PREVALENCE AND GENETIC DIVERSITY OF SIMIAN IMMUNODEFICIENCY VIRUS INFECTING FREE-RANGING NON-HUMAN PRIMATES IN KENYAN URBAN CENTRES

NYAMOTA RICHARD

A Thesis Submitted to the Graduate School in Partial Fulfilment of the Requirements for the Master of Science Degree in Biochemistry of Egerton University

EGERTON UNIVERSITY

FEBRUARY 2020

DECLARATION AND RECOMMENDATION

Declaration

This thesis is my original work and has not been presented in this university or any other for the award of a degree to my knowledge.

Signature: ..... Date: .....

Richard Nyamota

SM14/23502/14

Recommendation

This thesis has been submitted with our approval as University supervisors.

Signature: ..... Date: .....

Dr. Vincent A. Owino, PhD
Department of Biochemistry and Molecular Biology
Egerton University

Signature: ..... Date: .....

Dr. Maamun Jeneby, PhD
Department of Tropical and Infectious Diseases
Institute of Primate Research

DEDICATION

To the family and friends for whose support contributed to completion of this work.

ACKNOWLEDGEMENTS

My foremost profound gratitude goes to God Almighty who has granted me good health, strength and ability to accomplish this study and more so during the challenging academic times. Second, I am grateful to my supervisors Dr. Maamun Jeneby and Dr. Vincent Owino Adung'a for their time, guidance, dedication and involvement in my academic pursuits. Special thanks to Dr. Maamun and Dr. Mercy Akinyi who were kind enough to offer an opportunity to conduct a research project for my thesis under their guidance both in the laboratory and fieldwork. I would also like to thank Dr. Vincent Owino Adung'a whose expertise in molecular cell biology played a big role in shaping and implementation of this study. I am grateful to Dr. Moses Otiende for approving the use of KWS laboratory facilities and consumables. Lastly, I acknowledge Dr. Edwin Kimathi on immense support and expertise for the computational/bioinformatics work, starting from enlightening me on the tools to use for sequence analysis, guidance on phylogenetic analysis and selection pressure. I extend my gratitude to the Institute of Primate Research, International Centre of Insect Physiology and Ecology (*icipe*) and Kenya Wildlife Services for permitting use of their facilities and resources during my research. I also extend my acknowledgments to the World Federation of Scientists (WFS) for the scholarship that facilitated undertaking of this study at *icipe*.

ABSTRACT

Africa green monkeys (AGMs, *Chlorocebus aethiops*) (Gray, 1870) and olive baboons (*Papio anubis anubis*) (Lesson, 1827) are common non-human primates (NHPs) found within major urban centres in Kenya. The widely distributed AGMs are naturally infected with simian immunodeficiency virus (SIV) of the genus *Lentivirus*. While a baboon specific SIV has not been reported so far, studies have demonstrated that baboons are non-receptive to SIV infection leading to low prevalence in these species. Due to enhanced human-wildlife interactions, SIV can potentially infect humans during aggressive encounters resulting in injuries from bites and scratches and potential exposure to infected blood and other body fluids making SIV a virus of medical importance. Despite its zoonotic potential, a comprehensive investigation of SIVs in free-ranging Kenyan monkeys has not been undertaken. Therefore, this study sought to investigate the diversity of SIV strains infecting AGMs and olive baboons found within selected Kenyan urban centres using molecular approaches. Free-ranging NHPs (124 AGMs and 65 olive baboons) from within Mombasa, Kisumu, and Naivasha towns were trapped, and SIV identified by PCR targeting a partial *pol* and *env* gene fragments. Polymerase chain reaction-high resolution melt (PCR-HRM) analysis of *pol* amplicons revealed distinct melt profiles illustrating diverse virus strains. Detected SIV genetic fragments were further characterised by sequencing and phylogenetic analysis. We detected SIV in 32% (39/124) and 3% (2/65) of AGMs and baboons, respectively. Phylogenetic analysis of the *pol* and *env* gene sequences demonstrated that diverse host species-specific SIV (SIVagm) strains infect AGMs populations without definite phylogeographical groupings. Moreover, analysis of the evolutionary selection demonstrated signatures of episodic and pervasive diversification on the *env* gene suggesting continuous SIV evolution within the natural host which is crucial for a virus to be able to cross the species barrier and infect a new host. Notably, for the first time, this study partially characterised a strain of SIVagm infecting olive baboons indicating putative cross-infection among sympatric NHP species. Additional elaborate studies are required to conclusively decipher the prevalence, pathogenesis and immunological response associated with SIVagm infection of baboons. Better understanding of prevalence and diversity of potentially zoonotic SIV strains circulating in NHP hosts is crucial in controlling emergence of infections in human.

TABLE OF CONTENTS

DECLARATION AND RECOMMENDATION ..... **ii**
COPYRIGHT ..... **iii**
DEDICATION ..... **iv**
ACKNOWLEDGEMENTS ..... **v**
ABSTRACT ..... **vi**
TABLE OF CONTENTS ..... **vii**
LIST OF TABLES ..... **ix**
LIST OF FIGURES ..... **x**
LIST OF ABBREVIATIONS AND ACRONYMS ..... **xi**
CHAPTER ONE **1**
INTRODUCTION **1**
1.1 Background information 1
1.2 Statement of the problem 4
1.3 Objectives 4
1.3.1 General objective 4
1.3.2 Specific objectives 4
1.4 Hypotheses 5
1.5 Justification of the study 5
CHAPTER TWO **6**
LITERATURE REVIEW **6**
2.1 Emerging infections 6
2.2 Viral emerging infectious diseases 7
2.3 Origin of Human Immunodeficiency Virus (HIV) 8
2.4 Simian immunodeficiency virus (SIV) 9
2.4.1 Lineages of primate lentiviruses 9
2.4.2 Pathogenesis and transmission 10
2.4.3 Lentiviruses structure and replication mechanism 11
2.4.4 Recombination 12
2.4.5 Epidemiology and prevalence of SIV 14
2.4.6 Genetic diversity 14
2.5 Non-human primates as a source of emerging infections 15
CHAPTER THREE **17**
MATERIALS AND METHODS **17**

3.1 Ethics statement and permits 17
3.2 Summary of experimental design 17
3.3 Sampling sites 18
3.4 Sampling of NHPs 19
3.5 DNA extraction 19
3.6 Molecular identification of SIVs 19
3.7 Purification of PCR amplicons 22
3.8 Phylogenetic and evolutionary selection analyses 22
3.8.1 Phylogenetic analyses 22
3.8.2 Analyses of evolutionary selection 22
CHAPTER FOUR 24
RESULTS 24
4.1 Species of monkeys and their distribution 24
4.2 Prevalence of SIV in AGMs and baboons 25
4.3 High resolution melting profiles 26
4.4 SIV genetic diversity 28
4.4.1 Identification of SIV sequence lineages 28
4.4.2 Phylogenetic analyses of SIVagm infecting AGMs and baboons in Kenya 33
4.4.3 SIV *pol* and *env* protein diversity 39
4.5 Evolutionary selection 41
4.6 Summary 42
CHAPTER FIVE 43
DISCUSSION 43
CHAPTER SIX 50
CONCLUSIONS AND RECOMMENDATIONS 50
6.1 Conclusions 50
6.2 Recommendations 50
REFERENCES 51
APPENDICES 66
Appendix 1: Supplementary tables 66
Appendix 2: Supplementary figures 70
Appendix 3: Publication 73

LIST OF TABLES

Table 1: The major SIV lineages that have been characterised in Africa. 9
Table 2: The distribution and demographics of non-human primates (NHPs) sampled from the various sites of urban and peri-urban centres in Kenya. 24
Table 3: Prevalence of SIV infection in NHPs (AGMs and olive baboons) sampled from Kisumu, Naivasha and Mombasa towns in Kenya. 26
Table 4: SIV *env* gene codons under evolutionary selection pressure. 41
Table S1: Prevalence of SIV infection in AGMs and olive baboons from Mombasa, Naivasha and Kisumu as detected by PCR. 66
Table S2: Percent SIV *pol* amino acid identity matrix 67
Table S3: Percent SIV *env* amino acid identity matrix 68
Table S4: Analysis of the synonymous and non-synonymous substitution rates (dS and dN) for the *env* protein. 68
Table S5: Sample identification for NHPs that were positive for SIV and their respective accession numbers for sequences submitted to NCBI's GenBank database. 69

LIST OF FIGURES

Figure 1: A global map of emerging infectious diseases. 7
Figure 2: Proposed origin of Human Immunodeficiency Virus.. 13
Figure 3: Summary of the experimental pipeline for the study. 17
Figure 4: Urban and peri-urban sampling sites.. 18
Figure 5: SIV genome architecture and the PCR amplification strategy.. 21
Figure 6: Detection of SIV in NHPs by PCR-HRM. 25
Figure 7: Multiple sequence alignment for the 650 bp partial *pol* gene 27
Figure 8: SIV percentage sequence identity range matrices. 29
Figure 9: Percentage identity matrix of SIV partial *pol* sequence alignments for identification of SIV lineage.. 32
Figure 10: Percentage identity matrix of SIV partial *env* sequence alignments for identification of SIV lineage. 33
Figure 11 Maximum likelihood tree of the partial *pol* gene (650bp) from across Africa. 34
Figure 12: Maximum likelihood tree of the partial *env* gene (900 bp) from across Africa. 36
Figure 13: Maximum likelihood phylogenetic tree generated from partial *pol* gene (A) and *env* gene (B) from Kenya only. 38
Figure 14: SIV genome architecture and partial *env* protein domains 40
Figure S1: Multiple sequence alignments for the *pol* and *env* p acid residues 70
Figure S2: Maximum likelihood tree generated from partial *pol* protein 71
Figure S3: Maximum likelihood tree generated from partial *env* protein. 72

LIST OF ABBREVIATIONS AND ACRONYMS

AIDS	Acquired Immunodeficiency Syndrome
AGM	African Green Monkey
BLAST	Basic Local Alignment Search Tool
CA	California
cDNA	Complementary Deoxyribonucleic Acid
CRFs	Circulating Recombinant Forms
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide Triphosphates
EI	Emerging Infection
GAC	Government Administration Compound
GC	Golf Course
HIV	Human Immunodeficiency Virus
icipe	International centre of insect physiology and ecology
ISERC	Institutional Scientific Review Committee
IPR	Institute of Primate Research
KAJ	Kajulu
KWS	Kenya Wildlife Services
MBK	Mbaraki
MERS	Middle East Respiratory Syndrome
NEB	New England Biolabs
NCBI	National Centre for Biotechnology
NHP	Non-Human Primate
PCR-HRM	Polymerase Chain Reaction- High Resolution Melting
RNA	Ribonucleic acid
SARS	Severe Acute Respiratory Syndrome
SIV	Simian Immunodeficiency Virus
TB	Tuberculosis
TRIM5α	Tripartite Motif-containing protein 5
TWS	Tudor water sport
UNAIDS	United Nations Programme on HIV/AIDS
VPU	Viral Protein U
WFS	World Federation of Scientists
WHO	World Health Organization

CHAPTER ONE
INTRODUCTION

1.1 Background information

Simian immunodeficiency virus (SIV) belongs to the *Lentivirus* genus *Retroviridae* family (Desrosiers, 1990) and naturally infects over 30 species of African anthropoid non-human primates (NHPs) including chimpanzees and gorilla (Apes), and simian primates (monkeys) (Peeters *et al*., 2002). It has been shown that SIVcpz strain originating from chimpanzees (*Pan troglodytes*) and SIVsmm strain from sooty mangabeys (*Cercocebus atys*) represents zoonotic cross-species infections that produced the human immunodeficiency virus type 1 (HIV-1) and HIV-2 respectively (Chen *et al*., 1996; Sharp *et al*., 2005). HIV is one of the most devastating infectious disease in humans responsible for 1.1 million deaths in 2018 while an approximately annual budget of 26.2 billion dollars is spent on acquired immunodeficiency syndrome (AIDS) response (UNAIDS, 2019). The HIV virus first discovered in 1985 by Gallo & Montagnier (2003), is thought to have emerged from zoonotic cross-species transmission from NHPs (Sharp & Hahn, 2011) and was considered an emerging infection (EI) in the twentieth century. This demonstrates the human risk of contracting deadly infectious pathogens that originate from wild and domestic animals.

More than 60.3% of emerging infections (EIs) are caused by pathogens that infect animals and can be transmitted into humans i.e. zoonotic and 71% are of wildlife origin (Jones *et al*., 2008; Petersen *et al*., 2018). They are responsible for more than 25% of 57 million annual deaths (Wang *et al*., 2016) and approximately 30% of disability adjusted life years globally (Chua & Gubler, 2013; James *et al*., 2018). In addition, economic losses of approximately 120 billion dollars in the past decade has been incurred due to EIs alone (Smith *et al*., 2019). Due to the nature of EIs, their direct and indirect impact on socio-economic activities will increase. For example, most of these diseases are zoonotic and with increasing human population necessitating change of land use, increased interactions between people and animals is inevitable, resulting into acquisition of new infections that can be rapidly translocated through international travel and trade (Rhyan & Spraker, 2010). Further, the poor state of public health infrastructure in some regions does not allow efficient control and management of existing and new infections. An excellent example is the recent (2019) Ebola outbreak in Democratic Republic of Congo (DRC). This therefore calls for improved knowledge to inform on appropriate management and control of known EIs and/or emergent ones. In part, it is important to understand potential zoonotic pathogens in their natural non-human hosts to aid in predicting the potential risks of cross-species transmission (Rhyan & Spraker, 2010).

A potentially high risk virus is SIV, a single stranded RNA virus is known to naturally infect 45 different species of NHPs including apes and monkeys in Africa (Peeters *et al*., 2012). Several studies have demonstrated that SIV has crossed and infected various NHPs such as chimpanzee and sooty mangabeys on multiple occasions (Gao *et al*., 1999; Sharp & Hahn, 2011; D'arc *et al*., 2015), and possibly evolved into HIV that caused the AIDS pandemic (Sharp & Hahn, 2010). This strongly suggests that active transmission of retroviruses from NHP to humans (Wolfe *et al*., 2004; Locatelli & Peeters, 2012; Peeters & Delaporte, 2012) is possible, and is a potential source of novel EIs. Therefore, there is need to identify and understand microbes of NHPs, specifically viruses that can evolve into highly infectious human pathogens.

