory data were available for 1900 anonymised cases collected between January 2016 and November 2019. Four hundred fifty-nine (459) samples were excluded due to incomplete data such as age, gender, region, and missing date of sample collection. Complete demographic and laboratory data were available for 1,441 results, of which 137 (9.5%) had a positive IgM result for an arbovirus. Of the 1,441 samples, 862 (59.8%) samples were collected from the central region, 472 (32.7%) from the Western, 73 (5.1%) from the Northern, and 34 (2.4%) from the Eastern region.

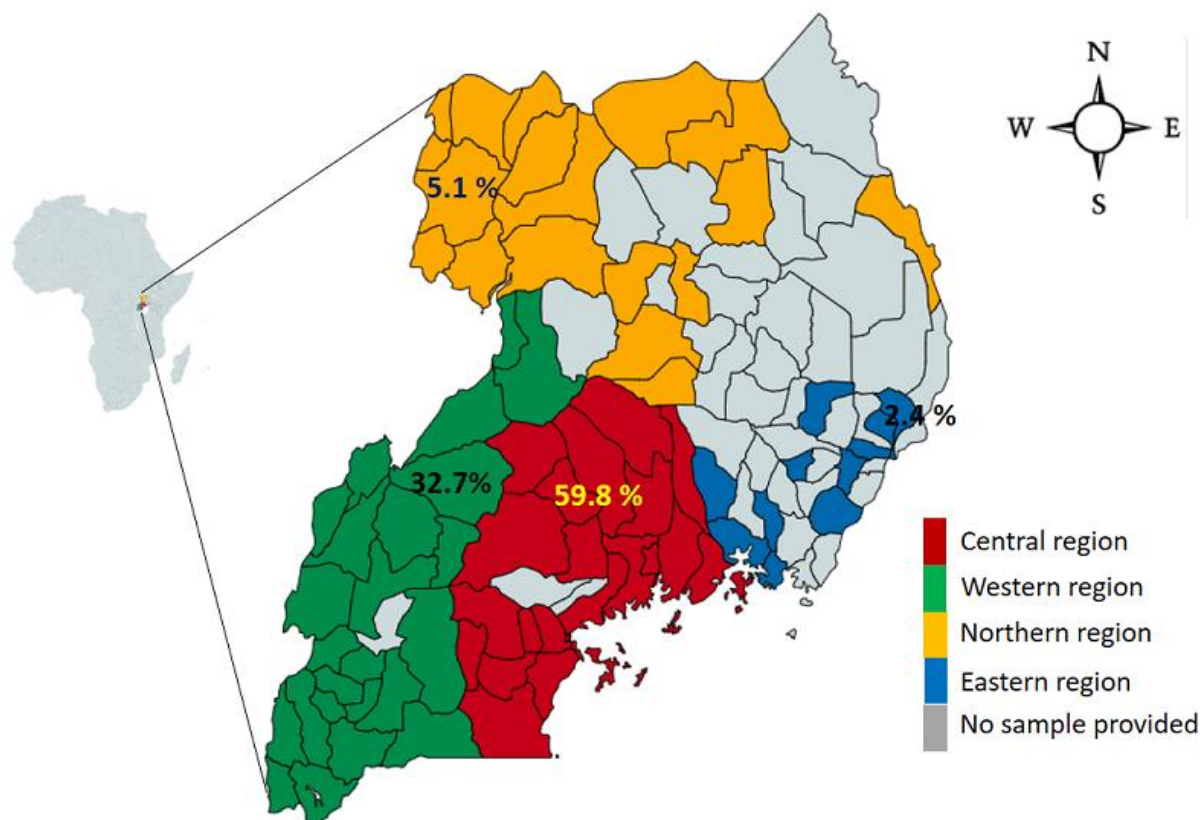


Figure 5. Sample collection by region: *Extract of the map of Uganda showing sample collection by region, further divided into districts. The orange represents the north, green for the west, red for the central, blue for the east, and grey shows areas with no sample collection. The number of samples collected from each region is represented as a percentage of the total collection. Map not drawn to scale.*

Of the 137 positives, 72 (52.6%) were confirmed positive by the PRNT. Among the PRNT confirmed cases, 53 (73.6%) were from the central, 13 (18.1%) were from the western, whilst six (8.3%) were from the eastern and northern regions (Figure 6).

Flaviviruses accounted for 55.6% of the total PRNT confirmed cases, whilst the rest (44.4%) were confirmed as alphavirus infections. YFV had the highest prevalence of the total positives with 20 cases (27.8%), while the PRNT test confirmed no DENV.

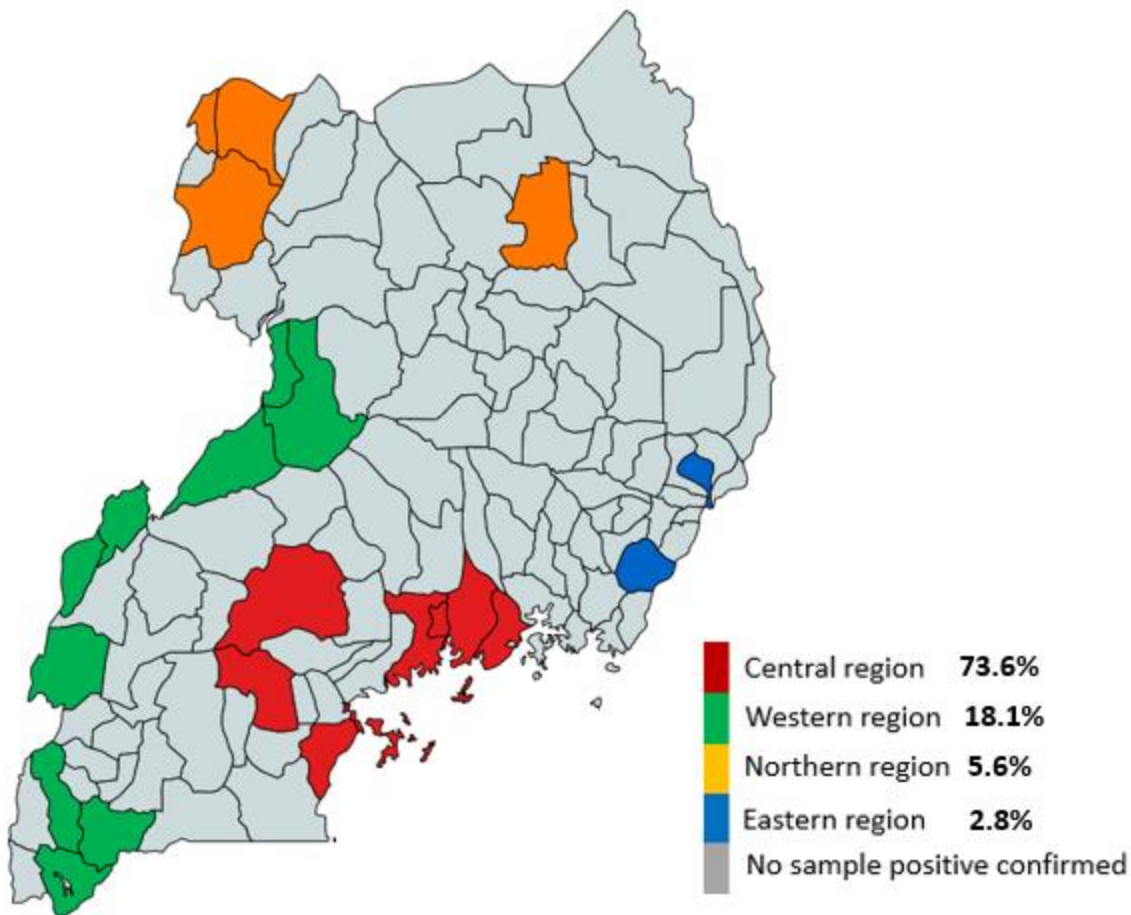


Figure 6. Regions with arboviruses confirmed by PRNT. *The map divisions represent districts and are coloured to represent geographical regions. The uncoloured districts represent areas where no arbovirus has been confirmed. The map is not drawn to scale.*

### 3.2.2 Seasonality of arboviruses

The East African region has two main rain seasons from March-May and October to December [352]. Arbovirus activity occurred throughout the study period, with higher peaks consistently observed during March-May than in the October – December rainy season (Figure 7). The observed flavivirus peak in March-May 2016 (Figure 7) is consistent with the YFV outbreak in central and western regions (Figure 8) [342] that

continued until June 2016. In addition, three previously unregistered YFV outbreaks occurred, with the first occurring in the central region in September-October 2017 (Masaka district) and the second in April 2018 (Wakiso district), while the third spread across the central, north, and west regions from March-September 2019. Generally, alphaviruses circulated throughout the year without a specific pattern (Figure 7). Only one CHIKV and seven (7) ONNV -positives were confirmed by PRNT whilst 24 cases were unspecific alphaviruses. All PRNT confirmed arboviruses, their distribution across age groups, geographical regions and gender are shown in Table 10.

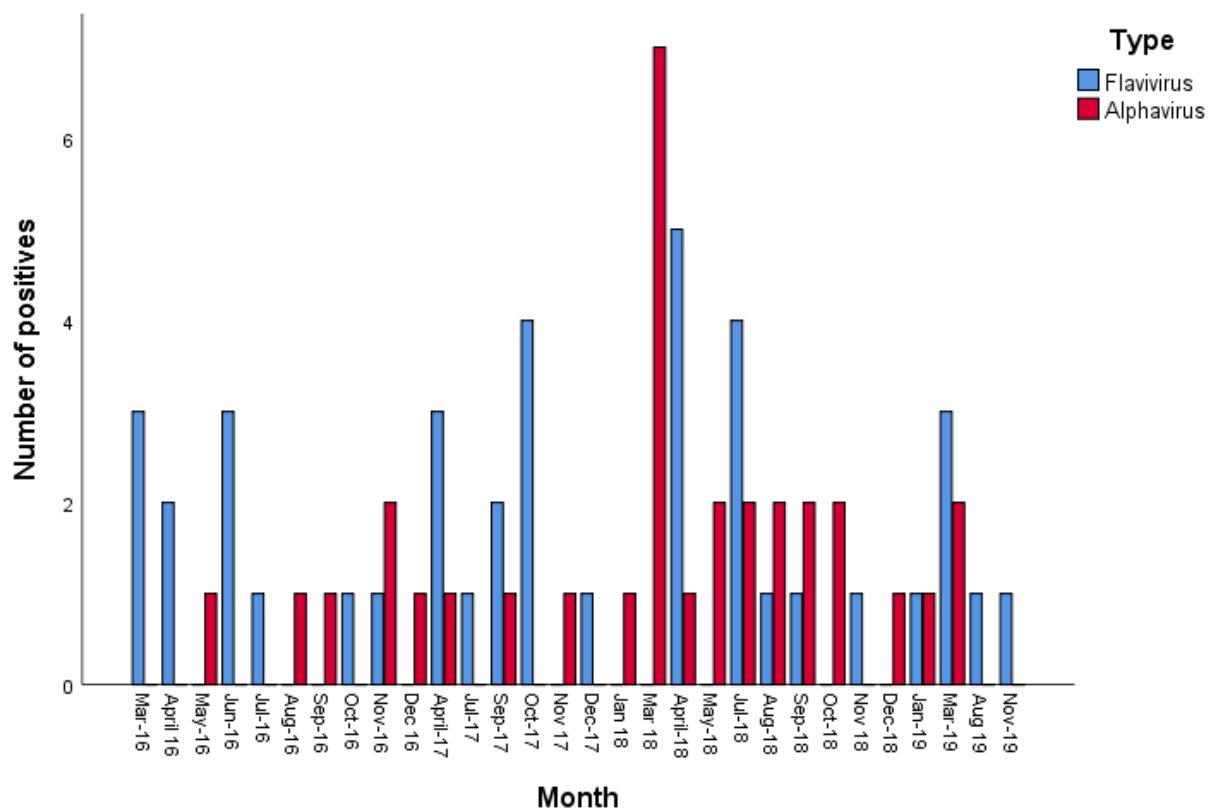


Figure 7. Number of arbovirus positives by month in Uganda from March 2016 to November 2019. *The alphaviruses (red) and flaviviruses (blue) incidences peaked between March and May every year. The March-May 2016 and 2019 peaks coincide with the YFV outbreak in Masaka and Koboko districts.*

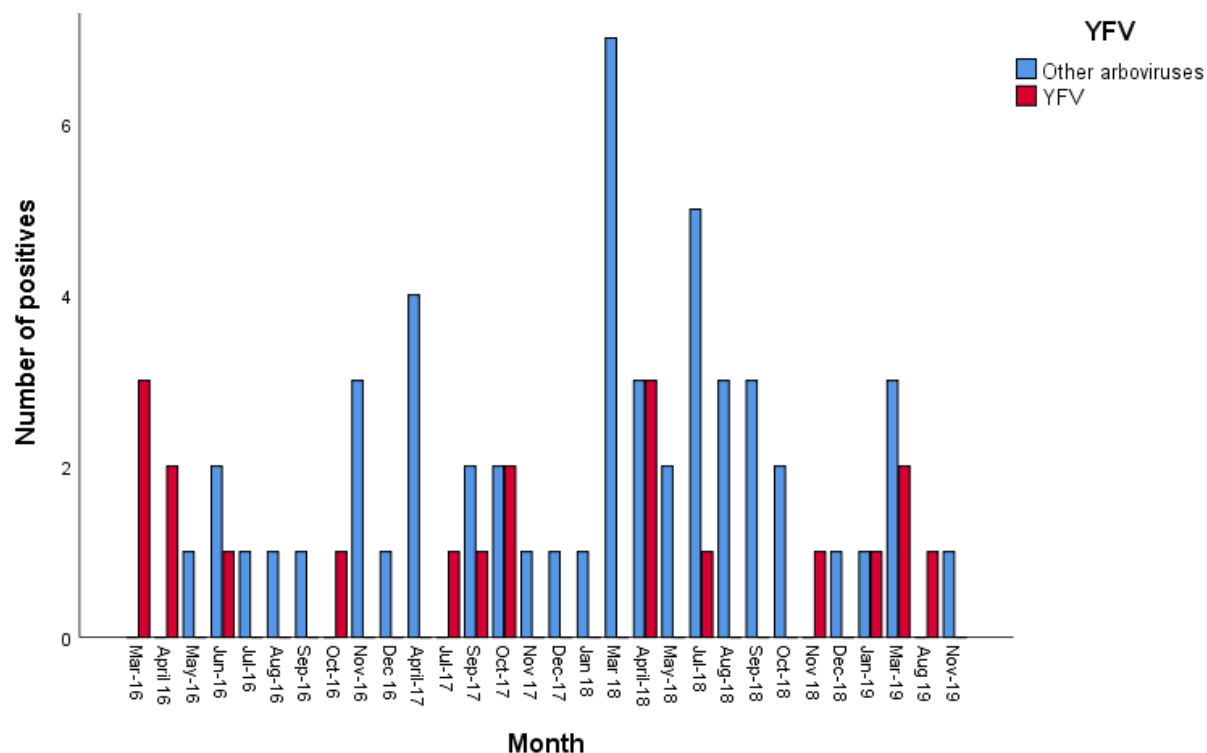


Figure 8. PRNT confirmed YFV (red) and other arboviruses (blue) cases observed in Uganda from 2016 to 2019. *Four distinct YFV outbreaks were observed, with the first occurring in the central region March-May 2016 in the central and western regions. The second was in September-October 2017 (Masaka district) and the third in April 2018 (Wakiso district), while the fourth spread across the central, north, and west regions from March to September 2019.*

Table 10. PRNT confirmed arboviruses and their distribution across age groups, geographical regions and gender. Abbreviations: YFV- Yellow fever virus, ZIKV- Zika virus, CHIKV- Chikungunya virus, ONNV- Onyong'nyong virus, USFV- Unspecific flaviviruses, and USAV- Unspecific alphaviruses.

|       | REGION  |      |       |      | AGE GROUP |       |       |       |     | GENDER |    |
|-------|---------|------|-------|------|-----------|-------|-------|-------|-----|--------|----|
|       | Central | West | North | East | 5-14      | 15-34 | 35-54 | 55-64 | 65+ | M      | F  |
| YFV   | 16      | 2    | 2     |      | 1         | 12    | 6     |       | 1   | 8      | 12 |
| WNV   |         | 1    |       |      |           | 1     |       |       |     | 1      |    |
| ZIKV  | 3       |      |       |      |           | 1     | 1     |       | 1   | 2      | 1  |
| CHIKV | 1       |      |       |      |           | 1     |       |       |     | 1      |    |
| ONNV  | 4       | 3    |       |      |           | 5     | 1     | 1     |     | 3      | 4  |
| USFV  | 15      |      |       | 1    | 1         | 8     | 6     | 1     |     | 8      | 8  |
| USAV  | 14      | 7    | 2     | 1    | 1         | 10    | 9     | 1     | 3   | 11     | 13 |

### 3.2.3 Risk factors for Arbovirus infections.

On univariate analysis, no risk was observed due to gender among the patients (Table 2). Among the 1-4-year-old age group, no arbovirus IgM antibodies were observed. However, the risk of an arbovirus infection, especially flaviviruses, was significant in the 5-14 age group  $F(1-1439) = 8.2, p = 0.004$ . In addition, patients in the 5-14 age group were four times more likely to be infected with an arbovirus  $p=0.003, 4.1 (1.3- 12.3)$ . Workers within the 55-64 age group were 1.4 times more likely to test positive for an arbovirus infection. Despite IgM antibody and PRNT confirmation of arboviruses within the 15-34, 35-54 and over the 65 years, these patients had a lower risk of arbovirus infection.

Patients within the central and western regions had a significantly high likelihood of arboviral infection  $F(1, 1439) = 6.0, p=0.014$  and  $F(1, 1439) = 7.5, p=0.006$ , respectively, than patients from the northern and western regions. Furthermore, patients from the central region were 1.9 times 1.9 (1.1- 3.1, CI 95%),  $p=0.006$  more likely to have an arboviral infection than other regions.

Patients who were farmers, fishmongers, security personnel, butchery operators, builders, soldiers, and business persons were categorised as the "outdoor workers". Patients whose occupation was officer-related, housewives, students, teachers, children, bar and hotel maids, tailors, and waiters were grouped in the "Indoors workers" category. Furthermore, patients whose occupation was nursing, veterinary officers, clinical and medical officers, and laboratory technicians were grouped within the Healthcare worker (HCW) category. Of the 72 PRNT confirmed positives, 44 (61.1%)

were patients described as outdoor workers ( $P 0.05$ ), and 13 (18.1%) were Indoor workers ( $p0.07$ ). In addition, Outdoor workers were 1.6 times (1.2- 2.9, CI 95%) more likely to have an arboviral disease than any other workers. Furthermore, the univariable analysis by linear regression showed that the type of occupation contributed significantly to infection (Table 11).

Table 11. Epidemiological characteristics of arboviruses. Abbreviation: HCW- Healthcare workers.

|                             | Total<br>positive<br>{n=<br>(%)} | Total<br>negative<br>{n=<br>(%)} | Univariate                      | CHI Square test:<br>pValue, OR (95% CI) |
|-----------------------------|----------------------------------|----------------------------------|---------------------------------|-----------------------------------------|
| <b>Prevalence (n= 1441)</b> | <b>72</b>                        | <b>1369,</b>                     |                                 |                                         |
| <b>GENDER</b>               |                                  |                                  |                                 |                                         |
| Female                      | 39 (54.2)                        | 670 (48.9)                       |                                 |                                         |
| Male                        | 33 (45.8)                        | 732 (51.1)                       |                                 |                                         |
| <b>AGE GROUP</b>            |                                  |                                  |                                 |                                         |
| 0-4 yrs                     | Nil                              | 83 (6.1)                         |                                 |                                         |
| <b>5-14 yrs</b>             | <b>3 (4.2)</b>                   | <b>231 (16.9)</b>                | <b>F(1,1439)= 8.2, p=0.004</b>  | <b>0.003, 4.1 (1.3- 12.3)</b>           |
| 15-34 yrs                   | 37 (51.4)                        | 584 (42.7)                       |                                 | 0.08 (0.6- 1.0)                         |
| 35-54 yrs                   | 24 (33.3)                        | 322 (23.5)                       |                                 | 0.7 (0.5- 1.0)                          |
| 55-64 yrs                   | 3 ( 4.2)                         | 81 (5.9)                         |                                 | <b>1.4 (0.5- 4.4)</b>                   |
| 65+ yrs                     | 5 ( 6.9)                         | 68 (4.9)                         |                                 |                                         |
| <b>REGION</b>               |                                  |                                  |                                 |                                         |
| <b>Central region</b>       | <b>53 (73.6)</b>                 | <b>809 (59.1)</b>                | <b>F(1, 1439)= 6.0, p=0.014</b> | 0.8 (0.6-0.93)                          |
| <b>Western region</b>       | <b>13 (18.1)</b>                 | <b>459 (33.5)</b>                | <b>F(1, 1439)= 7.5, p=0.006</b> | <b>p0.006, 1.9 (1.1- 3.1)</b>           |
| Northern region             | 4 (5.5 )                         | 69 (5)                           |                                 | 0.9 (0.3-2.4)                           |
| Eastern region              | 2 (2.8)                          | 32 ( 2.4)                        |                                 | 0.8 (0.2-3.1)                           |
| <b>OCCUPATION</b>           |                                  |                                  |                                 |                                         |
| Outdoors                    | 44 (61.1)                        | 674 (49.2)                       | <b>F(1,1439)= 3.9, p=0.05</b>   | <b>p0.05, 1.6 (1.1- 2.6)</b>            |
| Indoors                     | 13 (18.1)                        | 526 (38.4)                       | F(1,1439)= 3.4, p=0.07          | p0.08, 0.6 (0.3- 1.1)                   |
| HCW                         | 3 (4.2)                          | 28 (2)                           |                                 | <b>1.6 (0.2- 11.4)</b>                  |
| Other                       | 12 (16.7)                        | 141 (10.3)                       |                                 | 0.6 (0.4- 1.1)                          |

### 3.2.4 Spatial and spatiotemporal analysis of Arboviruses

Spatial and spatiotemporal analyses were used to understand arboviruses' transmission patterns by analysing districts with more than expected arbovirus incidences. Arboviral incidences were analysed by applying a cluster restriction of three (3) cases for areas with high rates and a maximum spatial cluster size set at a radius of 50km. Upon spatial analysis, 23 districts with arboviral activity and five distinct clusters in 6 districts were recorded (Figure 9-I). The average population in the districts with arboviral activity was 1,528,539 persons, with an average annual case incidence of 1.3 per 100,000. Of the five clusters, the first cluster, A, in Buikwe district (geographical Coordinates 0.3012 N, 33.012 E) was observed within an average population of 194,530. A total of 15 cases differing from the 11.7 cases expected occurred in Buikwe district, giving an annual incidence of 1.6 cases per 100,000. Two clusters, B and C, overlapped between Masaka (geo coordinates 0.341100 S, 31.7361 E) and Sembalule districts (0.0666 S, 31.4833 E), with an average radius of 41.47 km. Sixteen cases were detected (differing from the expected 8.4 cases) with an average population of 142,091. Therefore, the arboviral incidence was 2.4 per 100,000 with a relative risk of 2.15. Three spatial clusters (D- F) were observed in the south western region of Uganda, comprising districts, *i.e.* Ntungamo (0.881 S, 30.25 E), Rubanda (1.183 S, 29.85 E) and Rukungiri (0.783 S, 29.33E), respectively. These clusters covered an average radius of 45.4 km. A total of four cases were detected (differing from the expected 1.7 cases) in an average population of 23,481, giving an annual incidence of 2.9 cases per 100,000 and a relative risk of 2.41.

The spatial-temporal (space-time) scan for clusters applies a cylinder whose circular base covers the geographical area analysed whilst the height measures the time frame of infection activity in the given area [305]. Overall, 6 clusters were detected, of which only two were associated with high rates (Figure 9-II). Clusters A to D were associated with low rates and occurred in Buikwe (A), Mukono (B- coordinates, 0.481 N, 32.771 E), Kampala (C- coordinates 0.363 N, 32.61 E), and Wakiso ( D- 0.167 N, 32.5 E) districts. On the contrary, clusters E and F were associated with high rates and were observed in Masaka and Sembabule districts, respectively. The cluster in Masaka (cluster E) had an average population of 136,181 from March to May 2016. Three cases were detected (differing from the expected 0.18), giving an annual incidence of 21.2 cases per 100,000 and a relative risk of 17.6 ( $p= 0.024$ ). The second cluster occurred from September to November 2017 and overlapped between Masaka and Sembabule districts, with five cases (from an expected 0.84). Given an average population of 142,091, the annual arboviral incidence in the two districts was 7.5 cases per 100,000 and a relative risk of 6.35 ( $p=0.059$ ).

### *3.2.3.1 Spatial and Spatio-temporal distribution of YFV in Uganda*

The arboviral activity was further disintegrated to analyse the role of each viral species relative to districts and time. In Uganda, the disease threshold for a YFV alert is actioned when there is one suspected or confirmed case in the country

(<http://library.health.go.ug/publications/disease-surveillance-outbreaks/case-definitions-and-epidemic-thresholds-integrated>). For this reason, whilst setting up the data analysis

parameters for YFV activity, we considered a cluster to contain at least two PRNT confirmed cases. As a result, between March 2016 and August 2019, twenty (20) YFV confirmed cases were detected in 10 districts with an average population of 1,005,462, indicating an incidence of 0.6 per 100,000 (Figure 9-III). Furthermore, within the ten districts, one significant YFV cluster was identified in Masaka district (C), with 6 cases (expected cases were 0.8) and an average population of 41,344. The YFV incidence in Masaka was 4.1 per 100,000, with a relative risk of YFV infection of 9.99 ( $p<0.001$ ). The cluster in Wakiso district involved four (expected number, 8.6) YFV cases in an average population of 434,585. Thus, Wakiso district experienced 0.3 cases per 100,000, which was statistically insignificant ( $p=0.18$ ). Kampala district recorded only one case (expected 4.1) in an average population of 208 242 and an incidence of 0.1 per 100,000 ( $p= 0.26$ ). Of interest, YFV activity was detected in all but the eastern region.

Only the Masaka cluster was statistically significant in space-time analysis of YFV activity in Uganda. The cluster occurred from March to May 2016, with 3 (of the 0.08 expected) cases in an average population of 41,344. Thus the YFV incidence was estimated at 22.2 cases per 100,000 and a relative risk of 45.78 ( $p<0.001$ ).

### *3.2.3.2 Spatial and Spatio-temporal distribution of Alphaviruses in Uganda*

Uganda has no alert and action level to track alphavirus outbreaks. Therefore, Uganda's alphaviruses' spatial and space-time distribution was analysed by restricting high rate clusters to at least 4 cases.

On spatial analysis, 16 locations with 32 cases were detected between May 2016 and March 2019. Low rates of alphavirus activity were detected between Wakiso and Kampala districts, having 6 (expected 14.5) cases and covering an average radius distance of 25km. Thus, the alphavirus incidence was estimated at 0.4 cases per 100,000 and a relative risk of 0.28 ( $p=0.019$ ). In addition, one cluster was observed in Hoima district (1.43 N, 31.35 E) in the western region (Figure 9-IV). Two (of the 0.36 expected) cases were detected in an average population of 14,772. Thus the annual incidence in Hoima district was 5.8 cases per 100,000 and a relative rate of 5.86 ( $p=0.53$ ).

On space-time analysis, the cluster observed in Wakiso district occurred between August and December 2017, while the Hoima district cluster occurred from January to May 2018.

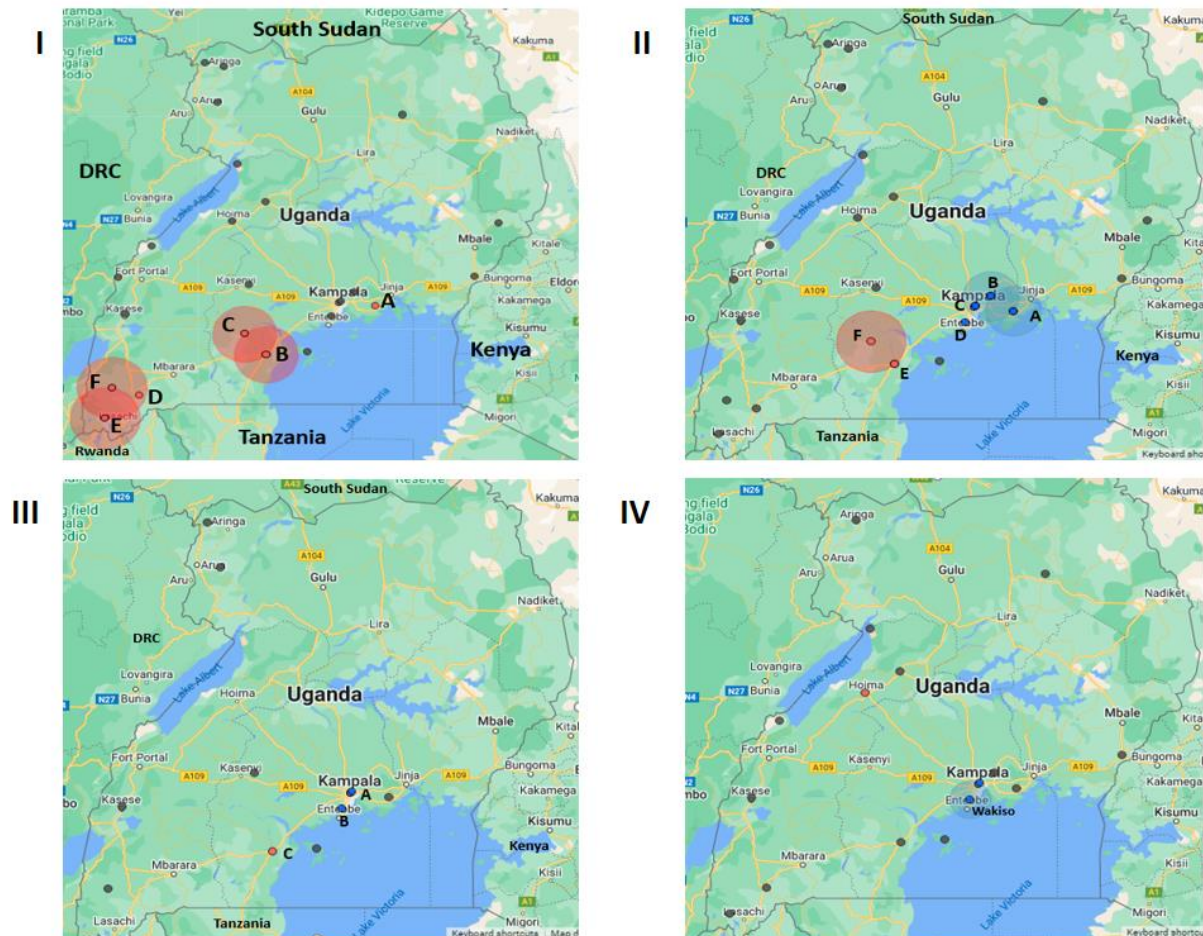


**Figure 9. Spatio-temporal distribution of arboviruses in Uganda in clock-wise direction.** **I.** Spatial analysis of arboviruses showing their distribution in Uganda. Arboviral incidences were observed in 23 districts (black and red dots). Red circles represent clusters in districts from which more than two arboviruses were detected, indicated as loci A- F. Districts A-F are: A. Buikwe, B. Masaka, C. Semababule, D. Ntungamo, E. Rubanda, and F. Rukungiri. Abbreviation: DRC- Democratic Republic of Congo. **II.** Space-time scan of arbovirus activity in Uganda. The blue, black and red spots show areas where arbovirus activity was observed. Two clusters E (Masaka) and F (Sembabule) had high arboviral rates and occurred from March-May 2016 in Masaka district. Cluster F overlapped with E but occurred from September to February 2017. The blue dots represent districts with low arboviral rates. A- Buikwe, B- Mukono, C- Kampala and D – Wakiso. **III.** Spatial Analysis of YFV activity. The blue, black and red spots show areas where YFV activity was observed. Cluster C (in Masaka) represent significant YFV activity, whilst clusters A (Kampala) and B (Wakiso) showed low activity. All but the eastern region showed YFV activity. **IV.** Spatial distribution of alphaviruses May 2016 - March 2019. The black, grey and red spots show the 16 locations where alphaviruses were detected. The red spot shows the Hoima district cluster with a high rate of alphavirus activity, whilst the blue spots represent areas of low virus rates in Kampala and Wakiso.

### 3.3 Discussion

This study describes the epidemiological pattern of arboviruses, highlighting the role of YFV in morbidity among patients in Uganda. In addition, this study describes the seasonality of arboviruses, defines the patient demographics during the acute phase of illness, and determines the country's high-risk areas.

The general prevalence of arboviruses from 2016 to 2019 was defined based on the PRNT results, highlighting potential outbreaks. Several previous studies have analysed arbovirus seroprevalence based on IgG studies [353],[354],[355], and a few more have attempted to analyse recent arboviral infections using IgM results [356],[357]. However, unlike these studies, this study provides more accurate epidemiological, and public health information on the nature of the outbreaks using PRNT confirmed results. It is unclear why the central region had a higher prevalence of arbovirus infection, however, one possibility is that Lake Victoria may play a significant role in arboviral prevalence in the region. Secondly, the trend of increasing temperatures observed in Uganda over the past two decades [358], [359], which may promote the survival of arboviral vectors [360] and, in turn, justify the endemicity of arboviruses within the region. Such an effect would not be unprecedented; a study in Brazil identified that environmental factors such as temperature played a vital role in the distribution of arboviruses in high-risk areas [361]. However, why the increasing temperature does not have an equal effect on other regions in the country is unknown. Furthermore, the factors that favoured the dominance of YFV and its expansion from the central and western regions into the northern region by 2019 are unknown, despite sharing a common vector with ZIKV, DENV, and CHIKV.

High arbovirus infection levels among the outdoor workers (farmers, fishmongers, market vendors and builders) may be associated with the vectors' feeding and resting behaviour patterns [362]. In addition, climate change has affected farmers' agricultural patterns [358],[359], which would probably impact the length and period of exposure to arboviral vectors whilst on the fields. In our study, the Arboviral incidence among health workers cannot be concluded as a work-related occurrence. However, before the 1970s, frequent arbovirus infection of laboratory personnel involved in studies manipulating arboviruses was common, with over 2700 cases and 107 fatalities [363]. Patients aged 1-4 years reported no arbovirus infection, contrasting with a Puerto Rico report that recorded CHIKV infections in infants below one year [364]. Our findings agree that arboviral infections have been documented in patients five years and older [364],[365].

Spatial and spatio-temporal analysis mapped the arbovirus distribution and time pattern of YFV. The spatial and temporal distribution of WNV, ZIKV, and DENV were not analysed due to insufficient data to support statistical analysis. The arbovirus species' definitive spatial and spatio-temporal distribution could not be fully defined due to the existence of unspecified alpha and flaviviruses. The diagnosis of unspecific alpha and flaviviruses highlights the importance of early detection by applying assays such as molecular detection by PCR. However, spatiotemporal analysis using IgM results identified a possible CHIKV outbreak in Buikwe district co-occurring with the YFV outbreak in Masaka and Rukungiri districts in February-May 2016. Simultaneous circulation of CHIKV and YFV has been reported elsewhere [366],[367], yet the notifiable YFV outbreak overshadows CHIKV activities. It is important to note that climatic conditions may play a role in the spatio-temporal distribution of arboviral infection. However, a study in Mexico

found that socio-economic attributes were larger than climatic attributes regarding the spatio-temporal distribution of arboviral infections [368]. A more detailed study focusing on the impact of climate on the distribution of arboviruses in Uganda would require more climatic data on temperature, precipitation, wind, atmospheric pressure, solar radiation and humidity. This study did not investigate such climatic data.

YFV cases were detected throughout the study period (2016-2019), of which 80% were observed within the central region. The reduction in YFV cases between 2017 to 2019 is attributed to mass vaccination [369], [370], although new YFV activity was detected in northern Uganda. However, according to our study, these outbreaks occurred where no previous YFV incidences or mass vaccinations had been conducted. In addition, because the northern and eastern regions have limited sentinel sites, active capture of arboviral outbreaks in the regions is dependent on incidence reporting by referral hospitals. Thus, vital information on the epidemiology of outbreaks may be missed out.

## 3.4 Conclusion

Arboviruses, mainly yellow fever virus, occur during the rainy seasons. There is a need to reduce the transmission of arboviruses among outdoor workers (farmers, fishmongers, market vendors, and builders) through mass vaccination for preventable arboviruses. Early mass vaccination of the 5-14 age group will significantly lower YFV cases in the country.

## 4.0 Screening for viral AFI in IgM - negative using RT-PCR

### 4.1 Introduction

The definition of an AFI is a fever spanning up to 7 days without a known cause of infection. Any febrile illness lasting more than seven days with a temperature (usually above 38°C) despite inpatient investigations qualifies as a fever of unknown origin (FUO) [371], [372]. Viral AFI is part of the broad aetiologies of AFI, which includes other causes such as bacteria, fungi and parasites.

All serum and whole blood samples received at the department of Arbovirology, UVRI, are tested for YFV, WNV, DENV, ZIKV, and CHIKV antibodies. The laboratory diagnosis involves testing for antibodies using IgM specific ELISA as described by Kayiwa *et al.*, confirmation by plaque reduction neutralisation test (PRNT) where possible, and PCR [300]. The specimens are collected from patients presenting with AFI signs and symptoms such as fever and a recorded temperature  $\geq 38^{\circ}\text{C}$  at sentinel sites. More than 80% of the samples are negative for arboviruses, which compelled us to explore other causes of viral AFI.