SIV poses a threat of transmission to humans because of hunting and consumption of NHPs as bushmeat (Friant *et al*., 2015; Steve *et al*., 2017). Handling and processing of NHP bushmeat exposes humans to infected blood and other body fluids thereby serving as major route of human infection with SIV. Phylogenetic analysis of SIV in African NHPs has demonstrated species specificity where viruses cluster according to their respective host-species (Locatelli & Peeters, 2012). This has been documented in viruses infecting *Chlorocebus* genus that includes AGMs (Wertheim & Worobey, 2007) and the arboreal *Cercopithecus* genus (Bibollet-Ruche *et al*., 2004). Further, epidemiological studies have revealed high SIV prevalence in wild AGM populations in West and Southern Africa (Ma *et al*., 2013). High prevalence of SIV and diversity in NHPs sharing a habitat (sympatric) has been associated with cross-species transmission of the virus. Examples include cross-species transmission between sympatric AGMs and chacma baboons (*Papio ursinus*) in Southern Africa (van Rensburg *et al*., 1998), yellow baboons (*Papio cynocephalus*) in Tanzania (Jin *et al*., 1994) and patas monkey (*Cercopithecus aethiops sabaeus*) in West Africa (Bibollet-Ruche *et al*., 1996). Despite the high prevalence of natural SIV infections in different NHPs species, comprehensive surveys and molecular characterisation of this lentivirus, especially from East African monkeys remain limited.

AGM/vervet monkey (*Chlorocebus aethiops*), Sykes' monkey (*Cercopithecus mitis*) and olive baboon (*Papio anubis anubis*) are common NHPs found within most urban and peri-urban environments in East Africa (Newman *et al.*, 2004; de Jong & Butynski, 2010). These are the most widely distributed NHPs in Kenya and regarded as pests in most towns and villages (Jeneby *et al*., 2011; Kagira *et al*., 2017; Mutemi *et al*., 2017). As primates of diverse habitats, synanthropic or urban-restricted free ranging monkeys are potential sources of zoonotic viral diseases of human health importance such as chikungunya (Tsetsarkin *et al*., 2016), dengue, yellow fever, and Zika viruses (Weaver *et al*., 2018). Despite their importance as reservoir

hosts of zoonotic viral diseases, surveys for SIV infections in Kenyan monkeys have been opportunistic, and mostly on captive AGMs (Otsyula *et al*., 1996) and Sykes' monkeys (Ellis *et al*., 2004). In addition, comprehensive molecular epidemiology studies of SIV strains circulating in wild monkey populations in Kenya have not been performed to date. Opportunistic sampling of captive animals, difficulties in collecting and storing field samples have largely contributed to the understudy of SIV epidemiology in NHPs found within Kenya and East Africa at large.

Natural SIV infections of AGMs and Sykes' monkeys generally do not progress to AIDS like symptoms, contrary to SIV experimental infection of Asian macaque monkeys (*Macaca mulatta*) (Silvestri *et al*., 2007) and HIV infection in humans (Hirsch *et al*., 1995). These observations suggest the possibility of African NHPs being natural hosts to species-specific SIV lineages as demonstrated by phylogenetic clustering of known strains (Peeters *et al*., 2002; Ma *et al*., 2013; Steve *et al*., 2017). The AGMs are a super-species comprising of several phenotypically and geographically distinct subspecies of monkeys (Groves, 2016). These subspecies harbour distinct, but related SIV strains considered as subtypes that cluster according to their respective hosts within the SIVagm lineage (Johnson *et al*., 1989; Johnson *et al*., 1990; Hirsch *et al*., 1993; Carr *et al*., 2017). The AGMs are of specific interest in SIV epidemiology given their status as the largest known reservoirs of the virus with reported prevalence of up to 60% (Hirsch *et al*., 1995; Vanderford *et al*., 2006). Unlike AGMs, baboon-specific SIV strain has not been documented. However, previous seroprevalence survey of yellow baboons in Tanzania showed exposure to SIV (Jin *et al*., 1994). Follow-up PCR amplification and sequence analyses of these specimens reported infection by SIVagm originating from AGMs (Jin *et al*., 1994). Another study isolated and sequenced SIVagm strain from a healthy chacma baboon (*Papio ursinus*) sampled in South Africa (van Rensburg *et al*., 1998). Both studies suggest that cross-species transmission of SIV from AGMs to sympatric populations of baboons is occurring due to exposure to infected AGMs during predation. However, SIV status in olive baboons (*Papio anubis*) remains unknown despite reports of SIVagm exposure in its counterparts, yellow and chacma baboons.

Transmission of SIV from NHPs to humans can be influenced by the frequency of contacts and prevalence of infection in the animal host among other factors (Steve *et al*., 2017). In Kenya, there has been increased frequency of contacts between humans and NHPs mainly through increased illegal bush-meat trade, which involves diverse monkey species (Kimwele *et al*., 2012). However, investigations on SIVs distribution in NHPs found near large urban populations has been scarce. Most studies on SIVs are focused on surveillance of the virus in

African NHPs inhabiting rural and forest habitats (Ma *et al*., 2013; Steve *et al*., 2017). Nonetheless, urban-restricted NHPs can also harbour these pathogens, with the potential of cross-species transmission into naïve human population.

Therefore, this study sought to determine the prevalence and diversity of SIV strains recovered from common free-ranging NHPs in Kenya. The targeted NHPs included AGMs and olive baboons within urban and peri-urban centres in Kenya, where interaction with human is potentially high. Molecular techniques including polymerase chain reactions (PCR) coupled with high-resolution melting (HRM) analysis, gel electrophoresis and sequence analysis were used to screen blood samples from the targeted NHP species for SIV.

1.2 Statement of the problem

Molecular characteristics of pathogens such as SIV from free-ranging NHPs in Kenya remain poorly understood despite their immense risk of zoonotic transmission due to enhanced human-NHP interactions. This interaction usually occur during the process of butchering and preparation of bushmeat thereby increasing the risk of human infection. This phenomenon necessitates molecular detection, characterisation and phylogenetic analysis of this potentially zoonotic virus in order to determine their prevalence, diversity and its evolutionary changes. Currently, SIV prevalence has largely been determined from serological surveys on a few opportunistic samples from wild caught NHPs and those from captive animals. Comprehensive data on molecular prevalence and diversity of SIV variants circulating in common and widely distributed monkeys in East Africa is also scarce.

1.3 Objectives

1.3.1 General objective

This study investigated the prevalence and diversity of SIV strains infecting free-ranging *Chlorocebus aethiops* (Africa green monkeys – AGMs) and *Papio anubis anubis* (olive baboons) from selected urban centres in Kenya.

1.3.2 Specific objectives

i. To determine the prevalence of SIV in free-ranging AGMs and olive baboons inhabiting Mombasa, Naivasha and Kisumu urban environs.
ii. To determine genetic diversity of SIV in free-ranging AGMs and olive baboons in Mombasa, Naivasha and Kisumu urban environs.
iii. To identify SIV evolutionary patterns associated with host-species.

1.4 Hypotheses

The following were the null hypothesis for this study

i. There is no difference in SIV prevalence and genetic diversity among non-human primates found in Mombasa, Naivasha and Kisumu geographical regions.
ii. There is no genetic diversity, cross-species transmission and recombination of SIVs among sympatric primate species from urban ecological zones in Kenya.
iii. There is no SIV evolutionary patterns associated with the host-species in Kenya.

1.5 Justification of the study

This study proposes to bridge the knowledge gap on comprehensive molecular survey of SIV diversity in common East African monkeys, *Chlorocebus aethiops* (Africa green monkeys - AGMs) and *Papio anubis anubis* (olive baboons) found in densely populated urban centres in Kenya. Free-ranging populations of NHPs are regarded as potential sources of pathogens that can cause infectious diseases in humans. One such potential pathogen is SIV discovered in AGMs and considered as a human zoonotic virus similar to HIV (Peeters *et al*., 2014). Molecular analysis of SIV strains in free-ranging NHPs and their evolutionary trends are important epidemiological factors. First, frequent human-NHP interactions increase the risk of SIV transmission to humans and probably initiate new HIV or HIV-like pandemic. Secondly, extensive SIV transmission and strains co-circulation in urban-restricted NHPs may lead to virus recombination resulting into new strains capable of infecting other hosts species including humans. This potential risk can only be demonstrated if our knowledge on SIV prevalence in monkeys found near human settlements is well understood with detailed analysis of the virus genetic diversity and its evolution. Lastly, knowledge on NHP retroviral reservoir will enable us to update, adapt and improve the current diagnostic tools and thus be able to identify a wide diversity of simian retroviruses at interfaces where the emergence of new zoonotic diseases is likely to occur.

CHAPTER TWO
LITERATURE REVIEW

2.1 Emerging infections

Emerging infections (EIs) are newly recognised infections or those that previously existed, but whose incidence is rapidly increasing or expanding in geographic boundaries and host range (Morse, 2001). EIs are grouped into two, namely, emerging infectious disease (EIDs) and re-emerging infectious diseases (re-EIDs). EIDs are defined as infectious diseases whose incidence is increasing following its first introduction into a new host population or whose incidence is increasing in an existing host population (Engering *et al*., 2013). Examples include HIV/AIDS, Ebola haemorrhagic fever and severe acquired respiratory syndrome (SARS), Middle East respiratory syndrome (MERS) and H1N1 influenza (Swine flu) amongst others. Re-EIDs are those that had been considered eradicated or controlled, but their incidences are threatening to increase or reappear in new geographical locations or in drug resistant forms due to mutations (Morens & Fauci, 2013). Examples of re-EIDs include multi-drug resistant tuberculosis, cholera, dengue fever and polio. EIs are of global health concern, with the 2014 - 2016 Ebola outbreak in West Africa, and present cases in DRC demonstrating their global impact. Therefore, their control and management is of great importance.

Globally, emerging infections directly account for 15 million deaths (Wang *et al.*, 2016), and 30% of disability adjusted life years annually (Chua & Gubler, 2013; James *et al*., 2018) with up to 120 billion dollars in economic loses (Smith *et al*., 2019). Developing countries have been majorly impacted with the burden of morbidity and mortality, where more than 3 million children succumb annually due to EIs alone (WHO, 2016) despite the risk of EIDs/re-EIDs being globally distributed (Figure 1).

Reasons for increased risk from EIs include increasing human population (Peeters *et al*., 2002; Rhyan & Spraker, 2010) that has necessitated changes in land use thus increasing human-wildlife interactions and chances of zoonotic transmission of pathogens (Pigott *et al*., 2015). Trade globalisation on the other hand propagates spread of the diseases (Morens *et al.*, 2004).

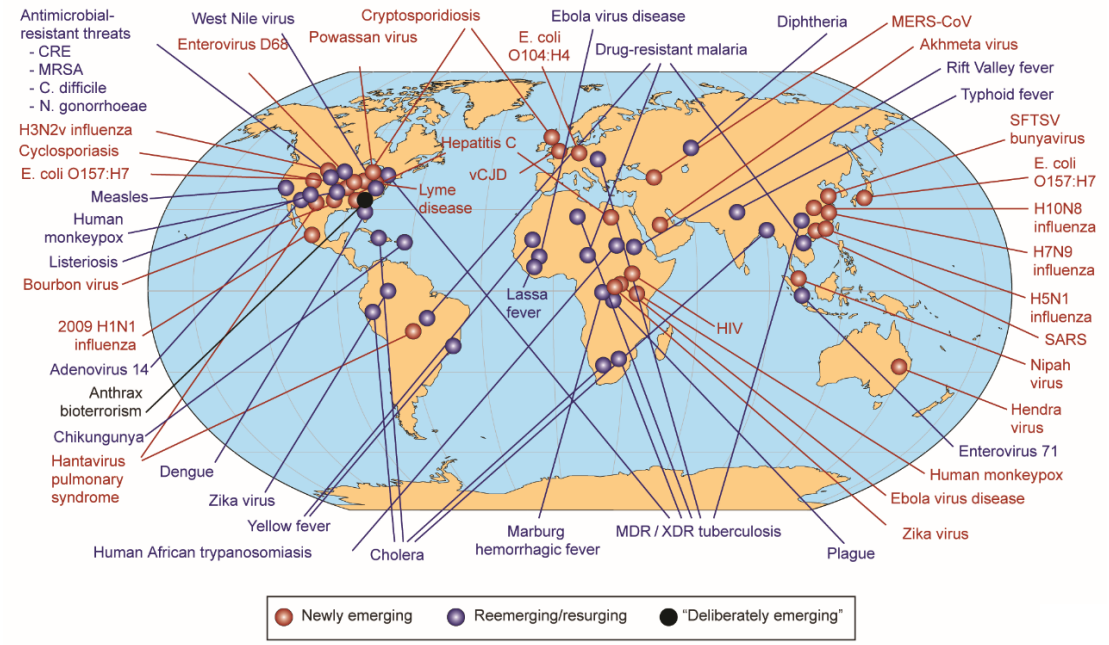


Figure 1: A global map of emerging infectious diseases. The global distribution of emerging, re-emerging and deliberately emerging pathogens of viral, bacterial, fungal and protozoan origins are shown. The figures is adapted from Paules *et al*., (2017).

Additionally, high pathogen prevalence in animals that are in frequent interaction with humans have also been implicated as a factor enhancing zoonotic transfer of pathogens into human population (Keesing *et al*., 2010). Poor public health infrastructures in developing countries hamper the efforts to efficiently control and manage present and new infections. Some pathogens that have crossed species barrier and caused diseases in humans include bacteria, viruses, protozoan parasites and helminths (Bengis *et al*., 2004). Pathogens with high mutational rates such as RNA viruses and those capable of acquiring foreign genetic materials such as bacteria and pathogens infecting multiple hosts are more likely to turn into an emerging disease agent (Engering *et al*., 2013).

2.2 Viral emerging infectious diseases

Viruses accounts for over 50% of zoonotic pathogens with more than a third of EIs emanating from viruses with RNA genomes (Howard & Fletcher, 2012). Ebola haemorrhagic fever, H1N1 influenza, SARS, MERS, Zika virus and HIV/AIDS are examples of viral EIs of significant public health concern and a considerable global socioeconomic impact. On average, at least two new human viruses are reported annually further increasing the existing disease burden (Woolhouse *et al*., 2008; 2012). Examples of newly reported viruses include human bocavirus, parvovirus 4 (Woolhouse *et al*., 2012), human hepegivirus 1 (HHpgV-1) and Merkel

Cell Polyomavirus (Kapoor *et al*., 2015), human Usutu virus (Simonin *et al*., 2018), porcine enteric alphacoronavirus (Fu *et al*., 2018) and Matryoshka RNA virus 1 (Charon *et al*., 2019). Emergence of pathogenic human viruses correlate to factors such as broad host range and a conserved host receptor across diverse potential hosts that permit invasion (Woolhouse *et al*., 2012). Bats, rodents and NHPs are important sources of zoonotic viruses (Peeters & Delaporte, 2012; Friant *et al*., 2015) thus making increased human interactions with these animals a potential source of human infections. Consequently, upon zoonotic acquisition of the virus, urbanisation and increased transnational travels have enhanced human-human transmission on a global scale (Engering *et al*., 2013).

2.3 Origin of Human Immunodeficiency Virus (HIV)

Acquired immunodeficiency syndrome (AIDS) is a worldwide pandemic in humans caused by two strains of HIV virus (HIV-1 and HIV-2) suspected to have emerged from cross-species infection of humans with SIV (Sharp & Hahn, 2011). AIDS is one of the most devastating infectious disease in the modern world associated with 1.2 million deaths and annual losses of USD 21.7 billion (UNAIDS, 2019). Developing countries have experienced the greatest HIV/AIDS morbidity and mortality, with the highest prevalence rates recorded in young adults in sub-Saharan Africa (Dwyer-Lindgren *et al*., 2019). Although AIDS related deaths have been reduced by antiviral administration (Barouch, 2008; Richman *et al*., 2009), the realisation of curative therapy for HIV eradication is still elusive. Both HIV-1 and HIV-2 comprise several subgroups that probably emerged from independent cross-species transmission events of SIV from NHPs into human (Van Heuverswyn *et al*., 2007).

Despite infecting different hosts and existence of differences in some of their regulatory genes, both SIV and HIV commonly share structural genes also found in other lentiviruses thus suggesting their relatedness (Coffin, 1992). Phylogenetic analyses of various HIV subgroups have consistently shown the virus interspersing between respective SIV lineages thus suggesting the origin of HIV to be cross-species transmission of SIV from NHPs (Sharp & Hahn, 2011). Studies have shown that over 45 species of old world monkeys are infected with SIV without pathogenic outcomes leading to conclusion that these species are natural hosts of SIV (Peeters *et al*., 2014). Transmission of SIV to human must have occurred through exposure to infected ape tissues and body fluids, which are common in bush-meat hunting and human wildlife conflicts (Peeters *et al*., 2002). Monkey bushmeat hunting and consumption has been reported in Kenya (Kimwele *et al*., 2012; Mwangi *et al*., 2016), thus implying that the risk of SIV transmission into humans is high. Recently, repeated SIV cross-species transmission into

humans leading to emergence of a divergent HIV-2 strain was documented in Central Africa (Ayouba *et al*., 2013). The AIDS pandemic can be managed by preventing transmission of the existing strains, while also actively monitoring SIV to avert potential risk of cross-species transmission of the virus into humans.

2.4 Simian immunodeficiency virus (SIV)

Simian immunodeficiency virus (SIV) is a single stranded RNA virus that naturally infects over 36 different species of non-human primates (NHPs) including apes and monkeys in Africa. SIV and HIV are closely related species of viruses with taxonomy classifying them within *lentivirus* genus of *Orthroretrovirinae* subfamily and *Retroviridae* family (Petropoulos, 1997). It is estimated that SIV has been infecting NHPs for over 32,000 years before being recently introduced to humans as HIV via cross-species transmission (Klatt *et al*., 2012).

2.4.1 Lineages of primate lentiviruses

Based on comparative nucleotide sequences, primate lentiviruses are categorised into different lineages (Table 1) though continued characterisation of viruses from more NHPs has resulted in an increase of these lineages.

Table 1: The major SIV lineages that have been characterised in Africa.

Primate Species	Lineage	Reference
Chimpanzees (*Pan troglodytes troglodytes*)	SIVcpz	Gao *et al*., (1999)
Sooty mangabeys (*Cercocebus atys*)	SIVsmm	Hirsch *et al*., (1995)
African green monkeys (*Chlorocebus* genus)	SIVagm	Jin *et al*., (1994)
Sykes' monkeys (*Cercopithecus albogularis*)	SIVsyk	Locateli (2008)
L'Hoest monkeys (*Cercopithecus lhoesti*), sun-tailed monkeys (*Cercopithecus solatus*)	SIVlhoest/sun	Hirsch *et al*., (1995)
Colobus monkey (*Colobus guereza*)	SIVcol	Locatelli *et al*., (2008)
Greater spot-nosed monkey	SIVgsn	Courgnaud *et al*., (2002)
De'Brazza monkey	SIVdeb	Bibollet-Ruche *et al*., (2004)
Dent's Mona monkey (*Cercopithecus denti)*	SIVmon	Dazza *et al*., (2005)
Mustached monkey (*Cercopithecus cephus*)	SIVmus	Aghokeng *et al*., (2007)
Black and white colobus monkey (*Colobus guereza*)	SIVcol	Lauck *et al*., (2013)

Due to co-evolution and co-speciation with the host, SIV lineages have been observed to be host specific where closely related host species might also share related SIV lineages. Studies have demonstrated that these various lineages are approximately equidistant from each other, sharing between 40 - 50% identity in the most conserved *gag* and *pol* genes, while also exhibiting a common genomic organisation encoding structural and enzymatic proteins (Klatt *et al*., 2012). However, these viruses also might selectively contain other unique genes such as *vpu* and *vpx* that are not common to all primate lentiviruses. Despite these viruses being categorised under different lineages, consistent clustering together upon phylogenetic analysis (Bibollet-Ruche *et al*., 2004), thus suggesting existence of an ancient SIV lineage in the distant past has implied a common evolutionary history. Existence of more than 60 species of old world monkeys necessitates continued SIV characterisation in order to understand their zoonotic potential.

2.4.2 Pathogenesis and transmission

Non-human primates that naturally harbour SIV do not exhibit disease conditions (Chakrabarti *et al*., 1991; Ma *et al*., 2014), while AIDS-like symptoms are observed in NHPs that do not naturally harbour the virus such as the Asian macaques (Pandrea *et al*., 2009). This phenomenon has been attributed to prolonged association between the virus and the host leading to evolution of a weaker immune response protecting the host against immunopathological damage (Seki & Matano, 2012). Down-regulation of immune response during acute infection, which is readily reversed by the end of the acute phase, has been observed in NHPs (Liovat *et al*., 2009). Despite an overall non-pathogenicity in their natural host species, studies have revealed that SIVs can potentially cause AIDS upon accidental or experimental introduction into unnatural hosts such rhesus macaques (Apetrei *et al*., 2005). Pathogenicity of SIV in susceptible hosts has also been attributed to allelic variations in virus restriction factors such as Trim5α in rhesus macaques (Kirmaier *et al*., 2010).

Within species, there are various routes for SIV transmission, with sexual route being the main mode of transmission amongst feral NHPs (Ma *et al*., 2013; 2014); few cases of mother to child transmission can also occur (Chahroudi *et al*., 2011; 2014). However, low levels of CCR5 receptors expressed in CD4+ cells confer SIV resistance to breastfeeding infants in AGMs and Mandrills (Sodora *et al*., 2009). Secondly, oral viral transmission is also possible in the course of cleaning wounds sustained from injuries inflicted from biting or fighting (Ma *et al*., 2013). Injuries sustained during aggressive confrontations amongst sympatric NHPs have also resulted in cases of cross-species transmission (Jin *et al*., 1994). Some apes and

monkeys harbour unique SIVs capable of replicating in human cells *in vitro* (Wolfe *et al*., 2004) suggesting potential of infecting humans. Thus, enhanced human-NHP interaction is a possible source of new human infections. Moreover, high pathogen prevalence in NHPs in addition to viral and host factors that are important in invasion, establishment of infection and consequently disease (Aghokeng *et al*., 2010) enhances the zoonotic risk.

2.4.3 Lentiviruses structure and replication mechanism

SIV is a spherical retrovirus whose structure comprises a viral envelope enclosing a dense core composed of two identical single-stranded RNA molecules, viral enzymes and structural proteins (Coffin, 1992). The viral gene is flanked by LTR (long terminal repeat) region at both 3'and 5'ends (Shah *et al.*, 2014). The 5' LTR regions is critical for transcription of viral genes since it codes for the transcriptional promotor. The viral genetic material codes for nine genes encoding 19 proteins. Three of these genes encode for structural proteins including the group-specific antigen (*gag*), envelope polyprotein (*env*) and the polymerase (*pol*) that are important for virus replication (Coffin, 1992). *Gag* forms the viral core protein that include the capsid protein (CA, P24), nucleocapsid (NC, P7) and outer core membrane protein (MA, P17) (Coffin, 1992). The *pol* polyprotein is cleaved into reverse transcriptase, integrase and protease enzymes.

Six of the remaining genes code for proteins that play regulatory roles and hence control virus tropism and replication (Faust *et al.*, 2017). These include *tat* that initiates viral genes transcription, *rev* that facilitates mRNA translocation from the nucleus into the cytoplasm during replication the *nef* protein that regulates expression of surface receptors on virus-infected. The other include *vif* protein that enhances viral infectivity, the *vpr* that facilitates translocation of the viral core from the cytoplasm into the nucleus and *vpu* that enhances virion release from the host-cell (Faust *et al.*, 2017). Other proteins include the reverse transcriptase enzyme that converts viral RNA into viral cDNA that is subsequently integrated into the host genome by integrase enzyme. This enables the virus to utilise the host's replication machinery (Luciw & Leung, 1992). The extreme error-rate exhibited by reverse transcriptase due to lack of proof-reading activity alongside the high chances of recombination during replication of viral RNA results in high genetic diversity of retroviruses. The envelope glycoprotein is the main viral structural protein that facilitate attachment and host-cell entry.

SIV entry into host-cells is initiated by the virus binding to cluster differentiation antigen 4 (CD4) receptors on the T-cells using viral receptors on the *env* glycoprotein (gp120) (Petrovas *et al.*, 2012). CCR5 is the main chemokine coreceptors that enhance SIV binding to

CD4 receptors (Edinger *et al.*, 1999). Following attachment, the SIV *env* glycoprotein undergo conformational changes leading to exposure of the variable domains (V3 loop) that activates proteolytic enzymes necessary for virus and host-cell membrane fusion (Prabakaran *et al*., 2007). After fusion, the capsid (which contains the RNA, reverse transcriptase, proteases, ribonuclease, and integrase) is injected into the host cells' cytoplasm (Engelman & Cherepanov, 2012). Inside the host-cell, P24 facilitates viral RNA dissociation from the capsid which is then reverse transcribed into a double stranded cDNA molecule. The generated cDNA is translocated into the nucleus where it is integrated into the host's genome by integrase and awaits transcription (Asante-Appiah & Skalka, 1997).

As the infected host cell transcribes its genetic material, it also transcribes the integrated viral genetic materials into mRNA (Luciw & Leung, 1992). Earlier mRNAs to be transcribed are major viral regulatory genes that include *tat*, *rev* and *nef* (Faust *et al.*, 2017). The mRNA transcripts encoding viral structural proteins such as envelope polyprotein (gp160) and the *gag-pol* polyprotein are translated into respective proteins using the host-cell machinery. The gp160 is then translocated through the endoplasmic reticulum into the Golgi complex, where it is cleaved by cellular proteases into gp41and gp120 proteins that are subsequently glycosylated (Luciw & Leung, 1992). As the virion matures, the *gag-pol* polyprotein is cleaved into individual SIV core proteins P24, P4 and P7 by viral proteases. It is during this cleavage process that reverse transcriptase, integrase and protease enzymes are also formed. Assembly of viral proteins alongside two genomic viral RNA molecules is catalysed by *vpu* and *vif* on the host-cell's plasma membrane after which the mature SIV virion buds off and can infect other cells (Luciw & Leung, 1992).

2.4.4 Recombination

The rates of recombination in RNA viruses such as SIV are known to be high due to extensive cross species transmission and co-infection with related strains in each species. Recombination plays a role in generation of new SIV variants in NHPs including strains with zoonotic potential (Chen *et al*., 1996). Divergence of recombinant strains from the representative SIV lineage strains has hampered the efforts of establishing their parental strains upon comprehensive sequence analysis (Klatt *et al*., 2012). The most notable examples of recombinant lentivirus strains include HIV-1 in humans and its intermediate ancestor SIVcpz found in chimpanzee (Sharp & Hahn, 2011). Recombination of viruses from *Cercopithecinae* family such as SIVrcm, and SIVmus (Figure 2) is thought to have generated SIVcpz infecting chimpanzees (Bailes *et al*., 2003) which further evolved into HIV-1 (Sharp & Hahn, 2010). To

successfully infect human host, SIV underwent evolutionary changes such as loss of viral protein X (VPX) followed by generation a unique viral protein R (VPR) (Figure 2) that was crucial in overcoming human restriction factors such as tetherin and TRIM5α proteins (Neil *et al*., 2008; Roesch, 2018). Therefore, characterisation of SIV infecting NHPs from *Cercopithecinae* family is therefore particularly important due to their recombination potential and the capability of crossing species barriers. It has also been reported that sequential or co-infection with different HIV-1 strains in human results in recombinant viruses known as circulating recombinant forms (CRFs), which result in enhanced viral genetic diversity (Peeters *et al*., 2013). Enhanced genetic diversity has been implicated as the main cause of antiretroviral drug resistance whose complexity is increasing (Santoro & Perno, 2013). The complexity of enhanced genetic diversity in SIV could also be increasing owing to recombination between SIV CRFs as seen in HIV (Peeters *et al*., 2013). This therefore necessitates comprehensive investigation into the genetic diversity of SIVs circulating in NHPs.

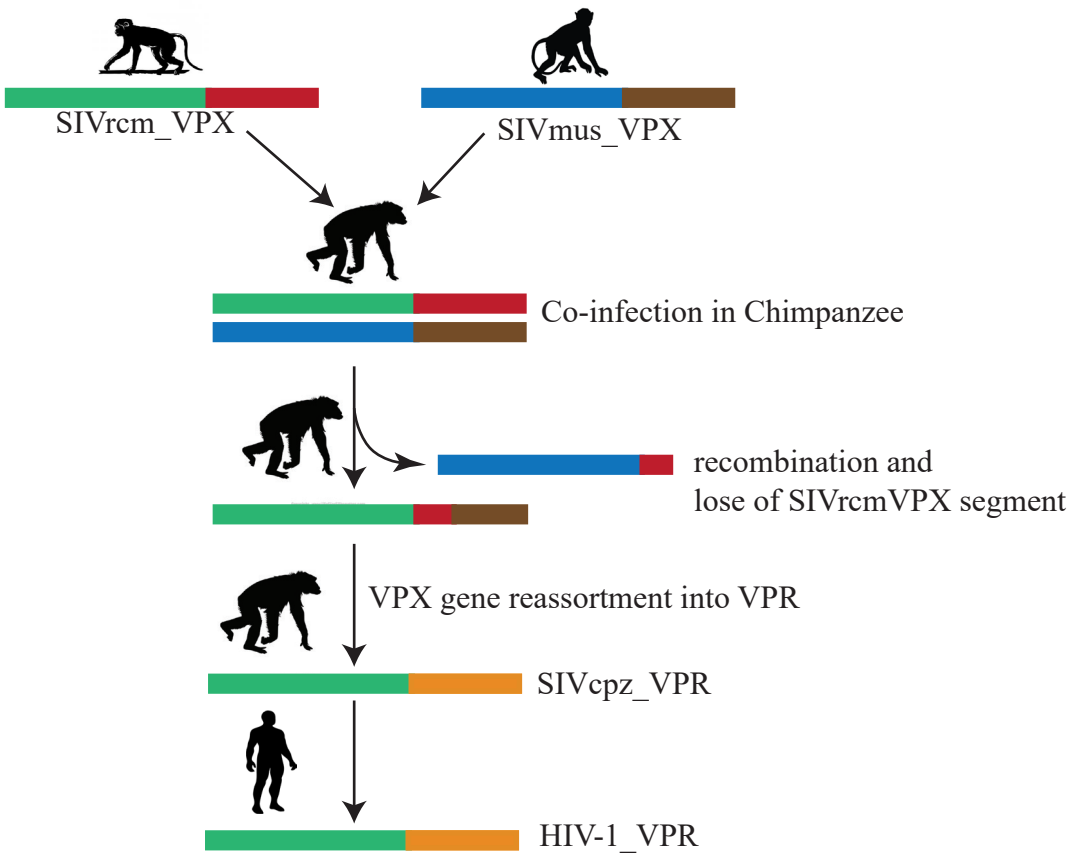


Figure 2: Proposed origin of human immunodeficiency virus (HIV). Cross-species transmission of SIVrcm from red-capped mangabeys and SIVmus from mustached monkey resulted in co-infection of the two strains in a new chimpanzee host. Both SIV from red-capped mangabeys and mustached monkey encode for their species-specific a viral infectivity protein (VPX) highlighted in red and brown respectively. Co-infection in a chimpanzee led to recombination of the two strains during replication. The VPX segment from the SIVrcm was

lost during the recombination process. Through genetic recombination, a unique viral infectivity gene (VPR) was generated in the chimpanzee host. This unique VPR gene was likely critical for viral adaptation to replicate in chimpanzees and humans by overcoming viral restriction factors such a TRIM5α. This figure was modified from Roesch, (2018).