Polymerase chain reaction (PCR) is an efficient and reliable method for the molecular detection of pathogens present in patient specimens. Numerous studies have applied the PCR assay to identify viruses in specimens successfully [373]. Where several subtypes or strains of a virus exist, degenerate primers targeting consensus regions within the genome can effectively identify and detect novel and existing strains [374]. Designing degenerate primers requires knowledge of the most conserved region of a

given virus and identifying loci of potential mismatches for which a degenerate base is required.

This study used previously published primer sets targeting viral families or genera to identify the aetiologies of AFI in the described arbovirus IgM negative cohort.

## 4.2 Results

### 4.2.1 Designing the positive control concatemer

In the absence of positive control material (which can be derived from RNA or DNA) from pathogens, artificial positive controls (APC) can be designed as alternatives [375]. The positive control included 14 viruses, *Plasmodia* spp. and two bacterial pathogens (*Brucella* and *Leptospira* spp.). The sources of the primers used to design the APC are elaborated in chapter two, section 2.2.2. In addition to the 14 viruses, primers targeting conserved regions of *Plasmodia* spp., *Leptospira* spp. and *Brucella* spp. were added. The pan –*Plasmodia* spp. and the *Leptospira* spp. primers targeting the 18s and the 16s rRNA genes, respectively, were adopted from Waggoner *et al.* [321]. Primers for the detection of *Brucella* spp. targeted *bcs*p31 gene as described by Probert *et al.* [322]. These non-viral pathogens were added onto the concatemer due to their endemicity in the region and further rule them out as causes of AFI. The APC was comprised of representative sequences of particular pathogens within viral families or genera (Figure 10). The order of the APC concatemer, starting from the 5' end, was Metapneumovirus, Lassa, *Brucella* ssp., Avulavirus/ Rubula viruses, Henipa/Morbillo/Respiroviruses, Enterovirus, Filoviruses, CCHFV, Malaria, Bwamba V, Rhabdoviruses, Hepatitis E Virus, Bunyamwera, YFV, Chikungunya virus, Rift valley fever, and *Leptospira*. The arrangement of the 3' end of the pathogens followed in a non-homogenous manner dependent on the expected size (bp) of the PCR product, as shown in Table 12.

5' ACTGATCTcAGcAAgTTcAAcCagGCAGAATTAGTGAAAGGGAGAGCAATTCGCTCGGTTGCCAATATCAATGCGGTTATCCT  
CATCTCTTTGAGTGGATTCAGCCATAATTTGTGGAATAATAATCAACGGACATGGTGTGAAGAGTCTATTAGCTGCCTTCCTA  
GTAACATGATGGTCTTCTGGCGACGCTTTACCCAGAAATACAAAGTCCCTGGTCGCGCAAGCGAGAAAAGTTTCTTCTCAATGCC  
GCAAAGGCCAAAACAGGGCYACAARGGTGGTCGATTTTTTGTCTCTAATGACCTTCGTGCTTGTGGTCCGTGGTGTGCA  
GGGAGATAATCAAGCAATTGCTTCTTGGATGGTGTATGCATGTTACGCTTAACCTTGGGGGCTGGAGATCACAGCAAATGGGGA  
CCAATGATGTGTGTGTCTATAACATGATGGGGCGCCATGATGAAATCTGGGTCCACAAcTTTTGgCACCAcCCtTCCCAATAG  
CTGAATCAAATGGGCTTATCAAGGATTGTATGCTGGACCAAGCAATTGATGTCGGGAGGCAGCAGTTAAGAAATTTGCTAAAAATG  
ACTTACAAAAATGAGTGGGAAGAAACACTCGGGTGAACCCGGTGGATGAACATGGAAGTGAAGATCAGGGTGCACAGATTGACACA  
AAATGTCAAAAAGAGGGCAAGGACACAGCAGGATGGGACACACAAAAGGGCTGGAGTCTAGTCGTATGGTGCAAGCGTTGTT  
GGAAGAAAACCAATGTTAG. 3'

Figure 10: Order of the APC concatemer. The APC was designed to comprise pathogens representative of viral families/ genera. The order of the pathogens according to their designated colours is **Metapneumovirus**, **Lassa**, **Brucella spp.**, **Avulavirus/ Rubula viruses**, **Henipa/Morbillo/Respiroviruses**, **Enterovirus**, **Filoviruses**, **CCHFV**, **Malaria**, **Bwamba V**, **Rhabdoviruses**, **Hepatitis E Virus**, **Bunyamwera**, **Yellow F.V.**, **Chikungunya virus**, **Riftvalley fever**, and **Leptospira**.

The sizes (bp) of some PCR amplicons were altered, differing from the expected size upon the design of the APC, as shown in Table 12. However, this did not affect the credibility of the PCR assay in detecting the presence of pathogens.

Table 12. Table of pathogens used to design the APC in their synthetic order. The PCR amplicon size (bp) differed from the expected size.

|    | Virus / Parasite                 | Expected Size (bp) | APC size (bp) |
|----|----------------------------------|--------------------|---------------|
| 1  | Metapneumoviruses                | 500                | 498           |
| 2  | Lassa fever virus                | 392                | 405           |
| 3  | Brucella spp.                    | 150                | 149           |
| 4  | Avulavirus/ Rubulavirus          | 270                | 284           |
| 5  | Henipa/Morbillo/Respirovirus     | 490                | 501           |
| 6  | Enteroviruses                    | 150                | 150           |
| 7  | Filovirus                        | 400                | 400           |
| 8  | CCHFV                            | 488                | 472           |
| 9  | Malaria ( <i>Plasmodia</i> spp.) | 158                | 158           |
| 10 | Bwamba virus                     | 270                | 279           |
| 11 | Rhabdoviruses                    | 466                | 453           |
| 12 | Hepatitis E virus                | 336                | 327           |
| 13 | Bunyamwera virus                 | 465                | 417           |
| 14 | Yellow fever virus               | 270                | 283           |
| 15 | Chikungunya virus                | 205                | 202           |
| 16 | Rift valley fever virus          | 172                | 176           |
| 17 | Leptospira                       | 195                | 197           |

#### 4.2.1.2 Analysis of the positive control concatemer

A 1750ng positive control concatemer/ artificial positive control (APC) of 775bp was synthesised and later reconstituted in 100µL of Diethyl pyrocarbonate (DEPC) water. Upon quantification using the Whelan *et al.* 2005 method for concentration estimation of plasmids [376], the APC yielded approximately  $2.06 \times 10^{10}$  copies/ µL. The detectable number of viral copies applied to each PCR was estimated by carrying out a two-fold dilution of the positive control. This was confirmed using the Flavivirus and Alphavirus PCR assays [310], [312] ( Figure 11).

Since this research section was carried out from Uganda where logistics constrained further analysis of each pathogen positive control material, a concentration of  $10^2$  APC copies/ µL was adopted as a PCR positive control for the rest of the other pathogens. This concentration was considered adequate based on the observable PCR gel products of the Flavivirus ( $10^2$  per µL, Figure 11A) and Alphavirus (1 copy per µL, Figure 11B) APCs.

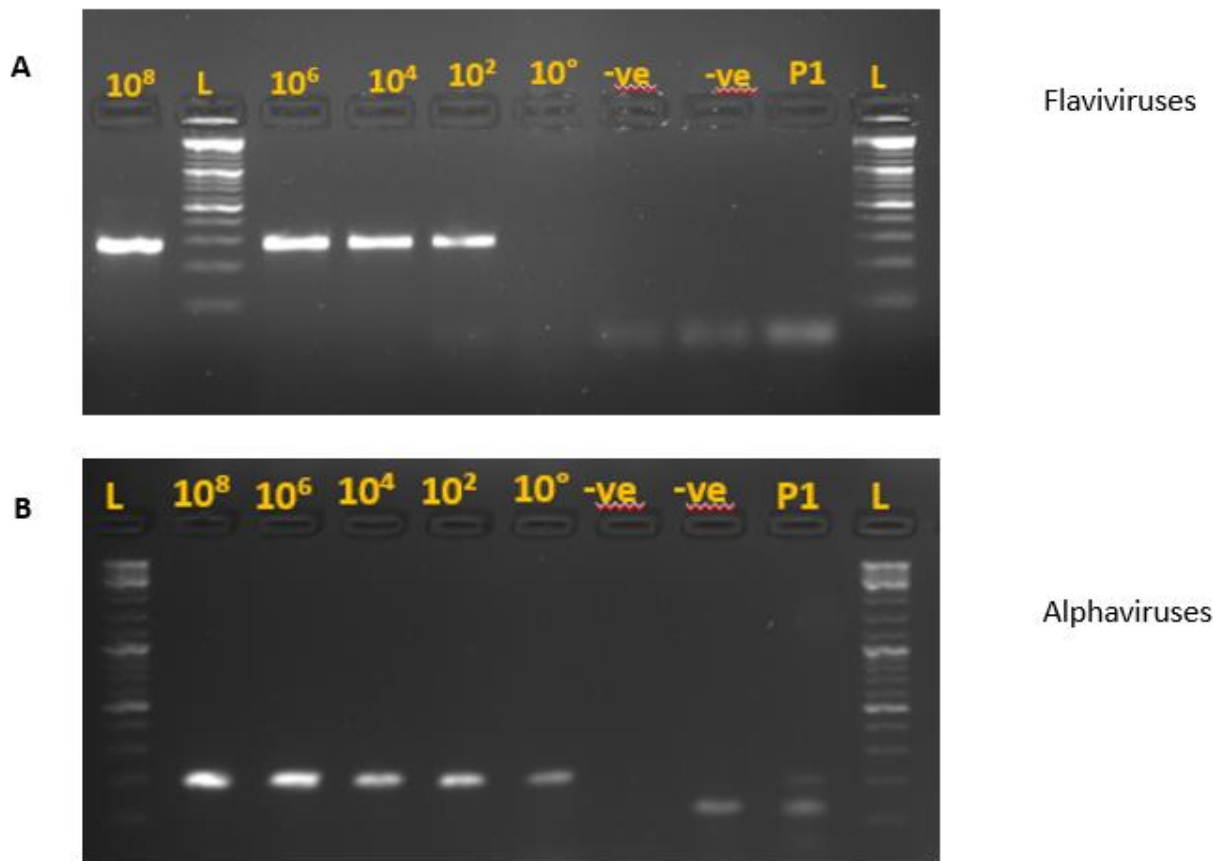


Figure 11. Gel electrophoresis of APC amplicons of PCR assay after a two-fold serial dilution. A). *Flaviviruses*: Lane 1. The PCR product for plasmid control at the concentration of  $10^8$  copies/ $\mu$ L, lanes 2 and 10 for the DNA ladder, 3- 6 for plasmid concentrations  $10^6$  - $10^2$  copies/ $\mu$ L, 7-8 DEPC water negative controls, 9 for negative serum control. The PCR product size was as expected ~270bp. B). *Alphaviruses* Lane 1 and 10 labelled as L, represent the DNA Ladder, 2-6 for alphavirus APC in descending order;  $10^8$ -  $10^0$  copies/ $\mu$ L, 7-8 DEPC water negative controls, 9 for negative serum control. The PCR product size ( ~200bp) was as expected.

#### 4.2.2 General analysis of viral and non-viral AFI.

Of the 1,441 samples discussed in chapter 3, only 137 had a laboratory diagnosis of arbovirus infection. Therefore, this implied that 1,304 samples were eligible for further screening to identify any aetiologies of AFI. Unfortunately, 214 samples had an

insufficient volume which limited further investigations. Thus, virus screening was attempted in 1,090 samples pooled by sample type ( i.e. five samples per pool of either sera or whole blood) to yield 218 pools. The aetiologies of non-viral AFI analysed in this study included *Plasmodia* spp., *Brucella* spp., and *Leptospira* spp. Of the 218 pools, 31 (14.2%) pools had viral infections, 18 (8.3%) had a *Leptospira* spp. infection, and in 29 (13.3%) pools, *Plasmodia* spp. was detected. No *Brucella* spp. infections were detected among the patient samples. Having identified samples with non-viral AFI and confirmed that these had no co-infections with viruses, no further analyses were carried out on non-viral AFI.

#### 4.2.3 Detection of Enteroviruses in Serum and whole blood.

Initial screening for enteroviruses was done using the Dierssen *et al.*[309] primers targeting the 5' UTR of the genome. Five (three from sera and two from whole blood) enterovirus PCR products of 150bp were observed on gel electrophoresis. These were confirmed as positives by the BLASTn function search of the returned Sanger sequences on the NCBI database. Serotyping the positive samples using the Nix *et al.* 2006 primers and assay was 80% successful, with sample 202 yielding no band despite three attempts (Figure 12). The four samples that yielded PCR products visible under U.V. light were identified as human Coxsackie B1 virus (CV-B1) in two samples, one Coxsackie A4 virus (CV-A4) and Human echovirus E4 (E4) upon blast searching the Sanger sequences.



Figure 12. Gel electrophoresis image of enterovirus-positive samples using the Nix et al. assay and method. *Gel electrophoresis was carried out for the second PCR product targeting the VP1 gene. The amplification size of the target DNA was ~350 -400bp due to variations in VP1 gene length across serotypes. Columns L represent the DNA ladder, Columns 202,2222, 2307, and 1700 represent sample I.D.s, column +VE represents the positive control (obtained from a known Enterovirus Coxsackie virus A4 sample), whilst column –VE represents the negative control. Sample 202 was positive for the enterovirus 5'UTR gene but negative thrice on repeat using the Nix et al.*

#### 4.2.3.1 Epidemiological and Phylogenetic analysis of the enteroviruses

The three coxsackieviruses were identified in males aged 14-38 years with signs including difficulty breathing and bleeding, whilst the E4 was identified in a 6-year-old female with no observable symptoms (Table 13). On blast analysis, the CV-A4 (ID. Entero 2313) was 86-90% identical to previously sequenced VP1 genes. Furthermore, phylogenetic analysis of the CV-A4 revealed that it clustered with previous isolates in Uganda (accession number MT661842), Figure 13. Nucleotide blast analysis of the CV-B1 serotypes showed these strains were 80-84% identical to other CV-B1 serotypes.

Echovirus E4 (ID. 1700) had 81-85% nucleotide identity compared to other echovirus E4 serotypes (Figure 13).

Table 13. Clinical signs and symptoms of serotyped enteroviruses

| <b>I.D.</b> | <b>Age<br/>(years)</b> | <b>Virus serotype<br/>(Date of sample<br/>collection)</b> | <b>Clinical signs and symptoms</b>                                                                     |
|-------------|------------------------|-----------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| 1700        | 6,<br>female           | Human echovirus E4<br>(4 <sup>th</sup> , July 2018)       | Fever, abdominal pain, and<br>headache                                                                 |
| 2222        | 14,<br>male            | Coxsackievirus B1<br>(15 <sup>th</sup> , April 2019)      | Fever, fatigue, anorexia,<br>abdominal pain, headache,<br>cough, difficulty breathing and<br>skin rash |
| 2307        | 38,<br>male            | Coxsackievirus B1<br>(18 <sup>th</sup> , July 2019)       | Fever and bleeding                                                                                     |
| 2213        | 20,<br>male            | Coxsackievirus A4<br>(19 <sup>th</sup> , July 2019)       | Vomiting, diarrhoea, fatigue, joint<br>pain, headache, cough, and<br>bleeding                          |

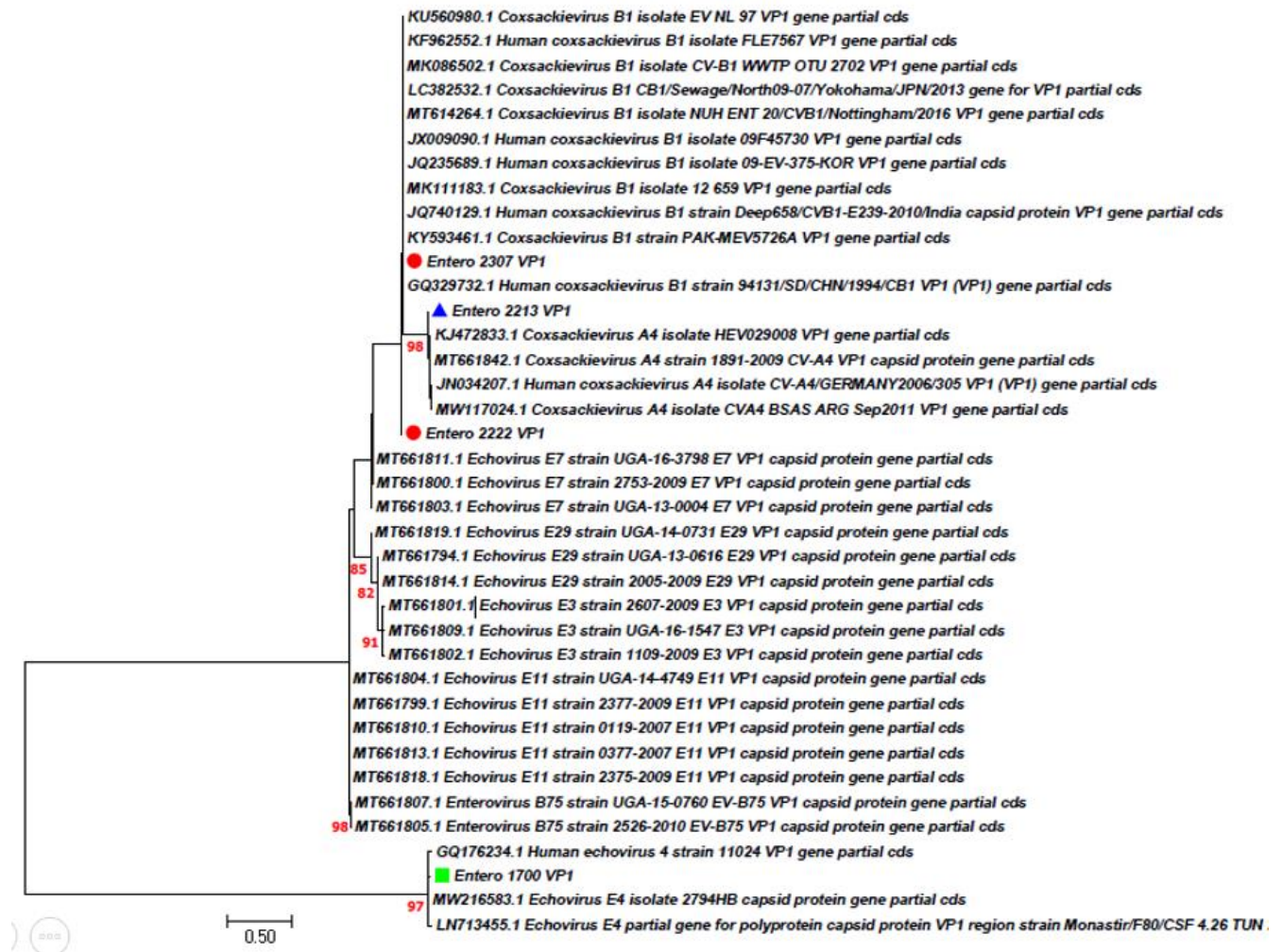


Figure 13. Phylogenetic tree showing selected enteroviruses following the amplification of the VP1 region. *Echoviruses* E7, E29, E3, E11 and *Enteroviruses* B75 are previous isolates detected by Tushabe et al., 2021. *Enteroviruses* marked with red, blue and green dots were detected in this study. The tree was inferred using the maximum likelihood method based on the JTT matrix-based model [377]. Bootstrapping was done for 1000 replications. Percentages greater than 70 are shown next to the branches for the associated serotypes. The tree is drawn to scale, and discrete Gamma distribution was used to model the evolutionary rate differences.

#### 4.2.4 Detection of Measles virus in whole blood

Initial screening of the measles virus (MeV) was done using the pan-Avula primers of the family Paramyxoviridae designed by Tong *et al.* 2008 [308]. Of the 1900 samples, only one positive was detected. A nucleotide sequence blast showed the virus was 94% identical to MeV genotype B3 strains from the United Kingdom (U.K.) and the United States of America (USA). The patient was a 47-year-old male farmer from Kagadi, in western Uganda. On triage, the patient presented with fever, vomiting, diarrhoea, fatigue, anorexia, abdominal, muscle and joint pain, headache, cough, shortness of breath, jaundice, skin rash and hiccups. Whole blood was taken from the patient on 1<sup>st</sup> August 2019. Unfortunately, the patient died days after hospitalisation.

##### *4.2.4.1 Analysis of the Measles Nucleoprotein 450 gene*

Using primers designed by Bankamp *et al.*, 2013, the MeV was amplified and the N450 gene sequenced. On phylogenetic analysis, the genotype clustered with various strains from Kenya, South Africa, Zambia, and Zimbabwe (Figure 14). It is important to note that, according to our phylogenetic analysis, previous genotype B3 strains of Uganda clustered with strains from Rwanda.

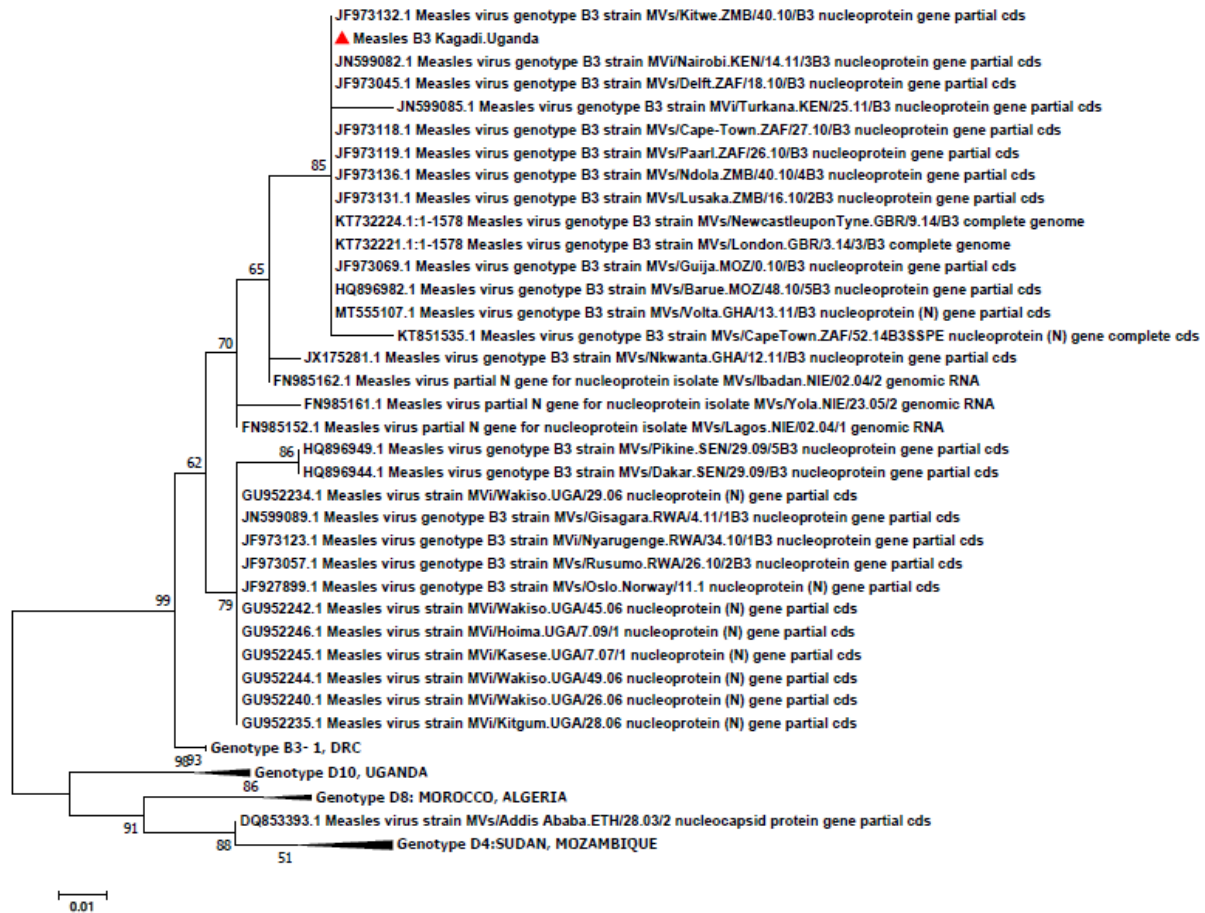


Figure 14. Phylogenetic tree of Measles viruses in Uganda and select sequences from elsewhere, following the amplification of PCR product of circa 600bp in the N450 region (genomic position 1054 – 1629, ref strain KT732221.1). The red triangle shows the strain detected in this study. Measles virus genotypes D4, D8 and D10 are compressed into subtrees. The tree was inferred using the maximum likelihood method based on the JTT matrix-based model [377]. Bootstrapping was done for 1000 replications. The percentages of trees in which the associated genotypes clustered together are shown next to the branches. The tree is drawn to scale, and discrete Gamma distribution was used to model the evolutionary rate differences.

#### 4.2.4.2 Analysis of the Measles Hemagglutinin gene

Primers were designed using the Primal Scheme tool as described in section 2.2.2.2. Upon visualisation under U.V. light, amplicons of approximately 600bp were observed (Figure 15).

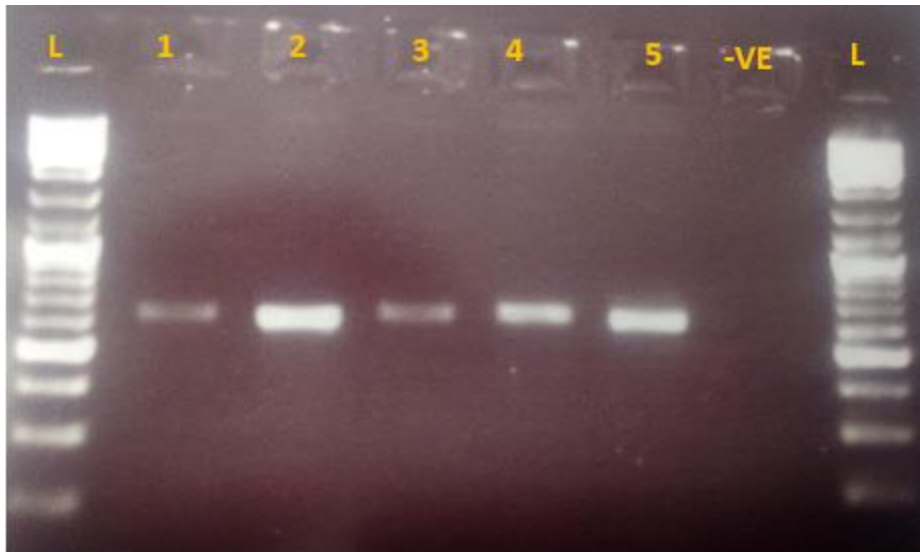


Figure 15 Visualisation of the H gene's PCR amplicons (approximately 600bp) after gel electrophoresis. Lanes L and –VE represent the DNA ladders and the negative, respectively, whilst lanes 1-5 show PCR amplicons of the H gene. The PCR product was approximately 600bp and covered the H gene from genome position circa 8854-9384, of reference strain accession number KT732214.1. The PCR conditions are described in 2.2.6.3.

The H gene sequences were aligned to form a single contig sequence of 1854bp, using MEGA v7.0.26. It is worth noting that the final length of the H gene is much shorter (1854bp) than the observed (3000bp) because during sequence generation overlapping primers were applied and the final product was trimmed of its redundant fragments. The recovered sequence was then BLASTn searched against nucleotide sequences in the

NCBI database. This analysis revealed a 99% nucleotide identity match to MeV genotype B3 from the U.K., Italy, Taiwan, South Korea, the USA and South Africa. Based on phylogenetic analysis, the MeV clustered with strains from Kenya, South Africa, Italy, and the U.K. (Figure 16).

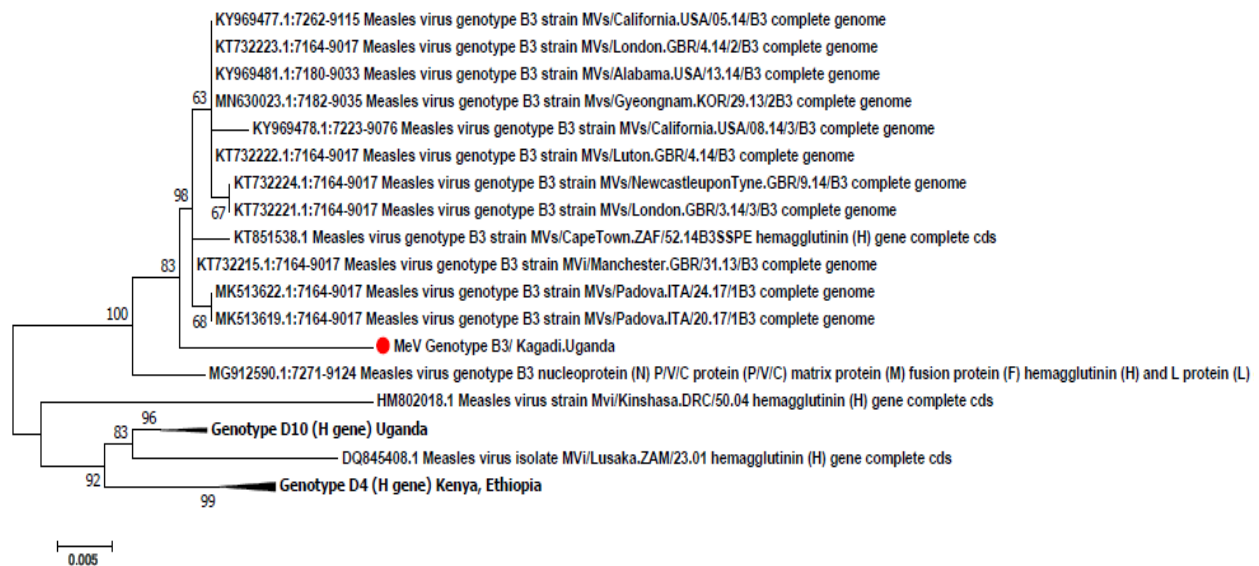


Figure 16 Phylogenetic tree for MV B3 genotype after amplification of the H- gene. The red dot shows the strain sequenced during this study. The tree was inferred using the maximum likelihood method based on the JTT matrix-based model [377]. Bootstrapping was done for 1000 replications. The percentages of trees in which the associated genotypes clustered together are shown next to the branches. The discrete Gamma distribution was used to model the evolutionary rate differences. Genotypes D4 from Kenya and D10 from Uganda are collapsed as shown above.

#### 4.2.5 Screening for Arenaviridae, Filoviridae, Hepeviridae, Nairoviridae, Peribunyaviridae, Phenuiviridae, Pneumovirinae, Togaviridae, and Rhabdoviridae.

No viruses from the above families/ subfamilies were detected in this study. Primers targeting the Bwamba Orthobunyavirus and Nairoviruses (CCHFV) were excluded from

further analysis since assay optimisations were not carried out. On the other hand, no appropriate specimens were available to carry out the detection of Rhabdoviruses (Lyssavirus) [378], [379]; thus, these were also excluded from further investigations.

#### 4.2.6 Screening for Flaviridae

Investigation of flaviruses by RT-PCR is fully described in chapter five.

#### 4.2.7 Quality control of samples shipped from Uganda to the United Kingdom

All PCR products that showed a presumptive positive band upon visualisation under U.V. light were shipped to the U.K for confirmation by Sanger sequencing.

All samples analysed in this study had an average RNA purity of approximately 1.9 at 260/280 absorbance and an average RNA concentration of 10ng/μL. However, this study did not analyse the gradual degradation of nucleic acid between the site of sample origin (UVRI) and transportation to the final destination (UoN) since no records of concentrations were available from the site of origin.

## 4.3 Discussion

We describe the detection of viral AFI in samples that received a negative test result for arbovirus Immunoglobulin M (IgM). The results described in this chapter were acquired by applying the conventional PCR method to 1090 serum and whole blood samples.

This chapter provides evidence of missed viral infections during routine diagnosis of pathogens associated with acute febrile illnesses. Over the years, research involving pathogens not routinely diagnosed has been made possible by the ability to design artificial positive controls (APC) [380]. This chapter detected enteroviruses, the measles virus genotype B3, *Plasmodium*, and *Leptospira* spp. by applying the APC construct as a PCR positive control. Thus, non-endemic pathogens [381], epidemics, internationally restricted pathogen material due to bioterrorism concerns, rarely detected pathogens and the risk of inconclusive results due to contaminated commercial kits [382]; can all be authentically diagnosed using APC. However, the fact that the APC did not undergo an equivalent sample processing procedure such as extraction and reverse transcription makes it less precise in analysing the efficiency of the PCR. Therefore, to improve accuracy of these quantification experiments, RNA could be transcribed from the control plasmid, spiked into control biological specimens, and then extracted. This would significantly improve the understanding of the sensitivity of these RT-PCR reactions.

It is no surprise to detect malaria among the AFI samples. For example, Kigozi *et al.*, 2020 reported that over 38 million cases were registered in the national health management information system (HMIS) data, with a monthly incidence rate of 20.4

cases per 1000 [383]. However, due to the nature of this study, we limited research to viral AFI; therefore, we could not determine the type species of Plasmodia or Leptospira involved in the non-viral infections.

This chapter identified enterovirus serotypes previously unscreened in Uganda, coxsackievirus B1 and echovirus E4. Therefore, this research added more data on the circulating enterovirus serotypes in Uganda, supplementing work done by Tushabe *et al.* 2021[155]. Despite a safe and efficient vaccine, measles virus strains, including genotype B3 identified in this study, are still endemic in the SSA region and other parts [47], [384]. Whilst previous MeV genotype B3 clustered with isolates from Rwanda, which neighbours Uganda in the southwest, current strains clustered with isolates from Kenya, neighbouring Uganda to the East. The MeV genomic regions applied in this study to detect MeV were the N450 and the hemagglutinin (H). Although a combination of these proteins helps detect new genotypes [385], these regions are not sufficient to determine the importation or exportation of the virus in geographical locations. Thus, this study also highlights the need to establish a variable genomic region that can be applied to monitor the epidemiology of MeV. A suggested region is the Matrix and Fusion protein genes, which are thought to provide more information on the source and thus virus elimination [386]. The measles virus identified in this study was from a 47-year-old male individual who later died in hospital. The sample was collected on 1<sup>st</sup> August 2019, amid a widespread MeV epidemic in Uganda, which probably lasted for over five years (2015- 2019) [387],[188, 388], [187],[389]. It is unknown whether the individual was suffering from any other immunocompromising disease that could have synergised the effect of MeV. However, available data shows that life-long immunity is

guaranteed once an individual recovers from MeV [390]. Infection and succumbing to the MeV by a 47-year-old individual may only point to missed childhood vaccinations and MeV- induced increased susceptibility to infection with bacteria or viruses [391]. However, the failure to clear the Measles virus in immunocompromised patients is associated with lung disease [392], which could have manifested as the shortness of breath observed in this patient.

This study did not detect any viruses within the arenaviridae, filoviridae, hepeviridae, peribunyaviridae, phenuiviridae, and togaviridae viral families. Data on the prevalence of arenaviruses in the region are scarce. However, an ongoing study by De Feo *et al.* assessing arenaviruses in odontogenic tumours may provide an update [393].

Surprisingly, seroprevalence for other viruses is high, with 31-38% for Togaviridae (CHIKV and ONNV) [344], [345]; 4-8.9% for filoviridae (Ebola virus) in southwestern Uganda [394] and 47-64% for Hepeviridae (Hepatitis virus E) [395], [396]. However, it is important to note that the data on Ebola and Hepatitis E viruses were collected during epidemics and, therefore, may not represent an accurate picture of the disease epidemiology in the country. Peribunyaviridae is thought to cause asymptomatic; thus, a 66-81% IgG seroprevalence in northern Uganda would be expected in outbreaks [344].

Nucleic acids were detectable from samples after a shipment at ambient temperature between Uganda and the U.K under the DNA/RNA shield. The recovery of nucleic acids from samples protected by the DNA/RNA shield at room temperature after more than three days in transit is a game-changer for resource-limited countries. Although, the lack of matching comparison data for the same samples before shipping from Uganda suggests that care should be taken whilst interpreting the DNA/RNA shield potential.

Furthermore, it eliminates the biosafety hazards associated with liquid nitrogen and dry ice on flights [397].

## 4.4 Conclusion

This study supplements available data on the enterovirus serotypes prevalent in Uganda. Detecting viral AFI by applying the conventional PCR method is achievable and is still a viable method, especially in resource-constrained settings. In addition, artificial positive control constructs are reliable and efficient as positive controls in the absence of pathogen-related material.

## 5.0 Detecting West Nile Virus after 80 years of silent transmission in Uganda

### 5.1 Introduction

The West Nile Virus (WNV), an arbovirus belonging to the family Flaviviridae and genus flavivirus, causes the WNV disease. WNV was first detected in 1937 in a patient presenting with febrile illness in the West-Nile region of Uganda [398]. The WNV is phylogenetically divided into two major lineages. However, some literature suggests lineages 3 and 4 isolated from the Czech Republic and Russia, respectively [399], [400], lineage 6 from Spain [401], lineages 7 and 8 from Senegal [402] and lineage 5 (or 1c) from India [403]. Lineages 1, 3, 4 and 5, together with a putative non-pathogenic subtype known as the Kunjin virus, have been reported in Africa, Asia, Europe, North America, and Australia [404],[403],[400]. Lineages 3-8 are considered as sub-lineages resulting from separate WNV introductions to various geographical regions [405]. Lineage 2 was initially thought to be exclusive to the sub-Saharan African (SSA) region of East Africa [404], [406]; however, several studies have confirmed its presence in Europe since 2004 [407], [408], [409]. Experimental inoculation indicated that mosquito species from the *Aedes* and *Culex* genera are vectors for WNV transmission [410], [411], [412]. Birds then act as natural reservoirs for the virus [413], with sporadic spillovers in equids [414], [415] and humans acting as dead-end hosts. Disease associated with the WNV is generally mild and asymptomatic. However, with several disease outbreaks reported in Europe,

Asia, and the Americas, the virus has established itself as a public and a farm animal threat [416], [417], [418]. In severe or fatal cases, infected individuals may present with neurological pathologies [414], which include encephalitis, meningitis, and acute flaccid paralysis [419], [420], [421].