2.4.5 Epidemiology and prevalence of SIV

Serological and molecular surveys of bush meat, wild caught primates as well as applying non-invasive sampling techniques in Africa has shown that SIV infection is widespread and prevalence may vary depending on species and geographic origin (Aghokeng *et al*., 2010). Wide distribution of NHPs that harbour a plethora of potentially zoonotic pathogens enhances the risk of cross-species transmission and super-infection with multiple virus strains with possibilities of virus recombination generating new strains that can infect human (Peeters *et al*., 2002). SIV infection has been reported to be uniformly distributed amongst feral NHPs such as AGMs, mandrills and sooty mangabeys due to high-steady state viral replication facilitating sexual transmission (Ma *et al*., 2014). SIV prevalence ranging between 46-50% has been reported, and is lower in males than females (Ma *et al*., 2013). This is probably because most male NHPs are expected to have significantly lower exposure, as only a few attain the dominance status necessary for sexual access to breeding females (Galat-Luong *et al*., 1994). Studies have shown that SIV prevalence increases with age, but is uniformly low in juveniles and infants (Ma *et al*., 2013), thus further corroborating suggestions that sexual routes as well as aggression could play important role in SIV transmission. For instance, uneven prevalence across different age groups with 28% infants and juveniles being infected compared to 67% infection in sexually mature adults has been reported (Ma *et al*., 2014). High SIV prevalence could be an important factor that drives cross-species transmission, and therefore necessitates evaluation of viral prevalence in monkeys from Kenya to estimate zoonotic risk more so in regions experiencing regular human-NHP interactions.

2.4.6 Genetic diversity

Phylogenetic analysis has revealed high SIV genetic diversity from geographically isolated NHPs of the same species such as chimpanzees and AGMs. For instance, both West African chimpanzees (*Pan troglodytes troglodytes*) and East African chimpanzees (*Pan troglodytes schweinfurthii*) are infected with viruses belonging to the same SIV lineage (Santiago *et al*., 2003) though they exhibited 30-50% sequence divergence (Haesevelde *et al*., 1996). This divergence could have occurred due to viruses infecting the two species having evolved in isolation over time. Phylogeographic diversity of up to 19% (Bibollet-Ruche *et al*.,

2004) has also been reported for SIVdeb despite these strains clustering together phylogenetically, hence suggesting existence of intra-clade divergence. Analysis of the *vpu* gene in SIVden virus infecting Dent's Mona monkey revealed a significant divergence from other virus lineages known to harbour this gene where SIVdens' *vpu* was found to be truncated at its C-terminal end (Dazza *et al*., 2005). Moreover, the *env* genes in SIVden were also shown to overlap as opposed to the genes being separate in HIV-1/SIVcpz lineage (Dazza *et al*., 2005). Mustached monkeys were also shown to be infected by multiple genetically distinct SIVmus strains, one being a recombinant between SIVmus and SIVgsn from cohabiting greater spot nosed monkeys while the second strain did not cluster with previously reported SIVmus strains (Aghokeng *et al*., 2010).

Wide genetic diversity of SIV and functional motifs variations suggests a complex evolutionary history and strong indication that co-infections are common. Super-infection of NHPs with genetically diverse virus strains is a potential avenue for recombination events that can lead to emergence of new viruses that could infect humans (Smith *et al*., 2005). Newly uncharacterised viruses are of particular interest as their zoonotic risks are unknown. Recent discovery of HIV-1 subtype P isolated from Cameroonian patients in 2009 (Peeters *et al*., 2013) and the possible new groups of HIV-2 reported in 2013 (Ayouba *et al*., 2013) clearly reveals that our knowledge on lentivirus genetic diversity and cross-species transmission is still incomplete. Hence, there is need for wider geographical sampling of NHPs and determination of prevalence and molecular characteristics of the viruses they harbour.

2.5 Non-human primates as a source of emerging infections

Close synanthropic interaction between humans and animals regarded as potential sources of zoonotic pathogens increases the risk of human retroviruses emergence (Wolfe *et al*., 2004; Peeters & Delaporte, 2012). Humans, being closely related in physiologic and genetic characteristics to NHPs, (Wolfe *et al*., 1998) can therefore be equally susceptible to many pathogens found in NHPs. Pathogens infecting NHPs such as bacteria, haemoprotozoan parasites, ectoparasites and viruses are known to be closely related to human pathogens (Wolfe *et al*., 1998). Significantly high prevalence of zoonotic pathogens such as African trypanosomes, piroplasms and even leishmania have previously been reported in wild caught NHPs including baboons and AGMs from Kenya (Jeneby *et al*., 2002; 2011). On the other hand, studies have revealed that apes are infected with a plethora of diverse *Plasmodium* species closely related to malaria-causing *Plasmodium falciparum* (Krief *et al*., 2010).

Phylogenetic analysis of these strains seemed to suggest that *Plasmodium* infecting humans could be having a simian origin (Liu *et al.*, 2010).

In general, over half of emerging pathogens in humans are viruses and a large number of them have been isolated from NHPs (Jones *et al*., 2008). Over the last decade, viral diseases associated with African NHPs as reservoirs or incidental hosts have emerged as major disease outbreaks in humans (Wolfe *et al.*, 2007). These include a series of Ebola virus outbreaks in West and Central Africa (Saéz *et al*., 2015) and a case of Marburg zoonotic infection (Nyakarahuka *et al*., 2017) in Uganda, East Africa. High prevalence of viruses such as SIV and simian T-cell lymphotropic virus (STLV) has been documented in many species of old world monkeys (Aghokeng *et al*., 2010; Liégeois *et al*., 2012) thus increasing the risk of zoonotic transmission into humans. Zoonotic transfer of these pathogens occurs through exposure to infected tissues and fluids (Peeters *et al*., 2002). This therefore makes NHPs important species for sentinel surveillance for emergence of pathogens that can potentially cross the species barrier into humans. Additionally, bats that also share habitats with free-ranging NHPs have been identified as the potential source of Ebola virus and Nipah virus (Friant *et al*., 2015; Saéz *et al*., 2015). Therefore, it is important to determine the molecular characteristics of these zoonotic pathogens in their natural host to enable prediction of cross-species transmission. Finally, failure of standard assays to detect newly emerged pathogens in humans (Peeters *et al*., 2002; Ndembi *et al*., 2009) calls for continued characterisation of zoonotic pathogens to facilitate development of specific assays capable of detecting potentially new infections emerging from zoonotic events before they can progress to the epidemic magnitudes (WHO, 2004).

Therefore, this study determined prevalence, and molecular characteristics of SIV in free-ranging NHPs that are in close proximity to large human populations in Mombasa, Naivasha and Kisumu urban centres with the aim of estimating the magnitude of zoonotic risk due to human exposure to these NHPs.

CHAPTER THREE
MATERIALS AND METHODS

3.1 Ethics statement and permits

The Institutional Scientific Review Committee (ISERC) at the Institute of Primate Research (IPR), Kenya, reviewed the experimental design and methods used in this study. Authorisation permit (ISERC/04/18) for *in situ* sampling of wildlife species was acquired from Kenya Wildlife Service (KWS). See appendix for the permit details.

3.2 Summary of experimental design

This was a cross-sectional study. NHPs were captured with care to ensure minimal bruises or injuries and sampled once before they were released back into their natural habitats. The experimental procedures carried out in this study are summarised in Figure 4.

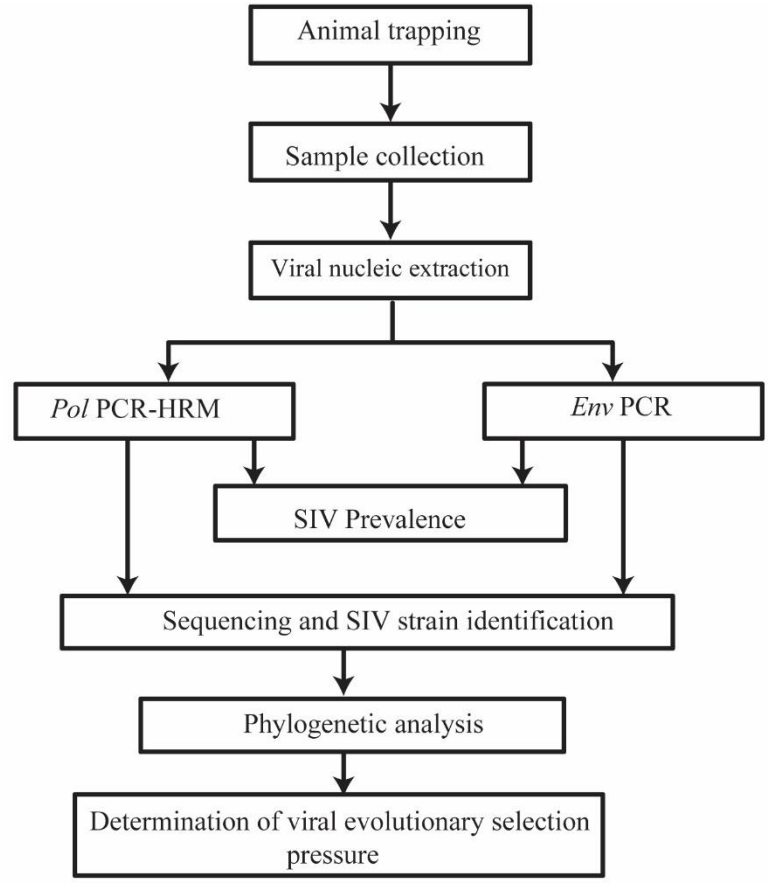


Figure 3: Summary of the experimental pipeline for the study. A workflow showing the sequence of various tasks and activities undertaken in the study

3.3 Sampling sites

The urban and peri-urban centres targeted for sampling were Mombasa, Kisumu and Naivasha in Kenya (Figure 3) where urban-dwelling NHPs can be found. The sampling sites in Mombasa Island were the Tudor water sports (TWS) facility located northwest of the Island, Mbaraki (MBK) in the south of the island, and government administration compounds (GAC) along the eastern parts of the Island. Also targeted was Mazeras (3°58′S, 39°33′E) on the mainland and within the outskirts of Mombasa town. In Kisumu (0° 6′ S, 34°45′E), a city located along the shores of Lake Victoria in Western Kenya three sites were targeted. These are the Impala park zoo within the town centre, residential golf course (GC) area near the Kisumu international airport and Kajulu (KAJ), a peri-urban centre located 5 km from the city. Kimana area, an eco-tourism destination at the shores of Lake Naivasha was the only sampling site considered in Naivasha (0°43′S, 36°26′E), located in the Great Rift Valley.

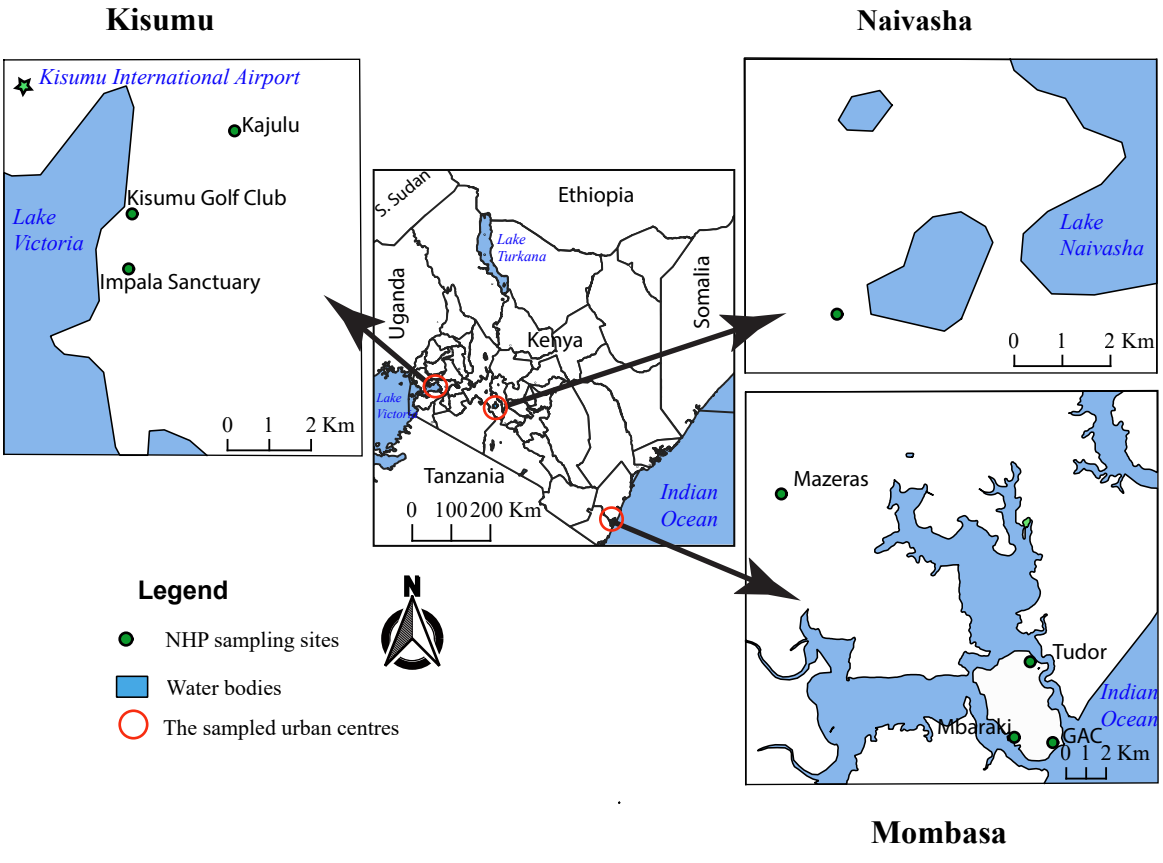


Figure 4: Urban and peri-urban sampling sites. Map of Kenya (center) showing the geographical locations of the three study urban and peri-urban centres (red circles) namely Kisumu, Mombasa and Naivasha where the non-human primates were trapped. The three sampling centres are shown offset with sampling sites represented by the green circles.