Except for the 1937 WNV genome, there is no other recorded genomic sequence of WNV isolated from humans in Uganda. However, WNV lineage 2 has been isolated from the *Culex neavei* species collected from the Mweya Queen Elizabeth National park, a game reserve located in Western Uganda [422]. In addition, serological studies indicating acute viral infection have been recorded [344]. Nonetheless, these studies provide no information on the genetic characteristics of the virus. Despite continuous surveillance activities of arboviruses in Uganda, molecular detection of WNV using contemporary and commonly affordable molecular tests such as polymerase chain reaction (PCR) has remained problematic. Furthermore, despite the advent of alternative molecular diagnostic methods such as next-generation sequencing (NGS), developing countries still need reliable and sustainable diagnostic methods. In this study, pan- flavivirus genus primers, as described by Moureau *et al.* [310], were used with the intent to capture known and unknown flaviviruses. Conventional reverse transcriptase (RT) PCR and Sanger sequencing were carried out to analyse the feasibility of applying such methods in routine molecular diagnostics in a resource-limited setting. Detecting viral nucleic acids using conventional rtPCR and genotypic confirmation by Sanger sequencing remains unrivalled and could also translate efficiently to third-generation portable sequencing platforms such as the minION.

## 5.2 Results

### 5.2.1 PCR, Gel electrophoresis and Sanger sequencing of Flaviviruses

Screening for flaviviruses was carried out retrospectively using primers as described by Moureau *et al.* [310] to detect flaviviruses from the synthesised cDNA samples at the genus level. PCR products were visualised using ethidium bromide staining under UV-light. One sample (1633) produced PCR products of the expected size (approximately 270bp), and these products were sent for Sanger sequencing using the forward PCR primer. BLASTn analysis of the resulting sequence showed 97.54% nucleotide identity to a 2009 mosquito strain (KY523178) of West Nile Virus, corresponding to genome nucleotide positions 9188 to 9309.

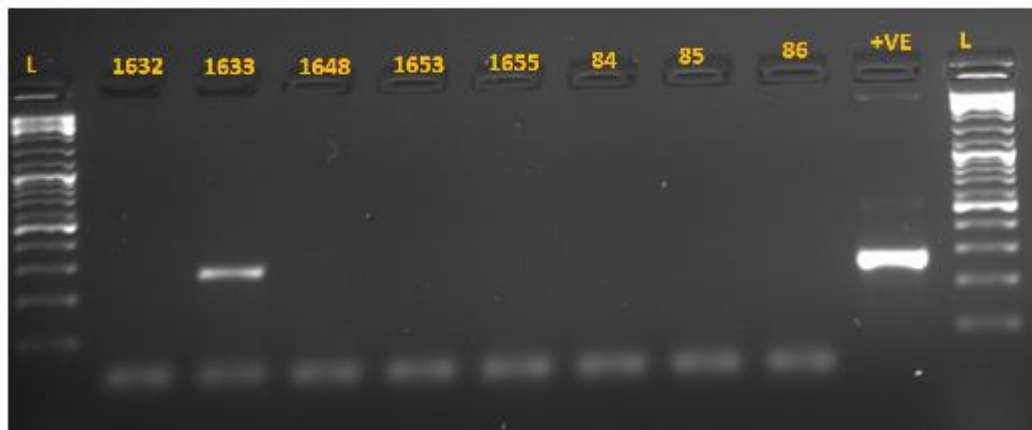


Figure 17: Visualisation of flavivirus: A band of the expected size (270bp) was observed in sample 1633 after PCR using pan-flavivirus primers adopted from Moureau *et al.* The samples identified as (1632, 1633, 1648, 1653, 1655, 84, 86) were loaded in between two DNA ladders (L) expressing between 100bp to 1kb. The positive control was the positive control concatemer.

## 5.2.2 Analysis of West Nile virus primers

The WNV genome accession number KY523178.1, isolated from mosquitoes in the western region of Uganda, was the most similar to the sequence detected in this study [422]. Therefore, the genomic sequence of KY523178.1 was downloaded from GenBank and used to design WNV primers. Primers were designed by uploading the KY523178.1 genome to the Primal Scheme tool [319] to design primers that generated (overlapping) amplicons of approximately 800bp. Of the 16 sets of primers generated, only seven (43%) produced a product visible on gel electrophoresis. We assumed nucleic acid degradation, given that the source of RNA was an archived clinical specimen collected three years ago and stored under sub-optimal conditions (recurrent power cuts). To recover the rest of the genome not recovered by the 800bp fragments, primer sets generating shorter overlapping fragments were subsequently tested. Of the 13 sets of primers generating the 400bp fragments, 10 (76.9%) produced a product visible under UV light and confirmed by Sanger sequencing as WNV positive. Of the six sets of primers designed to generate 200bp fragments, only 3 (50%) produced a product confirmed by Sanger sequencing as WNV. Four gaps equating to 550 nucleotides of unrecovered sequences were identified (Figure 18). Thus, the overall efficiency of the Primal Scheme tool in full genome recovery was 94.7%. However, due to the limited amount of RNA, the 5' and 3' ends of the virus were not recovered.

### 5.2.3 Analysis of West Nile Virus genome recovery

The coding sequences (CDS) of the recovered WNV were aligned in MEGA7 software, and a near-complete WNV genome (Figure 18B) of 9.8kb was recovered. Compared to the WNV strain KY523178 (Figure 18A), two gaps of 19bp (Capsid region) and 658bp (NS5 region) in the recovered genome were observed on the 5' and the 3' end, respectively. The Non-structural protein 3 (NS3) region contained two gaps, one with 39 missing base pairs between positions 5455 – 5494. The second gap of 275 bp was observed between positions 6083 – 6358. The third and fourth gaps of 84 and 107 base pairs were observed in the NS5 region between positions 8638 – 8722 and 8822 – 8929, respectively. The recovered near-complete WNV genome was further compared to the ancestor strain M12294 (Figure 18C) and was found to have acquired an insertion of 5 amino acids (NYSSQ) from position 154 to 158 in the envelope (E) protein.

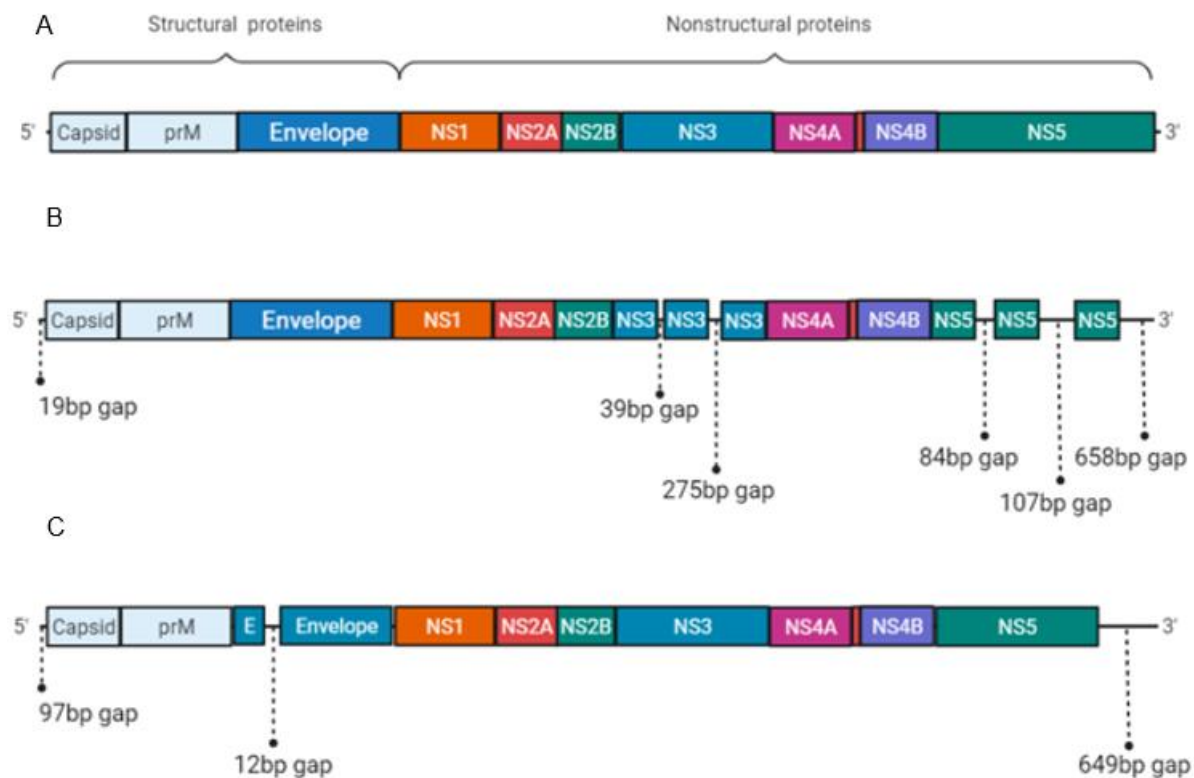


Figure 18. Cartoon of WNV genome: A. Complete cartoon of the WNV genome showing the structural and non-structural proteins (NS). B. Cartoon of the near-complete WNV virus recovered in this study, showing gaps at the 5' and 3' end, NS3 and NS5 regions. C. Cartoon of the WNV strain accession number M12294.2, showing a deletion within the envelope gene. Illustration not drawn to scale. The cartoons were constructed using Bio Render online app.

## 5.2.4 Phylogenetic analysis of WNV

Having recovered a near-complete WNV genome, we sought to determine its phylogeny. First, we downloaded 1,930 complete WNV genomes from GenBank as of 5<sup>th</sup> June 2021 and analysed them together with the WNV recovered in this study (now referred to as West Nile virus Kasese) using the MUSCLE function of MEGA 7. Phylogenetic analysis

showed that the WNV Kasese clustered with Lineage 2 viruses. However, it was observed that despite clustering within lineage 2 viruses, the recovered WNV Kasese did not cluster with the ancestral 1937 WNV isolates (AY532665 and M12294), as shown in Figure 19. The WNV Kasese clustered with virus strains, including accession number KY523178.1 (isolated from mosquitoes in Uganda) [422], LC318700 (isolated from mosquitoes in Zambia) [423] and JN393308.1 (isolated from a horse's brain with fatal encephalitis in South Africa). On the other hand, the 1937 WNV isolate from Uganda (M12294.2 and AY532665.1) clustered with DQ318019 from Senegal and JX041631 collected from Ukraine in 1980.

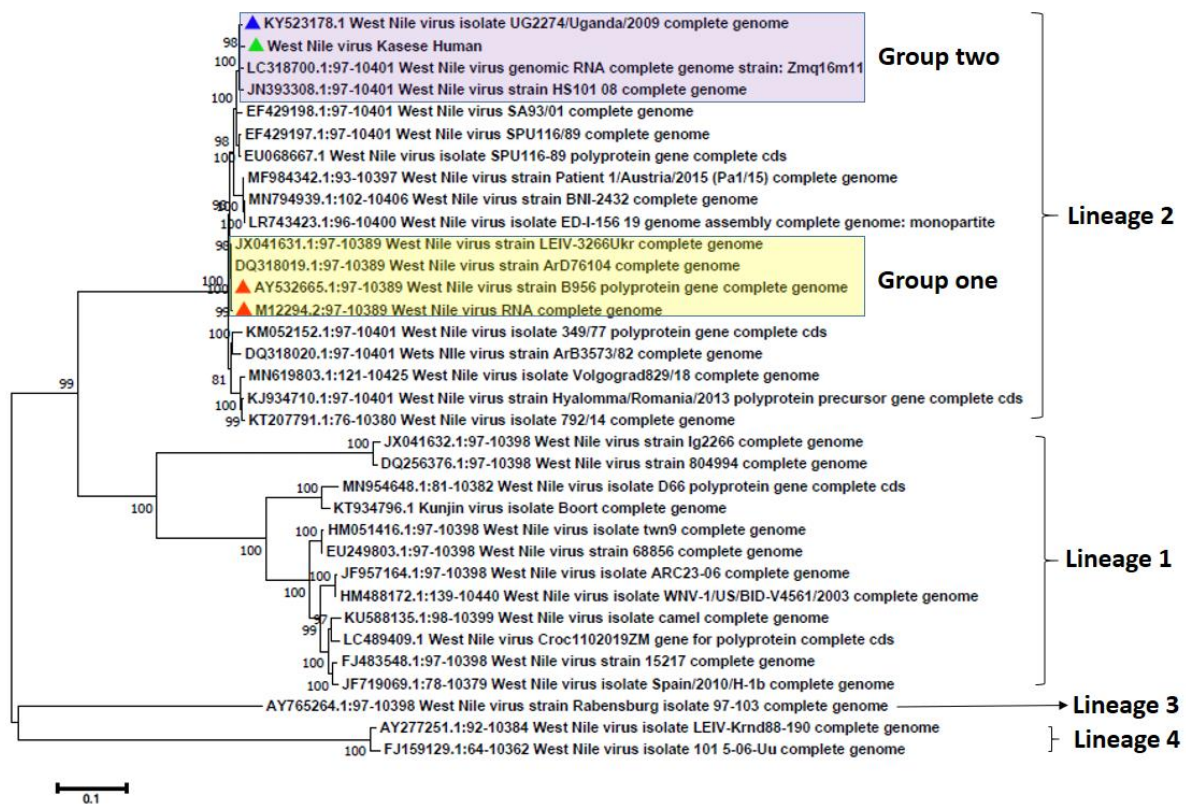


Figure 19: Molecular phylogenetic analysis of WNV using the maximum-likelihood method with GTR model. Analysis was done for 34 WNV nucleotide sequences selected to represent lineages 1- 4 without gaps. Bootstrapping based on 1000 replications is shown next to each branch at a scale of 0.1. The recovered WNV was named after the district of sample collection as West Nile virus Kasese, shown with a green triangle. Red triangles represent the ancestral WNV strains, whilst KY523178 from mosquitoes is represented by a blue triangle. The two groups highlighted in light purple and orange show the distinct clustering of WNV Kasese to M12294 within lineage 2.

## 5.2.5 Analysis of the genetic diversity of WNV

The first ever isolated WNV strain was the 1937 WNV from Uganda. This isolate underwent several cell culture passages, with the earliest sequences available identified as accession numbers M12294 and AY532665. Due to several passages during cell culture, the virus is thought to have picked up mutations that would affect its genetic and phenotypic characteristics. However, the WNV isolated from a bird's sample (accession number JX041631.1) collected in 1980 was similar to sequences M12294 and AY532665, implying few cell culture-related mutations. We, therefore, sought to compare the genetic diversity of the WNV Kasese to its ancestor, the 1937 WNV (M12294.2 and AY532665), by taking advantage of the two cluster groups observed within lineage 2 (Figure 19). Group one comprised the known ancestral WNV strains M12294.2 and AY532665.1 (WNV 1937), DQ318019.1 from Senegal, and JX041631.1 collected in 1980 from a bird in Ukraine. The second group (WNV Kasese) comprised WNV Kasese, JN393308.1 (horse isolate), KY523178.1, and LC318700.1. (Mosquito isolates). These two groups'

complete coding sequences (CDS) were then uploaded to MEGA v7, where they were translated into protein sequences. The protein sequences were uploaded and analysed for amino acid distance using the Simple Sequence Editor (SSE) v1.4 [324].

Over 26 mutations were observed within the two groups (Figure 20). Of the 26, at least eight were within the WNV Kasese group (denoted as red-green spikes), 17 were within the WNV 1937 group (denoted as blue-green), and five were distinctly detected between the WNV Kasese versus the WNV 1937 group (green spikes). The envelope (E) protein accounted for two of the five distinct mutations. First, the WNV Kasese group contained an insertion of 5 aa (NYSTQ) between positions 154-158, which was analysed as a single block deletion in the WNV 1937 group. The second mutation was a change from Arginine to Glycine at position 603 (R603G) in the WNV Kasese group. The other three distinct mutations included a substitution of Arginine for Lysine in the NS2A protein at position 1241, Methionine for Threonine at position 2515 in the NS4B protein and valine to Isoleucine at position 3061 in the NS5 protein.

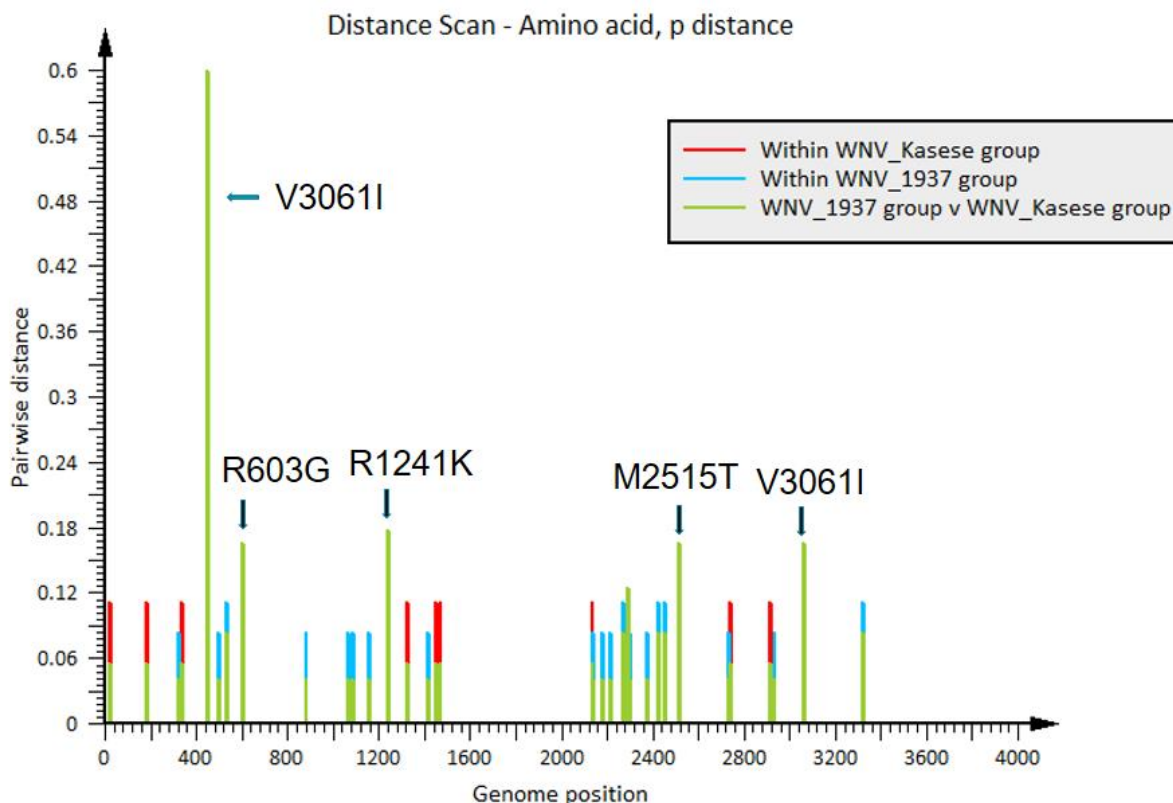


Figure 20. Sequence distance comparison between WNV Kasese versus WNV1937 group. The green lines represent amino acid (aa) substitution between WNV 1937 versus WNV Kasese group, the blue lines represent aa substitution within WNV 1937 group, while the red lines represent aa substitution within the WNV Kasese group.

#### 5.2.5.1 Analysis of the Pre- membrane and Envelope proteins for glycosylation sites

The pre-membrane (Pre-M) and Envelope (E) protein sequences of the WNV Kasese group were compared to the WNV 1937 group using a web-based tool that predicts the N-glycosylation of an amino acid in a protein residue [325]. A common N-linked glycosylation site was observed at position 138 (NAT) in both WNV Kasese and WNV

1937 groups (Figure 21). In addition, an additional N-glycosylation site was observed in the WNV Kasese group at position 444 (NYS).

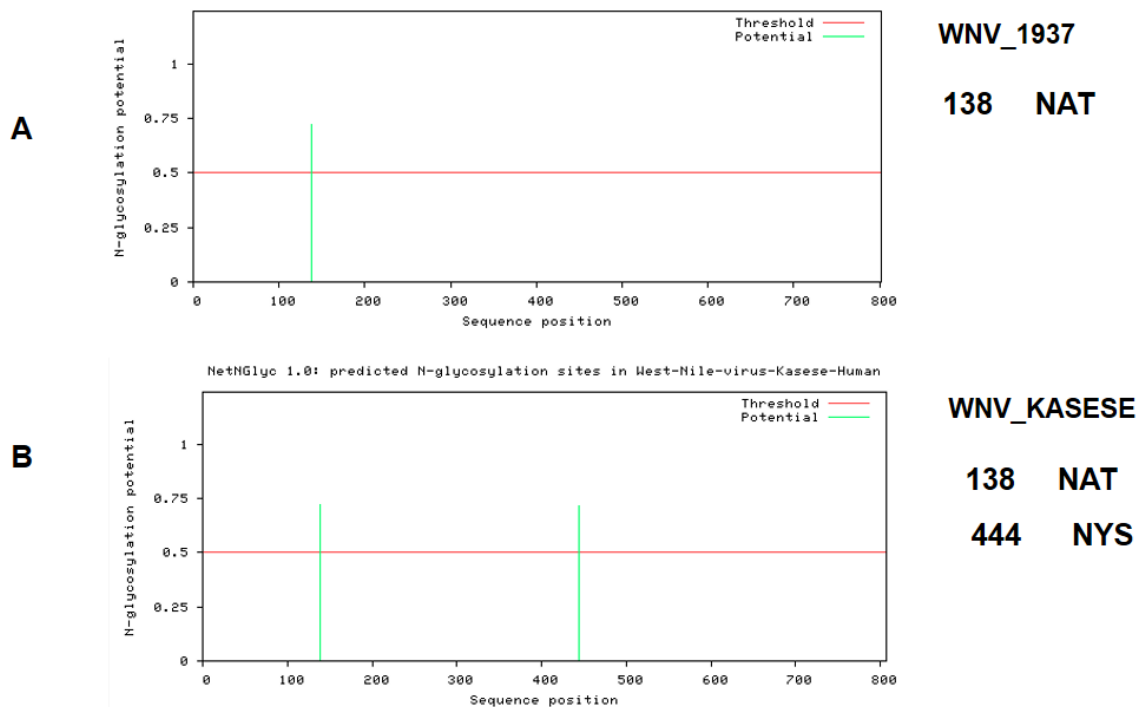


Figure 21. N- glycosylation of PrM and Envelope proteins: A. *Potential N-glycosylation site observed at position 138 for aa NAT in the WNV 1937 group. In this figure, the group is presented by WNV strain M12294.1.* B. *Two potential N-glycosylation sites at position 138 (aa NAT) and 444(aa NYS) were observed in the WNV kasese group, represented by KY523178.1. The red line represents the threshold at which a potential N-glycosylation is detected. . Potential N-linked glycosylation was observed if the threshold value was above 0.5.*

We further compared the WNV Kasese group to the lineage 1 viruses by applying isolate accession number HM488172 to the N-glycosylation tool (Figure 22). Potential N-linked glycosylation sites similar to the WNV Kasese group were observed.

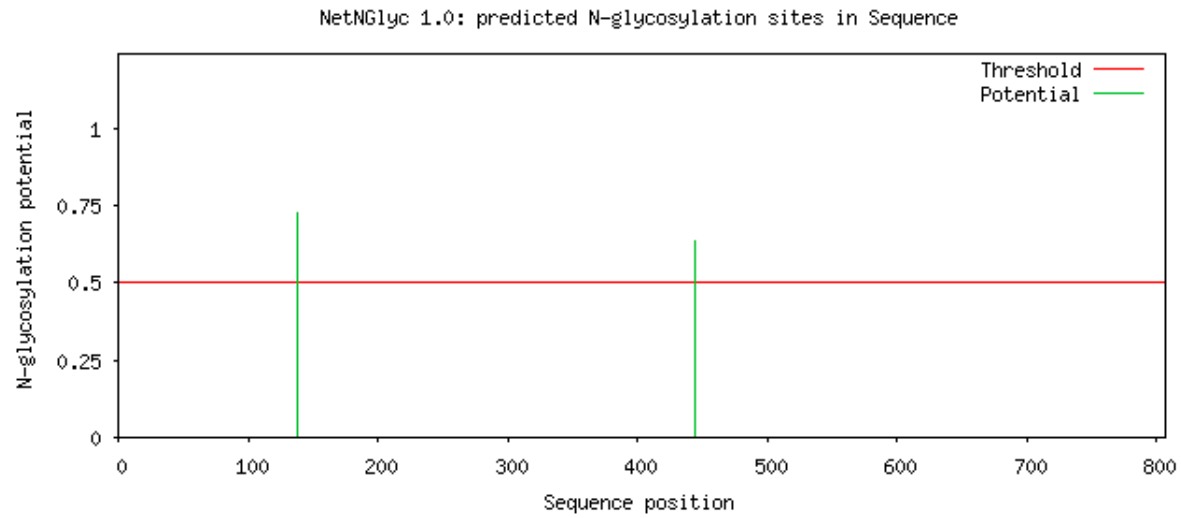


Figure 22. N-linked glycosylation of the lineage 1 virus. *Accession number HM488172 showed similar potential glycosylation sites to the WNV Kasese group. Potential N-linked glycosylation was observed if the threshold value was above 0.5.*

## 5.3 Discussion

This chapter describes the successful detection and recovery of WNV and draws a genetic comparison to its ancestor, the 1937 WNV from Uganda. Molecular detection of WNV RNA in Uganda has not been attained for over 80 years since its first detection in 1937. We report the successful detection of WNV nucleic acids by conventional RT-PCR using primers designed by Moureau *et al.* and genome recovery using the Primal Scheme tool and confirmation by Sanger sequencing. These methods are feasible for viral surveillance in a resource-limited setting. Genome recovery of 89% was probably the result of factors such as primer mismatch at the 3' end [424], [425],[426] and high RNA degradation. However, these factors were mitigated by designing primers generating shorter fragments of circa 200bp, diligent primer analysis using MEGA 7 and online applications such as Oligo Calc (<http://biotools.nubic.northwestern.edu/OligoCalc.html>) and Primer 3.

Demographic data showed that the sample was collected in 2018 from the western region of Uganda. The patient was a 20-year-old farmer from western Uganda, Kasese district. Clinical history showed that he was hospitalised on the 22<sup>nd</sup> of June 2018 with fever, an axillary temperature of 38.4°C, vomiting, nausea, diarrhoea, intense fatigue, headache, skin rash, abdominal, muscle, joint and chest pain. Epidemiological data indicated that he had a history of direct contact with apparently healthy pigs, chickens, goats, and rabbits.

Most of the serological tests done elsewhere reported previous exposure to WNV in the population by IgG antibody presence [427]. Where IgM antibodies have been reported, no WNV RNA was recovered [428]. Serological testing of the patient's serum was

negative for WNV IgM antibodies. The negative IgM test is probably attributed to the time lag between virus infection dissemination and the production of IgM antibodies by the adaptive immune system. For example, a report by Diamond et al., 2003, showed that WNV RNA was detected in mice sera on day two post-infection whilst the IgM antibodies were observed on day 4 post-infection [429].

One of the plausible sources of WNV infection in the patient may have been from his farm animals, as WNV seropositivity and RNA in goats and chicken has been described elsewhere [430], [431]. However, the second more likely source is from mosquitoes since strong evidence shows the presence of WNV in *Cx neavei* in Uganda, as previously described by Mossel *et al.* [422]. Moreover, mosquitoes from the genus *Culex* were described as the second most abundant species in two Ugandan districts, Arua and Kasese, the latter with higher mosquito abundance [432].

The isolated WNV strain phylogenetically grouped with other strains of lineage 2 as expected. Based on the phylogenetic tree, the WNV genome recovered from the patient showed minimal divergence compared to its ancestor. However, it still showed a more remarkable similarity to the WNV genome isolated from the horse's brain in South Africa and mosquitoes in Uganda and Zambia. The fact that the patient was hospitalised proves that lineage 2 strains are capable of a dramatic prognosis. For instance, in 2010, Greece's WNV lineage 2 outbreak was associated with 197 neuroinvasive cases, claiming 33 lives [433].

Differing from its ancestor, the WNV Kasese group showed a new N-linked glycosylation site in the envelope protein. This glycosylation site was similar to what is present in the lineage 1 virus strains. Research shows that the N-linked glycosylation of the E protein

influences particle assembly and infectivity [434], impacts strain neuroinvasiveness [435], increases viral multiplication [436] and transmission [437]. Several amino acid substitutions were observed between the WNV Kasese group versus the WNV 1937 group. The observation of these non-synonymous mutations could probably contribute to the increased virulence of the lineage as observed in Greece and the horses in South Africa. In addition, several studies have shown the effects of mutations (such as virus attenuation and virulence) in the lineage one strains, using the WNV isolated from New York in 1999 [438], [439]. This study did not attempt to analyse the differential codon usage among synonymous mutations within the different WNV strains (i.e. WNV\_Kasese, KY523178.1 and M12294). Differential codon usage analysis would provide more insight into codon bias among the WNV strains analysed in this study [440]. This would further be used to provide information on gene and protein expression [441], [442].

In addition, the 12 nucleotides (or four amino acids) deletion observed in the ancestral WNV (1937 WNV group) is not a surprise as it has been reported by Berthet *et al.* 1997 [443]. However, the novelty is the acquisition of the 12 nucleotides associated with the NYST glycosylation in WNV lineage one strains and other members of the Flavivirus genus. All these mutations, deletions and insertions occur within the ORF of WNV as notable that WNV has one ORF encoding a single polyprotein. Another method that can be used to confirm the glycosylation site at the NYST motif is Western Blotting as previously applied by Adams *et al.*, 1995 [444].

It is also unclear why a low virus activity (1 case in 1090 samples) was observed, given this study's moderately large sample size. In South Africa, low virus activity in the human population was recorded at 5-15 cases per year [445].

## 5.4 Conclusion

Together with work done by Mossel et al., this study provides evidence of existing and emerging viral infections in Uganda. More resources to support rigorous laboratory-based diagnostics are needed to uncover Uganda's underlying WNV and other viral burdens associated with acute febrile illnesses.

## **6.0 Screening samples for viruses using metagenomics Next-Generation Sequencing**

### **6.1 Introduction.**

Next-Generation Sequencing (NGS) or High Throughput Sequencing (HTS) can be used for the unbiased detection of pathogens in a specimen without the need for specific PCR- or probe-based enrichment [446]. Metagenomics NGS (mNGS) involves sequencing samples (such as environmental, human, and other mammalian samples) that contain various microbes. The applications of NGS are numerous, ranging from forensic science [447] to clinical metagenomics [281], patient-personalised treatment of cancers [448], ecological and population genetic conservation [449], [450], among others. This chapter focuses on applying NGS in virology to detect viral populations missed during routine diagnosis and explores its potential for novel virus discovery.

Because NGS generates large volumes of data, the earliest challenges were organising, analysing, and meaningful data interpretation. These data include raw DNA sequences, genomes, protein sequences, macromolecular structures, and gene expression, to mention but a few. Bioinformatics, therefore, emerged to provide storage and computational solutions to this dilemma and thus cuts across several data sources in biology. The storage challenge has been solved using resources such as the National Centre for Biotechnology Information (NCBI), which comprises several databases that organise and store generated data. In addition, many institutional-based resources (operated by private or institutional servers) exist.

The raw DNA sequence reads generated by NGS platforms (described in chapter one) are available as FASTQ files for Illumina platforms or unmapped binary alignment map (uBAM) files for Ion Torrent. Processing these files can be done by software designed to execute related programmes following specified command lines. The analysis depends on factors such as the bioinformatics expertise of the technician, competency in the command line tools, and the availability of storage and memory on the computer devices used. These factors, in turn, determine the choice of analytical platform applied to the data, either web-based (cloud-based) platforms such as the IDseq [326] and the Genome detective [451] or the manually installed software such as Genious Prime, CLC Bio workbench [452], among others.

Whether using the cloud-based or server-supported platforms, the analysis pipelines are designed for (meta) genomic assembly, host sequence filtering, gene prediction, data visualisation and quality control [326], [453]. Metagenomic assembly can be made by either mapping sequences against a reference genome or de novo assembly. De novo assembly tools utilise the overlap layout consensus (OLC) or the de Bruijn graph methods where all reads are compared pairwise to find overlapping regions (as seen in the OLC method) or reading the consecutive kmers (nucleotide sequence of a certain length) for each read (de Bruijn method) [453], [454]. In addition, some platforms such as IDseq have incorporated tools that phylogenetically analyse the outputted data. These tools are further supported by applications such as the Molecular Evolutionary Genetics Analysis (MEGA) tool [323] can easily be installed on computers for phylogenetic analysis of detected sequences. Thus, these tools have simplified sequence analysis and data interpretation.

Previous reports have shown that cloud-based bioinformatics platforms have a short turnaround time in analysing FASTQ files [455], [451] compared to other methods [456]. Furthermore, these platforms can be utilised by researchers and clinicians with minimal bioinformatics expertise or budget/infrastructure-constrained laboratories. The genome detective website offers a free service for non-routine use and evaluation. However, this service is limited to approximately 11GB of input data. Higher volumes of data are processed at a cost upon subscription. Unlike the genome detective website, the IDseq website is free to use upon creating an account. The downside of the IDseq website is publicising output data (only pathogen-related data after host filtration) to all account holders. Furthermore, its pipeline only supports Illumina metagenomic data, with the exception of nanopore data from the unprecedented SAR-CoV-2 pandemic. This chapter focuses on the generation of libraries from RNA collected from undiagnosed clinical samples suiting within the AFI definition. We applied mNGS to sequence the RNA libraries and uploaded the generated FASTQ files onto the IDseq platform to inform disease diagnosis in clinical samples.

## 6.2 Results

### 6.2.1 Sample selection

A total of 185 samples were selected based on the quality control procedure described in section 6.2.2 and processed for library preparation. Samples were pooled as indicated in Table 14. Briefly, pools 202 - 619 consisted of 10 samples per pool and were named after the last three digits of the first sample in the pool for easy identification. Pools J7 to J10 comprised 12 – 20 samples. No aetiologic agent had been identified in 180 samples by conventional PCR. We included five samples in pools J7 – J10, whose aetiology had been identified by the conventional PCR method. In addition, RNA from a known beta Polyomavirus (PyV) positive sample was included as a positive control. Importantly, because the samples had been shipped to the U.K. from Uganda, protected by the DNA/ RNA shield at ambient temperatures, fresh serum was acquired from a donor and processed with the NGS samples for comparative purposes. The serum from the donor was excluded from sequencing at the donor's request. A comparison of the total nucleic acids is shown in Table 15 for before (RNA concentration) and after (DNA concentration) library preparation.

The clinical signs and symptoms associated with the samples were vomiting, diarrhoea, fatigue, anorexia, abdominal, muscle, joint pains, headache, cough, difficulty breathing, jaundice, rash, and bleeding (Figure 23). Headaches and hiccups were the most and

least reported symptoms, respectively. Among the 185 patients, a 2% mortality rate was recorded.

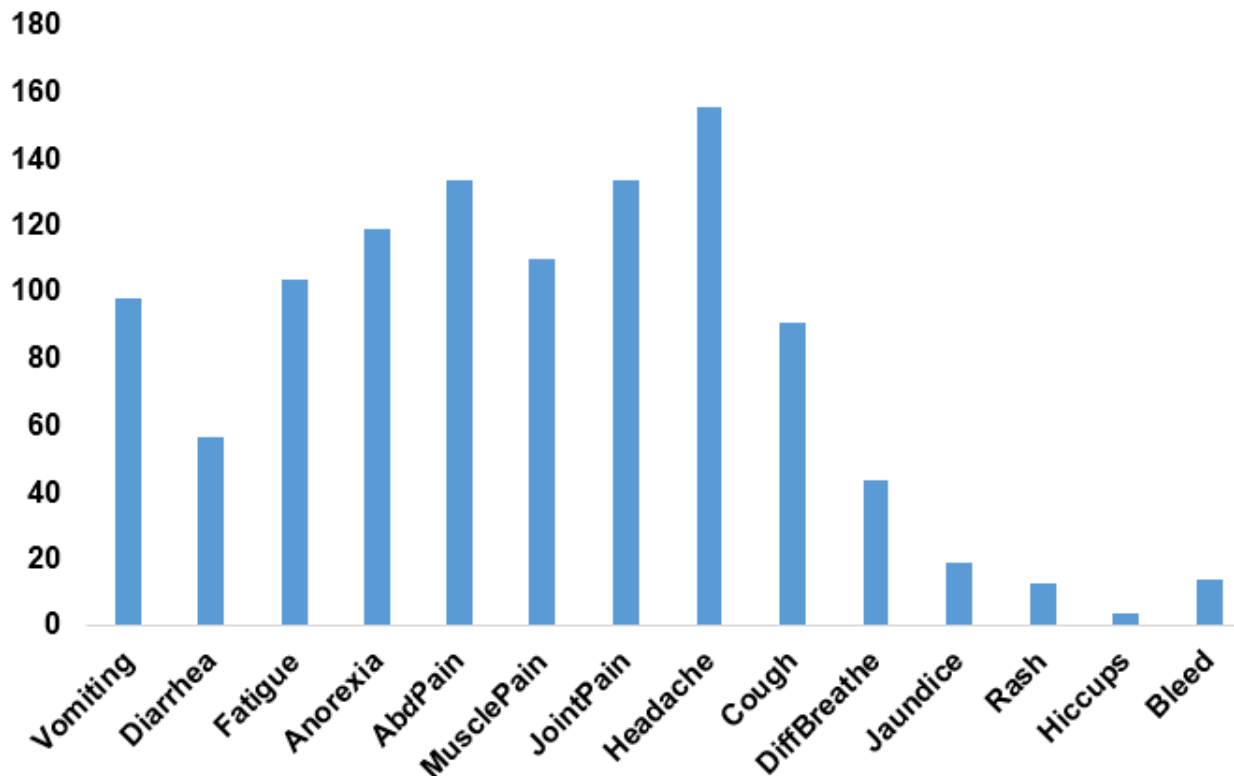


Figure 23. The signs and symptoms associated with the patient samples selected for NGS. The Y-axis represents the number of samples associated with the sign/ symptom on the X-axis

## 6.2.2 Quality control of the laboratory procedure

Since the Nanodrop One spectrophotometer RNA concentrations are less accurate for concentrations below 20 ng/μL, the RNA was quantified using the Agilent 4200

Tapestation (Agilent). The Tapestation provides more accurate RNA quantities and vital information on the extent of sample degradation (represented as RNA integrity number,

RIN). Thus, samples with an initial RNA concentration input below 0.83ng/μL (as required by the manufacturer's protocol) were not considered for further NGS library preparation (Table 14). The exception was sample (202), which was positive for Enterovirus on conventional PCR. Of the 18 RNA quantified pools, only 9 (50%) had a sufficient amount of RNA to proceed to library preparation. In addition, except for the freshly acquired donor sample, all pools showed a high level of RNA degradation as indicated by the RIN value (1 for highly degraded samples and 10 for nearly intact RNA), as shown in Table 14. In total, ten pools were processed, including pool 202. As expected, the donor's sample had a higher RIN, implying that the RNA in the sample was intact.

Table 14. Table showing concentration and corresponding RIN for whole blood specimens. \* exempted due to positive PCR result described in section 4.2.3, and \*\* the Tapestation could not calculate sample RIN.

| <b>Specimen<br/>pool (ID)</b> | <b>Nanodrop<br/>RNA (ng/μL)</b> | <b>Tapestation<br/>RNA (ng/μL)</b> | <b>RIN</b> | <b>Number<br/>of<br/>samples/<br/>pool</b> | <b>QC<br/>pass?<br/>YES/NO</b> |
|-------------------------------|---------------------------------|------------------------------------|------------|--------------------------------------------|--------------------------------|
| 202                           | 15.9                            | 0.18                               | 1          | 10                                         | NO*                            |
| 298                           | 3.3                             | 0.12                               | 1.6        | 10                                         | NO                             |
| 308                           | 1.8                             | 0.19                               | 0          | 10                                         | NO                             |
| <b>319</b>                    | <b>7.7</b>                      | <b>6.54</b>                        | <b>2.5</b> | <b>10</b>                                  | <b>YES</b>                     |
| <b>330</b>                    | <b>2.7</b>                      | <b>2.87</b>                        | <b>2.8</b> | <b>10</b>                                  | <b>YES</b>                     |
| <b>340</b>                    | <b>3.8</b>                      | <b>4.62</b>                        | <b>2</b>   | <b>10</b>                                  | <b>YES</b>                     |
| 517                           | 3.4                             | 0.05                               | -**        | 10                                         | NO                             |

|                     |             |             |                 |           |            |
|---------------------|-------------|-------------|-----------------|-----------|------------|
| 527                 | 7.0         | 0.09        | 1.3             | 10        | NO         |
| 548                 | 3.6         | 0.09        | 1.4             | 10        | NO         |
| 558                 | 4.4         | 0.08        | 1.5             | 10        | NO         |
| <b>609</b>          | <b>13.3</b> | <b>4.03</b> | <b>1.2</b>      | <b>10</b> | <b>YES</b> |
| <b>619</b>          | <b>2.2</b>  | <b>0.87</b> | <b>2</b>        | <b>10</b> | <b>YES</b> |
| J7                  | 5.5         | 0.44        | 1.7             | 15        | NO         |
| J8                  | 3.3         | 0.19        | _ <sup>**</sup> | 20        | NO         |
| <b>J9</b>           | <b>5.5</b>  | <b>1.78</b> | <b>2.2</b>      | <b>16</b> | <b>YES</b> |
| <b>J10</b>          | <b>10.2</b> | <b>5.0</b>  | <b>1</b>        | <b>12</b> | <b>YES</b> |
| <b>Donor sample</b> | <b>90</b>   | <b>27.6</b> | <b>6.1</b>      | <b>1</b>  | <b>YES</b> |
| <b>FFPE PyV</b>     | <b>12.6</b> | <b>6.34</b> | <b>1</b>        | <b>1</b>  | <b>YES</b> |

Four libraries were excluded from further processing because the sequencing pipeline required at least  $\geq 2\text{nM}$  of input DNA (Table 15). In addition, the donor sample was excluded at the donor's request. Thus, only 5 of the 10 pools were sent for sequencing.

Table 15. Library concentrations (ng/ $\mu\text{L}$ ) before and after HTS library preparation procedure.

| Specimen ID | Before Library Preparation             |          | After library preparation              | Proceed to NGS? |
|-------------|----------------------------------------|----------|----------------------------------------|-----------------|
|             | RNA Concentration (ng/ $\mu\text{L}$ ) | RIN      | DNA Concentration (ng/ $\mu\text{L}$ ) |                 |
| <b>202</b>  | <b>1.8</b>                             | <b>1</b> | <b>1.63</b>                            | <b>YES</b>      |

|                 |             |            |             |            |
|-----------------|-------------|------------|-------------|------------|
| 319             | 6.54        | 2.5        | 0.092       | NO         |
| <b>330</b>      | <b>2.87</b> | <b>2.8</b> | <b>6.76</b> | <b>YES</b> |
| 340             | 4.26        | 2          | -           | NO         |
| <b>609</b>      | <b>4.03</b> | <b>1.2</b> | <b>1.98</b> | <b>YES</b> |
| 619             | 0.674       | 2          | 0.0412      | NO         |
| J9              | 1.78        | 2.2        | 0.059       | NO         |
| <b>J10</b>      | <b>5.0</b>  | <b>1</b>   | <b>1.09</b> | <b>YES</b> |
| Donor<br>sample | 27.6        | 6.1        | 35          | NO*        |
| <b>PyV</b>      | <b>6.34</b> | <b>1</b>   | <b>35.9</b> | <b>YES</b> |

\* Excluded from NGS due to donor's request.

### 6.2.3 Analysis of the positive control sample (Polyomavirus)

To ascertain successful library preparation and HTS, we used known polyomavirus RNA acquired from the University of Nottingham laboratory archives. The positive control library was prepared and sequenced simultaneously with the other pools. HTS generated 150,000,000 reads. The IDseq platform generated results within 3 hours, and bioinformatics analysis showed that 20,998,496 (14%) reads passed the host and quality filters. Genomic assembly was carried out using the Human polyomavirus 1 (Betapolyomavirus), GenBank accession number LR216248.1. A full genome (5kb) with

100% coverage breadth was recovered from 3 contigs mapping to the polyomavirus (Figure 24).

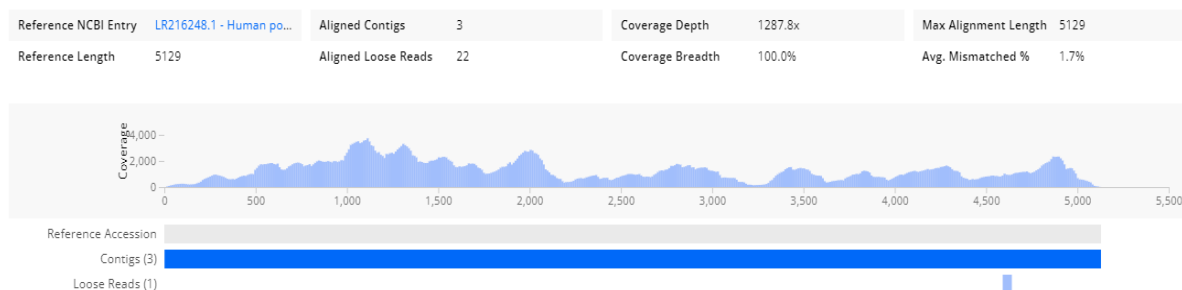


Figure 24. Assembled PyV positive control recovered by IDseq pipeline following mapping to the reference genome LR216248.1 (5kb long). Sample coverage depth was 1287.8X, with an average mismatch of 1.7%.

## 6.2.4 Sequence analysis using the IDseq platform

From pool 609, HTS generated 150,000,000 with 10M (6.7%) reads passing the host and QC filters (Figure 25).

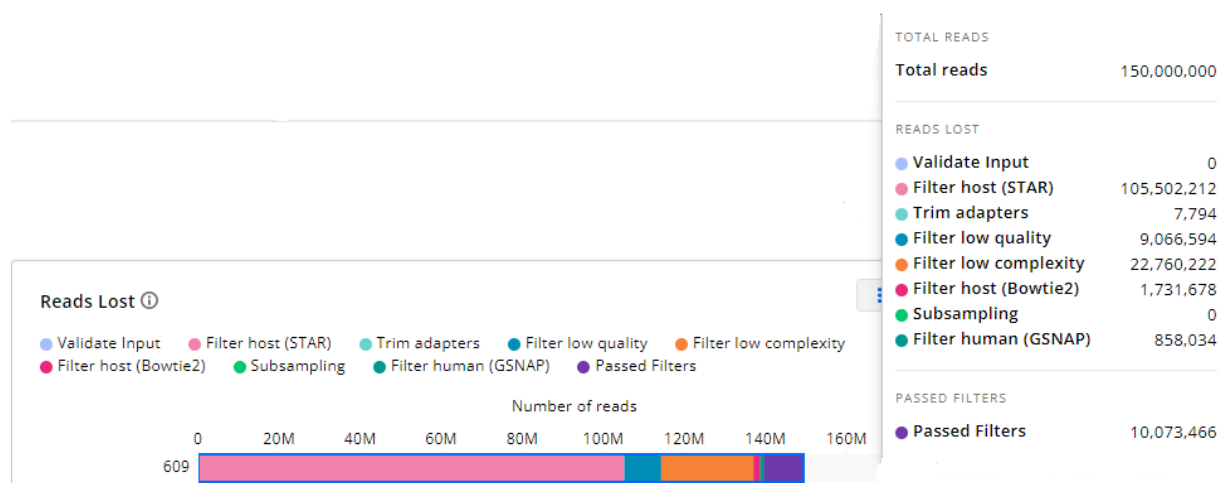


Figure 25. Analysis of pool 609 using the IDseq bioinformatics pipeline. Except for the purple colour, the different colours indicate the number of reads during sample processing.

Pool 202 comprised 150M reads, of which only 9.81M (6.54%) passed the host and QC filters. Pool 330 comprised 150M reads, although only 2.988M (1.99%) reads passed through the filters. Lastly, pool J10 contained 132M reads, of which only 23,030 (0.02%) reads passed the filters. The total time taken to process the pools ranged from 1 to 7 hours (Table 16)

Table 16. Viruses identified by NGS/HTS after bioinformatics analysis of IDseq data. For viruses enterovirus B and genus Orbivirus, only non-redundant reads were available, whilst contigs were available for the rest.

| Pool        | HTS reads (M) | Passed host and QC filters (%) | Time taken (Hours) | Virus identified                                                    |
|-------------|---------------|--------------------------------|--------------------|---------------------------------------------------------------------|
| PyV control | 150           | 14                             | 3                  | Polyomavirus                                                        |
| 202         | 150           | 6.54                           | 7                  | Pegivirus A,<br>Pegivirus C, Beta<br>Polyomavirus,<br>enterovirus B |
| 330         | 150           | 1.99                           | 3                  | Human alpha<br>herpesvirus 1                                        |
| 609         | 150           | 6.7                            | 4                  | Pegivirus A,<br>Hepatitis C virus                                   |
| J10         | 132           | 0.02                           | 1                  | Measles virus,<br>Orbivirus                                         |

#### 6.2.4.1 Analysis and PCR confirmation of HCV

Three hundred eleven (311) reads mapped against the HCV isolate GenBank accession number MH742361.1 isolated from Uganda, Rakai district. Four (4) aligned contigs covered 29.9% of the reference genome breadth, with a 7.5% average mismatched identified.

The pool was de-convoluted, and HCV was confirmed by conventional PCR in a sample collected from a 36-year-old male farmer from the western region of Uganda. At the time of triage, the patient presented with a fever and complained of headache, fatigue, anorexia, and abdominal pain. In addition, clinical examination indicated the patient to have jaundice and hiccups.

PCR screening of the HCV NS5 region for the remaining pools identified one more HCV-positive sample. This pool was de-convoluted, and the sample was identified to belong to a male patient, a hunter from the western region of Uganda. The clinical signs and symptoms included a fever, nausea and vomiting, diarrhoea, fatigue, anorexia, joint pain, headache, cough, difficulty breathing and cough.

HCV was first detected in pool 609 by HTS and confirmed by PCR using primers targeting the NS5B gene as described by Morice *et al.* [457]. Further screening of the 218 pools identified a second sample. Overall, there was a 0.18% HCV prevalence.

## 6.2.4.2 Phylogenetic analysis of Hepatitis C virus

Initial attempts to subtype the HCV were 50% successful, with only one sample (HCV 609) subtyping as genotype 4k. Therefore, a phylogenetic tree was constructed using the RNA-dependent RNA polymerase region based on 38 translated nucleotide sequences (Figure 26). The second sample (HCV 68) remained inconclusive as it clustered within genotypes 4s and 4t.

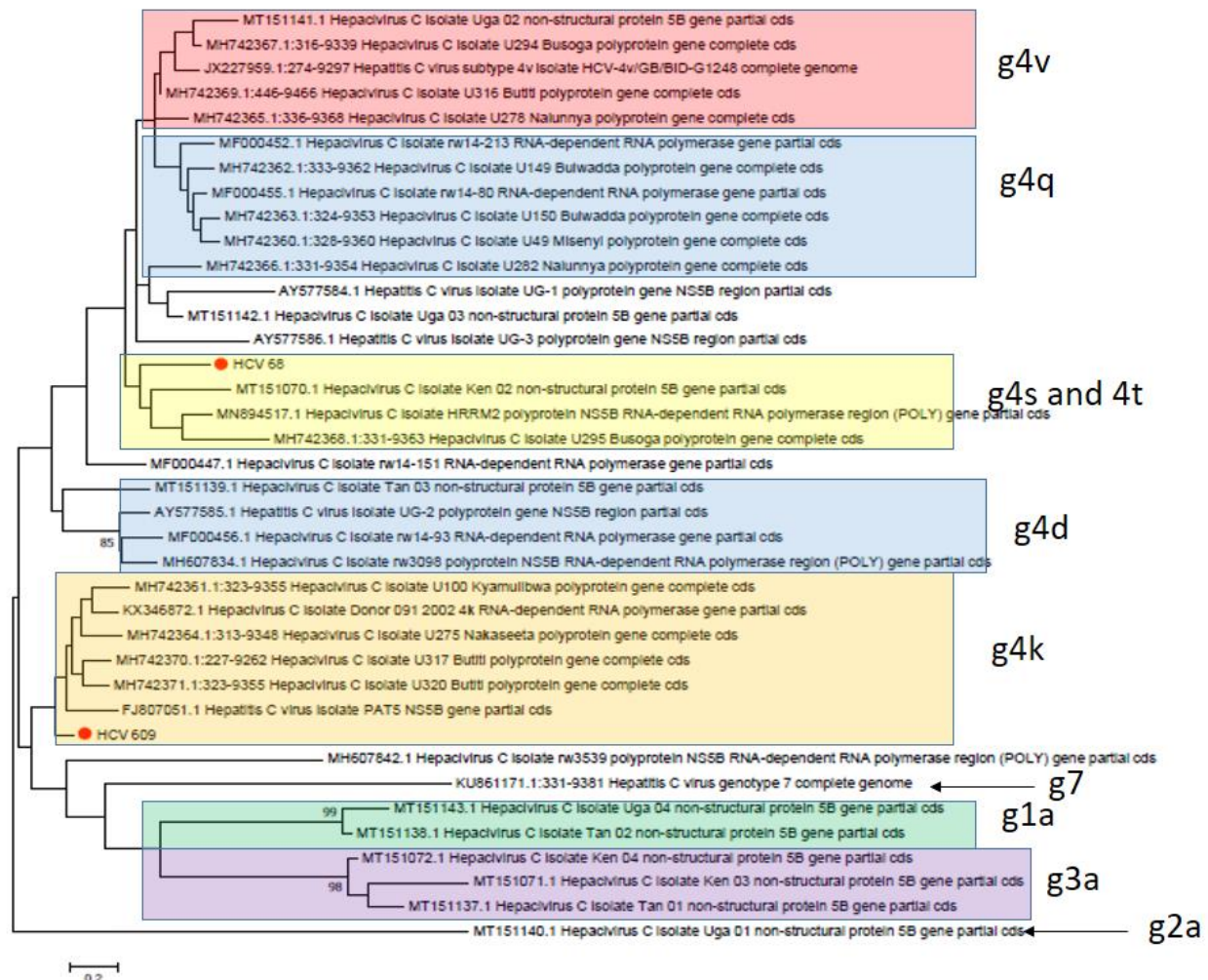


Figure 26. Phylogenetic tree showing 38 selected HCV from the East African region except accession number MN894517, from Egypt. *The red dots show viruses detected*

*in this study (HCV 68 and HCV 609). Genotypes are shown next to the coloured boxes. The tree was constructed based on the NS5B protein (amino acid positions 2536-2959 of the H77C genotype 1a, accession number AF011751). The tree was inferred using the maximum likelihood method based on the general reversible model from an initial tree for the heuristic search obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model [377]. Bootstrapping was done for 1000 replications. Percentages greater than 60 are shown next to the branches for the associated serotypes. The tree is drawn to scale, and discreet Gamma distribution was used to model the evolutionary rate differences.*

## 6.2.5 Analysis of Pegiviruses using the IDseq platform

Pegiviruses were identified in pools 202 and 609. Pool 202 contained Pegiviruses A and C, whilst pool 609 contained only Pegivirus A.

In pool 202, 11 individual reads and six (6) contigs mapped to the GB genome accession GenBank number KC618398.1 were identified as Pegivirus C ( Figure 27). A 37.7% reference sequence coverage was achieved, with a 94.6% nucleotide similarity average.

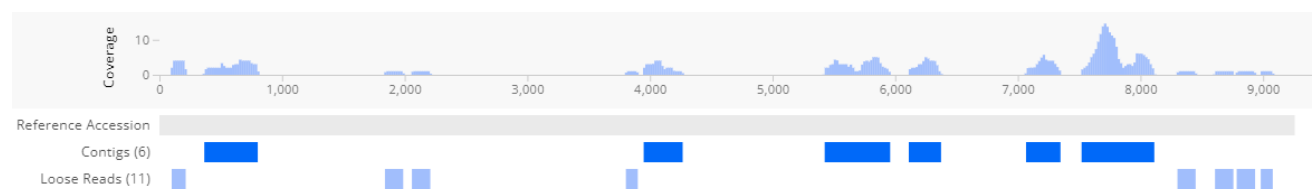


Figure 27. Cartoon of Pegivirus C reads and contigs identified in pool 202. *The dark blue bands indicate pegivirus C contigs, while the light blue bands indicate the reads and their position parallel to the reference accession sequence. The coverage depth of each contig and read band is represented by a 'bump' on a scale of 1-10X.*

For Pegivirus A, a single contig was mapped to the human pegivirus isolate, GenBank accession number MH053116.1, with a reference genome coverage of 6% and 88.7% nucleotide similarity average (Figure 28).

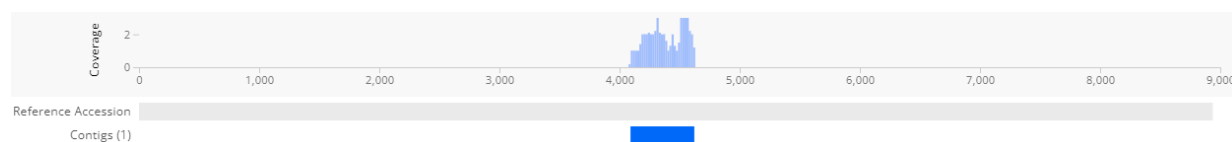


Figure 28. A single pegivirus A contig was identified in pool 202. *The dark blue band indicate pegivirus C contig and its position parallel to the reference accession sequence. The coverage depth of the contig band is represented by a 'bump' on a scale of 0 -2X.*

In pool 609, pegivirus A reads and contigs were mapped to the human pegivirus strain GenBank accession number LT009490.1. Overall, two contigs and six (6) individual reads were aligned parallel to the reference accession with 100% coverage and 92.8% average nucleotide sequence identity (Figure 29).

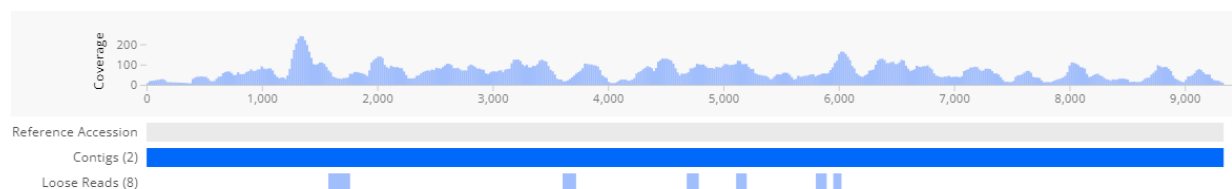


Figure 29. Pegivirus A in pool 609. *Two contigs (dark blue) and 8 individual reads (light blue) were mapped against the reference accession (LT009490.1). The coverage depth (65.1X) of the reads and contigs band is represented by a 'bump' on a scale of 0 -200X.*

PCR confirmation and screening of unsequenced pools were done using primer sets described in section 2.4.5.1. Of the 90 pools screened, 43 (47.8%) pools were positive for pegiviruses.

#### *6.2.5.1 Phylogenetic analysis of Pegiviruses*

The recovered pegiviruses A and C genomic sequences were downloaded from IDseq and aligned with MEGA v.7. Phylogenetic tree analysis identified all sequences as pegivirus C viruses after 1000 bootstrap replications (Figure 30). Both pegiviruses A and C clustered within Pegivirus C viruses.

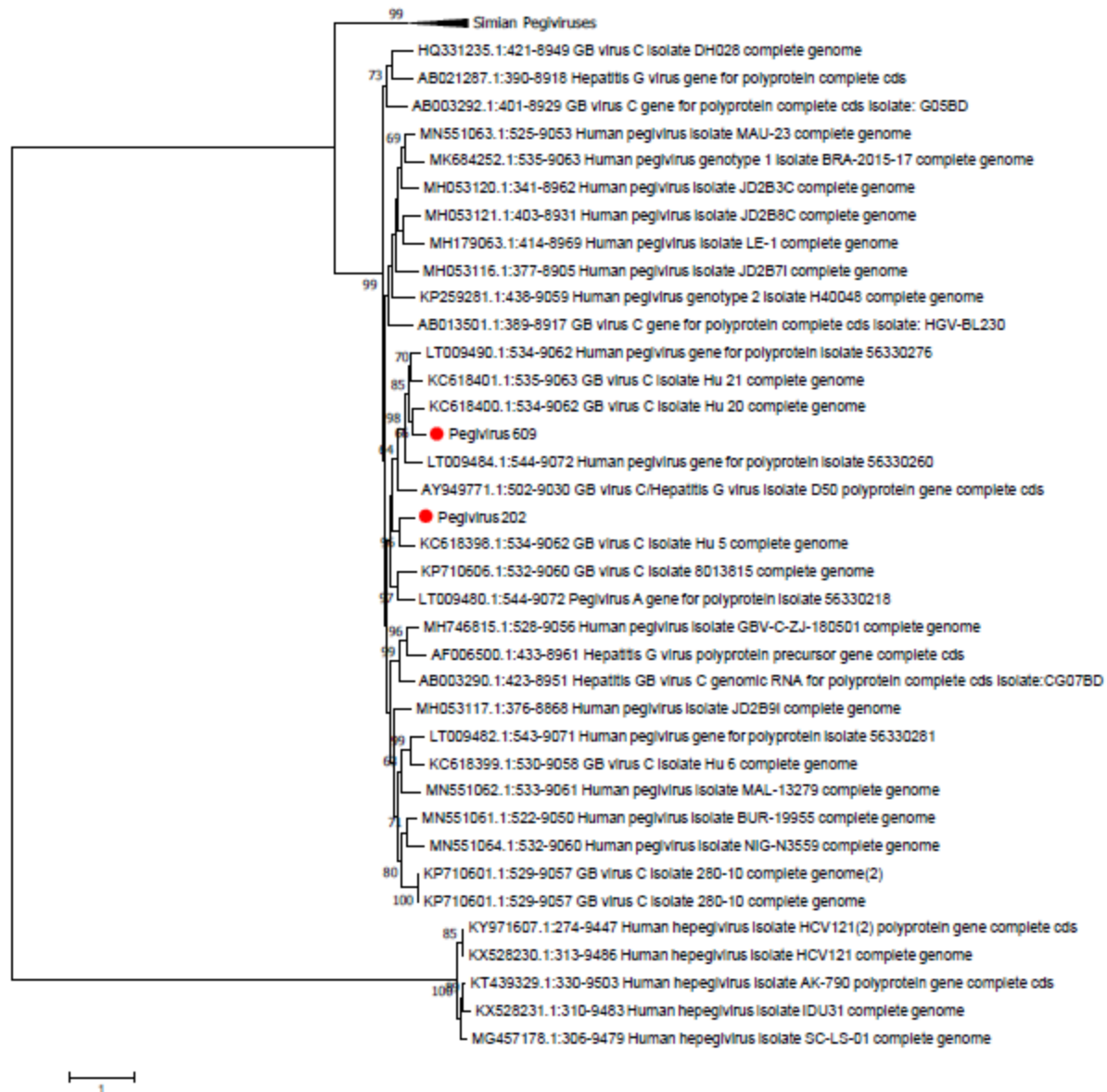


Figure 30. Phylogenetic tree showing 41 selected pegiviruses complete genomes. The red dots show viruses detected in this study (Pegivirus 202 and 609). The branch containing simian pegiviruses is condensed. The tree was inferred using the maximum likelihood method based on the general time reversible model from an initial tree for the heuristic search obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model [377]. Bootstrapping was done for 1000 replications. Percentages greater than 60 are shown next to the branches for the associated serotypes. The tree is drawn to scale, and discrete Gamma distribution was used to model the evolutionary rate differences.

## 6.2.6 Analysis of Measles virus using the IDseq platform

The MeV described in section 4.2.4 was pooled with 15 other whole blood samples and processed using the library preparation procedure for NGS (section 2.4.2). The recovered MeV was mapped to measles virus genotype B3, GenBank accession number KY969477.1. In total, only five (5) contigs and eight (8) individual reads were mapped to the reference sequence (KY969477.1), equating to 14.3% coverage of the reference genome (Figure 31).

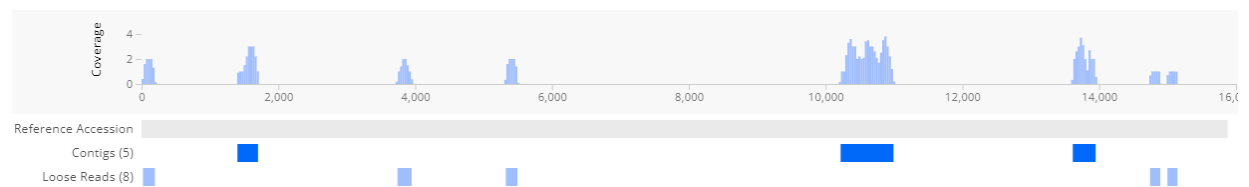


Figure 31. Cartoon of the MeV genome recovered using IDseq. *The five contigs and eight reads were mapped on a reference strain (MeV genotype B3, accession number KY969477.1). The contigs had a 14.3% coverage with an average mismatch of 1.2% to the reference strain (KY969477.1).*

## 6.2.7 Analysis of the Polyomavirus

Two contigs in pool 202 were mapped against GeneBank reference genome, accession number AB263920.1. The contigs and individual reads were identified as Beta Polyomavirus by the IDseq bioinformatics pipeline and covered 13.6% of the reference genome with a 99.9% nucleotide identity (Figure 32).

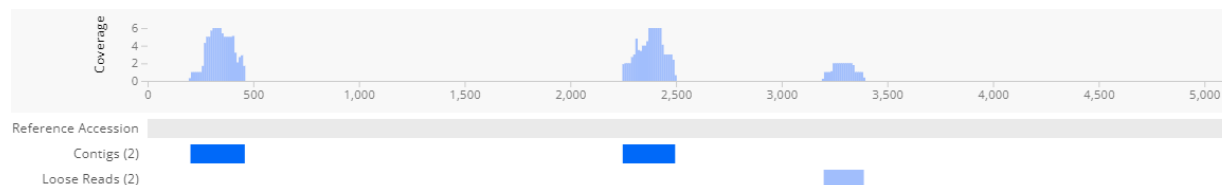


Figure 32. Cartoon of the PyV genome recovered using IDseq. *Two contigs and two individual reads were mapped on a reference strain accession number AB263920.1. The heat map represents the percentage coverage of each contig, equating to an average of 14.3%.*

The polymerase chain reaction to recover PyV was carried out using primers described in section 2.4.3.5.1. No PyV genome was recovered.

## 6.2.8 Analysis of the non-redundant nucleotide reads

The IDseq pipeline identified viral sequences (reads) for which no contigs were formed. These sequences were referred to as non-redundant reads (NR), which were confirmed by the BLASTN or BLASTX functions of the NCBI website. Non-redundant were identified in the genera enterovirus, alpha herpesvirus, and Orbivirus.

The ID seq bioinformatics pipeline identified 24 NR in pool 202 as enterovirus B species. All 24 reads were downloaded and BLASTN. There was no enterovirus confirmation from all the 24 non-redundant sequences identified by IDseq.

In addition, two non-redundant reads were identified as the Human alpha herpesvirus 1. These were downloaded, and BLASTN searched on the NCBI website for confirmation. However, none of the reads was identified as the human alpha herpesvirus 1.

Furthermore, the IDseq bioinformatics pipeline identified four (4) non-redundant translated reads as viruses from the genus Orbivirus. These were downloaded, and BLASTX searched on the NCBI website for confirmation. The sequences were confirmed related to the African horse sickness virus (AHSV) and the Umatilla virus.

Having done an in-silico confirmation, we attempted confirmation by conventional PCR using primers described by Palacios *et al.* [329] (section 2.4.6). However, attempts to detect the virus failed. Therefore, we attempted Orbivirus detection by designing in-house primers from the IDseq translated nucleotide orbivirus reads, as described in section 2.4.3.4.1. No Orbivirus genome was detected.

## 6.3 Discussion

We sought to uncover other aetiologies of AFI in undiagnosed clinical samples using unbiased high-throughput sequencing (HTS). Initially, other methods such as Enzyme-linked immunosorbent assay (ELISA), Plaque Reduction Neutralisation assay (PRNT), and Polymerase Chain Reaction (PCR) were attempted, producing varying degrees of success. This chapter describes the detection of viruses such as HCV and Pegiviruses that are rarely captured during routine surveillance in Uganda. It also illuminates the possibility of novel viral introductions (Orbiviruses) into the human population.

The capability of HTS to sequence pathogen genomes from degraded samples is very promising, whereby all viruses recovered in this study were from samples with very poor provenance. Compared to our study, Manso *et al.*, 2017 showed that viruses could be recovered from degraded clinical samples using available Illumina sequencing protocols; however, the authors observed higher viral loads by simply modifying reagent volumes of the manufacturer's protocols [458]. Sample degradation would result from intermittent power shortages, a common phenomenon in developing countries.

In our study, we applied a positive control (Polyomavirus) to standardise clinical diagnosis using HTS as elucidated by Gu *et al.*, 2019 [281]. In this study, we did not apply the use of a non-template control due to limited resources. However, as described by Asplund *et al.* 2019, using negative controls was important in informing on the presence and potential sources of contaminants [459]. Asplund *et al.* 2019, further described the occurrence of sequences which are of non-human origins, such as bacteriophages, as laboratory

component-associated (LCA) viral sequences. Similar to our findings, bacteriophages were constantly observed in our sequence data. In their findings, LCA viral sequences were observed in library preparation components such as AMPure, ScriptSeq v2, ScriptSeq Gold, Nextera, and Platinum Taq, to mention a few. The observation of such LCA viral sequences in our data concurs with their findings since the library preparation method followed the Illumina sequencing protocols.

Using HTS, this study detected the presence of Hepatitis C and pegiviruses within the samples. The HCV prevalence/ positivity rate (0.18%) detected by PCR in this study is comparable to the 0.26% (20 cases of 7,751) observed by Davis *et al.*, 2019 [460]. However, their serological HCV prevalence was estimated at 3.2%. Our findings concur with the low prevalence of HCV estimated in the Sub-Saharan region (SSA) [80]. Based on phylogenetic analysis, our findings suggest the two genotype 4 HCV in our data belong to two separate subtypes. The observation of HCV 68 phylogenetically clustering within genotypes 4s and 4t may suggest a possible recombination event or a plausible novel subtype in the region. An example of both events was observed in a study by Gededzha *et al.*, 2012 in South Africa [461]. Such recombinations have also been observed globally, especially in areas where multiple HCV genotypes are known to circulate [462], [463] and [464]. Further evaluation of this inconclusive subtype by high throughput sequencing would provide a clearer perspective.

The high rate (47.8%) of pegiviruses in our study is higher than previously observed in other studies [465]. Although the plethora of asymptomatic and healthy individuals with pegiviruses ruthlessly outweighs the causal relationship of the virus to disease [466], [467], [468], [469], our study concurs with Kufner *et al.* 2019, and Bukowska *et al.* 2018,

where pegivirus C was detected in clinically morbid patients [470], [471]. Although we could not confirm the presence of orbiviruses in our samples, the presence of non-redundant translated sequences could be an indicator of probable infection. The viral sequences detected in our study were more closely related to the African horse sickness and the Wallal virus. This finding may point to a novel human orbivirus in the human population in the SSA. Further work would involve applying laboratory assays that are known to support the extraction of double-stranded RNA, such as the RNASwift method [472] and thus aid the proper identification of the detected virus. The number of Measles virus (MeV) sequence reads, contigs, and coverage recovered by HTS was fewer than expected. This low MeV sequence recovery may result from low virus abundance in the pool (a mix of 15 whole blood samples) and sample degradation.

Attempts to confirm polyomavirus, human herpes simplex virus, and enterovirus B were unsuccessful in this study. One possible explanation could be the lack of sequence similarity between primers and template sequences since the primers used for PCR detection were not necessarily within the region identified by NGS analysis. The primers used in this study were previously published sets targeting conserved regions within these viral genera and therefore, could potentially miss out on any orphan reads in especially degraded samples.

On the other hand, since library prep was done simultaneously, the polyomavirus contigs and to a certain degree, the Pegivirus A contigs, observed could have been contamination due to the positive control's high viral load. In this regard, while sample pooling is cost-effective and time-efficient [473], it also has some disadvantages such as reduced sequencing depth, unequal coverage (due to differences in levels of DNA concentration

and quality which can result in unequal representation in the pooled library) and lastly, cross-contamination. This may be the case in the observed FFPE PyV control applied in this study. Analysis of Table 16 shows that the percentage of PyV reads that passed the host and QC filter pipeline was twice more compared to the next abundant pathogen reads. This phenomenon of 'read swamping' can lead to uneven coverage across the other genomes as a result of a higher concentration of the control and thus impact downstream analysis. This can also lead to misinterpretation of results.

While identifying the source of contamination in pooled samples can be a hurdle, effective methods such as designing primers specific to the observed reads for PCR can narrow down the likely sources. Another important measure is to compensate the libraries for differences in input concentration by adjusting the amount of DNA from each sample to ensure equal representation in the pooled library. By doing so, the sequencing depth and coverage can be more even, reducing the impact of sample-to-sample variation on downstream analysis.

Another argument is broadening the list of false positives (or contaminants) to include the human herpes simplex virus, Orbivirus and enterovirus B reads. In agreement with this conclusion, the human herpes simplex virus and enterovirus B reads were probably part of the LCA viral sequences since they were similar to what was observed by Asplund et al., 2019 [459]. Laurence et al., 2014 also observed the genus *Bradyrhizobium* as a common contaminant observed in sequences from various centres that worked on the 1000 Genome Project [474].

A potential source of false positives is the quality of the reference databases used by the IDSeq pipeline. The accuracy of the taxonomic and functional assignment by the pipeline

depends on the completeness and quality of the reference databases. If the reference databases are incomplete or contain errors, the accuracy of the analysis may be compromised.

Therefore, to mitigate the potential for false sequencing hits in IDSeq and probably other online bioinformatics pipelines, it is important to perform proper quality control of the sequencing data, including filtering out contaminants and low-quality reads. It is also essential to use high-quality reference databases and to regularly update them to ensure their accuracy. In addition, Laurence *et al.*, 2014 [474] recommended the use of UV irradiation of water, cross-validation of independent techniques, and complete genome sequencing of contaminants with the aim of computational subtraction during bioinformatics analysis.

Additionally, the quality of the sequencing reads themselves can also affect the accuracy of the analysis. Sequencing errors or low-quality reads can lead to an incorrect taxonomic and functional assignment. However, this was accounted for in the IDseq pipeline whereby samples were filtered of low-quality and low-complexity reads. It is important to note that while all HTS analyses applied the IDseq platform, other platforms such as the Genome Detective, ExpressionPlot [475], GenePattern, Galaxy, and ZENBU [476] among others. This implies that these bioinformatics tools would probably provide pipelines with better computation power, reference databases and less bias compared to IDseq.