3.4 Sampling of NHPs

Free ranging NHPs were baited with fresh carrots, bananas and maize cobs and captured using wire mesh cages fitted with self-triggering slide doors as previously described (Jeneby *et al*., 2011). Individual animals were anaesthetised by administration of xylazine (2%): ketamine hydrochloride (10%) (Sigma Aldrich, Darmstadt, Germany) in the ratio of 1:3 and a dosage of 0.1ml/kg body weight. Age was determined using dental eruption (Akinyi *et al*., 2013) before blood was collected once *in situ*. Whole blood (10 ml) was drawn from the femoral vein by venipuncture using 21 gauge needle and 10 ml syringe by an authorised wildlife veterinarian from KWS. The 21 gauge needle was chosen according to the body size of the NHPs captured. A bigger sized needle (<21 gauge) was avoided to minimize injury to the animal while drawing blood, whereas a smaller needle (>21 gauge) was avoided to minimize hemolysis. An aliquot (5 ml) of each blood sample was preserved in sterile vacutainer tubes with EDTA as anticoagulant (10 ml BD vacutainer© K2E (EDTA, BD-Plymouth UK), while the other 5 ml was preserved in serum tubes (6.0 ml BD vacutainer©, BD-Plymouth UK). The blood samples were transported in dry ice to the laboratory and stored at -80° C. Two drops of blood were spotted onto Whatman FTA cards (Sigma Aldrich, Darmstadt, Germany) (Strøm *et al*., 2014), air dried and stored at room temperature. For infants, drops of blood from sterilised fingertips was used to alleviate stressful venipuncture sampling. These drops of blood were spotted onto Whatman FTA cards, air dried, and stored at room temperature. Sampled animals were monitored until they regained full consciousness before release.

3.5 DNA extraction

Total viral genomic material (viral RNA and proviral DNA) was extracted from 200 µL of each EDTA-blood sample using MagNA 96 Pure DNA and Viral NA Small Volume Kit (Roche Applied Science, Penzberg, Germany) in a MagNA Pure 96 (Roche Applied Science, Germany) automated extractor as previously described (Villinger *et al*., 2017). Briefly, 200 µL of each EDTA blood samples was transferred into MagNA pure 96 processing cartridge (Cat No. 06 241 603 001). The cartridge was then loaded into the MagNA Pure 96 automated extractor and programmed to elute the recovered genomic materials in 50 µL elution buffer.

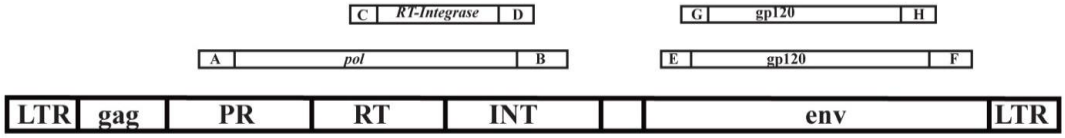
3.6 Molecular identification of SIVs

Screening for SIV was done by nested PCR-HRM analysis targeting the proviral DNA (Ayouba *et al*., 2015) to detect and differentiate primate lentiviruses naturally infecting NHPs. The primary amplification was done using NDR1-POLOR1 primer set (Figure 4) in a Veriti™ thermocycler (Life technologies holdings PTE LTD, Singapore). Each 25 µL reaction

contained 2.5 µL of 10x standard buffer (NEB Inc., Massachusetts, USA), 0.5 µL 10mM dNTPs, 0.5 µL of 10 µM forward primer and reverse primer each and 0.125 µL of 5U Taq polymerase (New England Biolabs Inc., Ipswich, Massachusetts, United States), 1 µL of template, and topped up with nuclease free PCR water. Amplification conditions included an initial denaturation at 95°C for 3 min followed by 35 cycles at 95°C for 30 secs, 45°C for 30 secs, and 72°C for 2 min and a final extension at 72°C for 10 min. Secondary PCRs were performed in a Rotor-Gene Q thermocycler (QIAGEN, Germany) using Polis4/Uni2 primers targeting a 650-nucleotide (nt) fragment *pol* gene broadly conserved in lentiviruses. The 10 µL reaction volumes contained 2 µL of 5X HRM EvaGgreen Mastermix (Solis Biotype, Riia Estonia), 0.5 µL of each 10 µM primer and 1 µL DNA template as described by (Villinger *et al*., 2017). Amplification conditions included an initial denaturation at 95°C for 15 min followed by 35 cycles at 95°C for 30 secs, 55°C for 30 secs, and 72°C for 30 secs and a final extension at 72°C for 5 min. Subsequent HRM analysis was conducted as previously described (Villinger *et al*., 2017). Briefly, PCR products were subjected to further denaturisation at 95 °C for 1 min, annealing at 40 °C for 1 min and equilibrating at 75 °C for 90 secs, and then increasing the temperature in 0.1°C increments up to 90 °C, with fluorescence acquisition after 2 secs incremental holding periods. Distinct HRM profiles, normalized in the range of 80-90°C, were visually determined for each reaction after completion of HRM data acquisition.

A

B



Primer name	Primer sequence 5'- 3'	Target gene	Product size	References
A. NDR1	TRGAYACAGGRGCWGAYGA	*pol*	2.7kb	Aghokeng *et*
B. POLOR1	ACBACYGCNCCTTCHCCTTTC			*al*., (2007)
C. Polis4	CCAGCNCACAAAGGNATAGGA	*RT-integrase*	650bp	Aghokeng *et*
D. Uni2	CCCCTATTCCTCCCCTTCTTTTA			*al*., (2007)
E. ENVA	GAAGCTTGTGATAAAACATATT	*env(gp120)*	1000bp	Ma *et al*.,
F. ENVB	AGAGCTGTGACGCGGGCATTG			(2013)
G. ENVC	GTGCATTGTACAGGGTTAATG	*env(gp120)*	900bp	Ma et al.,
H. ENVD	TTCTTCTGCTGCAGTATCCCAG			(2013)

Degenerate nucleotides: R = A or G ; Y = C or T; B = G or T; W = A or T; H = A, C, or T and N = A, C, G or T

Figure 5: SIV genome architecture and the PCR amplification strategy. A. A cartoon representing SIV genome architecture showing respective genes. The genes targeted for amplification are shown by the horizontal bars and include *pol*, *RT-integrase* and two regions of gp120. The primers flanking the target gene fragments are represented by the alphabets (A-H) at the terminal ends of each fragment. **B**. The primers used for amplification include A (NDR1) and B (POLOR1) targeting a 2.7kb of the *pol* gene fragment, primer sets C (Polis4) and D (Uni2) used in nested PCR to amplify the 650 bp fragment of integrase gene within the pol gene. The primer sets E (ENVA) and F (ENVB) targeting a 1kb fragment of the *env* gene while nested PCR using primer sets G (ENVC) and H (ENVD) amplified a 900bp partial *env* gene. (LTR - Long terminal repeat, gag - gene specific antigen, PR - protease, RT- reverse transcriptase, INT - integrase, env - envelope).

SIV strains were characterised by amplifying a 650 nt *pol* gene fragment (spanning reverse transcriptase and integrase regions) and a 900 nt envelope (*env*) gp120 fragment encompassing V3-V5 regions. The *pol* gene (Figure 4) was amplified using NDR1-POLOR1 primers for primary amplification, followed by nested PCR using the Polis4/Uni2 primers (Figure 4) in a Veriti™ thermo cycler (Life technologies holdings PTE LTD, Singapore). The reaction composition and amplification conditions for the primary reaction are as described in section 3.2. Each 25 µL secondary reaction contained 2.5 µL of 10x standard buffer (NEB Inc., Massachusetts, USA), 0.5 µL 10mM dNTPs, 0.5 µL of 10 µM forward primer and reverse primer each and 0.125 µL of 5U Taq polymerase (NEB Inc., Massachusetts, USA). The primary PCR products (1 µL) was used as the template for secondary PCR amplification. The, cycling conditions applied were denaturation at 95°C for 3 min, followed by 35 cycles at 95°C for 30 secs, annealing was done at 55°C for 30 secs, extension at 72°C for 2 min and a final extension was done at 68°C for 10 min.

For further characterisation of the *env* gp120 gene, we amplified longer fragments by nested PCR (Ma *et al.*, 2013) using the ENVA/ENVB (E-F) primers for the primary amplification and ENVC/ENVD (G-H) primers for the second nested amplification (Figure 4). Each 20 µL PCR mixture for primary amplification contained; 10 µL 2x Multiplex PCR Plus Kit Mastermix (QIAGEN, Hielden, Germany), 2 µL Q-Solution and 0.5 µL of 10 µM each forward and reverse primers, 5 µL of nuclease free PCR water, and 2 µL of template DNA. Nested amplifications also contained the same PCR components as described above but 1 µL of the primary amplicon was used as template. Amplification reactions were done in a SimpliAmp™ thermocycler (Applied Biosystems, Branchburg, NJ). The cycling conditions included an initial incubation at 95°C for 5 min, 35 cycles of extension at 95°C for 30 secs,

primer annealing at 55°C for 30 secs, extension at 72°C for 1 min and final extension at 68°C for 10 min for both primer sets. Nested PCR products were visualised under UV after electrophoresis on a 1.5% agarose gel stained with 0.5 µM/µL ethidium. Amplicons with the expected fragment size of 650 nt and 900 nt were purified using ExoSAP-IT protocol (USB® Products AffyInc., Ohio, USA) as described below.

3.7 Purification of PCR amplicons

Representative samples that generated the expected fragment size of 650bp and 900bp were purified using ExoSAP-IT protocol (USB® Products AffyInc., Ohio, USA) as described by the manufacturer. Briefly, 5 µL of post-PCR reaction products were mixed with 2 µL of ExoSAP-IT™ reagent in sterile 0.2 mL PCR tubes. The mixture was vortexed briefly and centrifuged before incubation at 37°C for 15 min. The reaction was terminated by incubation at 85°C for 15 min, and the resulting purified samples submitted for sequencing. This service was outsourced at Macrogen™ (Seoul, South Korea).

3.8 Phylogenetic and evolutionary selection analyses

3.8.1 Phylogenetic analyses

Consensus sequences generated using Geneious v10 (Kearse *et al*., 2012) were aligned using MUSCLE (Edgar, 2004). Phylogenetic reconstruction was performed by the maximum likelihood approach of PhyML (Guindon & Gascuel, 2003). The robustness of phylogenetic trees' topology was ascertained by bootstrapping using 1000 replicates. Trees were visualized using iTOL v3 (Letunic & Bork, 2016).

3.8.2 Analyses of evolutionary selection

Analysis of the evolutionary selection pressure on the SIV *env* gene was carried out using various statistical methods implemented on the adaptive evolution server (www.datamonkey.org) (Delport *et al*., 2010). These methods included single likelihood ancestor counting (SLAC), Fixed Effects Likelihood (FEL) (Pond & Frost, 2005), Fast Unconstrained Bayesian Approximation (FUBAR) (Murrell *et al*., 2013), mixed effects model of evolution (MEME) (Murrell *et al*., 2012), and Internal Fixed Effects Likelihood (IFEL) (Pond *et al*., 2006). FEL and SLAC methods combine maximum-likelihood (ML) and counting approaches to infer nonsynonymous (dN) and synonymous (dS) substitution rates on a per-site basis for a given coding alignment and corresponding phylogeny while assuming selection pressure for each site is constant along the entire phylogeny. However, FEL uses Likelihood Ratio Test, to ascertain if dN is significantly greater than dS while SLAC ascertains the

significance at each site using an extended binomial distribution approach. Under MEME, the hypothesis that individual sites are subject to episodic positive or diversifying selection is tested using a mixed-effects maximum likelihood approach. On the other hand, FUBAR infers the substitution rates per site of the provided codon alignment using the Bayesian approach with the assumption of constant selection pressure on the entire phylogeny. For SLAC, MEME, IFEL and FEL, results obtained were considered significant at 0.05 level of significance while for FUBAR, a posterior probability > 0.9 was deemed significant.

CHAPTER FOUR
RESULTS

4.1 Species of monkeys and their distribution

One hundred and twenty-four urban restricted AGMs (53 females, 71 males) and 65 (21 females, 44 males) peri-urban olive baboons of different age groups were trapped with care to minimise bruises (Table 2) and sampled under anaesthesia. Troops of AGMs were found within urban peri-urban areas while baboons were exclusively located in peri-urban sites. The ages of the captured NHPs were estimated based on external features such as body carriage, dental eruption and pelage condition by an experienced wildlife veterinarian (Akinyi *et al*., 2013). The sampled AGMs included 72 adults, 36 subadults and 16 infants while olive baboons included 48 adults, 6 subadults and 13 infants.

Table 2: The distribution and demographic of non-human primates (NHPs) sampled from the various sites of urban and peri-urban centre namely Mombasa, Naivasha and Kisumu in Kenya.

		AGMs			**Olive baboons**		
Sampling Towns	**Sampling Sites**	**Infants**	**Sub-adults**	**Adults**	**Infants**	**Sub-adults**	**Adults**
		T[M/F]	T[M/F]	T[M/F]	T[M/F]	T[M/F]	T[M/F]
Kisumu	Kajulu	-	10[4/6]	11[7/4]	-	2[2/0]	-
	Golf Club	-	1[1/0]	4[3/1]	-	-	-
	Impala Park	1[1/0]	10[7/3]	14[8/6]	-	-	-
	Subtotal	**1[1/0]**	**21[12/9]**	**29[18/11]**		**2[2/0]**	**8[6/2]**
Naivasha	Kimana	-	**4[2/2]**	**18[11/7]**	-	-	-
Mombasa	Mbaraki	6[3/3]	8[1/7]	9[5/4]	-	-	-
	Tudor	3[2/1]	2[1/1]	7[5/2]	-	-	-
	GAC	6[3/3]	1[1/0]	9[6/3]	-	-	-
	Mazeras	-	-	-	13[4/9]	4[3/1]	38[24/14]
	Subtotal	**15[8/7]**	**11[3/8]**	**25[16/9]**	**13[4/9]**	**4[3/1]**	**38[24/14]**
	Grand total	**16[9/7]**	**36[17/19]**	**72[45/27]**	**13[4/9]**	**6[5/1]**	**46[30/16]**

Key**: T**= Total, **M** = Male, **F** = Female

4.2 Prevalence of SIV in AGMs and baboons

SIV infection status was determined using a combination of PCR-HRM of the partial *pol* gene and nested PCR amplification of partial *env* gene. The fragment sizes of amplified PCR products were confirmed by agarose gel electrophoresis (Figure 6) and sequencing.