As previously stated, this study applied the use of MEGA v7 in determining the best-fit evolutionary model for phylogenetic analyses. It is worth noting that other tools such as

the ModelTest [477], PartitionFinder [478], IQ-TREE [479], and PhyML [480] may provide better alternative evolutionary models.

## 6.4 Conclusion

Metagenomic HTS can detect viral pathogens in highly degraded samples. However, caution should be taken whilst interpreting results to identify contaminants such as cloning vectors and LCA viral sequences. These can be identified by taking into account the most common pathogens or organisms a laboratory handles regularly and confirmation of HTS viral hits using PCR. There is a need to establish routine surveillance of HCV and thus track the circulating genotypes within the region.

## 7.0 Final discussion and conclusions

This thesis describes the epidemiology of arboviruses (flaviviruses) and further attempts to determine the aetiology of viral acute febrile illness (AFI) of unknown cause in patients attending general and regional hospitals in Uganda. The study applied a retrospective approach utilising serum and or whole blood collected from patients attending general and regional referral hospitals in Uganda. The hospitals were limited to sentinel sites only, that is, hospitals with which the UVRI has a research/ diagnosis partnership. Where such collaborations did not exist, patient specimens were acquired in response to emergencies such as suspicious fevers associated with haemorrhage. Arbovirus negative samples were assessed by a combination of conventional PCR and high throughput sequencing (HTS) where possible.

Since routine sample collection was only carried out from selected sentinel hospitals, this study cannot estimate the average annual burden of AFI, let alone viral AFI in Uganda. In addition, there is non-blood borne viral AFI which falls out of the scope of this study (majorly focused on blood-stream AFI). Estimating the burden of AFI would require information on the patient's utilisation of health services at hospitals and also account for the factors associated with the general underreporting of AFI in the SSA.

Firstly, the underreporting of febrile illnesses in Africa is a complex issue that has multiple factors contributing to it. Some of the main reasons for underreporting are inadequate healthcare infrastructure, lack of access to healthcare, limited resources, and poor surveillance systems. Briefly, Inadequate healthcare infrastructure, particularly in rural areas comprises a shortage of trained medical personnel, a lack of medical supplies, and

insufficient funding for healthcare facilities. Thus, the underreporting of AFI since many cases go undiagnosed and untreated. Another factor contributing to underreporting is the lack of access to healthcare. This includes populations in hard-to-reach areas with inaccessible tributary roads. Where the cost of medical care is not government provided, it is often viewed as a deterrent and very expensive and thus many people may opt not to seek medical attention. Limited resources also play a significant role in underreporting. Many healthcare facilities in Africa lack essential medical equipment, diagnostic tools, and medications. This can make it difficult for healthcare workers to diagnose and treat febrile illnesses. Moreover, many healthcare facilities cannot collect and report data on febrile illnesses, which can result in underreporting. Poor surveillance systems are another factor contributing to underreporting. Many SSA countries do not have robust surveillance systems that can accurately capture the incidence of febrile illnesses. Moreover, even when surveillance systems are in place, they may not be effectively implemented or monitored, leading to inaccurate data reporting. Therefore, the underreporting of febrile illnesses in Africa is a multifaceted issue that requires a comprehensive approach to address. Improving healthcare infrastructure, increasing access to healthcare, providing adequate resources, and implementing effective surveillance systems can help to address this problem and ensure that cases of febrile illnesses are accurately reported and treated.

The utilisation of health services depends on several factors, such as the availability of essential medicines and distance or accessibility to health centres [481],[482]. Unpublished data from Uganda showed that fewer patients were likely to attend health centres if the distance to the nearest health centre was reasonably far or if there was a

shortage in the supply of essential drugs such as anti-malarials. The challenge of patient non-attendance to clinics can probably be likened to missed appointments in developed countries. In addition to distance to health centres, researchers in the Sultanate of Oman reported that age, gender, service cost and waiting time were major predictors of missed appointments. Among patients that attended clinics, other factors that could have influenced a missed diagnosis may include but are not limited to; the low scope of laboratory diagnostic arsenal, cost of laboratory tests and health services [483], inexhaustive triage (the result of a high patient-to clinician ratio) [484], [485] and probably low numbers of infectious disease experts.

Our findings show that surveillance in the northern and eastern regions was less extensive than that performed in the central and western regions. Often, the samples received from the northern and eastern regions result from a physician/ clinical request or the suspicion of a viral haemorrhagic disease. However, physician laboratory requests do not cover all AFI; therefore, establishing more sentinel sites in the North and the East would provide a clearer perspective on these regions' viral aetiology and epidemiology. However, one of the biggest challenges to a regular surveillance system in the north and the eastern regions has been inaccessible road tributaries to remote areas and generally poor infrastructure.

The methods applied in this study included determining the proportion of patients with an arboviral infection by analysing existing data at the Department of Arbovirology, Emerging and Re-emerging Infectious Diseases at the Uganda Virus Research Institute (UVRI). Further, conventional PCR and unbiased HTS were applied to detect and attempt virus discovery within the undiagnosed clinical AFI samples. The laboratory methods employed

by the UVRI to detect the presence of an arboviral infection comprised IgM serology and confirmation by plaque reduction neutralisation test (PRNT). Although PRNT is considered a confirmatory test for the presence of infection, a majority (55.5%) of the confirmed results in this study remained as unspecific alpha or flavivirus infections. This is a challenge in patient management as the physician must draw a treatment or clinical management plan for flaviviruses such as ZIKV [486] and DENV cases [487], [488] from somewhat inconclusive laboratory findings. Furthermore, these laboratory results are usually delayed due to the length of the laboratory assays.

As previously discussed, PCR and HTS have strengths and weaknesses; however, as observed in this study, HTS can complement PCR in the surveillance and detection of infectious diseases. Unfortunately, regularly scanning samples in low and middle-income countries (LMIC) using this approach is not readily feasible. However, collaborations and research partner organisations such as the Centre for Disease Control and prevention (CDC), USA and the Medical Research Council (MRC, U.K) have enabled such activities to be carried out at research and academic institutes in LMIC. Better still, the sequencing costs are steadily getting cheaper based on sample pooling approaches [489], [490].

Among arboviruses, YFV was dominant (20 cases), with at least one identified case in all four regions. ZIKV (3 cases) was only confirmed in the central region in a location 65km away from the Zika forest. Only two incidences of WNV were confirmed; one by PRNT and another by PCR, both occurring in the west, over 500km away from Arua (West Nile region), where the first case (WNV 1947) was identified. Among alphaviruses, a single CHIKV and six ONNV were confirmed. No DENV infection was confirmed; however, neutralising antibodies were detected among the 16 unspecific flavivirus cases, implying

DENV infection cannot be ruled out in the region. Furthermore, all four DENV serotypes have been reported circulating in the East African region, namely, Kenya [491], Tanzania [492], [493], [494], and the Democratic Republic of Congo [495] by the PCR method.

In Uganda, the *Aedes (stegomyia) africanus* Theobald is a common vector for YFV [496], ZIKV [497], and CHIKV [346]. Therefore, all the viruses would theoretically have an equal chance of transmission using the same vector. However, the factors that favour YFV dominance over ZIKV and CHIKV are not fully understood in Uganda. Perhaps YFV may have more than one mosquito species involved in the transmission; for example, Kirya *et al.* 1977, reported the isolation of YFV from *Coquillettidia fuscopennata* [496]. This may suggest the probable dominance of YFV among arboviruses in Uganda, which might be supported by studies on mosquito abundance in the country [498], [499], [45]. However, no experiments were carried out to confirm *Coquillettidia fuscopennata* as a competent vector for YFV. Therefore, it would be worth exploiting factors such as the role of host-virus (mosquito-virus) interactions in determining virus (YFV) dominance. For example, studies determining the YFV's ability to replicate in *C. fuscopennata*'s midgut, disseminate to other tissues and, ultimately, the salivary glands, where the mosquito must be capable of expectorating the virus [500].

This study detected and recovered a near-complete WNV genome in humans for the first time in 85 years. Furthermore, phylogenetic analysis clustered the recovered WNV (WNV Kasese) among lineage two viruses, together with the mosquito isolate by Mossel *et al.* and the 1937 WNV ancestral strain (M12249). Despite clustering with its' ancestral 1937 strain, we showed that the recovered lineage two WNV had acquired amino acid changes in the envelope protein and an N- glycosylation motif, which was missing in the ancestral

strain. These acquired features are associated with increased virulence and pathogenicity, as observed in lineage one viruses. Because only one of the 1,090 samples tested positive for WNV, this highlighted the need to acquire specimens other than blood. For example, several studies with varying results have reported persistent shedding of WNV in patients' urine for at least three weeks after the onset of symptoms [501], [502]. Therefore, WNV and probably other arbovirus surveillance such as YFV [503], RVFV [504], ZIKV [505], [506] would be done by applying such a non-invasive sample collection technique.

Detecting MeV, enteroviruses and other bloodborne pathogens (malaria and Leptospirosis) among febrile patients underscores the importance of continuous surveillance. Among the viruses detected in this study, MeV would require immediate interruption using available control measures such as vaccination. The enteroviruses would require close monitoring to detect existing or emerging highly pathogenic species such as EV-D68 [507] and EV- 71 [508]. Malaria would require a relentless effort by the government and the Ministry of Health to control and perhaps eradicate it. The control methods include the provision of mosquito nets, the availability of anti-malarial drugs, community education and the availability of a trained scientific workforce, among others [509]. Recently, the World Health Organisation (WHO) recommended the use of the RST, S/AS01, a subunit malaria vaccine targeting the circumsporozoite protein (CSP) on *Plasmodium falciparum* sporozoites [510], [511]. Routine treatment of Leptospirosis probably benefits from broad-spectrum antibiotics as there is no routine surveillance to determine the extent of the problem in Uganda.

Including positive and negative controls while carrying out HTS procedures and sequencing provides a better perspective on results interpretation [281], [459]. Unfortunately, due to a limited budget, only a positive control was included in this study. Future studies on virus discovery and patient virome projects could benefit from including both controls as these provide insight into potential contaminants or artefacts.

Of the 1,090 samples that were taken for further laboratory testing in this study, only 185 samples were processed for HTS. A large proportion of the samples was not considered for HTS due to poor sample quantity and / or quality. For this reason, despite detecting malarial (29 of 218 pools), leptospiral (18 pools) and viral (31 pools) infections, this study could not confirm the AFI aetiology in approximately 64% of samples. Factors that could limit establishing an aetiology in the 64% of the samples include a short/ narrow viraemia window for viruses such as Sindbis virus [512]. In addition, Miller *et al.* noted that haemolysis, excessive host nucleic acids (high background noise), and weak positives by PCR (low viral loads,  $ct \leq 37$ ) negatively impacted mNGS by reducing assay sensitivity [513].

Many arboviruses previously observed to circulate or are endemic to the East African region [514], [515] were not detected. One of the reasons was probably the use of previously published primers, targeting regions prone to mutations [314]. Therefore, future studies should consider consistent surveillance and improved diagnostic methods, such as adapting multiplex PCRs that code for complete sequence genome [319] or use molecular beacons [516] catering for potential mutations or substitutions and thus eliminating challenges of gene dropouts at the most conserved genomic positions in viral sequences to improve detection.

In conclusion, high arboviral activity was observed during the rainy seasons. Yellow fever virus was observed in all regions except the East. The general arboviral activity in Uganda requires establishing routine sentinel sites in the northern and eastern regions to provide an accurate description. The application of PCR in routine surveillance is useful in detecting pathogens; however, it benefits from frequent screening using HTS.

## 8.0 Future Work

One of the activities that can be done with the availability of resources such as time and funding is surveillance of viral AFI in specimens other than whole-blood or serum. With a constrained budget, different viral families can be screened using conventional PCR by designing primers targeting their conserved genes. Previously, viral surveillance was done by Cummings *et al.* on respiratory samples [51]; however, it can be extended to non-invasive specimens such as urine.

Ascertaining the presence of Orbiviruses in human samples is another project worth pursuing. Orbiviruses are associated with fever (Kemerovo, Orungo and Changuinola viruses), whereas severe disease such as encephalitis has been observed in the Colorado tick fever virus infections. Interestingly, since its first detection in 1959 in Uganda [517], no reports exist on surveillance or detection of Orungo virus in neither mosquitoes nor humans.

## 9.0 References.

- [1] R. Daumas, P. Brasil, C. Bressan, R. Oliveira, B. Carvalho, D. Carneiro, S. Passos, Interobserver agreement on signs and symptoms of patients with acute febrile illness, *Infection* 39(2) (2011) 135-140.
- [2] F. Ricapa-Antay, K. Diaz-Melon, W. Silva-Caso, L.J. del Valle, M.A. Aguilar-Luis, F. Vasquez-Achaya, C. Palomares-Reyes, P. Weilg, D. Li, C. Manrique, Molecular detection and clinical characteristics of *Bartonella bacilliformis*, *Leptospira* spp., and *Rickettsia* spp. in the Southeastern Peruvian Amazon basin, *BMC infectious diseases* 18(1) (2018) 618.
- [3] C.G. Albariño, Novel Paramyxovirus Associated with Severe Acute Febrile Disease, South Sudan and Uganda, 2012-Volume 20, Number 2—February 2014-Emerging Infectious Disease journal-CDC, (2014).
- [4] R.G. PETERSDORF, P.B. BEESON, Fever of unexplained origin: report on 100 cases, *Medicine* 40(1) (1961) 1-30.
- [5] C.f.D. Control, Prevention, Outbreak of acute febrile illness among athletes participating in triathlons-Wisconsin and Illinois, 1998, *MMWR. Morbidity and mortality weekly report* 47(28) (1998) 585.
- [6] N. Prasad, D.R. Murdoch, H. Reyburn, J.A. Crump, Etiology of severe febrile illness in low-and middle-income countries: a systematic review, *PLoS One* 10(6) (2015) e0127962.
- [7] M.E. Reller, A.M. de Silva, J.J. Miles, R.S. Jadi, A. Broadwater, K. Walker, C. Woods, O. Mayorga, A. Matute, Unsuspected dengue as a cause of acute febrile illness in children and adults in western Nicaragua, *PLoS neglected tropical diseases* 10(10) (2016) e0005026.
- [8] M.E. Reller, U. Akoroda, A. Nagahawatte, V. Devasiri, W. Kodikaarachchi, J.J. Strouse, R. Chua, Y.a. Hou, A. Chow, O.M. Sessions, Chikungunya as a cause of acute febrile illness in southern Sri Lanka, *PloS one* 8(12) (2013) e82259.
- [9] M.R. Capeding, M.N. Chua, S.R. Hadinegoro, I.I. Hussain, R. Nallusamy, P. Pitisuttithum, K. Rusmil, S.J. Thomas, N.H. Tran, D.N. Wirawan, Dengue and other common causes of acute febrile illness in Asia: an active surveillance study in children, *PLoS neglected tropical diseases* 7(7) (2013) e2331.
- [10] O.D. Lorenzi, C.J. Gregory, L.M. Santiago, H. Acosta, I.E. Galarza, E. Hunsperger, J. Munoz, D.M. Bui, M.S. Oberste, S. Penaranda, Acute febrile illness surveillance in a tertiary hospital emergency department: comparison of influenza and dengue virus infections, *The American journal of tropical medicine and hygiene* 88(3) (2013) 472-480.
- [11] W. de Jong, M. Rusli, S. Bhoelan, S. Rohde, F.A. Rantam, P.A. Noeryoto, U. Hadi, E.C.v. Gorp, M. Goeijenbier, Endemic and emerging acute virus infections in Indonesia: an overview of the past decade and implications for the future, *Critical reviews in microbiology* 44(4) (2018) 487-503.
- [12] W.H. Organization, Ageing, (2020).
- [13] S. Vanderschueren, D. Knockaert, T. Adriaenssens, W. Demey, A. Durnez, D. Blockmans, H. Bobbaers, From prolonged febrile illness to fever of unknown origin: the challenge continues, *Archives of internal medicine* 163(9) (2003) 1033-1041.
- [14] D. Heymann, J. Weisfeld, P. Webb, K. Johnson, T. Cairns, H. Berquist, Ebola hemorrhagic fever: Tandala, Zaire, 1977–1978, *Journal of Infectious Diseases* 142(3) (1980) 372-376.
- [15] N.I. Paton, Y.S. Leo, S.R. Zaki, A.P. Auchus, K.E. Lee, A.E. Ling, S.K. Chew, B. Ang, P.E. Rollin, T. Umapathi, Outbreak of Nipah-virus infection among abattoir workers in Singapore, *The Lancet* 354(9186) (1999) 1253-1256.
- [16] J. Lednicky, V.M.B. De Rochars, M. Elbadry, J. Loeb, T. Telisma, S. Chavannes, G. Anilis, E. Cella, M. Ciccozzi, B. Okech, Mayaro virus in child with acute febrile illness, Haiti, 2015, *Emerging infectious diseases* 22(11) (2016) 2000.

- [17] J.A. Lednicky, S.K. White, C.N. Mavian, M.A. El Badry, T. Telisma, M. Salemi, B.A. O'Kech, V.M. Beau De Rochars, J.G. Morris Jr, Emergence of Madariaga virus as a cause of acute febrile illness in children, Haiti, 2015-2016, *PLoS neglected tropical diseases* 13(1) (2019) e0006972.
- [18] B.M. Forshey, C. Guevara, V.A. Laguna-Torres, M. Cespedes, J. Vargas, A. Gianella, E. Vallejo, C. Madrid, N. Aguayo, E. Gotuzzo, Arboviral etiologies of acute febrile illnesses in Western South America, 2000–2007, *PLoS Negl Trop Dis* 4(8) (2010) e787.
- [19] M. Limper, I. Gerstenbluth, A. Duits, E. van Gorp, Epidemiology of febrile diseases in the emergency department of a Caribbean island: the Curaçao experience, *West indian medical journal* 61(1) (2012) 76-80.
- [20] A. Leelarasamee, C. Chupaprawan, M. Chenchittikul, S. Udompanthurat, Etiologies of acute undifferentiated febrile illness in Thailand, *J Med Assoc Thai* 87(5) (2004) 464-72.
- [21] A. Chrispal, H. Boorugu, K.G. Gopinath, S. Chandy, J.A.J. Prakash, E.M. Thomas, A.M. Abraham, O. Abraham, K. Thomas, Acute undifferentiated febrile illness in adult hospitalized patients: the disease spectrum and diagnostic predictors—an experience from a tertiary care hospital in South India, *Tropical doctor* 40(4) (2010) 230-234.
- [22] M.R. Kasper, P.J. Blair, S. Touch, B. Sokhal, C.Y. Yasuda, M. Williams, A.L. Richards, T.H. Burgess, T.F. Wierzbza, S.D. Putnam, Infectious etiologies of acute febrile illness among patients seeking health care in south-central Cambodia, *The American journal of tropical medicine and hygiene* 86(2) (2012) 246-253.
- [23] M.R. Capeding, M.N. Chua, S.R. Hadinegoro, I.I. Hussain, R. Nallusamy, P. Pitisuttithum, K. Rusmil, S.J. Thomas, N.H. Tran, D.N. Wirawan, Dengue and other common causes of acute febrile illness in Asia: an active surveillance study in children, *PLoS Negl Trop Dis* 7(7) (2013) e2331.
- [24] M.E. Reller, U. Akoroda, A. Nagahawatte, V. Devasiri, W. Kodikaarachchi, J.J. Strouse, R. Chua, A. Chow, O.M. Sessions, T. Østbye, Chikungunya as a cause of acute febrile illness in southern Sri Lanka, *PloS one* 8(12) (2013) e82259.
- [25] M.E. Reller, C. Bodinayake, A. Nagahawatte, V. Devasiri, W. Kodikara-Arachichi, J.J. Strouse, A. Broadwater, T. Østbye, A. de Silva, C.W. Woods, Unsuspected dengue and acute febrile illness in rural and semi-urban southern Sri Lanka, *Emerging infectious diseases* 18(2) (2012) 256.
- [26] C.N. Ma'roef, R. Dhenni, D. Megawati, A. Fadhillah, A. Lucanus, I.M. Artika, S. Masyeni, A. Lestarini, K. Sari, K. Suryana, Japanese encephalitis virus infection in non-encephalitic acute febrile illness patients, *PLoS neglected tropical diseases* 14(7) (2020) e0008454.
- [27] L.G. Tillekeratne, C.K. Bodinayake, R. Simmons, A. Nagahawatte, V. Devasiri, W. Kodikara Arachchi, B.P. Nicholson, L.P. Park, S. Vanderburg, R. Kurukulasooriya, Respiratory Viral Infection: An Underappreciated Cause of Acute Febrile Illness Admissions in Southern Sri Lanka, *The American journal of tropical medicine and hygiene* (2018) tpmd180699.
- [28] C. Mahende, B. Ngasala, J. Lusingu, A. Butichi, P. Lushino, M. Lemnge, Z. Premji, Aetiology of acute febrile episodes in children attending Korogwe District Hospital in north-eastern Tanzania, *PLoS One* 9(8) (2014) e104197.
- [29] L.K. Archibald, M.O.d. Dulk, K.J. Pallangyo, L.B. Reller, Fatal Mycobacterium tuberculosis bloodstream infections in febrile hospitalized adults in Dar es Salaam, Tanzania, *Clinical Infectious Diseases* 26(2) (1998) 290-296.
- [30] R.J. Schoepp, C.A. Rossi, S.H. Khan, A. Goba, J.N. Fair, Undiagnosed acute viral febrile illnesses, Sierra Leone, *Emerging infectious diseases* 20(7) (2014) 1176.
- [31] R.J. Schoepp, C.A. Rossi, S.H. Khan, A. Goba, J.N. Fair, Undiagnosed acute viral febrile illnesses, Sierra Leone, DTIC Document, 2014.
- [32] L.K. Archibald, M.O. den Dulk, K.J. Pallangyo, L.B. Reller, Fatal Mycobacterium tuberculosis bloodstream infections in febrile hospitalized adults in Dar es Salaam, Tanzania, *Clinical Infectious Diseases* 26(2) (1998) 290-296.

- [33] L.K. Archibald, L.C. McDonald, O. Nwanyanwu, P. Kazembe, H. Dobbie, J. Tokars, L.B. Reller, W.R. Jarvis, A hospital-based prevalence survey of bloodstream infections in febrile patients in Malawi: implications for diagnosis and therapy, *Journal of Infectious Diseases* 181(4) (2000) 1414-1420.
- [34] C. Hercik, L. Cosmas, O. Mogeni, N. Wamola, W. Kohi, C. Ochieng, C. Onyango, B. Fields, S. Mfinanga, J. Montgomery, A Diagnostic and Epidemiologic Investigation of Acute Febrile Illness (AFI) in Kilombero, Tanzania, *International Journal of Infectious Diseases* 53 (2016) 107-108.
- [35] J.A. Crump, A.B. Morrissey, W.L. Nicholson, R.F. Massung, R.A. Stoddard, R.L. Galloway, E.E. Ooi, V.P. Maro, W. Saganda, G.D. Kinabo, Etiology of severe non-malaria febrile illness in Northern Tanzania: a prospective cohort study, *PLoS neglected tropical diseases* 7(7) (2013) e2324.
- [36] A. Routray, A.P. Rath, S. Panigrahi, U.P. Lambe, S. Sahoo, S. Ganguly, RIFT Valley Fever: An Update, *International Journal of Contemporary Pathology* 3(1) (2017) 20-23.
- [37] L.K. McMullan, M. Frace, S.A. Sammons, T. Shoemaker, S. Balinandi, J.F. Wamala, J.J. Lutwama, R.G. Downing, U. Stroehrer, A. MacNeil, Using next generation sequencing to identify yellow fever virus in Uganda, *Virology* 422(1) (2012) 1-5.
- [38] L. Nyakarahuka, T.R. Shoemaker, S. Balinandi, G. Chemos, B. Kwesiga, S. Mulei, J. Kyondo, A. Tumusiime, A. Kofman, B. Masiira, Marburg virus disease outbreak in Kween District Uganda, 2017: Epidemiological and laboratory findings, *PLoS neglected tropical diseases* 13(3) (2019) e0007257.
- [39] P. GARNHAM, A. GLEDHILL, C. ANDREWES, I. MANN, A. ADAMS, DICK, EPIDEMIOLOGICAL NOTES ON SOME VIRUSES ISOLATED IN UGANDA-DISCUSSION, *TRANSACTIONS OF THE ROYAL SOCIETY OF TROPICAL MEDICINE AND HYGIENE* 47(1) (1953) 44-48.
- [40] J. HUNTER, An epidemic of bacillary dysentery in Kigezi District, Uganda, *East African medical journal* 24(2) (1947) 98-105.
- [41] C.J. Hackett, Yaws and Venereal Disease in Uganda, *British medical journal* 1(4489) (1947) 88.
- [42] C.f.D. Control, Prevention, Outbreak of Ebola hemorrhagic fever Uganda, August 2000-January 2001, *MMWR. Morbidity and mortality weekly report* 50(5) (2001) 73.
- [43] M. Lamunu, J. Lutwama, J. Kamugisha, A. Opio, J. Namboozee, N. Ndayimirije, S. Okware, Containing a haemorrhagic fever epidemic: the Ebola experience in Uganda (October 2000–January 2001), *International journal of infectious diseases* 8(1) (2004) 27-37.
- [44] L. Nyakarahuka, S. Balinandi, S. Mulei, J. Kyondo, A. Tumusiime, J. Klena, J. Lutwama, T. Shoemaker, Ten outbreaks of rift valley fever in Uganda 2016-2018: epidemiological and laboratory findings, *International Journal of Infectious Diseases* 79 (2019) 4.
- [45] T.R. Shoemaker, L. Nyakarahuka, S. Balinandi, J. Ojwang, A. Tumusiime, S. Mulei, J. Kyondo, B. Lubwama, M. Sekamatte, A. Namutebi, First Laboratory-Confirmed Outbreak of Human and Animal Rift Valley Fever Virus in Uganda in 48 Years, *The American journal of tropical medicine and hygiene* 100(3) (2019) 659-671.
- [46] W. Yang, M.J. Cummings, B. Bakamutumaho, J. Kayiwa, N. Owor, B. Namagambo, T. Byaruhanga, J.J. Lutwama, M.R. O'Donnell, J. Shaman, Transmission dynamics of influenza in two major cities of Uganda, *Epidemics* 24 (2018) 43-48.
- [47] M.J. Cummings, R. Tokarz, B. Bakamutumaho, J. Kayiwa, T. Byaruhanga, N. Owor, B. Namagambo, A. Wolf, B. Mathema, J.J. Lutwama, N.W. Schluger, W.I. Lipkin, M.R. O'Donnell, Precision Surveillance for Viral Respiratory Pathogens: Virome Capture Sequencing for the Detection and Genomic Characterization of Severe Acute Respiratory Infection in Uganda, *Clinical Infectious Diseases* 68(7) (2018) 1118-1125.
- [48] A.K. Mbonye, M. Sekamatte, Disease outbreaks and reporting in Uganda, *The Lancet* 392(10162) (2018) 2347-2348.
- [49] M. Sekamatte, V. Krishnasamy, L. Bulage, C. Kihembo, N. Nantima, F. Monje, D. Ndumu, J. Sentumbwe, B. Mbolanyi, R. Aruho, Multisectoral prioritization of zoonotic diseases in Uganda, 2017: A One Health perspective, *PloS one* 13(5) (2018) e0196799.

- [50] S. Balinandi, B. Bakamutumaho, J.T. Kayiwa, J. Ongus, J. Oundo, A.C. Awor, J.J. Lutwama, The viral aetiology of influenza-like illnesses in Kampala and Entebbe, Uganda, 2008, *African journal of laboratory medicine* 2(1) (2013).
- [51] M.J. Cummings, R. Tokarz, B. Bakamutumaho, J. Kayiwa, T. Byaruhanga, N. Owor, B. Namagambo, A. Wolf, B. Mathema, J.J. Lutwama, Precision surveillance for viral respiratory pathogens: virome capture sequencing for the detection and genomic characterization of severe acute respiratory infection in Uganda, *Clinical Infectious Diseases* 68(7) (2019) 1118-1125.
- [52] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, A novel coronavirus from patients with pneumonia in China, 2019, *New England journal of medicine* (2020).
- [53] T. Bousema, L. Okell, I. Felger, C. Drakeley, Asymptomatic malaria infections: detectability, transmissibility and public health relevance, *Nature Reviews Microbiology* 12(12) (2014) 833.
- [54] B. Namagambo, J. Kayiwa, T. Byaruhanga, N. Owor, I. Nabukenya, B. Bakamutumaho, J. Lutwama, Co-infection of malaria and influenza viruses in Uganda: A pilot study, *International Journal of Infectious Diseases* 21 (2014) 140.
- [55] A. Sow, C. Loucoubar, D. Diallo, O. Faye, Y. Ndiaye, C.S. Senghor, A.T. Dia, O. Faye, S.C. Weaver, M. Diallo, Concurrent malaria and arbovirus infections in Kedougou, southeastern Senegal, *Malaria journal* 15(1) (2016) 47.
- [56] S. Ahmad, M. Dhar, G. Mittal, N. Bhat, N. Shirazi, V. Kalra, H. Sati, V. Gupta, A comparative hospital-based observational study of mono-and co-infections of malaria, dengue virus and scrub typhus causing acute undifferentiated fever, *European Journal of Clinical Microbiology & Infectious Diseases* 35(4) (2016) 705-711.
- [57] N. Salam, S. Mustafa, A. Hafiz, A.A. Chaudhary, F. Deeba, S. Parveen, Global prevalence and distribution of coinfection of malaria, dengue and chikungunya: a systematic review, *BMC public health* 18(1) (2018) 710.
- [58] M. Theiler, W.G. Downs, The arthropod-borne viruses of vertebrates. An account of the Rockefeller Foundation virus program, 1951-1970, *The arthropod-borne viruses of vertebrates. An account of The Rockefeller Foundation Virus Program, 1951-1970.* (1973).
- [59] A. De Madrid, J. Porterfield, The flaviviruses (group B arboviruses): a cross-neutralization study, *Journal of General Virology* 23(1) (1974) 91-96.
- [60] C.H. Calisher, N. Karabatsos, J.M. Dalrymple, R.E. Shope, J.S. Porterfield, E.G. Westaway, W.E. Brandt, Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera, *Journal of general virology* 70(1) (1989) 37-43.
- [61] P.P. Marra, S.M. Griffing, R.G. McLean, West Nile virus and wildlife health, *Emerging infectious diseases* 9(7) (2003) 898.
- [62] R.G. McLean, West Nile Virus: emerging threat to public health and animal health, (2003).
- [63] E. Simulundu, K. Ndashe, H.M. Chambaro, D. Squarre, P.M. Reilly, S. Chitanga, K. Changula, A.N. Mukubesa, J. Ndebe, J. Tembo, West Nile Virus in Farmed Crocodiles, Zambia, 2019, *Emerging infectious diseases* 26(4) (2020) 811.
- [64] P.D. Burbelo, E.J. Dubovi, P. Simmonds, J.L. Medina, J.A. Henriquez, N. Mishra, J. Wagner, R. Tokarz, J.M. Cullen, M.J. Iadarola, Serology-enabled discovery of genetically diverse hepaciviruses in a new host, *Journal of virology* 86(11) (2012) 6171-6178.
- [65] A. Kapoor, P. Simmonds, G. Gerold, N. Qaisar, K. Jain, J.A. Henriquez, C. Firth, D.L. Hirschberg, C.M. Rice, S. Shields, Characterization of a canine homolog of hepatitis C virus, *Proceedings of the National Academy of Sciences* 108(28) (2011) 11608-11613.
- [66] J.N. Simons, T.J. Pilot-Matias, T.P. Leary, G.J. Dawson, S.M. Desai, G.G. Schlauder, A.S. Muerhoff, J.C. Erker, S.L. Buijk, M.L. Chalmers, Identification of two flavivirus-like genomes in the GB hepatitis agent, *Proceedings of the National academy of Sciences* 92(8) (1995) 3401-3405.

- [67] S.D. Sibley, M. Lauck, A.L. Bailey, D. Hyeroba, A. Tumukunde, G. Weny, C.A. Chapman, D.H. O'connor, T.L. Goldberg, T.C. Friedrich, Discovery and characterization of distinct simian pegiviruses in three wild African Old World monkey species, *PLoS One* 9(6) (2014) e98569.
- [68] J.F. Drexler, V.M. Corman, M.A. Müller, A.N. Lukashev, A. Gmyl, B. Coutard, A. Adam, D. Ritz, L.M. Leijten, D. Van Riel, Evidence for novel hepaciviruses in rodents, *PLoS Pathog* 9(6) (2013) e1003438.
- [69] P.-L. Quan, C. Firth, J.M. Conte, S.H. Williams, C.M. Zambrana-Torrel, S.J. Anthony, J.A. Ellison, A.T. Gilbert, I.V. Kuzmin, M. Niezgoda, Bats are a major natural reservoir for hepaciviruses and pegiviruses, *Proceedings of the National Academy of Sciences* 110(20) (2013) 8194-8199.
- [70] V.M. Corman, A. Grundhoff, C. Baechlein, N. Fischer, A. Gmyl, R. Wollny, D. Dei, D. Ritz, T. Binger, E. Adankwah, Highly divergent hepaciviruses from African cattle, *Journal of virology* 89(11) (2015) 5876-5882.
- [71] C. Baechlein, N. Fischer, A. Grundhoff, M. Alawi, D. Indenbirken, A. Postel, A.L. Baron, J. Offinger, K. Becker, A. Beineke, Identification of a novel hepacivirus in domestic cattle from Germany, *Journal of virology* 89(14) (2015) 7007-7015.
- [72] D.B. Smith, G. Meyers, J. Bukh, E.A. Gould, T. Monath, A.S. Muerhoff, A. Pletnev, R. Rico-Hesse, J.T. Stapleton, P. Simmonds, Proposed revision to the taxonomy of the genus Pestivirus, family Flaviviridae, *The Journal of general virology* 98(8) (2017) 2106.
- [73] P. Becher, M. Orlich, A. Shannon, G. Horner, H. Thiel, Phylogenetic analysis of pestiviruses from domestic and wild ruminants, *Journal of General Virology* 78(6) (1997) 1357-1366.
- [74] C.H. Calisher, E.A. Gould, Taxonomy of the virus family Flaviviridae, *Advances in virus research* 59 (2003) 1-20.
- [75] D.B. Smith, P. Becher, J. Bukh, E.A. Gould, G. Meyers, T. Monath, A.S. Muerhoff, A. Pletnev, R. Rico-Hesse, J.T. Stapleton, Proposed update to the taxonomy of the genera Hepacivirus and Pegivirus within the Flaviviridae family, *The Journal of general virology* 97(11) (2016) 2894.
- [76] P. Simmonds, P. Becher, J. Bukh, E.A. Gould, G. Meyers, T. Monath, S. Muerhoff, A. Pletnev, R. Rico-Hesse, D.B. Smith, ICTV virus taxonomy profile: Flaviviridae, *The Journal of general virology* 98(1) (2017) 2.
- [77] D.J. Gubler, Emerging vector-borne flavivirus diseases: are vaccines the solution?, *Expert review of vaccines* 10(5) (2011) 563-565.
- [78] J.D. Stanaway, D.S. Shepard, E.A. Undurraga, Y.A. Halasa, L.E. Coffeng, O.J. Brady, S.I. Hay, N. Bedi, I.M. Bensenor, C.A. Castañeda-Orjuela, The global burden of dengue: an analysis from the Global Burden of Disease Study 2013, *The Lancet infectious diseases* 16(6) (2016) 712-723.
- [79] W. WHO, Global hepatitis report 2017, Geneva: World Health Organization (2017).
- [80] S. Blach, S. Zeuzem, M. Manns, I. Altraif, A.-S. Duberg, D.H. Muljono, I. Waked, S.M. Alavian, M.-H. Lee, F. Negro, Global prevalence and genotype distribution of hepatitis C virus infection in 2015: a modelling study, *The lancet Gastroenterology & hepatology* 2(3) (2017) 161-176.
- [81] G. Umutesi, F. Shumbusho, F. Kateera, J. Serumondo, J. Kabahizi, E. Musabeyezu, A. Ngwije, N. Gupta, S. Nsanzimana, Rwanda launches a 5-year national hepatitis C elimination plan: A landmark in sub-Saharan Africa, *Journal of hepatology* 70(6) (2019) 1043-1045.
- [82] S. Makiala-Mandanda, F. Le Gal, N. Ngwaka-Matsung, S. Ahuka-Mundeke, R. Onanga, B. Bivigou-Mboumba, E. Pukuta-Simbu, A. Gerber, J.L. Abbate, D. Mwamba, High prevalence and diversity of hepatitis viruses in suspected cases of yellow fever in the Democratic Republic of Congo, *Journal of clinical microbiology* 55(5) (2017) 1299-1312.
- [83] M.J. Akiyama, C.M. Cleland, J.A. Lizcano, P. Cherutich, A.E. Kurth, Prevalence, estimated incidence, risk behaviours, and genotypic distribution of hepatitis C virus among people who inject drugs accessing harm-reduction services in Kenya: a retrospective cohort study, *The Lancet Infectious Diseases* 19(11) (2019) 1255-1263.

- [84] R.H. Kawambwa, M.V. Majigo, A.A. Mohamed, M.I. Matee, High prevalence of human immunodeficiency virus, hepatitis B and C viral infections among people who inject drugs: a potential stumbling block in the control of HIV and viral hepatitis in Tanzania, *BMC public health* 20(1) (2020) 177.
- [85] B.Y. Lee, J.A. Alfaro-Murillo, A.S. Parpia, L. Asti, P.T. Wedlock, P.J. Hotez, A.P. Galvani, The potential economic burden of Zika in the continental United States, *PLoS neglected tropical diseases* 11(4) (2017) e0005531.
- [86] T.H. Santa Rita, R.B. Barra, G.P. Peixoto, P.G. Mesquita, G.B. Barra, Association between suspected Zika virus disease during pregnancy and giving birth to a newborn with congenital microcephaly: a matched case–control study, *BMC research notes* 10(1) (2017) 1-7.
- [87] F.U. Ahmed, C.B. Mahmood, J.D. Sharma, S.M. Hoque, R. Zaman, M.S. Hasan, Dengue and Dengue Haemorrhagic Fever in children during the 2000 outbreak in Chittagong, Bangladesh, (2001).
- [88] D.J. Gubler, G.G. Clark, Dengue/dengue hemorrhagic fever: the emergence of a global health problem, *Emerging infectious diseases* 1(2) (1995) 55.
- [89] G. Malavige, S. Fernando, D. Fernando, S. Seneviratne, Dengue viral infections, *Postgraduate medical journal* 80(948) (2004) 588-601.
- [90] H.Y. Wang, T. Takasaki, S.H. Fu, X.H. Sun, H.L. Zhang, Z.X. Wang, Z.Y. Hao, J.K. Zhang, Q. Tang, A. Kotaki, Molecular epidemiological analysis of Japanese encephalitis virus in China, *Journal of General Virology* 88(3) (2007) 885-894.
- [91] M. Velasco, M.P. Sánchez-Seco, C. Campelo, F. de Ory, O. Martin, L. Herrero, O.J. Salmerón Béliz, T. Minguito, M. Campos, F. Molero, Imported Human West Nile Virus Lineage 2 Infection in Spain: Neurological and Gastrointestinal Complications, *Viruses* 12(2) (2020) 156.
- [92] J.H. Hoofnagle, Hepatitis C: the clinical spectrum of disease, *Hepatology* 26(S3) (1997) 15S-20S.
- [93] C.M. Meyers, L.B. Seeff, C.O. Stehman-Breen, J.H. Hoofnagle, Hepatitis C and renal disease: an update, *American journal of kidney diseases* 42(4) (2003) 631-657.
- [94] M.J. Alter, Epidemiology of hepatitis C, *Hepatology* 26(S3) (1997) 62S-65S.
- [95] E. D’Ortenzio, S. Matheron, X. de Lamballerie, B. Hubert, G. Piorkowski, M. Maquart, D. Descamps, F. Damond, Y. Yazdanpanah, I. Leparç-Goffart, Evidence of sexual transmission of Zika virus, *New England Journal of Medicine* 374(22) (2016) 2195-2198.
- [96] I.J. Motta, B.R. Spencer, S.G. Cordeiro da Silva, M.B. Arruda, J.A. Dobbin, Y.B. Gonzaga, I.P. Arcuri, R.C. Tavares, E.H. Atta, R.F. Fernandes, Evidence for transmission of Zika virus by platelet transfusion, *New England Journal of Medicine* 375(11) (2016) 1101-1103.
- [97] L.M. Katz, S.N. Rossmann, Zika and the blood supply: a work in progress, *Archives of pathology & laboratory medicine* 141(1) (2017) 85-92.
- [98] W.A. Sawyer, S. Kitchen, W. Lloyd, Vaccination of humans against yellow fever with immune serum and virus fixed for mice, *Proceedings of the Society for Experimental Biology and Medicine* 29(1) (1931) 62-64.
- [99] C. Durieux, Yellow fever vaccination, Geneva, Switzerland: World Health Organization (1956) 31-66.
- [100] M. Theiler, H.H. Smith, The use of yellow fever virus modified by in vitro cultivation for human immunization, *The Journal of experimental medicine* 65(6) (1937) 787.
- [101] H.H. Smith, H. Penna, A. Paoliello, Yellow Fever Vaccination with Cultured Virus (17D) without Immune Serum<sup>1</sup>, *The American Journal of Tropical Medicine and Hygiene* 1(5) (1938) 437-468.
- [102] C. Kunz, H. Hofmann, A. Sary, Field studies with a new tick-borne encephalitis (TBE) vaccine, *Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene. Erste Abteilung Originale. Reihe A: Medizinische Mikrobiologie und Parasitologie* 234(1) (1976) 141-144.
- [103] Y.Y. Xin, Z.G. Ming, G.Y. Peng, A. Jian, L.H. Min, Safety of a live-attenuated Japanese encephalitis virus vaccine (SA14-14-2) for children, *The American journal of tropical medicine and hygiene* 39(2) (1988) 214-217.

- [104] C.H. Hoke, A. Nisalak, N. Sangawhipa, S. Jatanasen, T. Laorakapongse, B.L. Innis, S.-o. Kotchasenee, J.B. Gingrich, J. Latendresse, K. Fukai, Protection against Japanese encephalitis by inactivated vaccines, *New England Journal of Medicine* 319(10) (1988) 608-614.
- [105] T. Ng, D. Hathaway, N. Jennings, D. Champ, Y. Chiang, H. Chu, Equine vaccine for West Nile virus, *Developments in biologicals* 114 (2003) 221-227.
- [106] H. El Garch, J. Minke, J. Rehder, S. Richard, C.E. Toulemonde, S. Dinic, C. Andreoni, J. Audonnet, R. Nordgren, V. Juillard, A West Nile virus (WNV) recombinant canarypox virus vaccine elicits WNV-specific neutralizing antibodies and cell-mediated immune responses in the horse, *Veterinary immunology and immunopathology* 123(3-4) (2008) 230-239.
- [107] R. Biedendbender, J. Bevilacqua, A.M. Gregg, M. Watson, G. Dayan, Phase II, randomized, double-blind, placebo-controlled, multicenter study to investigate the immunogenicity and safety of a West Nile virus vaccine in healthy adults, *Journal of Infectious Diseases* 203(1) (2011) 75-84.
- [108] T. Asselah, P. Marcellin, New direct-acting antivirals' combination for the treatment of chronic hepatitis C, *Liver International* 31 (2011) 68-77.
- [109] A. Aghemo, R. De Francesco, New horizons in hepatitis C antiviral therapy with direct-acting antivirals, *Hepatology* 58(1) (2013) 428-438.
- [110] J.D. Duncan, R.A. Urbanowicz, A.W. Tarr, J.K. Ball, Hepatitis C virus vaccine: Challenges and prospects, *Vaccines* 8(1) (2020) 90.
- [111] M. Lechmann, K. Murata, J. Satoi, J. Vergalla, T.F. Baumert, T.J. Liang, Hepatitis C virus-like particles induce virus-specific humoral and cellular immune responses in mice, *Hepatology* 34(2) (2001) 417-423.
- [112] S.-H. Jeong, M. Qiao, M. Nascimbeni, Z. Hu, B. Rehmann, K. Murthy, T.J. Liang, Immunization with hepatitis C virus-like particles induces humoral and cellular immune responses in nonhuman primates, *Journal of Virology* 78(13) (2004) 6995-7003.
- [113] S.E. Frey, M. Houghton, S. Coates, S. Abrignani, D. Chien, D. Rosa, P. Pileri, R. Ray, A.M. Di Bisceglie, P. Rinella, Safety and immunogenicity of HCV E1E2 vaccine adjuvanted with MF59 administered to healthy adults, *Vaccine* 28(38) (2010) 6367-6373.
- [114] L. Swadling, S. Capone, R.D. Antrobus, A. Brown, R. Richardson, E.W. Newell, J. Halliday, C. Kelly, D. Bowen, J. Fergusson, A human vaccine strategy based on chimpanzee adenoviral and MVA vectors that primes, boosts, and sustains functional HCV-specific T cell memory, *Science translational medicine* 6(261) (2014) 261ra153-261ra153.
- [115] L. Schmidt, O. Sehic, C. Wild, EU FP7 research funding for an orphan drug (Orfadin®) and vaccine (Hep C) development: a success and a failure?, *Journal of Pharmaceutical Policy and Practice* 14(1) (2021) 1-7.
- [116] M. Matičič, A. Lombardi, M.U. Mondelli, M. Colombo, Elimination of hepatitis C in Europe: can WHO targets be achieved?, *Clinical Microbiology and Infection* 26(7) (2020) 818-823.
- [117] I. Reimann, K. Depner, S. Trapp, M. Beer, An avirulent chimeric Pestivirus with altered cell tropism protects pigs against lethal infection with classical swine fever virus, *Virology* 322(1) (2004) 143-157.
- [118] T.B. Rasmussen, Å. Uttenthal, I. Reimann, J. Nielsen, K. Depner, M. Beer, Virulence, immunogenicity and vaccine properties of a novel chimeric pestivirus, *Journal of general virology* 88(2) (2007) 481-486.
- [119] A. Lwoff, P. Tournier, The classification of viruses, *Annual review of microbiology* 20 (1966) 45-74.
- [120] P.C. Woo, S.K. Lau, G.K. Choi, Y. Huang, J.L. Teng, H.-W. Tsoi, H. Tse, M.L. Yeung, K.-H. Chan, D.-Y. Jin, Natural occurrence and characterization of two internal ribosome entry site elements in a novel virus, canine picodistrovirus, in the picornavirus-like superfamily, *Journal of virology* 86(5) (2012) 2797-2808.
- [121] M.G. Rossmann, J.E. Johnson, Icosahedral RNA virus structure, *Annual review of biochemistry* 58(1) (1989) 533-569.
- [122] R. Zell, Picornaviridae—the ever-growing virus family, *Archives of virology* 163(2) (2018) 299-317.
- [123] S.M. Lemon, J.W. Leduc, L.N. Binn, A. Escajadillo, K.G. Ishak, Transmission of hepatitis A virus among recently captured Panamanian owl monkeys, *Journal of medical virology* 10(1) (1982) 25-36.

- [124] M. Yin-Murphy, J.W. Almond, Picornaviruses, Medical Microbiology. 4th edition (1996).
- [125] M.S. Jones, V.V. Lukashov, R.D. Ganac, D.P. Schnurr, Discovery of a novel human picornavirus in a stool sample from a pediatric patient presenting with fever of unknown origin, *Journal of clinical microbiology* 45(7) (2007) 2144-2150.
- [126] M.S. Oberste, M.M. Feeroz, K. Maher, W.A. Nix, G.A. Engel, K.M. Hasan, S. Begum, G. Oh, A.H. Chowdhury, M.A. Pallansch, Characterizing the picornavirus landscape among synanthropic nonhuman primates in Bangladesh, 2007 to 2008, *Journal of virology* 87(1) (2013) 558-571.
- [127] E.C. Scherbatskoy, K. Subramaniam, L. Al-Hussiney, K. Imnoi, P.M. Thompson, V.L. Popov, T.F.F. Ng, K.L. Kelley, R. Alvarado, J.C. Wolf, Characterization of a novel picornavirus isolated from moribund aquacultured clownfish, *Journal of General Virology* 101(7) (2020) 735-745.
- [128] S.K. Lau, P.C. Woo, C.C. Yip, K.S. Li, R.Y. Fan, R. Bai, Y. Huang, K.-H. Chan, K.-Y. Yuen, Chickens host diverse picornaviruses originated from potential interspecies transmission with recombination, *Journal of General Virology* 95(9) (2014) 1929-1944.
- [129] O.O. Alegbeleye, I. Singleton, A.S. Sant'Ana, Sources and contamination routes of microbial pathogens to fresh produce during field cultivation: a review, *Food microbiology* 73 (2018) 177-208.
- [130] R.D. Schubert, I.A. Hawes, P.S. Ramachandran, A. Ramesh, E.D. Crawford, J.E. Pak, W. Wu, C.K. Cheung, B.D. O'Donovan, C.M. Tato, Pan-viral serology implicates enteroviruses in acute flaccid myelitis, *Nature medicine* 25(11) (2019) 1748-1752.
- [131] L. Bubba, E.K. Broberg, A. Jasir, P. Simmonds, H. Harvala, M. Redlberger-Fritz, L. Nikolaeva-Glomb, M. Havlíčková, P. Rainetova, T.K. Fischer, Circulation of non-polio enteroviruses in 24 EU and EEA countries between 2015 and 2017: a retrospective surveillance study, *The Lancet Infectious Diseases* 20(3) (2020) 350-361.
- [132] A. Enfissi, M.-L. Joffret, D. Delaune, F. Delpeyroux, D. Rousset, M. Bessaud, Coxsackievirus A24 variant associated with acute haemorrhagic conjunctivitis cases, French Guiana, 2017, *Intervirology* 60 (2017) 271-275.
- [133] W.-C.G. Yeung, W.D. Rawlinson, M.E. Craig, Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies, *Bmj* 342 (2011).
- [134] O.H. Laitinen, H. Honkanen, O. Pakkanen, S. Oikarinen, M.M. Hankaniemi, H. Huhtala, T. Ruokoranta, V. Lecouturier, P. André, R. Harju, Coxsackievirus B1 is associated with induction of  $\beta$ -cell autoimmunity that portends type 1 diabetes, *Diabetes* 63(2) (2014) 446-455.
- [135] D.T. Haydon, R.R. Kao, R.P. Kitching, The UK foot-and-mouth disease outbreak—the aftermath, *Nature Reviews Microbiology* 2(8) (2004) 675-681.
- [136] D. Thompson, P. Muriel, D. Russell, P. Osborne, A. Bromley, M. Rowland, S. Creigh-Tyte, C. Brown, Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001, *Revue scientifique et technique (International Office of Epizootics)* 21(3) (2002) 675-687.
- [137] J.E. Salk, U. Krech, J. Youngner, B.L. Bennett, L. Lewis, P. Bazeley, Formaldehyde treatment and safety testing of experimental poliomyelitis vaccines, *American Journal of Public Health and the Nations Health* 44(5) (1954) 563-570.
- [138] A.B. Sabin, L. Boulger, History of Sabin attenuated poliovirus oral live vaccine strains, *Journal of Biological Standardization* 1(2) (1973) 115-118.
- [139] R. Li, L. Liu, Z. Mo, X. Wang, J. Xia, Z. Liang, Y. Zhang, Y. Li, Q. Mao, J. Wang, An inactivated enterovirus 71 vaccine in healthy children, *New England Journal of Medicine* 370(9) (2014) 829-837.
- [140] F. Zhu, W. Xu, J. Xia, Z. Liang, Y. Liu, X. Zhang, X. Tan, L. Wang, Q. Mao, J. Wu, Efficacy, safety, and immunogenicity of an enterovirus 71 vaccine in China, *New England Journal of Medicine* 370(9) (2014) 818-828.
- [141] K.L. Aw-Yong, N.M.N. NikNadia, C.W. Tan, I.C. Sam, Y.F. Chan, Immune responses against enterovirus A71 infection: Implications for vaccine success, *Reviews in medical virology* 29(5) (2019) e2073.

- [142] H. Cassidy, R. Poelman, M. Knoester, V. Leer-Buter, C. Coretta, H.G. Niesters, Enterovirus D68—the new polio?, *Frontiers in microbiology* 9 (2018) 2677.
- [143] A.L. Greninger, S.N. Naccache, K. Messacar, A. Clayton, G. Yu, S. Somasekar, S. Federman, D. Stryke, C. Anderson, S. Yagi, A novel outbreak enterovirus D68 strain associated with acute flaccid myelitis cases in the USA (2012–14): a retrospective cohort study, *The Lancet Infectious Diseases* 15(6) (2015) 671-682.
- [144] W. Dai, C. Zhang, X. Zhang, P. Xiong, Q. Liu, S. Gong, L. Geng, D. Zhou, Z. Huang, A virus-like particle vaccine confers protection against enterovirus D68 lethal challenge in mice, *Vaccine* 36(5) (2018) 653-659.
- [145] A. Raychoudhuri, A.K. Naru, S.R. Kanubothula, R. Uddala, Development of an experimental inactivated vaccine from Vero cell adapted Enterovirus D68, *Virus Research* (2021) 198528.
- [146] G.R. McLean, Developing a vaccine for human rhinoviruses, *Journal of vaccines & immunization* 2(3) (2014) 16.
- [147] S. Lee, M.T. Nguyen, M.G. Currier, J.B. Jenkins, E.A. Strobert, A.E. Kajon, R. Madan-Lala, Y.A. Bochkov, J.E. Gern, K. Roy, A polyvalent inactivated rhinovirus vaccine is broadly immunogenic in rhesus macaques, *Nature communications* 7(1) (2016) 1-7.
- [148] R. Clemens, A. Safary, A. Hepburn, C. Roche, W.J. Stanbury, F.E. André, Clinical experience with an inactivated hepatitis A vaccine, *Journal of infectious diseases* 171(Supplement\_1) (1995) S44-S49.
- [149] S. Parida, Vaccination against foot-and-mouth disease virus: strategies and effectiveness, *Expert review of vaccines* 8(3) (2009) 347-365.
- [150] K. Schaaf, Avian encephalomyelitis immunization with inactivated virus, *Avian Diseases* 3(3) (1959) 245-256.
- [151] I. Wanyama, S. Leitner, A review on health and environmental aspects of current manure management practices in pig production systems in Uganda, (2019).
- [152] E.J. Scully, S. Basnet, R.W. Wrangham, M.N. Muller, E. Otali, D. Hyeroba, K.A. Grindle, T.E. Pappas, M.E. Thompson, Z. Machanda, Lethal respiratory disease associated with human rhinovirus C in wild chimpanzees, Uganda, 2013, *Emerging infectious diseases* 24(2) (2018) 267.
- [153] E. O'Brien, J. Nakyzze, H. Wu, N. Kiwanuka, W. Cunningham, J.B. Kaneene, I. Xagorarakis, Viral diversity and abundance in polluted waters in Kampala, Uganda, *Water research* 127 (2017) 41-49.
- [154] O. Soběšlavský, S. Sebikari, P. Harland, N. Skrtić, O. Fayinka, A. Soneji, The viral etiology of acute respiratory infections in children in Uganda, *Bulletin of the World Health Organization* 55(5) (1977) 625.
- [155] P. Tushabe, W. Howard, J. Bwogi, M. Birungi, J.P. Eliku, P. Kakooza, H. Bukenya, P. Namuwulya, J. Gaizi, M. Tibanagwa, Molecular characterization of non-polio enteroviruses isolated from acute flaccid paralysis patients in Uganda, *Journal of Medical Virology* (2021).
- [156] A. Ramesh, S. Nakielny, J. Hsu, M. Kyohere, O. Byaruhanga, C. de Bourcy, R. Egger, B. Dimitrov, Y.-F. Juan, J. Sheu, Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda, *PLoS One* 14(6) (2019) e0218318.
- [157] B. Rima, A. Balkema-Buschmann, W.G. Dundon, P. Duprex, A. Easton, R. Fouchier, G. Kurath, R. Lamb, B. Lee, P. Rota, ICTV virus taxonomy profile: Paramyxoviridae, *The Journal of general virology* 100(12) (2019) 1593.
- [158] A.S. Samuel, S. Kumar, S. Madhuri, P.L. Collins, S.K. Samal, Complete sequence of the genome of avian paramyxovirus type 9 and comparison with other paramyxoviruses, *Virus research* 142(1-2) (2009) 10-18.
- [159] H.-J. Lee, J.-Y. Kim, Y.-J. Lee, E.-K. Lee, B.-M. Song, H.-S. Lee, K.-S. Choi, A novel avian paramyxovirus (putative serotype 15) isolated from wild birds, *Frontiers in microbiology* 8 (2017) 786.
- [160] M. Shi, X.-D. Lin, X. Chen, J.-H. Tian, L.-J. Chen, K. Li, W. Wang, J.-S. Eden, J.-J. Shen, L. Liu, The evolutionary history of vertebrate RNA viruses, *Nature* 556(7700) (2018) 197-202.
- [161] L.-F. Wang, E. Hansson, M. Yu, K. Chua, N. Mathe, G. Crameri, B. Rima, J. Moreno-Lopez, B. Eaton, Full-length genome sequence and genetic relationship of two paramyxoviruses isolated from bat and pigs in the Americas, *Archives of virology* 152(7) (2007) 1259.

- [162] K.S. Baker, S. Todd, G.A. Marsh, G. Cramer, J. Barr, A.O. Kamins, A.J. Peel, M. Yu, D.T. Hayman, B. Nadjm, Novel, potentially zoonotic paramyxoviruses from the African straw-colored fruit bat *Eidolon helvum*, *Journal of virology* 87(3) (2013) 1348-1358.
- [163] M. Kawano, H. Bando, T. Yuasa, K. Kondo, M. Tsurudome, H. Komada, M. Nishio, Y. Ito, Sequence determination of the hemagglutinin-neuraminidase (HN) gene of human parainfluenza type 2 virus and the construction of a phylogenetic tree for HN proteins of all the paramyxoviruses that are infectious to humans, *Virology* 174(1) (1990) 308-313.
- [164] C.G. Albariño, M. Foltzer, J.S. Towner, L.A. Rowe, S. Campbell, C.M. Jaramillo, B.H. Bird, D.M. Reeder, M.E. Vodzak, P. Rota, Novel paramyxovirus associated with severe acute febrile disease, South Sudan and Uganda, 2012, *Emerging infectious diseases* 20(2) (2014) 211.
- [165] Y.I. Wolf, D. Kazlauskas, J. Iranzo, A. Lucía-Sanz, J.H. Kuhn, M. Krupovic, V.V. Dolja, E.V. Koonin, Origins and evolution of the global RNA virome, *MBio* 9(6) (2018).
- [166] V.V. Dolja, E.V. Koonin, Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer, *Virus research* 244 (2018) 36-52.
- [167] S.L. Gorbach, S. Baron, *Medical microbiology, Microbiology of the gastrointestinal tract* (4th Edition, Chapter 95). University of Texas Medical Branch at Galveston, Galveston (TX), Retrieved from NCBI (1996).
- [168] G.Y.S. Park, K. Tishkowski, Paramyxovirus, *StatPearls* [Internet] (2021).
- [169] K.B. Chua, Nipah virus outbreak in Malaysia, *Journal of Clinical Virology* 26(3) (2003) 265-275.
- [170] P. Chatterjee, Nipah virus outbreak in India, *The Lancet* 391(10136) (2018) 2200.
- [171] H. Field, K. Schaaf, N. Kung, C. Simon, D. Waltisbuhl, H. Hobert, F. Moore, D. Middleton, A. Crook, G. Smith, Hendra virus outbreak with novel clinical features, Australia, *Emerging infectious diseases* 16(2) (2010) 338.
- [172] L.M. Looi, K.B. Chua, Lessons from the Nipah virus outbreak in Malaysia, *Malaysian Journal of Pathology* 29(2) (2007) 63-67.
- [173] E.G. Playford, B. McCall, G. Smith, V. Slinko, G. Allen, I. Smith, F. Moore, C. Taylor, Y.-H. Kung, H. Field, Human Hendra virus encephalitis associated with equine outbreak, Australia, 2008, *Emerging infectious diseases* 16(2) (2010) 219.
- [174] Y.C. Li, W.Z. Bai, T. Hashikawa, The neuroinvasive potential of SARS-CoV2 may play a role in the respiratory failure of COVID-19 patients, *Journal of medical virology* 92(6) (2020) 552-555.
- [175] B.S. Ang, T.C. Lim, L. Wang, Nipah virus infection, *Journal of clinical microbiology* 56(6) (2018).
- [176] K.N. Bossart, Z. Zhu, D. Middleton, J. Klippel, G. Cramer, J. Bingham, J.A. McEachern, D. Green, T.J. Hancock, Y.-P. Chan, A neutralizing human monoclonal antibody protects against lethal disease in a new ferret model of acute nipah virus infection, *PLoS pathogens* 5(10) (2009) e1000642.
- [177] K.N. Bossart, T.W. Geisbert, H. Feldmann, Z. Zhu, F. Feldmann, J.B. Geisbert, L. Yan, Y.-R. Feng, D. Brining, D. Scott, A neutralizing human monoclonal antibody protects african green monkeys from hendra virus challenge, *Science translational medicine* 3(105) (2011) 105ra103-105ra103.
- [178] H.T. Chong, A. Kamarulzaman, C.T. Tan, K.J. Goh, T. Thayaparan, S.R. Kunjapan, N.K. Chew, K.B. Chua, S.K. Lam, Treatment of acute Nipah encephalitis with ribavirin, *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* 49(6) (2001) 810-813.
- [179] S. Chakrabarti, K. Collingham, K. Holder, C. Fegan, H. Osman, D. Milligan, Pre-emptive oral ribavirin therapy of paramyxovirus infections after haematopoietic stem cell transplantation: a pilot study, *Bone marrow transplantation* 28(8) (2001) 759-763.
- [180] D.E. Griffin, Measles vaccine, *Viral immunology* 31(2) (2018) 86-95.
- [181] J. Wise, MMR vaccine: Johnson urges new impetus to increase uptake as UK loses measles-free status, *British Medical Journal Publishing Group*, 2019.
- [182] J. Gramigna, 4 European countries lose measles-free status; US on the brink, *Infectious Diseases in Children* 32(9) (2019) 7-7.

- [183] M.N. Litvoc, M.I.B.F. Lopes, From the measles-free status to the current outbreak in Brasil, *Revista da Associação Médica Brasileira* 65(10) (2019) 1229-1230.
- [184] A. Muwonge, M. Nanyunja, P.A. Rota, J. Bwogi, L. Lowe, S.L. Liffick, W.J. Bellini, S. Sylvester, New measles genotype, Uganda, *Emerging infectious diseases* 11(10) (2005) 1522.
- [185] J. Rota, L. Lowe, P. Rota, W. Bellini, S. Redd, G. Dayan, R. Van Binnendijk, S. Hahné, G. Tipples, J. Macey, Identical genotype B3 sequences from measles patients in 4 countries, 2005, *Emerging infectious diseases* 12(11) (2006) 1779.
- [186] F.N. Baliraine, J. Bwogi, H. Bukenya, R. Seguya, T. Kabaliisa, A. Kisakye, W.B. Mbabazi, S.B. Smit, Possible interruption of measles virus transmission, Uganda, 2006–2009, *Emerging infectious diseases* 17(1) (2011) 110.
- [187] R. Mafigiri, F. Nsubuga, A.R. Ario, Risk factors for measles death: Kyegegwa District, western Uganda, February–September, 2015, *BMC infectious diseases* 17(1) (2017) 1-7.
- [188] J. Nguna, J.K. Bbaale, D. Birungi, B. Kwesiga, D. Kadobera, B.T. Opar, I. Ampaire, H. Luuze, B. Kwesiga, W. Mwanje, Costs of Responding to a Measles Outbreak: Buvuma Islands, Lake Victoria, Uganda, February-May 2017, (2019).
- [189] C. Biribawa, J.A. Atuhairwe, L. Bulage, D.O. Okethwangu, B. Kwesiga, A.R. Ario, B.-P. Zhu, Measles outbreak amplified in a pediatric ward: Lyantonde District, Uganda, August 2017, *BMC infectious diseases* 20 (2020) 1-8.
- [190] D. Bishai, B. Johns, D. Nair, J. Nabyonga-Orem, B. Fiona-Makmot, E. Simons, A. Dabbagh, The cost-effectiveness of supplementary immunization activities for measles: a stochastic model for Uganda, *The Journal of infectious diseases* 204(suppl\_1) (2011) S107-S115.
- [191] C.H. Calisher, N. Karabatsos, Arbovirus serogroups: definition and geographic distribution, *The arboviruses: epidemiology and ecology*, CRC Press 2020, pp. 19-58.
- [192] S.C. Weaver, R. Winegar, I.D. Manger, N.L. Forrester, Alphaviruses: population genetics and determinants of emergence, *Antiviral research* 94(3) (2012) 242-257.
- [193] M.A. Zacks, S. Paessler, Encephalitic alphaviruses, *Veterinary microbiology* 140(3-4) (2010) 281-286.
- [194] A. Suhrbier, M.-C. Jaffar-Bandjee, P. Gasque, Arthritogenic alphaviruses—an overview, *Nature Reviews Rheumatology* 8(7) (2012) 420-429.
- [195] A. Delfraro, A. Burgueño, N. Morel, G. González, A. García, J. Morelli, W. Pérez, H. Chiparelli, J. Arbiza, Fatal human case of Western equine encephalitis, Uruguay, *Emerging infectious diseases* 17(5) (2011) 952.
- [196] P.N. Yergolkar, B.V. Tandale, V.A. Arankalle, P.S. Sathe, S. AB, S.S. Gandhe, M.D. Gokhle, G.P. Jacob, S.L. Hundekar, A.C. Mishra, Chikungunya outbreaks caused by African genotype, India, *Emerging infectious diseases* 12(10) (2006) 1580.
- [197] L.C. Ng, H.C. Hapuarachchi, Tracing the path of Chikungunya virus—evolution and adaptation, *Infection, Genetics and Evolution* 10(7) (2010) 876-885.
- [198] D.L. Posey, T. O'ROURKE, J.T. Roehrig, R.S. Lanciotti, M. Weinberg, S. Maloney, O'nyong-nyong fever in West Africa, *The American journal of tropical medicine and hygiene* 73(1) (2005) 32-32.
- [199] M. Brummer-Korvenkontio, O. Vapalahti, P. Kuusisto, P. Saikku, T. Manni, P. Koskela, T. Nygren, H. Brummer-Korvenkontio, A. Vaheri, Epidemiology of Sindbis virus infections in Finland 1981–96: possible factors explaining a peculiar disease pattern, *Epidemiology & Infection* 129(2) (2002) 335-345.
- [200] S. Naish, W. Hu, K. Mengersen, S. Tong, Spatio-temporal patterns of Barmah Forest virus disease in Queensland, Australia, *PloS one* 6(10) (2011) e25688.
- [201] A. Suhrbier, M. La Linn, Clinical and pathologic aspects of arthritis due to Ross River virus and other alphaviruses, *Current opinion in rheumatology* 16(4) (2004) 374-379.
- [202] J.-J. Hoarau, M.-C.J. Bandjee, P.K. Trotot, T. Das, G. Li-Pat-Yuen, B. Dassa, M. Denizot, E. Guichard, A. Ribera, T. Henni, Persistent chronic inflammation and infection by Chikungunya arthritogenic alphavirus in spite of a robust host immune response, *The Journal of Immunology* 184(10) (2010) 5914-5927.

- [203] N. Wauquier, P. Becquart, D. Nkoghe, C. Padilla, A. Ndjoi-Mbiguino, E.M. Leroy, The acute phase of Chikungunya virus infection in humans is associated with strong innate immunity and T CD8 cell activation, *Journal of Infectious Diseases* 204(1) (2011) 115-123.
- [204] M.P. Sánchez-Seco, D. Rosario, E. Quiroz, G. Guzmán, A. Tenorio, A generic nested-RT-PCR followed by sequencing for detection and identification of members of the alphavirus genus, *Journal of Virological Methods* 95(1-2) (2001) 153-161.
- [205] S. Torii, Y. Orba, B.M. Hang'ombe, A.S. Mweene, Y. Wada, P.D. Anindita, W. Phongphaew, Y. Qiu, M. Kajihara, A. Mori-Kajihara, Discovery of Mwinilunga alphavirus: a novel alphavirus in *Culex* mosquitoes in Zambia, *Virus research* 250 (2018) 31-36.
- [206] N.J. MacLachlan, A.J. Guthrie, Re-emergence of bluetongue, African horse sickness, and other orbivirus diseases, *Veterinary research* 41(6) (2010) 35.
- [207] P. ROY, Orbivirus structure and assembly, *Virology* 216(1) (1996) 1-11.
- [208] M. Henning, African Horse Sickness, Perdesiekte, Pestis Equorum. Animal Diseases of South. Africa, Central News Agency Ltd., Pretoria, 1956.
- [209] N.J. MacLachlan, W.C. Wilson, B.M. Crossley, C.E. Mayo, D.C. Jasperson, R.E. Breitmeyer, A.M. Whiteford, Novel serotype of bluetongue virus, western North America, *Emerging infectious diseases* 19(4) (2013) 665.
- [210] H. Attoui, M.R. Mendez-Lopez, S. Rao, A. Hurtado-Alendes, F. Lizaraso-Caparo, F.M. Jaafar, A.R. Samuel, M. Belhouchet, L.I. Pritchard, L. Melville, Peruvian horse sickness virus and Yunnan orbivirus, isolated from vertebrates and mosquitoes in Peru and Australia, *Virology* 394(2) (2009) 298-310.
- [211] F.J. Conraths, J.M. Gethmann, C. Staubach, T.C. Mettenleiter, M. Beer, B. Hoffmann, Epidemiology of bluetongue virus serotype 8, Germany, *Emerging infectious diseases* 15(3) (2009) 433.
- [212] N. Zhang, N.J. MacLachlan, K. Bonneau, J. Zhu, Z. Li, K. Zhang, F. Zhang, L. Xia, W. Xiang, Identification of seven serotypes of bluetongue virus from the People's Republic of China, *Veterinary Record* 145(15) (1999) 427-429.
- [213] H. Els, Electron microscopy of bluetongue virus RNA, (1973).
- [214] H. Yang, W. Gu, Z. Li, L. Zhang, D. Liao, J. Song, B. Shi, J. Hasimu, Z. Li, Z. Yang, Novel putative Bluetongue virus serotype 29 isolated from inapparently infected goat in Xinjiang of China, *Transboundary and Emerging Diseases* (2021).
- [215] P. Mertens, S. Maan, A. Samuel, H. Attoui, Orbivirus, reoviridae, *Virus taxonomy* 466 (2005) 483.
- [216] A. Theiler, African Horse Sickness (Pestis equorum), Union S. Africa Dept. Agric., Pretoria, Sci. Bull. (19) (1921).
- [217] O. Tomori, A. Fabiyi, Neutralizing antibodies to Orungo virus in man and animals in Nigeria, *Tropical and geographical medicine* 28(3) (1976) 233-238.
- [218] A.C. Fagre, J.S. Lee, R.M. Kityo, N.A. Bergren, E.C. Mossel, T. Nakayiki, B. Nalikka, L. Nyakarahuka, A.T. Gilbert, J.K. Peterhans, Discovery and characterization of Bukakata orbivirus (Reoviridae: Orbivirus), a novel virus from a Ugandan bat, *Viruses* 11(3) (2019) 209.
- [219] C. Mayo, J. Lee, J. Kopanke, N.J. MacLachlan, A review of potential bluetongue virus vaccine strategies, *Veterinary microbiology* 206 (2017) 84-90.
- [220] S.J. Dennis, A.E. Meyers, I.I. Hitzeroth, E.P. Rybicki, African horse sickness: a review of current understanding and vaccine development, *Viruses* 11(9) (2019) 844.
- [221] H. Yamaguchi, S. Kobayashi, A. Ishii, H. Ogawa, I. Nakamura, L. Moonga, B.M. Hang'ombe, A.S. Mweene, Y. Thomas, T. Kimura, Identification of a novel polyomavirus from vervet monkeys in Zambia, *Journal of General Virology* 94(6) (2013) 1357-1364.
- [222] R. Johne, H. Müller, The genome of goose hemorrhagic polyomavirus, a new member of the proposed subgenus Avipolyomavirus, *Virology* 308(2) (2003) 291-302.

- [223] A. Peretti, P.C. FitzGerald, V. Bliskovsky, D.V. Pastrana, C.B. Buck, Genome sequence of a fish-associated polyomavirus, black sea bass (*Centropristis striata*) polyomavirus 1, *Genome announcements* 3(1) (2015).
- [224] J.A. Dill, T.F. Ng, A.C. Camus, Complete sequence of the smallest polyomavirus genome, giant guitarfish (*Rhynchobatus djiddensis*) polyomavirus 1, *Genome announcements* 4(3) (2016).
- [225] J.M. Kean, S. Rao, M. Wang, R.L. Garcea, Seroepidemiology of human polyomaviruses, *PLoS Pathog* 5(3) (2009) e1000363.
- [226] A. Egli, L. Infanti, A. Dumoulin, A. Buser, J. Samaridis, C. Stebler, R. Gosert, H.H. Hirsch, Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors, *The Journal of infectious diseases* 199(6) (2009) 837-846.
- [227] R. Purighalla, R. Shapiro, J. McCauley, P. Randhawa, BK virus infection in a kidney allograft diagnosed by needle biopsy, *American journal of kidney diseases* 26(4) (1995) 671-673.
- [228] H.H. Hirsch, Polyomavirus BK nephropathy: a (re-) emerging complication in renal transplantation, *American journal of transplantation* 2(1) (2002) 25-30.
- [229] B. Padgett, G. Zurhein, D. Walker, R. Eckroade, B. Dessel, Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy, *The Lancet* 297(7712) (1971) 1257-1260.
- [230] A.A. Bagasi, T. Khandaker, G. Clark, T. Akagha, J.K. Ball, W.L. Irving, C.P. McClure, Trichodysplasia spinulosa polyomavirus in respiratory tract of immunocompromised child, *Emerging infectious diseases* 24(9) (2018) 1744.
- [231] F.H. Leendertz, N. Scuda, K.N. Cameron, T. Kidega, K. Zuberbühler, S.A.J. Leendertz, E. Couacy-Hymann, C. Boesch, S. Calvignac, B. Ehlers, African great apes are naturally infected with polyomaviruses closely related to Merkel cell polyomavirus, *Journal of virology* 85(2) (2011) 916-924.
- [232] G. Stoner, H. Agostini, C. Ryschkewitsch, M. Mazlo, F. Gullotta, W. Wamukota, S. Lucas, Detection of JC virus in two African cases of progressive multifocal leukoencephalopathy including identification of JCV type 3 in a Gambian AIDS patient, *Journal of medical microbiology* 47(8) (1998) 733-742.
- [233] W. Lang, J. Miklossy, J. Deruaz, G. Pizzolato, A. Probst, T. Schaffner, E. Gessaga, P. Kleihues, Neuropathology of the acquired immune deficiency syndrome (AIDS): a report of 135 consecutive autopsy cases from Switzerland, *Acta neuropathologica* 77(4) (1989) 379-390.
- [234] L. Vago, G. Trabattoni, A. Lechi, S. Cristina, H. Budka, Neuropathology of AIDS dementia: A review after 205 post mortem examinations, *Acta neurologica* (1990).
- [235] J.R. Berger, M. Concha, Progressive multifocal leukoencephalopathy: the evolution of a disease once considered rare, *Journal of neurovirology* 1(1) (1995) 5-18.
- [236] R. Newton, T. Ribeiro, E. Alvarez, J. Ziegler, D. Casabonne, L. Carpenter, V. Beral, E. Mbidde, D.M. Parkin, H. Wabinga, BK virus and cancer in Uganda, *European journal of cancer prevention* 15(4) (2006) 285-289.
- [237] W.H. Organization, World malaria report 2020: 20 years of global progress and challenges, (2020).
- [238] S. Yegorov, R.M. Galiwango, A. Ssemaganda, M. Muwanga, I. Wesonga, G. Miiro, D.A. Draji, K.C. Kain, N. Kiwanuka, B.S. Bagaya, Low prevalence of laboratory-confirmed malaria in clinically diagnosed adult women from the Wakiso district of Uganda, *Malaria journal* 15(1) (2016) 1-8.
- [239] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, *The lancet* 395(10223) (2020) 470-473.
- [240] M.R. Duffy, T.-H. Chen, W.T. Hancock, A.M. Powers, J.L. Kool, R.S. Lanciotti, M. Pretrick, M. Marfel, S. Holzbauer, C. Dubray, Zika virus outbreak on Yap Island, federated states of Micronesia, *New England Journal of Medicine* 360(24) (2009) 2536-2543.
- [241] R. Lowe, C. Barcellos, P. Brasil, O. Cruz, N. Honório, H. Kuper, M. Carvalho, The Zika virus epidemic in Brazil: from discovery to future implications, *International journal of environmental research and public health* 15(1) (2018) 96.