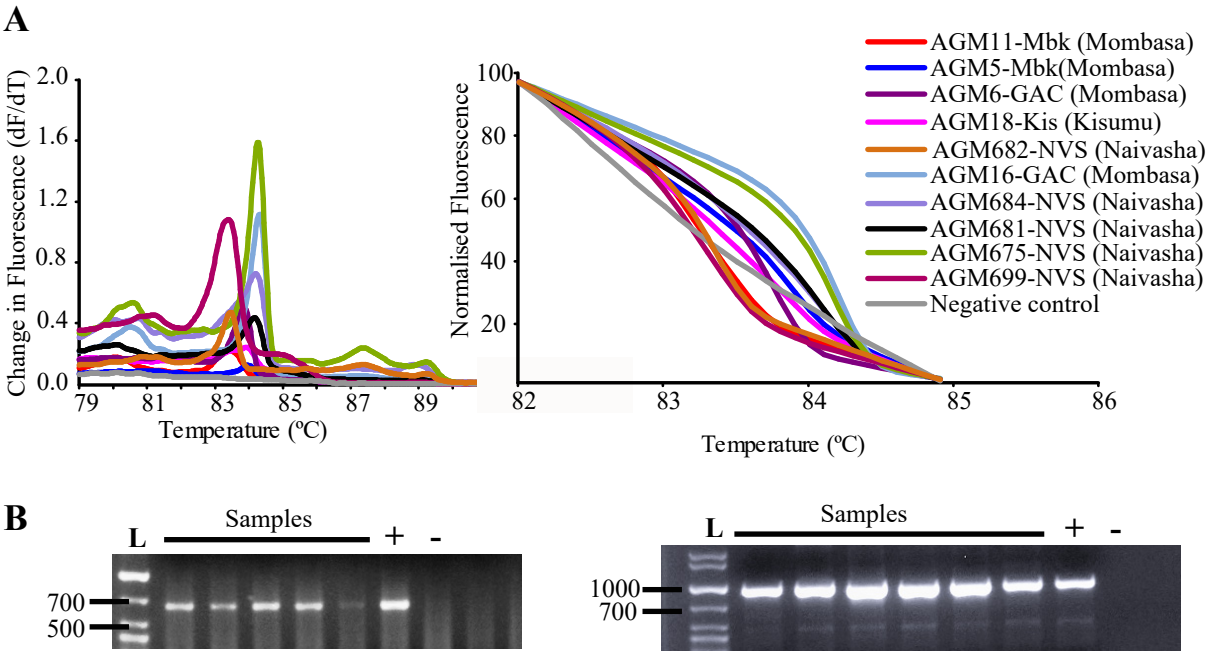


Figure 6: Detection of SIV in NHPs by a combination of PCR-HRM and gel electrophoresis. A. Representative melt peaks (left panel) and HRM melt profiles (right panel) generated by PCR-HRM of the partial *pol* gene (650 bp) from SIV positive samples. **B.** The amplified and electrophoresis size separated SIV partial *pol* gene (left panel) and *env* gene (right panel) fragments of 650bp and 900bp respectively. **L** represents the GelPilot 100 bp Plus molecular weight ladder (QIAGEN, Hielden, Germany) ; + and - represent SIV positive and negative controls respectively.

From these amplifications, SIV prevalence was 32% (39/124) in AGMs and 3% (2/65) in baboons (Table 3). Infection rates in male AGMs was 34% (24/71) and females 28% (15/53). Two male baboons, one from Kajulu in Kisumu and the other from Mazeras, in Mombasa were positive for SIV. Prevalence of SIV among age groups of NHPs was higher in adults (27 out of 72, 37%), followed by sub adults (9 out 36, 25%) and infants (3 out of 16, 19 %) (Table 3). The distribution of SIV infections in different sampling locations within urban and peri-urban location was uneven and ranged from 8% to 48% in AGMs (Table 3). Infections by sampling sites was highest in Mbaraki, Mombasa (10/22, 48%) followed by Kimana in Naivasha (17/51, 33%) and Impala park in Kisumu (7/25, 28%). The AGMs sampled in Tudor, Mombasa has the lowest SIV prevalence of 8% (1/12) (Table S1. Appendix).

Table 3: Prevalence of SIV infection in NHPs (AGMs and olive baboons) sampled from Kisumu, Naivasha and Mombasa towns in Kenya. Prevalence of SIV infection in NHPs (AGMs and olive baboons) sampled from Kisumu, Naivasha and Mombasa towns in Kenya.

	SIV positive /Total sampled (% Infected)					
	AGMs			**Olive Baboons**		
	Male	**Female**	**Total**	**Male**	**Female**	**Total**
Kisumu	10/31(32)	2/20(10)	12/51 (24)	1/8 (13)	0/2 (0)	1/10 (10)
Mombasa	8/27(30)	9/24 (38)	17/51 (33)	1/36 (3)	0/19 (0)	1/55 (1)
Naivasha	6/13(46)	4/9 (44)	10/22 (45)	-	-	-
Total	**24/71(34)**	**15/53 (28)**	**39/124 (31)**	**2/44 (5)**	**0/21 (0)**	**2/65 (3)**

4.3 High resolution melting profiles

HRM analysis of the PCR amplified SIV *pol* gene demonstrated distinct melt profiles (Figure 6A) illustrating either diversity of detected viruses or unknown product of the predicted sizes. These results most likely illustrate virus diversity but must be supported by empirical evidence from amplicon sequencing. The normalization regions used for the analysis of the amplified *pol* gene ranged from 82.0°C to 82.14°C in the leading range and 85.86°C to 86.0°C in the trailing range. Positive virus amplification from specimens MG590127-AGM675 and MG590131-AGM699from NHPs sampled in Naivasha resulted in two distinct melt profiles (Figure 6A) suggesting existence of virus divergence. In contrast, SIV isolates from MG590129-AGM682 and MG590127-AGM675 from Naivasha generated similar melt peaks but distinct melt profiles as well further illustrating existence of diversity. Amplifications of DNA specimens from MG590133-AGMKis18 from Kisumu, MG757169 AGMGAC6 from Mombasa, and MG590135 AGMMbk11 from Mombasa had distinct melt profiles from each other. The plausible divergence was confirmed by sequencing the amplicons and querying them against GenBank database (Clark *et al.,* 2016). This generated hits of 82% - 84% identity to the already characterised SIVagm strains (Figure 7). Generally, there was nucleotide sequence plasticity in this gene fragment upon sequence alignment also demonstrated by the percent identities of the gene products i.e. protein sequences (Table S2).

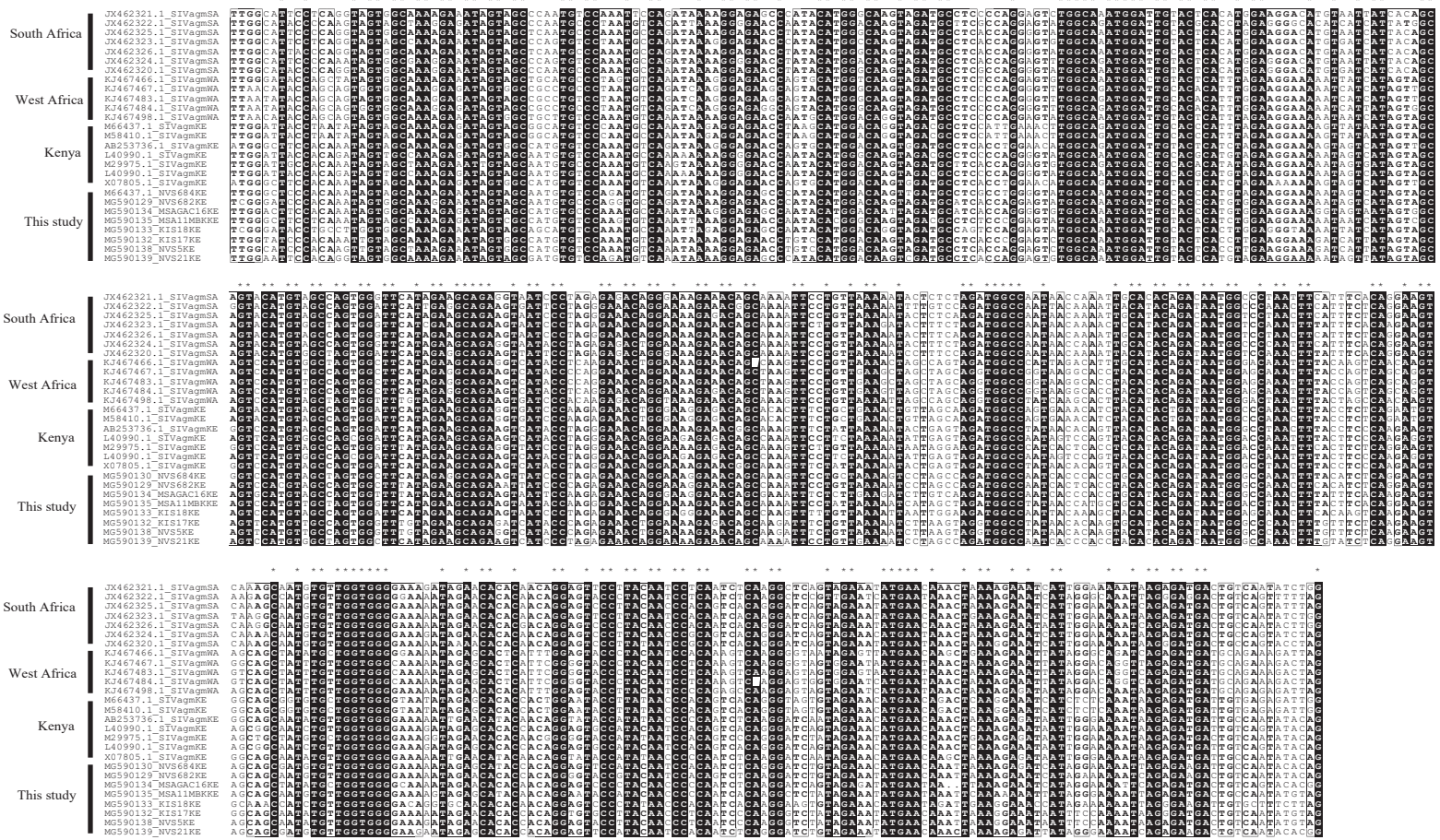


Figure 7. Multiple sequence alignment for the 650 bp partial *pol* gene amplified by PCR-HRM showing the distribution of conserved nucleotides. Amplified and sequenced gene fragments of SIV *pol* gene from AGMs were aligned, showing a range of between 75 to 83 % identity to our reference SIVagm sequences previously characterised from Kenya. The conserved nucleotides are highlighted with a black background and an asterisk (*) at the top of the alignment. Characterised SIVagm sequences from South Africa, West Africa and previously characterised sequences from Kenya were included in the alignment. The sequence identities including their accession numbers and the sampling region are shown on the left side of the alignment. Sequences were aligned using MUSCLE and the alignments visualised and exported for printing on ESPript (Robert & Gouet, 2014) and modified on Adobe Illustrator.

4.4 SIV genetic diversity

4.4.1 Identification of SIV sequence lineages

A stringent basic local alignment search (BLAST) (Altschul *et al.*, 1990) at an e- value of 0.0001 was used at NCBI's GenBank database (www.ncbi.nlm.nih.gov) to identify the strains of SIV sequences generated in this study. BLAST search using partial *pol* nucleotide sequences revealed '81-93%' sequence identity to the published SIVagm *pol* gene from AGMs isolated in Africa while the partial *env* nucleotide sequences showed identity of '79 - 85%' to "simian immunodeficiency virus isolate 9649env05" (GenBank Accession U93806) also isolated from an AGM. Further, these identities were confirmed by performing sequence alignments alongside reference sequences from GenBank.

Generally, taxonomic classification of SIV into strains/lineage correlate with their respective primate host species (Foley *et al.*, 2016). In addition, significant diversity has been documented in SIVs from the same NHP host from different geographic locations (Ma *et al*., 2013; 2014). Therefore, this diversity pose a limitation to the criterion of grouping SIV according to the host species. This is further complicated when cross-species transmission or co-infection with multiple strains occur. To interrogate the lineage identification further, we first determined the SIV sequence identity ranges for viruses from GenBank isolated from various NHPs species in different geographic locations namely West, Central, Eastern and Southern Africa.

These sequences included SIV isolated from 18 NHP species including chimpanzees (*Pan troglodytes*), sooty mangabeys (*Cercocebus atys*), mandrills (*Mandrillus sphinx*), colobus monkey (*Colobus guereza*), monkeys (*Cercopithecus albogularis*), De Brazza's monkeys (*Cercopithecus neglectus*), greater spot nosed monkeys (*Cercopithecus nictitans*), mona monkeys (*Cercopithecus mona*), moustached monkey (*Cercopithecus cephus*) and Africa green monkeys (*Chlorocebus aethiops*). The separate multiple sequence alignments for partial *pol* and *env* gene fragments were used to generate percent identity matrices for each species from different regions. Resultant matrices were used to derive SIV identities range per strain per region (Figure 8). The SIVs derived from AGMs (SIVagm) are widely distributed in Africa. Their sequence identities ranged from 81-98% in West Africa, 81% - 100% in Eastern Africa and 82-90% in Southern Africa. Only one SIVagm sequence was available from monkeys in Central Africa region, explaining the 100% identity observed. The comparison of SIVagm sequences from different regions in Africa showed the identities ranging from 71% - 83%.

These results supported the observation of nucleotide conservation plasticity upon alignment of SIV sequences amplified in NHPs from Kenya in this study (Figure 7). As illustrated in Figure 8, colobus monkeys had the lowest range of SIV pol identities (52% - 60%). In general, the identities of SIV sequences within an NHP species ranged between 70% and 85%. Amongst the 18 NHP species from which the SIV sequences were analysed, only the AGMs had representatives from all the four regions.