- [242] S. Baize, D. Pannetier, L. Oestereich, T. Rieger, L. Koivogui, N.F. Magassouba, B. Soropogui, M.S. Sow, S. Keita, H. De Clerck, Emergence of Zaire Ebola virus disease in Guinea, *New England Journal of Medicine* 371(15) (2014) 1418-1425.
- [243] B.M. Hause, E.A. Collin, R. Liu, B. Huang, Z. Sheng, W. Lu, D. Wang, E.A. Nelson, F. Li, Characterization of a novel influenza virus in cattle and swine: proposal for a new genus in the Orthomyxoviridae family, *MBio* 5(2) (2014) e00031-14.
- [244] G.W. Dick, S.F. Kitchen, A.J. Haddow, Zika virus (I). Isolations and serological specificity, *Transactions of the royal society of tropical medicine and hygiene* 46(5) (1952) 509-520.
- [245] H. Zhou, B. Eaton, Z. Hu, B. Arif, Accidental discovery and isolation of Zika virus in Uganda and the relentless epidemiologist behind the investigations, *Virologica Sinica* 31(4) (2016) 357-361.
- [246] P. Simmonds, F. Davidson, C. Lycett, L. Prescott, D. MacDonald, J. Ellender, P. Yap, G. Haydon, J. Gillon, L. Jarvis, Detection of a novel DNA virus (TT virus) in blood donors and blood products, *The lancet* 352(9123) (1998) 191-195.
- [247] J.F. Wamala, L. Lukwago, M. Malimbo, P. Nguku, Z. Yoti, M. Musenero, J. Amone, W. Mbabazi, M. Nanyunja, S. Zambwa, Ebola hemorrhagic fever associated with novel virus strain, Uganda, 2007–2008, *Emerging infectious diseases* 16(7) (2010) 1087.
- [248] N.S.-O.I.A.V.I. Team, Emergence of a novel swine-origin influenza A (H1N1) virus in humans, *New England journal of medicine* 360(25) (2009) 2605-2615.
- [249] M. Beer, F. Conraths, W. Van der Poel, 'Schmallenberg virus'—a novel orthobunyavirus emerging in Europe, *Epidemiology & Infection* 141(1) (2013) 1-8.
- [250] T.F.F. Ng, C. Manire, K. Borrowman, T. Langer, L. Ehrhart, M. Breitbart, Discovery of a novel single-stranded DNA virus from a sea turtle fibropapilloma by using viral metagenomics, *Journal of Virology* 83(6) (2009) 2500-2509.
- [251] G.A. Marsh, C. De Jong, J.A. Barr, M. Tachedjian, C. Smith, D. Middleton, M. Yu, S. Todd, A.J. Foord, V. Haring, Cedar virus: a novel Henipavirus isolated from Australian bats, (2012).
- [252] J. Waggoner, J. Brichard, F. Mutuku, B. Ndenga, C.J. Heath, A. Mohamed-Hadley, M.K. Sahoo, J. Vulule, M. Lefterova, N. Banaei, Malaria and chikungunya detected using molecular diagnostics among febrile Kenyan children, *Open forum infectious diseases*, Oxford University Press, 2017.
- [253] M. Ocan, E.A. Obuku, F. Bwanga, D. Akena, S. Richard, J. Ogwal-Okeng, C. Obua, Household antimicrobial self-medication: a systematic review and meta-analysis of the burden, risk factors and outcomes in developing countries, *BMC public health* 15(1) (2015) 742.
- [254] C.I. Chandler, C. Jones, G. Boniface, K. Juma, H. Reyburn, C.J. Whitty, Guidelines and mindlines: why do clinical staff over-diagnose malaria in Tanzania? A qualitative study, *Malaria journal* 7(1) (2008) 53.
- [255] G.A. Storch, Diagnostic virology, *Clinical infectious diseases* 31(3) (2000) 739-751.
- [256] T. Byaruhanga, B. Bagaya, J. Namulondo, J.T. Kayiwa, B. Namagambo, N. Owor, I. Nabukenya, B. Bakamuntumaho, J.J. Lutwama, The Correlation of IFN  $\gamma$  to the Preferential Isolation of Influenza Type B over Type A Viruses in Madin Darby Canine Kidney Cells, *Open Journal of Medical Microbiology* 7(01) (2017) 12.
- [257] L. Elliott, T. Ksiazek, P. Rollin, C. Spiropoulou, S. Morzunov, M. Monroe, C. Goldsmith, C. Humphrey, S. Zaki, J. Krebs, Isolation of the causative agent of hantavirus pulmonary syndrome, *The American journal of tropical medicine and hygiene* 51(1) (1994) 102-108.
- [258] G. Dick, S. Kitchen, A. Haddow, Zika virus (I). Isolations and serological specificity, *Transactions of the royal society of tropical medicine and hygiene* 46(5) (1952) 509-520.
- [259] G. Gard, R. Weir, S. Walsh, Arboviruses recovered from sentinel cattle using several virus isolation methods, *Veterinary microbiology* 18(2) (1988) 119-125.
- [260] K.G. Beavis, S.M. Matushek, A.P.F. Abeleda, C. Bethel, C. Hunt, S. Gillen, A. Moran, V. Tesic, Evaluation of the EUROIMMUN Anti-SARS-CoV-2 ELISA Assay for detection of IgA and IgG antibodies, *Journal of Clinical Virology* 129 (2020) 104468.

- [261] J.T. Kayiwa, A.M. Nankya, I. Ataliba, C.A. Nassuna, I.E. Omara, J.W. Koehler, J.M. Dye, E.C. Mossel, J.J. Lutwama, Dengue fever and chikungunya virus infections: identification in travelers in Uganda–2017, *Tropical Diseases, Travel Medicine and Vaccines* 5(1) (2019) 1-5.
- [262] S. Inoue, K. Morita, R. Matias, J. Tuplano, R. Resuello, J. Candelario, D. Cruz, C. Mapua, F. Hasebe, A. Igarashi, Distribution of three arbovirus antibodies among monkeys (*Macaca fascicularis*) in the Philippines, *Journal of medical primatology* 32(2) (2003) 89-94.
- [263] D. Wang, A. Urisman, Y.-T. Liu, M. Springer, T.G. Ksiazek, D.D. Erdman, E.R. Mardis, M. Hickenbotham, V. Magrini, J. Eldred, Viral discovery and sequence recovery using DNA microarrays, *PLoS biology* 1(2) (2003) e2.
- [264] P.C. Woo, S.K. Lau, C.-m. Chu, K.-h. Chan, H.-w. Tsoi, Y. Huang, B.H. Wong, R.W. Poon, J.J. Cai, W.-k. Luk, Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia, *Journal of virology* 79(2) (2005) 884-895.
- [265] T. Tsoleridis, O. Onianwa, E. Horncastle, E. Dayman, M. Zhu, T. Danjitrang, M. Wachtl, J. Behnke, S. Chapman, V. Strong, Discovery of novel alphacoronaviruses in European rodents and shrews, *Viruses* 8(3) (2016) 84.
- [266] C.W. Bachem, R.S. Van Der Hoeven, S.M. De Bruijn, D. Vreugdenhil, M. Zabeau, R.G. Visser, Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development, *The plant journal* 9(5) (1996) 745-753.
- [267] L. van der Hoek, K. Pyrc, M.F. Jebbink, W. Vermeulen-Oost, R.J. Berkhout, K.C. Wolthers, P.M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, B. Berkhout, Identification of a new human coronavirus, *Nature medicine* 10(4) (2004) 368.
- [268] N.L. Yozwiak, P. Skewes-Cox, M.D. Stenglein, A. Balmaseda, E. Harris, J.L. DeRisi, Virus identification in unknown tropical febrile illness cases using deep sequencing, *Plos neglected tropical diseases* 6(2) (2012) e1485.
- [269] A.W. Edridge, M. Deijs, R. Namazzi, C. Cristella, M.F. Jebbink, I. Maurer, N.A. Kootstra, L.R. Buluma, J.B. van Woensel, M.D. de Jong, Novel orthobunyavirus identified in the cerebrospinal fluid of a Ugandan child with severe encephalopathy, *Clinical Infectious Diseases* 68(1) (2018) 139-142.
- [270] F. Rohwer, R. Edwards, The Phage Proteomic Tree: a genome-based taxonomy for phage, *Journal of bacteriology* 184(16) (2002) 4529-4535.
- [271] S. Vieth, C. Drosten, O. Lenz, M. Vincent, S. Omilabu, M. Hass, B. Becker-Ziaja, J. ter Meulen, S.T. Nichol, H. Schmitz, RT-PCR assay for detection of Lassa virus and related Old World arenaviruses targeting the L gene, *Transactions of the Royal Society of Tropical Medicine and Hygiene* 101(12) (2007) 1253-1264.
- [272] F.F. Bier, F. Kleinjung, Feature-size limitations of microarray technology—a critical review, *Fresenius' journal of analytical chemistry* 371(2) (2001) 151-156.
- [273] R. Jaksik, M. Iwanaszko, J. Rzeszowska-Wolny, M. Kimmel, Microarray experiments and factors which affect their reliability, *Biology direct* 10(1) (2015) 1-14.
- [274] K.B. Mullis, The unusual origin of the polymerase chain reaction, *Scientific American* 262(4) (1990) 56-65.
- [275] J.M. Bartlett, D. Stirling, A short history of the polymerase chain reaction, *PCR protocols*, Springer 2003, pp. 3-6.
- [276] F. Sanger, S. Nicklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, *Proceedings of the national academy of sciences* 74(12) (1977) 5463-5467.
- [277] T. Nishizawa, H. Okamoto, K. Konishi, H. Yoshizawa, Y. Miyakawa, M. Mayumi, A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology, *Biochemical and biophysical research communications* 241(1) (1997) 92-97.
- [278] H. Okamoto, T. Nishizawa, A. Tawara, Y. Peng, M. Takahashi, J. Kishimoto, T. Tanaka, Y. Miyakawa, M. Mayumi, Species-specific TT viruses in humans and nonhuman primates and their phylogenetic relatedness, *Virology* 277(2) (2000) 368-378.

- [279] J.J. Lutwama, B. Bakamutumaho, J.T. Kayiwa, R. Chiiza, B. Namagambo, M.A. Katz, A.L. Geissler, Clinic-and hospital-based sentinel influenza surveillance, Uganda 2007–2010, *The Journal of infectious diseases* 206(suppl\_1) (2012) S87-S93.
- [280] M.R. Sandbulte, B.J. Gautney, M. Maloba, C. Wexler, M. Brown, N. Mabachi, K. Goggin, R. Lwembe, N. Nazir, T.A. Odeny, Infant HIV testing at birth using point-of-care and conventional HIV DNA PCR: an implementation feasibility pilot study in Kenya, *Pilot and feasibility studies* 5(1) (2019) 1-10.
- [281] W. Gu, S. Miller, C.Y. Chiu, Clinical metagenomic next-generation sequencing for pathogen detection, *Annual Review of Pathology: Mechanisms of Disease* 14 (2019) 319-338.
- [282] E.E. Schadt, S. Turner, A. Kasarskis, A window into third-generation sequencing, *Human molecular genetics* 19(R2) (2010) R227-R240.
- [283] A. Rhoads, K.F. Au, PacBio sequencing and its applications, *Genomics, proteomics & bioinformatics* 13(5) (2015) 278-289.
- [284] J.M. Rothberg, W. Hinz, T.M. Rearick, J. Schultz, W. Mileski, M. Davey, J.H. Leamon, K. Johnson, M.J. Milgrew, M. Edwards, An integrated semiconductor device enabling non-optical genome sequencing, *Nature* 475(7356) (2011) 348-352.
- [285] M. Jain, H.E. Olsen, B. Paten, M. Akeson, The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community, *Genome biology* 17(1) (2016) 1-11.
- [286] H. Lu, F. Giordano, Z. Ning, Oxford Nanopore MinION sequencing and genome assembly, *Genomics, proteomics & bioinformatics* 14(5) (2016) 265-279.
- [287] J.A. Reuter, D.V. Spacek, M.P. Snyder, High-throughput sequencing technologies, *Molecular cell* 58(4) (2015) 586-597.
- [288] J.C. Dohm, C. Lottaz, T. Borodina, H. Himmelbauer, Substantial biases in ultra-short read data sets from high-throughput DNA sequencing, *Nucleic acids research* 36(16) (2008) e105.
- [289] M.H. Stremlau, K.G. Andersen, O.A. Folarin, J.N. Grove, I. Odia, P.E. Ehiane, O. Omoniwa, O. Omoregie, P.-P. Jiang, N.L. Yozwiak, Discovery of novel rhabdoviruses in the blood of healthy individuals from West Africa, *PLoS neglected tropical diseases* 9(3) (2015) e0003631.
- [290] M.L. Robinson, Y.C. Manabe, Reducing uncertainty for acute febrile illness in resource-limited settings: the current diagnostic landscape, *The American journal of tropical medicine and hygiene* 96(6) (2017) 1285-1295.
- [291] P.C. Ng, E.F. Kirkness, Whole genome sequencing, *Genetic variation: Methods and protocols* (2010) 215-226.
- [292] C.G. Van El, M.C. Cornel, P. Borry, R.J. Hastings, F. Fellmann, S.V. Hodgson, H.C. Howard, A. Cambon-Thomsen, B.M. Knoppers, H. Meijers-Heijboer, Whole-genome sequencing in health care, *European Journal of Human Genetics* 21(6) (2013) 580-584.
- [293] K.E. Ormond, M.T. Wheeler, L. Hudgins, T.E. Klein, A.J. Butte, R.B. Altman, E.A. Ashley, H.T. Greely, Challenges in the clinical application of whole-genome sequencing, *The Lancet* 375(9727) (2010) 1749-1751.
- [294] F. Mertes, A. ElSharawy, S. Sauer, J.M. van Helvoort, P. Van Der Zaag, A. Franke, M. Nilsson, H. Lehrach, A.J. Brookes, Targeted enrichment of genomic DNA regions for next-generation sequencing, *Briefings in functional genomics* 10(6) (2011) 374-386.
- [295] R. Stark, M. Grzelak, J. Hadfield, RNA sequencing: the teenage years, *Nature Reviews Genetics* 20(11) (2019) 631-656.
- [296] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, *Nature reviews genetics* 10(1) (2009) 57-63.
- [297] J.A. Crump, S. Gove, C.M. Parry, Management of adolescents and adults with febrile illness in resource limited areas, *BMJ-British Medical Journal* 343(8) (2011) d4847.

- [298] J.J. Lutwama, B. Bakamutumaho, J.T. Kayiwa, R. Chiiza, B. Namagambo, M.A. Katz, A.L. Geissler, Clinic-and hospital-based sentinel influenza surveillance, Uganda 2007–2010, *Journal of Infectious Diseases* 206(suppl 1) (2012) S87-S93.
- [299] S. Balinandi, B. Bakamutumaho, J.T. Kayiwa, J. Ongus, J. Oundo, A.C. Awor, J.J. Lutwama, The viral aetiology of influenza-like illnesses in Kampala and Entebbe, Uganda, 2008, *African Journal of Laboratory Medicine* 2(1) (2013) 5 pages.
- [300] J.T. Kayiwa, A.M. Nankya, I.J. Ataliba, E.C. Mossel, M.B. Crabtree, J.J. Lutwama, Confirmation of Zika virus infection through hospital-based sentinel surveillance of acute febrile illness in Uganda, 2014–2017, *Journal of General Virology* 99(9) (2018) 1248-1252.
- [301] J.J. Lutwama, Uganda Arboviruses Surveillance Activities, *African Journal of Pharmacology and Therapeutics* 1(4) (2012).
- [302] S. UBOS, Uganda Bureau of Statistics, The National Population and Housing Census (2014).
- [303] D.A. Martin, D.A. Muth, T. Brown, A.J. Johnson, N. Karabatsos, J.T. Roehrig, Standardization of immunoglobulin M capture enzyme-linked immunosorbent assays for routine diagnosis of arboviral infections, *Journal of clinical microbiology* 38(5) (2000) 1823-1826.
- [304] M. Kulldorff, A spatial scan statistic, *Communications in Statistics-Theory and methods* 26(6) (1997) 1481-1496.
- [305] M. Kulldorff, W.F. Athas, E.J. Feurer, B.A. Miller, C.R. Key, Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico, *American journal of public health* 88(9) (1998) 1377-1380.
- [306] M.J. Cummings, B. Bakamutumaho, J. Kayiwa, T. Byaruhanga, N. Owor, B. Namagambo, A. Wolf, J.F. Wamala, S.S. Morse, J.J. Lutwama, Epidemiologic and spatiotemporal characterization of influenza and severe acute respiratory infection in Uganda, 2010-2015, *Annals of the American Thoracic Society* 13(12) (2016) 2159-2168.
- [307] M. Chan, B. Jiang, T.-Y. Tan, Using pooled recombinant plasmids as control materials for diagnostic real-time PCR, *Clin Lab* 62 (2016) 1893-901.
- [308] S. Tong, S.-W.W. Chern, Y. Li, M.A. Pallansch, L.J. Anderson, Sensitive and Broadly Reactive Reverse Transcription-PCR Assays To Detect Novel Paramyxoviruses, *Journal of Clinical Microbiology* 46(8) (2008) 2652-2658.
- [309] U. Dierssen, F. Rehren, C. Henke-Gendo, G. Harste, A. Heim, Rapid routine detection of enterovirus RNA in cerebrospinal fluid by a one-step real-time RT-PCR assay, *Journal of Clinical Virology* 42(1) (2008) 58-64.
- [310] G. Moureau, S. Temmam, J. Gonzalez, R. Charrel, G. Grard, X. De Lamballerie, A real-time RT-PCR method for the universal detection and identification of flaviviruses, *Vector-Borne and Zoonotic Diseases* 7(4) (2007) 467-478.
- [311] J.F. Drexler, A. Seelen, V.M. Corman, A.F. Tatenno, V. Cottontail, R.M. Zerbinati, F. Gloza-Rausch, S.M. Klose, Y. Adu-Sarkodie, S.K. Oppong, Bats worldwide carry hepatitis E virus-related viruses that form a putative novel genus within the family Hepeviridae, *Journal of virology* 86(17) (2012) 9134-9147.
- [312] B.W. Johnson, B.J. Russell, C.H. Goodman, Laboratory diagnosis of chikungunya virus infections and commercial sources for diagnostic assays, *The Journal of infectious diseases* 214(suppl\_5) (2016) S471-S474.
- [313] E.N. Yandoko, S. Gribaldo, C. Finance, A. Le Faou, B. Rihn, Molecular characterization of African orthobunyaviruses, *Journal of general virology* 88(6) (2007) 1761-1766.
- [314] A.J. Lambert, R.S. Lanciotti, Consensus amplification and novel multiplex sequencing method for S segment species identification of 47 viruses of the Orthobunyavirus, Phlebovirus, and Nairovirus genera of the family Bunyaviridae, *Journal of Clinical Microbiology* 47(8) (2009) 2398-2404.
- [315] O. Onianwa, Investigation of arenavirus and filovirus infections in rodents and non-human primates in the United Kingdom, University of Nottingham, 2018.

- [316] P.R. Heaton, P. Johnstone, L.M. McElhinney, R. Cowley, E. O'Sullivan, J.E. Whitby, Heminested PCR assay for detection of six genotypes of rabies and rabies-related viruses, *Journal of clinical microbiology* 35(11) (1997) 2762-2766.
- [317] W.A. Nix, M.S. Oberste, M.A. Pallansch, Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens, *Journal of clinical microbiology* 44(8) (2006) 2698-2704.
- [318] B. Bankamp, L.A. Byrd-Leotis, E.N. Lopareva, G.K. Woo, C. Liu, Y. Jee, H. Ahmed, W.W. Lim, N. Ramamurthy, M.N. Mulders, Improving molecular tools for global surveillance of measles virus, *Journal of Clinical Virology* 58(1) (2013) 176-182.
- [319] J. Quick, N.D. Grubaugh, S.T. Pullan, I.M. Claro, A.D. Smith, K. Gangavarapu, G. Oliveira, R. Robles-Sikisaka, T.F. Rogers, N.A. Beutler, Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples, *nature protocols* 12(6) (2017) 1261.
- [320] N. Nowotny, J. Kolodziejek, Middle East respiratory syndrome coronavirus (MERS-CoV) in dromedary camels, Oman, 2013, *Eurosurveillance* 19(16) (2014) 20781.
- [321] J.J. Waggoner, J. Abeynayake, I. Balassiano, M. Lefterova, M.K. Sahoo, Y. Liu, J.M. Vital-Brazil, L. Gresh, A. Balmaseda, E. Harris, Multiplex nucleic acid amplification test for diagnosis of dengue fever, malaria, and leptospirosis, *Journal of clinical microbiology* 52(6) (2014) 2011-2018.
- [322] W.S. Probert, K.N. Schrader, N.Y. Khuong, S.L. Bystrom, M.H. Graves, Real-time multiplex PCR assay for detection of *Brucella* spp., *B. abortus*, and *B. melitensis*, *Journal of clinical microbiology* 42(3) (2004) 1290-1293.
- [323] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, *Molecular biology and evolution* 33(7) (2016) 1870-1874.
- [324] P. Simmonds, SSE: a nucleotide and amino acid sequence analysis platform, *BMC research notes* 5(1) (2012) 1-10.
- [325] R. Gupta, S. Brunak, Prediction of glycosylation across the human proteome and the correlation to protein function, *Pac Symp Biocomput*, 2001, pp. 310-22.
- [326] K.L. Kalantar, T. Carvalho, C.F. de Bourcy, B. Dimitrov, G. Dingle, R. Egger, J. Han, O.B. Holmes, Y.-F. Juan, R. King, IDseq—an open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring, *Gigascience* 9(10) (2020) giaa111.
- [327] K. Adeboyejo, B.J. King, T. Tsoleridis, A.W. Tarr, J. McLauchlan, W.L. Irving, J.K. Ball, C.P. McClure, Hepatitis C subtyping assay failure in UK patients born in Sub-Saharan Africa: implications for global treatment and elimination, *medRxiv* (2021).
- [328] D. Bonsall, W.F. Gregory, C.L. Ip, S. Donfield, J. Iles, M.A. Ansari, P. Piazza, A. Trebes, A. Brown, J. Frater, Evaluation of viremia frequencies of a novel human pegivirus by using bioinformatic screening and PCR, *Emerging infectious diseases* 22(4) (2016) 671.
- [329] G. Palacios, C. Cowled, A.V. Bussetti, N. Savji, R. Weir, I. Wick, A. Travassos da Rosa, C.H. Calisher, R.B. Tesh, D. Boyle, Rapid molecular strategy for orbivirus detection and characterization, *Journal of Clinical Microbiology* 49(6) (2011) 2314-2317.
- [330] R. Johne, D. Enderlein, H. Nieper, H. Müller, Novel polyomavirus detected in the feces of a chimpanzee by nested broad-spectrum PCR, *Journal of virology* 79(6) (2005) 3883-3887.
- [331] H. Ketkar, D. Herman, P. Wang, Genetic determinants of the re-emergence of arboviral diseases, *Viruses* 11(2) (2019) 150.
- [332] A. LaBeaud, F. Bashir, C.H. King, Measuring the burden of arboviral diseases: the spectrum of morbidity and mortality from four prevalent infections, *Population health metrics* 9(1) (2011) 1-11.
- [333] C.J. Puntasecca, C.H. King, A.D. LaBeaud, Measuring the global burden of chikungunya and Zika viruses: A systematic review, *PLoS neglected tropical diseases* 15(3) (2021) e0009055.

- [334] A. Wilder-Smith, D.J. Gubler, S.C. Weaver, T.P. Monath, D.L. Heymann, T.W. Scott, Epidemic arboviral diseases: priorities for research and public health, *The Lancet infectious diseases* 17(3) (2017) e101-e106.
- [335] A. Green, Yellow fever continues to spread in Angola, *The Lancet* 387(10037) (2016) 2493.
- [336] M.U. Kraemer, N.R. Faria, R.C. Reiner Jr, N. Golding, B. Nikolay, S. Stasse, M.A. Johansson, H. Salje, O. Faye, G.W. Wint, Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study, *The Lancet infectious diseases* 17(3) (2017) 330-338.
- [337] P.F. Vasconcelos, J.E. Bryant, A.P.T. Da Rosa, R.B. Tesh, S.G. Rodrigues, A.D. Barrett, Genetic divergence and dispersal of yellow fever virus, Brazil, *Emerging infectious diseases* 10(9) (2004) 1578.
- [338] P. Van Der Stuyft, A. Gianella, M. Pirard, J. Cespedes, J. Lora, C. Peredo, J. Pelegrino, V. Vorndam, M. Boelaert, Urbanisation of yellow fever in Santa Cruz, Bolivia, *The Lancet* 353(9164) (1999) 1558-1562.
- [339] J.-P. Mutebi, A.D. Barrett, The epidemiology of yellow fever in Africa, *Microbes and infection* 4(14) (2002) 1459-1468.
- [340] K. Smithburn, A.J. Haddow, W. Lumsden, An outbreak of sylvan yellow fever in Uganda with *Aedes* (*Stegomyia*) *africanus* Theobald as principal vector and insect host of the virus, *Annals of Tropical Medicine & Parasitology* 43(1) (1949) 74-89.
- [341] A.F. Mahaffy, K.C. Smithburn, H.R. Jacobs, J.D. Gillett, Yellow fever in western Uganda, *Transactions of the Royal Society of Tropical Medicine and Hygiene* 36(1) (1942) 9-20.
- [342] L. Kwagonza, B. Masiira, H. Kyobe-Bosa, D. Kadobera, E.B. Atuheire, B. Lubwama, A. Kagirita, E. Katushabe, J.T. Kayiwa, J.J. Lutwama, Outbreak of yellow fever in central and southwestern Uganda, February–May 2016, *BMC infectious diseases* 18(1) (2018) 1-9.
- [343] J.F. Wamala, M. Malimbo, C.L. Okot, A.D. Atai-Omoruto, E. Tenywa, J.R. Miller, S. Balinandi, T. Shoemaker, C. Oyoo, E.O. Omony, Epidemiological and laboratory characterization of a yellow fever outbreak in northern Uganda, October 2010–January 2011, *International Journal of Infectious Diseases* 16(7) (2012) e536-e542.
- [344] B.E. Henderson, G.B. Kirya, L. Hewitt, Serological survey for arboviruses in Uganda, 1967-69, *Bulletin of the World Health Organization* 42(5) (1970) 797.
- [345] T.L. Clements, C.A. Rossi, A.K. Irish, H. Kibuuka, L.A. Eller, M.L. Robb, P. Kataaha, N.L. Michael, L.E. Hensley, R.J. Schoepp, Chikungunya and O'nyong-nyong Viruses in Uganda: Implications for Diagnostics, *Open forum infectious diseases*, Oxford University Press US, 2019, p. ofz001.
- [346] M. Weinbren, A.J. Haddow, M. Williams, The occurrence of chikungunya virus in Uganda I. Isolation from mosquitoes, *Transactions of the Royal Society of Tropical Medicine and Hygiene* 52(3) (1958) 253-262.
- [347] A. McCrae, B. Henderson, B. Kirya, S. Sempala, Chikungunya virus in the Entebbe area of Uganda: isolations and epidemiology, *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2 (1971).
- [348] N.P. Lindsey, J.E. Staples, K. Powell, I.B. Rabe, M. Fischer, A.M. Powers, O.I. Kosoy, E.C. Mossel, J.L. Munoz-Jordan, M. Beltran, Ability to serologically confirm recent Zika virus infection in areas with varying past incidence of dengue virus infection in the United States and US territories in 2016, *Journal of clinical microbiology* 56(1) (2018) e01115-17.
- [349] I.B. Rabe, J.E. Staples, J. Villanueva, K.B. Hummel, J.A. Johnson, L. Rose, S. Hills, A. Wasley, M. Fischer, A.M. Powers, Interim guidance for interpretation of Zika virus antibody test results, *Morbidity and mortality weekly report* 65(21) (2016) 543-546.
- [350] G.P. Göertz, C.B. Vogels, C. Geertsema, C.J. Koenraadt, G.P. Pijlman, Mosquito co-infection with Zika and chikungunya virus allows simultaneous transmission without affecting vector competence of *Aedes aegypti*, *PLoS neglected tropical diseases* 11(6) (2017) e0005654.
- [351] D. Diallo, G. Fall, C.T. Diagne, A. Gaye, Y. Ba, I. Dia, O. Faye, M. Diallo, Concurrent amplification of Zika, chikungunya, and yellow fever virus in a sylvatic focus of arboviruses in Southeastern Senegal, 2015, *BMC microbiology* 20(1) (2020) 1-11.

- [352] D. Conway, E. Allison, R. Felstead, M. Goulden, Rainfall variability in East Africa: implications for natural resources management and livelihoods, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 363(1826) (2005) 49-54.
- [353] C.C. Chisenga, S. Bosomprah, K. Musukuma, C. Mubanga, O.N. Chilyabanyama, R.M. Velu, Y.C. Kim, A. Reyes-Sandoval, R. Chilengi, Sero-prevalence of arthropod-borne viral infections among Lukanga swamp residents in Zambia, *Plos one* 15(7) (2020) e0235322.
- [354] S. Bailly, D. Rousset, C. Fritzell, N. Hozé, S. Ben Achour, L. Berthelot, A. Enfissi, J. Vanhomwegen, H. Salje, S. Fernandes-Pellerin, Spatial distribution and burden of emerging arboviruses in French Guiana, *Viruses* 13(7) (2021) 1299.
- [355] Y. Ushijima, H. Abe, G. Nguema Ondo, R. Bikangui, M. Massinga Loembé, V.R. Zadeh, J.G. Essimengane, A.V. Mbouna, E.B. Bache, S.T. Agnandji, Surveillance of the major pathogenic arboviruses of public health concern in Gabon, Central Africa: increased risk of West Nile virus and dengue virus infections, *BMC Infectious Diseases* 21(1) (2021) 1-11.
- [356] J.-P. Carrera, Z.M. Cucunubá, K. Neira, B. Lambert, Y. Pittí, J. Liscano, J.L. Garzón, D. Beltran, L. Collado-Mariscal, L. Saenz, Endemic and Epidemic Human Alphavirus Infections in Eastern Panama: An Analysis of Population-Based Cross-Sectional Surveys, *The American journal of tropical medicine and hygiene* 103(6) (2020) 2429.
- [357] L. Konongoi, V. Ofula, A. Nyunja, S. Owaka, H. Koka, A. Makio, E. Koskei, F. Eyase, D. Langat, R.J. Schoepp, Detection of dengue virus serotypes 1, 2 and 3 in selected regions of Kenya: 2011–2014, *Virology journal* 13(1) (2016) 1-11.
- [358] H. Osbahr, P. Dorward, R. Stern, S. Cooper, Supporting agricultural innovation in Uganda to respond to climate risk: linking climate change and variability with farmer perceptions, *Experimental agriculture* 47(2) (2011) 293-316.
- [359] J.S. Okonya, K. Syndikus, J. Kroschel, Farmers' perception of and coping strategies to climate change: Evidence from six agro-ecological zones of Uganda, *Journal of agricultural science* 5(8) (2013) 252.
- [360] E.A. Mordecai, S.J. Ryan, J.M. Caldwell, M.M. Shah, A.D. LaBeaud, Climate change could shift disease burden from malaria to arboviruses in Africa, *The Lancet Planetary Health* 4(9) (2020) e416-e423.
- [361] C. Lorenz, T.S. Azevedo, F. Virginio, B.S. Aguiar, F. Chiaravalloti-Neto, L. Suesdek, Impact of environmental factors on neglected emerging arboviral diseases, *PLoS neglected tropical diseases* 11(9) (2017) e0005959.
- [362] H. Delatte, A. Desvars, A. Bouétard, S. Bord, G. Gimonneau, G. Vourc'h, D. Fontenille, Blood-feeding behavior of *Aedes albopictus*, a vector of Chikungunya on La Réunion, *Vector-Borne and Zoonotic Diseases* 10(3) (2010) 249-258.
- [363] R. Hanson, S. Sulkin, E. Buescher, W.M. Hammon, R. McKinney, T. Work, Arbovirus Infections of Laboratory Workers: Extent of problem emphasizes the need for more effective measures to reduce hazards, *Science* 158(3806) (1967) 1283-1286.
- [364] C.H. Hsu, F. Cruz-Lopez, D. Vargas Torres, J. Perez-Padilla, O.D. Lorenzi, A. Rivera, J.E. Staples, E. Lugo, J. Munoz-Jordan, M. Fischer, Risk factors for hospitalization of patients with chikungunya virus infection at sentinel hospitals in Puerto Rico, *PLoS neglected tropical diseases* 13(1) (2019) e0007084.
- [365] M.L.B. Ferreira, C.A.A. de Brito, R.F. de Oliveira França, Á.J.P. Moreira, M.Í. de Moraes Machado, R. da Paz Melo, R. Medialdea-Carrera, S.D. Mesquita, M.L. Santos, R. Mehta, Neurological disease in adults with Zika and chikungunya virus infection in Northeast Brazil: a prospective observational study, *The Lancet Neurology* 19(10) (2020) 826-839.
- [366] S. Takaya, S. Kutsuna, E. Nakayama, S. Taniguchi, S. Tajima, Y. Katanami, K. Yamamoto, N. Takeshita, K. Hayakawa, Y. Kato, Chikungunya fever in traveler from Angola to Japan, 2016, *Emerging Infectious Diseases* 23(1) (2017) 156.
- [367] S. Makiala-Mandanda, S. Ahuka-Mundeke, J.L. Abbate, E. Pukuta-Simbu, J. Nsio-Mbeta, N. Berthet, E.M. Leroy, P. Becquart, J.-J. Muyembe-Tamfum, Identification of dengue and chikungunya cases among

- suspected cases of yellow fever in the Democratic Republic of the Congo, *Vector-Borne and Zoonotic Diseases* 18(7) (2018) 364-370.
- [368] B. Dong, L. Khan, M. Smith, J. Trevino, B. Zhao, G.L. Hamer, U.A. Lopez-Lemus, A.A. Molina, J. Lubinda, U.-S.D. Nguyen, Spatio-temporal dynamics of three diseases caused by Aedes-borne arboviruses in Mexico, *Communications Medicine* 2(1) (2022) 134.
- [369] J. Bagonza, E. Rutebemberwa, M. Mugaga, N. Tumuhamy, I. Makumbi, Yellow fever vaccination coverage following massive emergency immunization campaigns in rural Uganda, May 2011: a community cluster survey, *BMC Public Health* 13(1) (2013) 1-7.
- [370] W.H. Organization, Weekly Bulletin on Outbreak and other Emergencies: Week 21: 20-26 May 2019, *Weekly Bulletin on Outbreaks and other Emergencies* (2019) 1-27.
- [371] G. Haidar, N. Singh, Fever of Unknown Origin, *New England Journal of Medicine* 386(5) (2022) 463-477.
- [372] D. Durack, Fever of unknown origin-reexamined and redefined, *Current Clinical Topics in Infectious Disease* 11 (1991) 35-51.
- [373] H.-L. You, S.-J. Chang, H.-R. Yu, C.-C. Li, C.-H. Chen, W.-T. Liao, Simultaneous detection of respiratory syncytial virus and human metapneumovirus by one-step multiplex real-time RT-PCR in patients with respiratory symptoms, *BMC pediatrics* 17(1) (2017) 1-7.
- [374] R. Charrel, X. de Lamballerie, Zoonotic aspects of arenavirus infections, *Veterinary microbiology* 140(3-4) (2010) 213-220.
- [375] G. Smith, I. Smith, B. Harrower, D. Warrilow, C. Bletchly, A simple method for preparing synthetic controls for conventional and real-time PCR for the identification of endemic and exotic disease agents, *Journal of virological methods* 135(2) (2006) 229-234.
- [376] J.A. Whelan, N.B. Russell, M.A. Whelan, A method for the absolute quantification of cDNA using real-time PCR, *Journal of immunological methods* 278(1-2) (2003) 261-269.
- [377] D.T. Jones, W.R. Taylor, J.M. Thornton, The rapid generation of mutation data matrices from protein sequences, *Bioinformatics* 8(3) (1992) 275-282.
- [378] I. Casas, L. Powell, P. Klapper, G. Cleator, New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay, *Journal of virological methods* 53(1) (1995) 25-36.
- [379] J.E. Echevarría, A. Avellón, J. Juste, M. Vera, C. Ibáñez, Screening of active lyssavirus infection in wild bat populations by viral RNA detection on oropharyngeal swabs, *Journal of clinical microbiology* 39(10) (2001) 3678-3683.
- [380] D.R.J. Caasi, M. Arif, M. Payton, U. Melcher, L. Winder, F.M. Ochoa-Corona, A multi-target, non-infectious and clonable artificial positive control for routine PCR-based assays, *Journal of microbiological methods* 95(2) (2013) 229-234.
- [381] T. de Souza Rosés, A.P. Andreolla, L. de Figueiredo Soveral, M.I.B. Vieira, J.D. Kich, R. Frandoloso, L.C. Kreutz, Synthetic gene as target to assess the sensitivity of PCR to detect *Trichinella* spp. larvae in meat from a non-endemic region, *Tropical Animal Health and Production* 52(2) (2020) 619-623.
- [382] H. Cho, Y.H. Jung, H.B. Cho, H.-t. Kim, K.-s. Kim, Positive control synthesis method for COVID-19 diagnosis by one-step real-time RT-PCR, *Clinica Chimica Acta* 511 (2020) 149-153.
- [383] S.P. Kigozi, R.N. Kigozi, C.M. Sebuguzi, J. Cano, D. Rutazaana, J. Opigo, T. Bousema, A. Yeka, A. Gasasira, B. Sartorius, Spatial-temporal patterns of malaria incidence in Uganda using HMIS data from 2015 to 2019, *BMC public health* 20(1) (2020) 1-14.
- [384] W.H. Organization, New genotype of measles virus and update on global distribution of measles genotypes, *Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire* 80(40) (2005) 347-351.
- [385] W.H. Organization, Measles virus nomenclature update: 2012, *Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire* 87(09) (2012) 73-80.