A

AGMS	West	Central	East	South
West	81-98	81-83	71-76	71-75
Central	81-83	100	73-76	72-73
East	71-76	73.76	81-100	77-80
South	71-75	72-73	77-80	82-90

Gorilla	West	Central	East	South
West	-	-	-	-
Central	-	81-93	-	-
East	-	-	-	-
South	-	-	-	-

Piliocolobus	West	Central	East	South
West	-	-	-	-
Central	-	80-99	-	-
East	-	-	-	-
South	-	-	-	-

Spot nosed	West	Central	East	South
West	82-100	-	-	-
Central	-	-	-	-
East	-	-	-	-
South				

Talapoin	West	Central	East	South
West	-	-	-	-
Central	-	76-99	-	-
East	-	-	-	-

Chimpanzee	West	Central	East	South
West	-	-	-	-
Central	-	78-86	75-78	-
East	-	75-78	90	-
South	-	-	-	-

Lhoest	West	Central	East	South
West	-	-	-	-
Central	-	80-95	-	-
East	-	-	-	-
South	-	-	-	-

Red capped	West	Central	East	South
West	-	-	-	-
Central	-	84-100	-	-
East	-	-	-	-
South	-	-	-	-

Sykes	West	Central	East	South
West	-	-	-	-
Central	-	-	-	-
East	-	-	75-84	-
South				

Drill	West	Central	East	South
West	-	-	-	-
Central	-	85	-	-
East	-	-	-	-

Colobus	West	Central	East	South
West	-	-	-	-
Central	-	80-100	52-60	-
East	-	52-60	76-85	-
South	-	-	-	-

Mandrill	West	Central	East	South
West	-	-	-	-
Central	-	95-100	-	-
East	-	-	-	-
South	-	-	-	-

Red Tailed	West	Central	East	South
West	-	-	-	-
Central	-	74-99	71-77	-
East	-	71-77	96-97	-
South	-	-	-	-

Mustached monkey	West	Central	East	South
West	-	-	-	-
Central	-	71-99	-	-
East	-	-	-	-
South				

Suntailed	West	Central	East	South
West				
Central				
East	-	-	-	-

Debrazza	West	Central	East	South
West	-	-	-	-
Central	-	71-100	-	-
East	-	-	-	-
South	-	-	-	-

Mona	West	Central	East	South
West	-	-	-	-
Central	-	67-100	-	-
East	-	-	-	-
South	-	-	-	-

Sooty mangabey	West	Central	East	South
West	-	-	-	-
Central	-	84-97	-	-
East	-	-	-	-
South	-	-	-	-

B

AGMS	West	Central	East	South
West	71-96	67-70	68-72	68-73
Central	67-70	100	70-71	69-72
East	68-72	70-71	67-99	70-81
South	68-73	69-72	70-81	79-97

Gorilla	West	Central	East	South
West	-	-	-	-
Central	-	73-82	-	-
East	-	-	-	-
South	-	-	-	-

Piliocolobus	West	Central	East	South
West	-	-	-	-
Central	-	67-97	-	-
East	-	-	-	-
South	-	-	-	-

Spot nosed	West	Central	East	South
West	87	-	-	-
Central	-	-	-	-
East	-	-	-	-
South	-	-	-	-

Talapoin	West	Central	East	South
West	-	-	-	-
Central	-	83	-	-
East	-	-	-	-
South	-	-	-	-

Chimpanzee	West	Central	East	South
West	-	-	-	-
Central	-	67-99	55-70	-
East	-	55-70	57-68	-
South	-	-	-	-

Lhoest	West	Central	East	South
West	-	-	-	-
Central	-	81-95	-	-
East	-	-	-	-
South	-	-	-	-

Red capped	West	Central	East	South
West	-	-	-	-
Central	-	84-100	-	-
East	-	-	-	-
South	-	-	-	-

Sykes	West	Central	East	South
West	-	-	-	-
Central	-	-	-	-
East	-	-	73-76	-
South	-	-	-	-

Drill	West	Central	East	South
West	-	-	-	-
Central	-	100	-	-
East	-	-	-	-
South	-	-	-	-

Colobus	West	Central	East	South
West	-	-	-	-
Central	-	81-87	54-56	-
East	-	54-56	75	-
South	-	-	-	-

Mandrill	West	Central	East	South
West	-	-	-	-
Central	-	90-100	-	-
East	-	-	-	-
South	-	-	-	-

Red Tailed	West	Central	East	South
West	-	-	-	-
Central	-	74-99	63-78	-
East	-	63-78	94-95	-
South	-	-	-	-

Mustached monkey	West	Central	East	South
West	-	-	-	-
Central	-	68-98	-	-
East	-	-	-	-
South	-	-	-	-

Suntailed	West	Central	East	South
West	-	-	-	-
Central	-	100	-	-
East	-	-	-	-
South	-	-	-	-

Debrazza	West	Central	East	South
West	-	-	-	-
Central	-	72-83	-	-
East	-	-	-	-
South	-	-	-	-

Mona	West	Central	East	South
West	-	-	-	-
Central	-	57-72	-	-
East	-	-	-	-
South	-	-	-	-

Sooty mangabey	West	Central	East	South
West	-	-	-	-
Central	-	79-99	-	-
East	-	-	-	-
South	-	-	-	-

Figure 8: The SIV percentage sequence identity matrices generated from sequence comparison. A. Range of percentage identity of SIV partial *pol* gene and **B.** *env* gene fragments from different NHP species from different geographical regions (West, Central, East and Southern) in Africa. The SIV sequences were grouped according to the NHP host species and the geographical region of isolation. The sequence alignments were first compared within the NHP species within the same region and then across different regions. The percent identities

are recorded in the matrices as a range from the lowest to the highest identity observed for every sequence alignment done.

The results for the partial *env* gene were comparable to the partial *pol* gene. The identities for AGMs ranged from 71% -79% in West African region, 67% - 99% in Eastern Africa and 79% -97% in Southern Africa. Likewise, only one sample was available from Central Africa thus accounting for 100% identity observed in this region. For Central African suntailed monkeys and drill monkeys, single *env* gene sequence was available for each NHP species, representing the observed 100% identity. There was a significant difference in the identity ranges for the mona monkey (52% - 72% in the *env* gene and 67%-100% in the *pol* gene) from Central Africa suggesting a higher diversity.

Once the identities range was determined (Figure 8), representative SIV sequences from different NHP species were aligned with the SIV sequences recovered from AGMs in Kenya and used to generate percent identity matrices for partial *pol* and *env* gene fragments. From the percentage identities matrices, separate heat maps were generated for the partial gene fragments (Figures 9 and 10). Generated heat maps enabled placement of SIV sequences from this study into SIVagm lineage as shown in heat maps. The highest identities are highlighted in green thus, the sequences from this study were very close to AGM derived viruses than the mandrills and chimpanzees whose identities to AGM derived viruses was low (41 – 42%) as shown by the red highlights of the heat map.

From the study, the identities for SIVs to other characterised AGMs' isolates ranged from 75% to 83% for the partial *pol* gene and 62% to 80% for the partial *env* gene. Comparing SIV isolated from AGMs and Sykes in Kenya, whose geographic source is the same, the identities ranged from 63% to 70%. Moreover, in this study, the SIV sequenced from an archived sample of a Sykes monkey from IPR (MG590108.1_SIVsyk) had 64% to 65% identity to viruses from AGMs while the identity was 77% to 86% when compared with viruses from Sykes monkeys in GenBank (Figure 10). Comparing sequences derived from AGMs with other NHP species, the identities ranged from 63% - 72% against Chimpanzees, 63% - 68% against gorillas, 66% - 70% against greater spot-nosed monkey, 58% - 63% against colobus monkey, 65% - 70% against mona monkey, 65% - 71% against sooty mangabeys and 68% - 73% against the red-capped mangabeys. The SIV sequence identities remained marginally low amongst the different NHP species with overlapping geographic ranges. For instance, between the red-capped mangabeys and the mustached monkeys from Central Africa, the identities ranged

between 66% - 77%, colobus monkeys and the debrazza's monkeys had an identity range of 60% - 66% while the chimpanzee and colobus monkeys had an identity range of 55% - 62%. SIV sequences from evolutionarily related chimpanzee and gorilla had a range of 76% - 80%. The plausible explanation for this high identity could be that SIVs infecting gorillas originated from chimpanzees (Takehisa *et al.*, 2009). Collectively, these findings reveal highest identity of the sequences from this study to AGM derived SIVs. This observation indicates that the SIV amplified from AGMs qualify for classification within the SIVagm lineage. However, the interpretation of these identities needs careful consideration in cases of cross-species transmission and/or recombination. For instance, SIVs from red-capped mangabeys and chimpanzee shared 76% - 81% nucleotide identities for the partial *pol* gene. Notably, the *env* fragments for red-capped mangabeys were unavailable for comparison. These *pol* identities could probably be emanating from the possibility of SIV in chimpanzee being a recombinant involving the viruses from red-capped mangabeys (Bailes *et al*., 2003). This therefore necessitated phylogenetic analysis to determine the evolutionary relationship of SIVs within the AGM lineage.

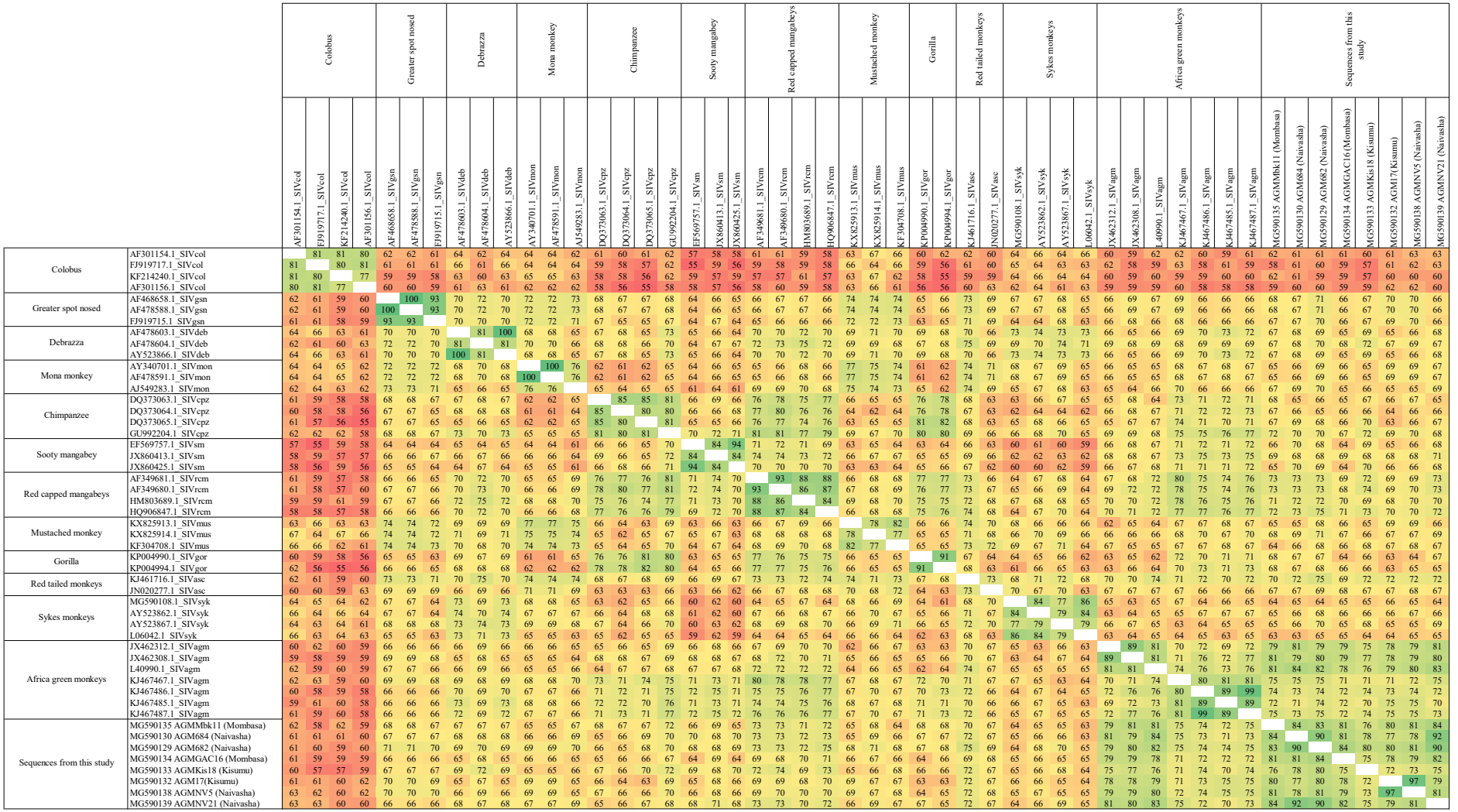


Figure 9. Percentage identity matrix of SIV partial *pol* sequence alignments for SIV lineage identification. Sequences of *pol* gene from SIV identified in the study were aligned against SIV sequences retrieved from GenBank database originating from different non-human primate species from Africa. The alignment was done using MAFFT run in Geneious software v10 (Kearse *et al*., 2012). The percent identities are shown in the table with the heat map indicating the identities of SIV sequences in a descending order with green representing most identical sequences and red showing the least identical sequences.

		Chimpanzee		Northern Talapoins		Greater spotnosed monkey		Mustached monkey		Mona monkey		Sykes monkey	Debrazza	Suntailed monkey		Mandrill		Africa green monkey (Vervet)		SIV sequences from this study						
		SIVcpz_FR686523.1	SIVcpz_FR686527.1	SIVtal_AM182197.1	SIVtal_AY655744.1	SIVgsn_AF468658.1	SIVgsn_AF468659.1	SIVmus_EF070330.1	SIVmus_EF070329.1	SIVmon_AJ549283.1	SIVmon_AY340701.1	SIVsyk_AY523867.1	SIVdeb_FJ919724.2	SIVsun_AF131870.1	SIVsun_FR751162.1	SIVmnd_AM158193.1	SIVmnd_AM396979.1	SIVagmVER_JX462426.1	SIVagmVER_JX462444.1	MG590115 AGM681 (Naivasha)	MG590118 AGM685 (Naivasha)	MG590132 AGM17(Kisumu)	MG590111 AGM5 (Kisumu)	MG590126 Baboon5 (Kisumu)	MG590123 AGM7GAC (Mombasa)	MG590124 AGM7Mbk (Mombasa)
Chimpanzee	SIVcpz_FR686523.1		99	39	40	46	47	46	44	46	46	42	37	35	34	28	27	41	41	38	38	37	37	37	38	36
	SIVcpz_FR686527.1	99		39	40	46	47	46	44	46	46	42	37	35	34	29	28	41	41	39	38	37	37	37	39	36
Northern Talapoins	SIVtal_AM182197.1	39	39		73	47	47	46	48	48	45	44	39	35	34	29	28	47	48	41	42	43	43	41	43	42
	SIVtal_AY655744.1	40	40	73		49	48	46	46	47	46	41	40	34	33	27	27	44	46	42	42	42	40	40	41	41
Greater spotnosed monkey	SIVgsn_AF468658.1	46	46	47	49		82	68	66	61	61	48	42	34	34	26	25	45	44	38	38	40	40	38	39	37
	SIVgsn_AF468659.1	47	47	47	48	82		65	65	57	58	48	43	33	34	28	27	42	43	38	38	40	40	38	38	39
Mustached monkey	SIVmus_EF070330.1	46	46	46	46	68	65		75	60	60	48	40	36	35	26	25	42	42	36	37	37	36	34	37	37
	SIVmus_EF070329.1	44	44	48	46	66	65	75		58	58	48	39	35	35	25	23	40	42	36	36	37	37	36	38	37
Mona monkey	SIVmon_AJ549283.1	46	46	48	47	61	57	60	58		67	46	40	34	34	24	23	43	43	39	39	41	41	39	39	40
	SIVmon_AY340701.1	46	46	45	46	61	58	60	58	67		46	38	33	32	25	25	41	43	38	38	38	37	37	35	37
Sykes monkey	SIVsyk_AY523867.1	42	42	44	41	48	48	48	48	46	46		37	34	33	34	32	38	41	37	37	34	34	33	36	37
Debrazza	SIVdeb_FJ919724.2	37	37	39	40	42	43	40	39	40	38	37		38	37	31	30	36	37	37	37	38	36	37	36	37
Suntailed monkey	SIVsun_AF131870.1	35	35	35	34	34	33	36	35	34	33	34	38		80	34	31	34	34	32	32	32	31	30	33	29
	SIVsun_FR751162.1	34	34	34	33	34	34	35	35	34	32	33	37	80		33	32	35	34	33	32	33	32	32	33	30
Mandrill	SIVmnd_AM158193.1	28	29	29	27	26	28	26	25	24	25	34	31	34	33		94	24	25	26	25	26	23	21	25	23
	SIVmnd_AM396979.1	27	28	28	27	25	27	25	23	23	25	32	30	31	32	94		24	25	25	25	26	22	22	26	23
Africa green monkey (Vervet)	SIVagmVER_JX462426.1	41	41	47	44	45	42	42	40	43	41	38	36	34	35	24	24		80	67	68	63	63	62	65	63
	SIVagmVER_JX462444.1	41	41	48	46	44	43	42	42	43	43	41	37	34	34	25	25	80		67	68	64	64	65	64	66
SIV sequences from this study	MG590115 AGM681 (Naivasha)	38	39	41	42	38	38	36	36	39	38	37	37	32	33	26	25	67	67		99	77	74	75	78	78
	MG590118 AGM685 (Naivasha)	38	38	42	42	38	38	37	36	39	38	37	37	32	32	25	25	68	68	99		77	75	75	78	78
	MG590132 AGM17(Kisumu)	37	37	43	42	40	40	37	37	41	38	34	38	32	33	26	26	63	64	77	77		81	82	76	76
	MG590111 AGM5 (Kisumu)	37	37	43	40	40	40	36	37	41	37	34	36	31	32	23	22	63	64	74	75	81		97	76	76
	MG590126 Baboon5 (Kisumu)	37	37	41	40	38	38	34	36	39	37	33	37	30	32	21	22	62	65	75	75	82	97		75	76
	MG590123 AGM7GAC (Mombasa)	38	39	43	41	39	38	37	38	39	35	36	36	33	33	25	26	65	64	78	78	76	76	75		75
	MG590124 AGM7Mbk (Mombasa)	36	36	42	41	37	39	37	37	40	37	37	37	29	30	23	23	63	66	78	78	76	76	76	75	