- [386] A.R. Penedos, R. Myers, B. Hadeef, F. Aladin, K.E. Brown, Assessment of the utility of whole genome sequencing of measles virus in the characterisation of outbreaks, *PloS one* 10(11) (2015) e0143081.
- [387] A.W. Walekhwa, M. Ntaro, P.C. Kawungezi, C. Achangwa, R. Muhindo, E. Baguma, M. Matte, R. Migisha, R. Reyes, P. Thompson, Measles outbreak in Western Uganda: a case-control study, *BMC infectious diseases* 21(1) (2021) 1-9.
- [388] G. De Broucker, A. Ssebagereka, R.R. Apolot, M. Aloysius, E.E. Kiracho, B. Patenaude, D. Constenla, The economic burden of measles in children under five in Uganda, *Vaccine: X* 6 (2020) 100077.
- [389] R.K. Majwala, L. Nakiire, D. Kadobera, A.R. Ario, J. Kusiima, J.A. Atuhairwe, J.K. Matovu, B.-P. Zhu, Measles outbreak propagated by children congregating at water collection points in Mayuge District, eastern Uganda, July–October, 2016, *BMC infectious diseases* 18(1) (2018) 1-7.
- [390] I.J. Amanna, N.E. Carlson, M.K. Slifka, Duration of humoral immunity to common viral and vaccine antigens, *New England Journal of Medicine* 357(19) (2007) 1903-1915.
- [391] A. Beckford, Kaschula, ROC, C. Stephen, Factors associated with fatal cases of measles-a retrospective autopsy study, *South African Medical Journal* 68(12) (1985) 858-863.
- [392] D.E. Griffin, Measles virus persistence and its consequences, *Current opinion in virology* 41 (2020) 46-51.
- [393] M. De Feo, C. De Leo, U. Romeo, P. Muti, G. Blandino, S. Di Agostino, Arenavirus as a potential etiological agent of odontogenic tumours in humans, *Journal of Experimental & Clinical Cancer Research* 39(1) (2020) 1-4.
- [394] T. Smiley Evans, L. Tutaryebwa, K.V. Gilardi, P.A. Barry, A. Marzi, M. Eberhardt, B. Ssebide, M.R. Cranfield, O. Mugisha, E. Mugisha, Suspected exposure to filoviruses among people contacting wildlife in southwestern Uganda, *The Journal of infectious diseases* 218(suppl\_5) (2018) S277-S286.
- [395] D. Boon, A.D. Redd, O. Laeyendecker, R.E. Engle, H. Nguyen, P. Ocama, I. Boaz, A. Ndyababo, V. Kiggundu, S.J. Reynolds, Hepatitis E virus seroprevalence and correlates of anti-HEV IgG antibodies in the Rakai District, Uganda, *The Journal of Infectious Diseases* 217(5) (2018) 785-789.
- [396] E.H. Teshale, C.M. Howard, S.P. Grytdal, T.R. Handzel, V. Barry, S. Kamili, J. Drobeniuc, S. Okware, R. Downing, J.W. Tappero, Hepatitis E epidemic, Uganda, *Emerging infectious diseases* 16(1) (2010) 123.
- [397] W.H. Organization, Guidance on regulations for the transport of infectious substances 2019–2020: applicable from 1 January 2019, World Health Organization, 2019.
- [398] K. Smithburn, T. Hughes, A. Burke, J. Paul, A neurotropic virus isolated from the blood of a native of Uganda1, *The American Journal of Tropical Medicine and Hygiene* 1(4) (1940) 471-492.
- [399] T. Bakonyi, Z. Hubálek, I. Rudolf, N. Nowotny, Novel flavivirus or new lineage of West Nile virus, central Europe, *Emerging infectious diseases* 11(2) (2005) 225.
- [400] D. Lvov, A. Butenko, V. Gromashevsky, A. Kovtunov, A. Prilipov, R. Kinney, V. Aristova, A. Dzharkenov, E. Samokhvalov, H. Savage, West Nile virus and other zoonotic viruses in Russia: examples of emerging-reemerging situations, *Emergence and control of zoonotic viral encephalitides*, Springer2004, pp. 85-96.
- [401] A. Vazquez, M.P. Sánchez-Seco, S. Ruiz, F. Molero, L. Hernandez, J. Moreno, A. Magallanes, C.G. Tejedor, A. Tenorio, Putative new lineage of West Nile virus, Spain, *Emerging Infectious Diseases* 16(3) (2010) 549.
- [402] G. Fall, M. Diallo, C. Loucoubar, O. Faye, A.A. Sall, Vector competence of *Culex neavei* and *Culex quinquefasciatus* (Diptera: Culicidae) from Senegal for lineages 1, 2, Koutango and a putative new lineage of West Nile virus, *The American journal of tropical medicine and hygiene* 90(4) (2014) 747.
- [403] V.P. Bondre, R. Jadhav, A. Mishra, P. Yergolkar, V. Arankalle, West Nile virus isolates from India: evidence for a distinct genetic lineage, *Journal of General Virology* 88(3) (2007) 875-884.
- [404] R. Lanciotti, J. Roehrig, V. Deubel, J. Smith, M. Parker, K. Steele, B. Crise, K. Volpe, M. Crabtree, J. Scherret, Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States, *Science* 286(5448) (1999) 2333-2337.

- [405] L.M. Hernández-Triana, C.L. Jeffries, K.L. Mansfield, G. Carnell, A.R. Fooks, N. Johnson, Emergence of West Nile virus lineage 2 in Europe: a review on the introduction and spread of a mosquito-borne disease, *Frontiers in public health* 2 (2014) 271.
- [406] M.L. Tantely, S.M. Goodman, T. Rakotondranaivo, S. Boyer, Review of West Nile virus circulation and outbreak risk in Madagascar: Entomological and ornithological perspectives, *Parasite* 23 (2016).
- [407] T. Bakonyi, É. Ivanics, K. Erdélyi, K. Ursu, E. Ferenczi, H. Weissenböck, N. Nowotny, Lineage 1 and 2 strains of encephalitic West Nile virus, central Europe, *Emerging infectious diseases* 12(4) (2006) 618.
- [408] A. Platonov, L. Karan, T. Shopenskaia, M. Fedorova, N. Koliashnikova, N. Rusakova, L. Shishkina, T. Arshba, V. Zhuravlev, M. Govorukhina, Genotyping of West Nile fever virus strains circulating in southern Russia as an epidemiological investigation method: principles and results, *Zhurnal mikrobiologii, epidemiologii i immunobiologii* (2) (2011) 29-37.
- [409] A. Papa, K. Xanthopoulou, S. Gewehr, S. Mourelatos, Detection of West Nile virus lineage 2 in mosquitoes during a human outbreak in Greece, *Clinical Microbiology and Infection* 17(8) (2011) 1176-1180.
- [410] Z. Hubálek, J. Halouzka, West Nile fever--a reemerging mosquito-borne viral disease in Europe, *Emerging infectious diseases* 5(5) (1999) 643.
- [411] N. Karabatsos, International catalogue of arboviruses including certain other viruses of vertebrates. San Antonio, J Am Soc Trop Med Hyg (1985).
- [412] H.S. Hurlbut, West Nile Virus Infection in Arthropods<sup>1</sup>, *The American journal of tropical medicine and hygiene* 5(1) (1956) 76-85.
- [413] N. Komar, S. Langevin, S. Hinten, N. Nemeth, E. Edwards, D. Hettler, B. Davis, R. Bowen, M. Bunning, Experimental infection of North American birds with the New York 1999 strain of West Nile virus, *Emerging infectious diseases* 9(3) (2003) 311.
- [414] S. Abutarbush, A. Al-Majali, West Nile virus infection in horses in Jordan: clinical cases, seroprevalence and risk factors, *Transboundary and emerging diseases* 61 (2014) 1-6.
- [415] M.L. Bunning, R.A. Bowen, B.C. Cropp, K.G. Sullivan, B.S. Davis, N. Komar, M. Godsey, D. Baker, D.L. Hettler, D.A. Holmes, Experimental infection of horses with West Nile virus, *Emerging infectious diseases* 8(4) (2002) 380.
- [416] B. Murgue, H. Zeller, V. Deubel, The ecology and epidemiology of West Nile virus in Africa, Europe and Asia, *Japanese encephalitis and West Nile viruses*, Springer 2002, pp. 195-221.
- [417] A.C. NOVELLO, West Nile virus in New York State: the 1999 outbreak and response plan for 2000, *Viral immunology* 13(4) (2000) 463-467.
- [418] T. Burki, Increase of West Nile virus cases in Europe for 2018, *The Lancet* 392(10152) (2018) 1000.
- [419] F. Bai, E.A. Thompson, P.J. Vig, A.A. Leis, Current Understanding of West Nile Virus Clinical Manifestations, Immune Responses, Neuroinvasion, and Immunotherapeutic Implications, *Pathogens* 8(4) (2019) 193.
- [420] C. Huang, B. Slater, R. Rudd, N. Parchuri, R. Hull, M. Dupuis, A. Hindenburg, First isolation of West Nile virus from a patient with encephalitis in the United States, *Emerging infectious diseases* 8(12) (2002) 1367.
- [421] M.Y. Chowers, R. Lang, F. Nassar, D. Ben-David, M. Giladi, E. Rubinshtein, A. Itzhaki, J. Mishal, Y. Siegman-Igra, R. Kitzes, Clinical characteristics of the West Nile fever outbreak, Israel, 2000, *Emerging infectious diseases* 7(4) (2001) 675.
- [422] E.C. Mossel, M.B. Crabtree, J.-P. Mutebi, J.J. Lutwama, E.M. Borland, A.M. Powers, B.R. Miller, Arboviruses isolated from mosquitoes collected in Uganda, 2008–2012, *Journal of medical entomology* 54(5) (2017) 1403-1409.
- [423] Y. Orba, B. Hang'ombe, A. Mweene, Y. Wada, P. Anindita, W. Phongphaew, Y. Qiu, M. Kajihara, A. Mori-Kajihara, Y. Eto, First isolation of West Nile virus in Zambia from mosquitoes, *Transboundary and emerging diseases* 65(4) (2018) 933-938.

- [424] J.-H. Wu, P.-Y. Hong, W.-T. Liu, Quantitative effects of position and type of single mismatch on single base primer extension, *Journal of microbiological methods* 77(3) (2009) 267-275.
- [425] D. Bru, F. Martin-Laurent, L. Philippot, Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example, *Applied and environmental microbiology* 74(5) (2008) 1660-1663.
- [426] R. Stadhouders, S.D. Pas, J. Anber, J. Voermans, T.H. Mes, M. Schutten, The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay, *The Journal of Molecular Diagnostics* 12(1) (2010) 109-117.
- [427] O. Cabre, M. Grandadam, J.-L. Marié, P. Gravier, A. Prangé, Y. Santinelli, V. Rous, O. Bourry, J.-P. Durand, H. Tolou, West Nile virus in horses, sub-Saharan Africa, *Emerging infectious diseases* 12(12) (2006) 1958.
- [428] W. Wang, F. Sarkodie, K. Danso, E. Addo-Yobo, S. Owusu-Ofori, J.-P. Allain, C. Li, Seroprevalence of west nile virus in ghana, *Viral immunology* 22(1) (2009) 17-22.
- [429] M.S. Diamond, E.M. Sitati, L.D. Friend, S. Higgs, B. Shrestha, M. Engle, A critical role for induced IgM in the protection against West Nile virus infection, *The Journal of experimental medicine* 198(12) (2003) 1853-1862.
- [430] B. Davoust, M. Maquart, C. Roqueplo, P. Gravier, M. Sambou, O. Mediannikov, I. Leparç-Goffart, Serological survey of West Nile virus in domestic animals from Northwest Senegal, *Vector-Borne and Zoonotic Diseases* 16(5) (2016) 359-361.
- [431] J. Amdouni, F. Monaco, O. Portanti, S. Sghaier, A. Conte, T.B. Hassine, A. Polci, F. Valleriani, A. Di Gennaro, M. Zoueri, Detection of enzootic circulation of a new strain of West Nile virus lineage 1 in sentinel chickens in the north of Tunisia, *Acta tropica* 202 (2020) 105223.
- [432] M.N. Mayanja, F.N. Mwiine, A. Kohl, E.C. Thomson, J.J. Lutwama, An update on the mosquito species composition and diversity in western and North Western Uganda, *Journal of Entomology and Zoology Studies* 8(2) (2020) 1074-1086.
- [433] K. Danis, A. Papa, G. Theocharopoulos, G. Dougas, M. Athanasiou, M. Detsis, A. Baka, T. Lytras, K. Mellou, S. Bonovas, Outbreak of West Nile virus infection in Greece, 2010, *Emerging Infectious Diseases* 17(10) (2011) 1868.
- [434] S.L. Hanna, T.C. Pierson, M.D. Sanchez, A.A. Ahmed, M.M. Murtadha, R.W. Doms, N-linked glycosylation of west nile virus envelope proteins influences particle assembly and infectivity, *Journal of virology* 79(21) (2005) 13262-13274.
- [435] K. Shirato, H. Miyoshi, A. Goto, Y. Ako, T. Ueki, H. Kariwa, I. Takashima, Viral envelope protein glycosylation is a molecular determinant of the neuroinvasiveness of the New York strain of West Nile virus, *Journal of general virology* 85 (2004) 3637-3645.
- [436] R. Murata, Y. Eshita, A. Maeda, J. Maeda, S. Akita, T. Tanaka, K. Yoshii, H. Kariwa, T. Umemura, I. Takashima, Glycosylation of the West Nile Virus envelope protein increases in vivo and in vitro viral multiplication in birds, *The American journal of tropical medicine and hygiene* 82(4) (2010) 696.
- [437] R.M. Moudy, B. Zhang, P.-Y. Shi, L.D. Kramer, West Nile virus envelope protein glycosylation is required for efficient viral transmission by *Culex* vectors, *Virology* 387(1) (2009) 222-228.
- [438] S.A. Langevin, R.A. Bowen, W.N. Ramey, T.A. Sanders, P.D. Maharaj, Y. Fang, J. Cornelius, C.M. Barker, W.K. Reisen, D.W. Beasley, Envelope and pre-membrane protein structural amino acid mutations mediate diminished avian growth and virulence of a Mexican West Nile virus isolate, *The Journal of General Virology* 92(Pt 12) (2011) 2810.
- [439] C.C. Andrade, P.D. Maharaj, W.K. Reisen, A.C. Brault, North American West Nile virus genotype isolates demonstrate differential replicative capacities in response to temperature, *The Journal of general virology* 92(Pt 11) (2011) 2523.
- [440] P.M. Sharp, W.-H. Li, An evolutionary perspective on synonymous codon usage in unicellular organisms, *Journal of molecular evolution* 24 (1986) 28-38.

- [441] T.E. Quax, N.J. Claassens, D. Söll, J. van der Oost, Codon bias as a means to fine-tune gene expression, *Molecular cell* 59(2) (2015) 149-161.
- [442] C. Gustafsson, S. Govindarajan, J. Minshull, Codon bias and heterologous protein expression, *Trends in biotechnology* 22(7) (2004) 346-353.
- [443] F.-X. Berthet, H.G. Zeller, M.-T. Drouet, J. Rauzier, J.-P. Digoutte, V. Deubel, Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses, *Journal of General Virology* 78(9) (1997) 2293-2297.
- [444] S. Adams, A. Broom, L. Sammels, A. Hartnett, M. Howard, R. Coelen, J. Mackenzie, R. Hall, Glycosylation and antigenic variation among Kunjin virus isolates, *Virology* 206(1) (1995) 49-56.
- [445] F.J. Burt, A.A. Grobbelaar, P.A. Leman, F.S. Anthony, G.V. Gibson, R. Swanepoel, Phylogenetic relationships of southern African West Nile virus isolates, *Emerging infectious diseases* 8(8) (2002) 820.
- [446] M.V. Phan, C.N. Agoti, P.K. Munywoki, G.P. Otieno, M. Ngama, P. Kellam, M. Cotten, D.J. Nokes, Identification of missed viruses by metagenomic sequencing of clinical respiratory samples from Kenya, *Scientific reports* 12(1) (2022) 1-11.
- [447] Y. Yang, B. Xie, J. Yan, Application of next-generation sequencing technology in forensic science, *Genomics, proteomics & bioinformatics* 12(5) (2014) 190-197.
- [448] S. Morganti, P. Tarantino, E. Ferraro, P. D'Amico, B.A. Duso, G. Curigliano, Next Generation Sequencing (NGS): a revolutionary technology in pharmacogenomics and personalized medicine in cancer, *Translational Research and Onco-Omics Applications in the Era of Cancer Personal Genomics* (2019) 9-30.
- [449] R. Ekblom, J. Galindo, Applications of next generation sequencing in molecular ecology of non-model organisms, *Heredity* 107(1) (2011) 1-15.
- [450] M.P. Tan, L.L. Wong, S.A. Razali, N. Afiqah-Aleng, S.A. Mohd Nor, Y.Y. Sung, Y. Van de Peer, P. Sorgeloos, M. Danish-Daniel, Applications of next-generation sequencing technologies and computational tools in molecular evolution and aquatic animals conservation studies: a short review, *Evolutionary Bioinformatics* 15 (2019) 1176934319892284.
- [451] M. Vilsker, Y. Moosa, S. Nooij, V. Fonseca, Y. Ghysens, K. Dumon, R. Pauwels, L.C. Alcantara, E. Vanden Eynden, A.-M. Vandamme, Genome Detective: an automated system for virus identification from high-throughput sequencing data, *Bioinformatics* 35(5) (2019) 871-873.
- [452] C.-H. Liu, Y.P. Di, Analysis of RNA sequencing data using CLC genomics workbench, *Molecular Toxicology Protocols*, Springer 2020, pp. 61-113.
- [453] D.D. Roumpeka, R.J. Wallace, F. Escalettes, I. Fotheringham, M. Watson, A review of bioinformatics tools for bio-prospecting from metagenomic sequence data, *Frontiers in genetics* 8 (2017) 23.
- [454] R. Pereira, J. Oliveira, M. Sousa, Bioinformatics and computational tools for next-generation sequencing analysis in clinical genetics, *Journal of clinical medicine* 9(1) (2020) 132.
- [455] J.G. Chappell, T. Byaruhanga, T. Tsoleridis, J.K. Ball, C.P. McClure, Identification of infectious agents in high-throughput sequencing data sets is easily achievable using free, cloud-based bioinformatics platforms, *Journal of Clinical Microbiology* 57(12) (2019) e01386-19.
- [456] A. Brinkmann, A. Andrusch, A. Belka, C. Wylezich, D. Höper, A. Pohlmann, T. Nordahl Petersen, P. Lucas, Y. Blanchard, A. Papa, Proficiency testing of virus diagnostics based on bioinformatics analysis of simulated in silico high-throughput sequencing data sets, *Journal of clinical microbiology* 57(8) (2019) e00466-19.
- [457] Y. Morice, D. Roulot, V. Grando, J. Stirnemann, E. Gault, V. Jeantils, M. Bentata, B. Jarrousse, O. Lortholary, C. Pallier, Phylogenetic analyses confirm the high prevalence of hepatitis C virus (HCV) type 4 in the Seine-Saint-Denis district (France) and indicate seven different HCV-4 subtypes linked to two different epidemiological patterns, *Journal of General Virology* 82(5) (2001) 1001-1012.
- [458] C.F. Manso, D.F. Bibby, J.L. Mbisa, Efficient and unbiased metagenomic recovery of RNA virus genomes from human plasma samples, *Scientific reports* 7(1) (2017) 1-13.

- [459] M. Asplund, K.R. Kjartansdóttir, S. Møllerup, L. Vinner, H. Fridholm, J.A. Herrera, J. Friis-Nielsen, T.A. Hansen, R.H. Jensen, I.B. Nielsen, Contaminating viral sequences in high-throughput sequencing viromics: a linkage study of 700 sequencing libraries, *Clinical Microbiology and Infection* 25(10) (2019) 1277-1285.
- [460] C. Davis, G.S. Mgomella, A. da Silva Filipe, E.H. Frost, G. Giroux, J. Hughes, C. Hogan, P. Kaleebu, G. Asiki, J. McLauchlan, Highly diverse hepatitis C strains detected in Sub-Saharan Africa have unknown susceptibility to direct-acting antiviral treatments, *Hepatology* 69(4) (2019) 1426-1441.
- [461] M.P. Gededzha, S.G. Selabe, T. Kyaw, J.N. Rakgole, J.T. Blackard, M.J. Mphahlele, Introduction of new subtypes and variants of hepatitis C virus genotype 4 in South Africa, *Journal of medical virology* 84(4) (2012) 601-607.
- [462] I. Moreau, S. Hegarty, J. Levis, P. Sheehy, O. Crosbie, E. Kenny-Walsh, L.J. Fanning, Serendipitous identification of natural intergenotypic recombinants of hepatitis C in Ireland, *Virology journal* 3 (2006) 1-7.
- [463] M.M. Farag, A.R. Sofy, A.A. Mousa, M.A. Ahmed, M.R. Alganzy, Molecular assay and genotyping of hepatitis C virus among infected Egyptian and Saudi Arabian patients, *Virology: research and Treatment* 6 (2015) VRT. S32016.
- [464] L. Duchesne, R. Njouom, F. Lissack, G.F. Tamko-Mella, S. Rallier, L. Poiteau, A. Soulier, S. Chevaliez, G. Vernet, N. Rouveau, HCV Ag quantification as a one-step procedure in diagnosing chronic hepatitis C infection in Cameroon: the ANRS 12336 study, *Journal of the International AIDS Society* 20(1) (2017) 21446.
- [465] K.-C. Luk, M.G. Berg, S.N. Naccache, B. Kabre, S. Federman, D. Mbanya, L. Kaptué, C.Y. Chiu, C.A. Brennan, J. Hackett Jr, Utility of metagenomic next-generation sequencing for characterization of HIV and human pegivirus diversity, *PloS one* 10(11) (2015) e0141723.
- [466] L.I. Compston, C. Li, F. Sarkodie, S. Owusu-Ofori, O. Opare-Sem, J.P. Allain, Prevalence of persistent and latent viruses in untreated patients infected with HIV-1 from Ghana, West Africa, *Journal of medical virology* 81(11) (2009) 1860-1868.
- [467] T.J. Tucker, S.J. Louw, S.C. Robson, S. Isaacs, R.E. Kirsch, High prevalence of GBV-C hepatitis G virus infection in a rural South African population, *Journal of medical virology* 53(3) (1997) 225-228.
- [468] C. Li, K. Danso, E. Addo-Yobo, A. Dompok, F. Sarkodie, S. Owusu-Ofori, J.-P. Allain, GB virus C genotype 1 is rarely transmitted vertically but acquired during infancy in West Africa, *Aids* 20(10) (2006) 1458-1460.
- [469] A. Casteling, E. Song, J. Sim, D. Blaauw, A. Heyns, R. Schweizer, L. Margolius, E. Kuun, S. Field, B. Schoub, GB virus C prevalence in blood donors and high risk groups for parenterally transmitted agents from Gauteng, South Africa, *Journal of medical virology* 55(2) (1998) 103-108.
- [470] V. Kufner, A. Plate, S. Schmutz, D.L. Braun, H.F. Günthard, R. Capaul, A. Zbinden, N.J. Mueller, A. Trkola, M. Huber, Two years of viral metagenomics in a tertiary diagnostics unit: evaluation of the first 105 cases, *Genes* 10(9) (2019) 661.
- [471] I. Bukowska-Ośko, K. Perlejewski, A. Pawełczyk, M. Rydzanicz, A. Pollak, M. Popiel, K.C. Cortés, M. Paciorek, A. Horban, T. Dzieciatkowski, Human pegivirus in patients with encephalitis of unclear etiology, Poland, *Emerging infectious diseases* 24(10) (2018) 1785.
- [472] A.O. Nwokeoji, P.M. Kilby, D.E. Portwood, M.J. Dickman, RNASwift: A rapid, versatile RNA extraction method free from phenol and chloroform, *Analytical Biochemistry* 512 (2016) 36-46.
- [473] C. Schlötterer, R. Tobler, R. Kofler, V. Nolte, Sequencing pools of individuals—mining genome-wide polymorphism data without big funding, *Nature Reviews Genetics* 15(11) (2014) 749-763.
- [474] M. Laurence, C. Hatzis, D.E. Brash, Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes, *PloS one* 9(5) (2014) e97876.
- [475] B.A. Friedman, T. Maniatis, ExpressionPlot: a web-based framework for analysis of RNA-Seq and microarray gene expression data, *Genome biology* 12(7) (2011) 1-12.

- [476] E. Afgan, D. Baker, M. Van den Beek, D. Blankenberg, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update, *Nucleic acids research* 44(W1) (2016) W3-W10.
- [477] D. Darriba, D. Posada, A.M. Kozlov, A. Stamatakis, B. Morel, T. Flouri, ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models, *Molecular biology and evolution* 37(1) (2020) 291-294.
- [478] A.J. Tighe, M.D. Gallagher, J. Carlsson, I. Matejusova, F. Swords, D.J. Macqueen, N.M. Ruane, Nanopore whole genome sequencing and partitioned phylogenetic analysis supports a new salmonid alphavirus genotype (SAV7), *Diseases of Aquatic Organisms* 142 (2020) 203-211.
- [479] L.-T. Nguyen, H.A. Schmidt, A. Von Haeseler, B.Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, *Molecular biology and evolution* 32(1) (2015) 268-274.
- [480] S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0, *Systematic biology* 59(3) (2010) 307-321.
- [481] B. Obrist, N. Iteba, C. Lengeler, A. Makemba, C. Mshana, R. Nathan, S. Alba, A. Dillip, M.W. Hetzel, I. Mayumana, Access to health care in contexts of livelihood insecurity: a framework for analysis and action, *PLoS medicine* 4(10) (2007) e308.
- [482] S.K. Bakeera, S.P. Wamala, S. Galea, S. Peterson, G.W. Pariyo, Community perceptions and factors influencing utilization of health services in Uganda, *International journal for equity in health* 8(1) (2009) 1-12.
- [483] G.W. Pariyo, E. Ekirapa-Kiracho, O. Okui, M.H. Rahman, S. Peterson, D.M. Bishai, H. Lucas, D.H. Peters, Changes in utilization of health services among poor and rural residents in Uganda: are reforms benefitting the poor?, *International Journal for Equity in Health* 8(1) (2009) 1-11.
- [484] S. Kizito, D. Mukunya, J. Nakitende, S. Nambasa, A. Nampogo, R. Kalyesubula, A. Katamba, N. Sewankambo, Career intentions of final year medical students in Uganda after graduating: the burden of brain drain, *BMC medical education* 15(1) (2015) 1-7.
- [485] D. Musoke, P. Boynton, C. Butler, M.B. Musoke, Health seeking behaviour and challenges in utilising health facilities in Wakiso district, Uganda, *African health sciences* 14(4) (2014) 1046-1055.
- [486] D. Baud, T. Van Mieghem, D. Musso, A.C. Truttmann, A. Panchaud, M. Vouga, Clinical management of pregnant women exposed to Zika virus, *The Lancet Infectious Diseases* 16(5) (2016) 523.
- [487] W.H. Organization, UNICEF, Handbook for clinical management of dengue, (2012).
- [488] C.P. Simmons, K. McPherson, N.V.V. Chau, D.H. Tam, P. Young, J. Mackenzie, B. Wills, Recent advances in dengue pathogenesis and clinical management, *Vaccine* 33(50) (2015) 7061-7068.
- [489] A. Srivathsan, E. Hartop, J. Puniamoorthy, W.T. Lee, S.N. Kutty, O. Kurina, R. Meier, Rapid, large-scale species discovery in hyperdiverse taxa using 1D MinION sequencing, *BMC biology* 17(1) (2019) 1-20.
- [490] J.J.M. Chang, Y.C.A. Ip, A.G. Bauman, D. Huang, MinION-in-ARMS: nanopore sequencing to expedite barcoding of specimen-rich macrofaunal samples from autonomous reef monitoring structures, *Frontiers in Marine Science* 7 (2020) 448.
- [491] D.M. Vu, N. Mutai, C.J. Heath, J.M. Vulule, F.M. Mutuku, B.A. Ndenga, A.D. LaBeaud, Unrecognized dengue virus infections in children, western Kenya, 2014–2015, *Emerging infectious diseases* 23(11) (2017) 1915.
- [492] M.L. Moi, T. Takasaki, A. Kotaki, S. Tajima, C.-K. Lim, M. Sakamoto, H. Iwagoe, K. Kobayashi, I. Kurane, Importation of dengue virus type 3 to Japan from Tanzania and Côte d'Ivoire, *Emerging infectious diseases* 16(11) (2010) 1770.
- [493] K. Okada, R. Morita, K. Egawa, Y. Hirai, A. Kaida, M. Shirano, H. Kubo, T. Goto, S.P. Yamamoto, Dengue virus type 1 infection in traveler returning from Tanzania to Japan, 2019, *Emerging infectious diseases* 25(9) (2019) 1782.

- [494] B. Chipwaza, R.D. Sumaye, M. Weisser, W. Gingo, N.K.-W. Yeo, S.N. Amrun, F.O. Okumu, L.F. Ng, Occurrence of 4 dengue virus serotypes and chikungunya virus in Kilombero Valley, Tanzania, during the dengue outbreak in 2018, *Open forum infectious diseases*, Oxford University Press US, 2021, p. ofaa626.
- [495] S. Proesmans, F. Katshongo, J. Milambu, B. Fungula, H. Muhindo Mavoko, S. Ahuka-Mundeki, R. Inocêncio da Luz, M. Van Esbroeck, K.K. Ariën, L. Cnops, Dengue and chikungunya among outpatients with acute undifferentiated fever in Kinshasa, Democratic Republic of Congo: A cross-sectional study, *PLoS neglected tropical diseases* 13(9) (2019) e0007047.
- [496] B. Kirya, L. Mukwaya, S. Sempala, A yellow fever epizootic in Zika forest, Uganda, during 1972: Part 1: Virus isolation and sentinel monkeys, *Transactions of the Royal Society of Tropical Medicine and Hygiene* 71(3) (1977) 254-260.
- [497] A.J. Haddow, M.C. Williams, J.P. Woodall, D.I. Simpson, L.K. Goma, Twelve isolations of Zika virus from *Aedes (Stegomyia) africanus* (Theobald) taken in and above a Uganda forest, *Bulletin of the World Health Organization* 31(1) (1964) 57.
- [498] M. Mayanja, J.-P. Mutebi, M.B. Crabtree, F. Ssenfuka, T. Muwawu, J.J. Lutwama, Studies on the species composition and relative abundance of mosquitoes of Mpigi District, central Uganda, *Journal of entomology and zoology studies* 2(5) (2014) 317.
- [499] J.-P. Mutebi, M. Crabtree, R. Kading, A. Powers, J. Lutwama, B. Miller, Mosquitoes of western Uganda, *Journal of medical entomology* 49(6) (2014) 1289-1306.
- [500] C. Rückert, G.D. Ebel, How do virus–mosquito interactions lead to viral emergence?, *Trends in parasitology* 34(4) (2018) 310-321.
- [501] L. Barzon, M. Pacenti, E. Franchin, S. Pagni, T. Martello, M. Cattai, R. Cusinato, G. Palù, Excretion of West Nile virus in urine during acute infection, *The Journal of infectious diseases* 208(7) (2013) 1086-1092.
- [502] S.A. Baty, K.B. Gibney, J.E. Staples, A.B. Patterson, C. Levy, J. Lehman, T. Wadleigh, J. Feld, R. Lanciotti, C.T. Nugent, Evaluation for West Nile Virus (WNV) RNA in urine of patients within 5 months of WNV infection, *Journal of Infectious Diseases* 205(9) (2012) 1476-1477.
- [503] C.M. Barbosa, N. Di Paola, M.P. Cunha, M.J. Rodrigues-Jesus, D.B. Araujo, V.B. Silveira, F.B. Leal, F.S. Mesquita, V.F. Botosso, P.M. Zanotto, Yellow fever virus DNA in urine and semen of convalescent patient, Brazil, *Emerging infectious diseases* 24(1) (2018) 176.
- [504] M. Li, B. Wang, L. Li, G. Wong, Y. Liu, J. Ma, J. Li, H. Lu, M. Liang, A. Li, Rift Valley Fever Virus and yellow fever virus in urine: a potential source of infection, *Virologica Sinica* 34(3) (2019) 342-345.
- [505] A.-C. Gourinat, O. O'Connor, E. Calvez, C. Goarant, M. Dupont-Rouzeyrol, Detection of Zika virus in urine, *Emerging infectious diseases* 21(1) (2015) 84.
- [506] R. de M Campos, C. Cirne-Santos, G.L. Meira, L.L. Santos, M.D. de Meneses, J. Friedrich, S. Jansen, M.S. Ribeiro, I.C. da Cruz, J. Schmidt-Chanasit, Prolonged detection of Zika virus RNA in urine samples during the ongoing Zika virus epidemic in Brazil, *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology* 77 (2016) 69-70.
- [507] C.C. Holm-Hansen, S.E. Midgley, T.K. Fischer, Global emergence of enterovirus D68: a systematic review, *The Lancet Infectious Diseases* 16(5) (2016) e64-e75.
- [508] C. Zhang, R. Zhu, Y. Yang, Y. Chi, J. Yin, X. Tang, L. Yu, C. Zhang, Z. Huang, D. Zhou, Phylogenetic analysis of the major causative agents of hand, foot and mouth disease in Suzhou city, Jiangsu province, China, in 2012–2013, *Emerging microbes & infections* 4(1) (2015) 1-10.
- [509] C. Shiff, Integrated approach to malaria control, *Clinical microbiology reviews* 15(2) (2002) 278-293.
- [510] O. mondiale de la Santé, W.H. Organization, Malaria vaccine: WHO position paper–March 2022–Rapport mensuel des cas de dracunculose, janvier 2022, *Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire* 97(09) (2022) 60-78.
- [511] J. Das, J.K. Fallon, C.Y. Timothy, A. Michell, T.J. Suscovich, C. Linde, H. Natarajan, J. Weiner, M. Coccia, S. Gregory, Delayed fractional dosing with RTS, S/AS01 improves humoral immunity to malaria via a balance of polyfunctional NANP6-and Pf16-specific antibodies, *Med* 2(11) (2021) 1269-1286. e9.

- [512] S. Kurkela, T. Manni, A. Vaheri, O. Vapalahti, Causative agent of Pogosta disease isolated from blood and skin lesions, *Emerging infectious diseases* 10(5) (2004) 889.
- [513] S. Miller, S.N. Naccache, E. Samayoa, K. Messacar, S. Arevalo, S. Federman, D. Stryke, E. Pham, B. Fung, W.J. Bolosky, Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid, *Genome research* 29(5) (2019) 831-842.
- [514] G.W. Dick, Epidemiological notes on some viruses isolated in Uganda (Yellow fever, Rift Valley fever, Bwamba fever, West Nile, Mengo, Semliki forest, Bunyamwera, Ntaya, Uganda S and Zika viruses), *Transactions of the Royal Society of Tropical Medicine and Hygiene* 47(1) (1953) 13-48.
- [515] K. Smithburn, A.J. Haddow, A. Mahaffy, A Neurotropic Virus isolated from Aedes Mosquitoes caught in the Semliki Forest, *Amer. J. Trop. Med.* 26(2) (1946) 189-208.
- [516] R.J. Dikdan, S.A. Marras, A.P. Field, A. Brownlee, A. Cironi, D.A. Hill, S. Tyagi, Multiplex PCR assays for identifying all major severe acute respiratory syndrome coronavirus 2 variants, *The Journal of molecular diagnostics* 24(4) (2022) 309-319.
- [517] S. Brown, H. Morrison, N. Karabatsos, D. Knudson, Genetic relatedness of two new Orbivirus serogroups: Orungo and Lebombo, *Journal of general virology* 72(5) (1991) 1065-1072.