Figure 10. Percentage identity matrix of SIV partial *env* sequence alignments for SIV lineage identification. Sequences of *env* gene from SIV identified in the study were aligned against SIV sequences retrieved from GenBank database originating from different non-human primate species from Africa. The alignment was done using MAFFT run in Geneious software v10 (Kearse *et al*., 2012). The percent identities are shown in the table with the heat map indicating the identities of SIV sequences in a descending order with green representing most identical sequences and red showing the least identical sequences.

4.4.2 Phylogenetic analyses of SIVagm infecting AGMs and baboons in Kenya

All samples determined to be SIV positive by PCR-HRM were amplified using primers targeting *pol* (19 samples) or *env* (22 samples) genes and sequenced. Out of these, 14 samples generated unambiguous sequences for *pol* and 17 for *env* gene fragments respectively. These sequences have been deposited in GenBank under accession numbers MG590108 - MG590139, MG757169-MG757171(Table S5). Phylogenetic analysis revealed that the partial *pol* sequences obtained in this study clustered in a monophyletic clade with those from SIVagm isolates previously obtained from Kenya (Figure 11). No clear segregation of the

Kenyan isolates of SIVagm according to the respective sampling towns was detected. Within the Kenyan clades, sub clusters had robust bootstraps support values greater than 90%. One isolate from an AGM, MG590133-AGMKis18, sampled from Kisumu is closely related to SIVagmTAN-1 (Accession number U58991) isolated from an AGM in Eastern Uganda (Figure S1).

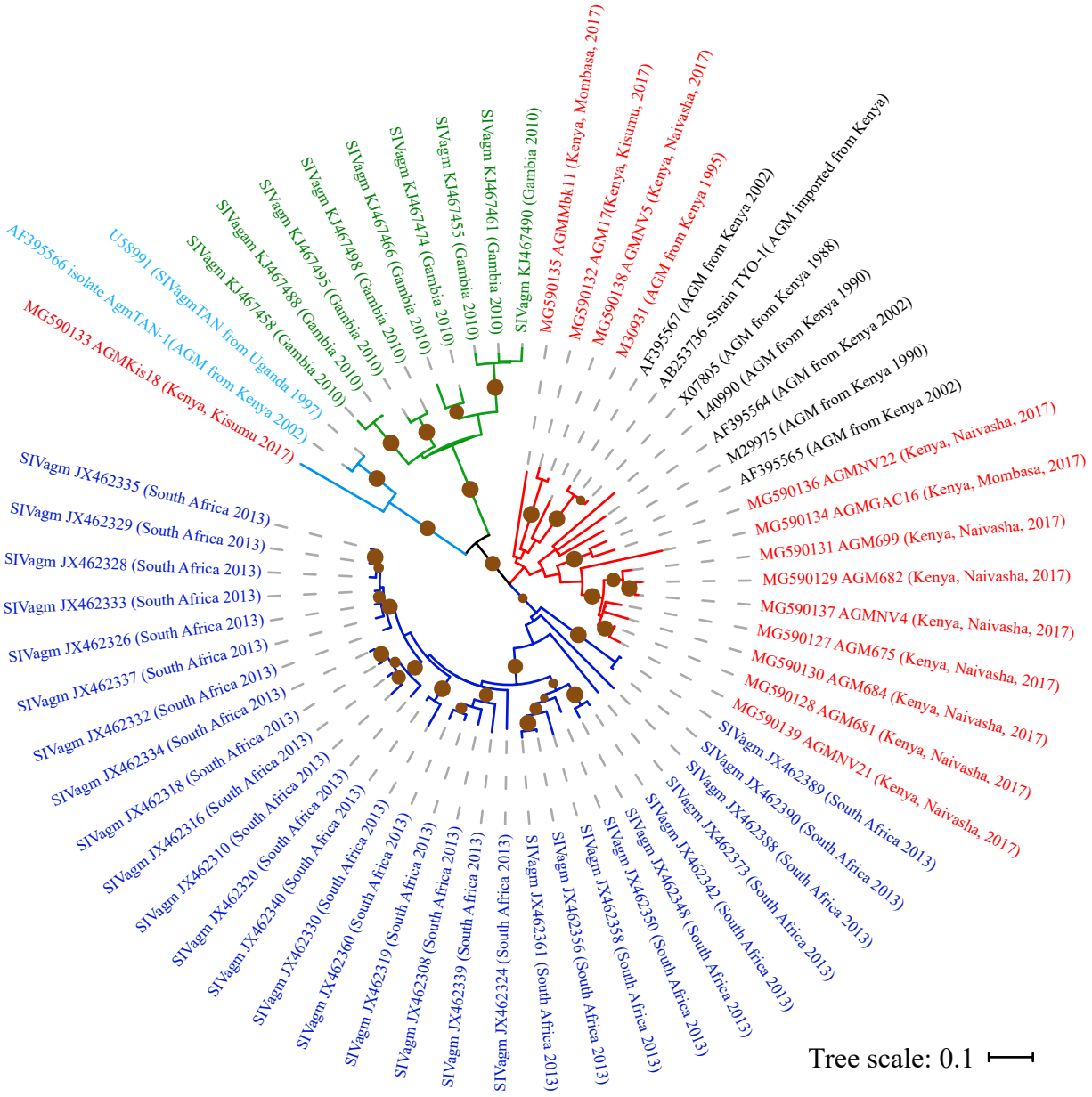


Figure 11. Maximum likelihood tree of the partial *pol* gene (650bp) illustrating genetic diversity of SIV from synanthropic (urban dwelling) AGMs from three urban centres in Kenya. Sequences from Kenya, South Africa and West Africa are highlighted in red, blue and green respectively. Sequences obtained from GenBank with the host species originating from Kenya are shown in black. The sequence obtained from tantalus subspecies of AGMs from Uganda is highlighted in lightblue. Solid brown circles represent bootstrap support values of >70 while the bar represents the number of nucleotide substitutions (0.1) per site. Naming of sequences generated in this study is according to the monkey species (AGM), the animal number and sampling site.

Phylogenetic results obtained for the *env* and *pol* genes show similarity in the geographical clustering of the sequences according to the country of origin (Figure 12). No clear segregation of SIV variants according to the geographical locations of host sampling points was evident within Kenya. SIV sequence isolated from an olive baboon sampled in Kisumu clustered with sequences from sympatric AGMs with high bootstrap support. Based on the branching pattern of the phylogenetic trees, SIV from animals sampled in Mombasa and Kisumu appeared to be ancestral to those infecting animals sampled in Naivasha. Epidemiologically linked SIVs (AGM682 and AGM-NV4, AGM - NV681 and AGM - NV2, AGM - GC5 and baboon - Kaj5) were identified in Naivasha and Kisumu (Figure 12).

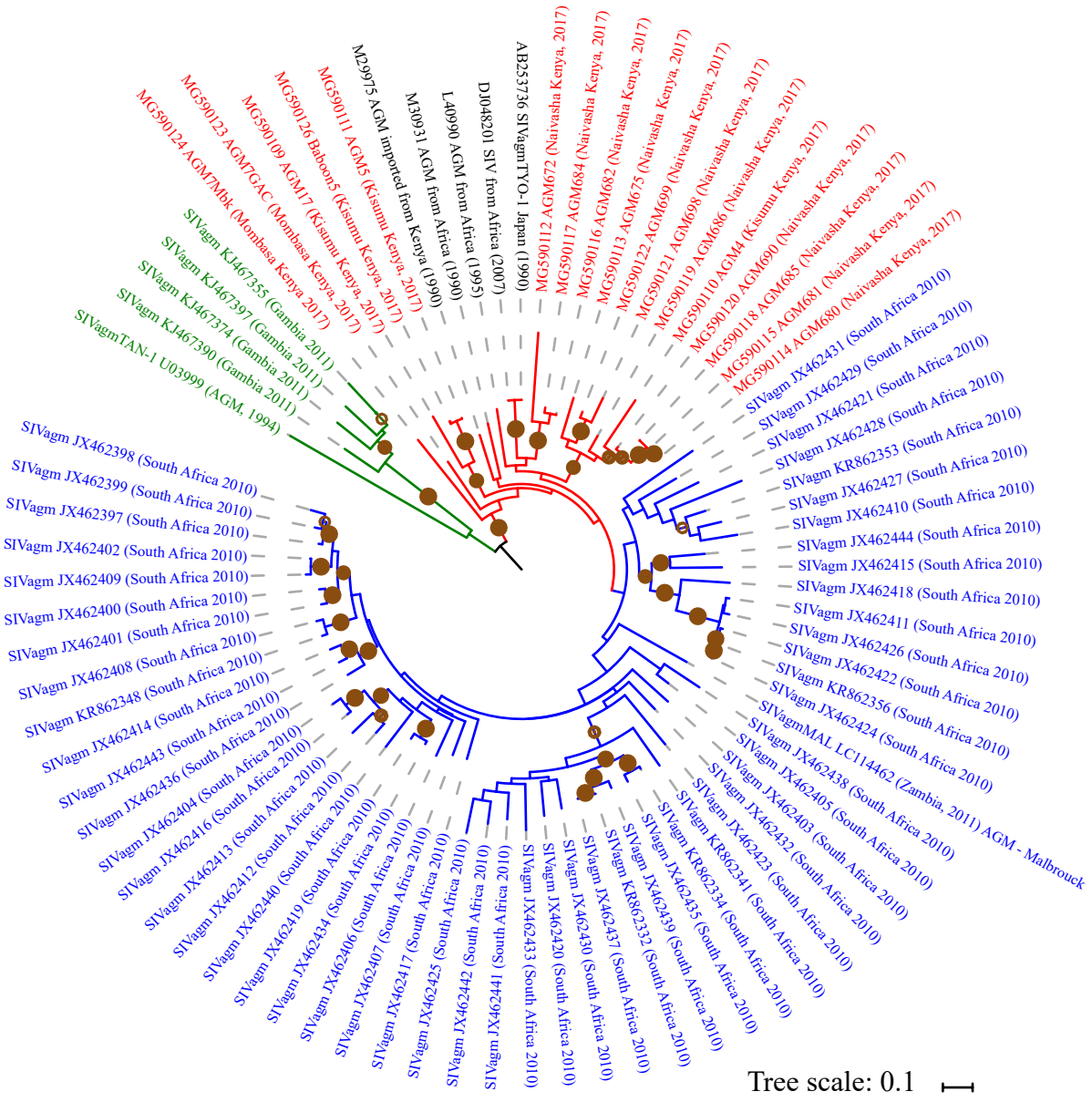


Figure 12: Maximum likelihood tree of the partial *env* gene (900 bp) illustrating genetic diversity of SIV infecting synanthropic AGMs from several urban centres in Kenya. Sequences from Kenya, South Africa and West Africa are highlighted in red, blue and green respectively. Sequences obtained from GenBank with the host species originating from Kenya are shown in black. Solid brown circles represent bootstrap support values of >70 while the bar represents the number of nucleotide substitutions (0.1) per site. Naming of sequences generated in this study is according to the monkey species (AGM), the animal number and sampling site.

Within the phylogenetic cluster of SIVs isolated from Kenya, there was no clear segregation according to the sampling towns amongst the *pol* and *env* sequences from this study (Figure 13). This topology indicates existence of SIVagm sequence divergence despite falling within the same Kenyan cluster. SIV sequences from the same urban centre for instance

Naivasha (Figure 13) clustered in phylogenetic branches with low bootstrap support values (<70), demonstrating that AGMs sharing a habitat are infected with highly divergent SIV strains. This observation is well supported by the results from HRM analysis which showed unique melt profiles and thus limitation in identifying the SIV strains by HRM alone. On the other hand, sequences with high bootstrap support values (100) illustrates existence of a common evolutionary history and possibilities of sharing a common ancestor in the recent past.

The phylogenetic topology exhibited by nucleotide sequences was also observed for amino acids where there was no clustering according to the location of the hosts in Kenya. Whereas, the *env* protein sequences from Kenya interspersed with sequences from Africa (Figure S2), the *pol* protein maintained the tree topology (Figure S3) observed in the nucleotide sequences (Figure 11). The diversity results from the protein sequences are in agreement with HRM and phylogenetic results where samples from Naivasha, (MG590129-AGM682 and MG590127-AGM675) with similar melt profiles had 99% identity. SIVs infecting MG590135-AGMMbk11 from Mombasa and MG590131-AGM699 from Naivasha had 90% identity, a trend illustrated by their melt profile differences. Whereas, SIV variants from MG590131-AGM699and MG590127-AGM675 having different melt profiles, had 97% in their pol amino acids sequences. On the other hand, the amino acids conservation was lower in the env protein compared to the *pol* protein as shown from the amino acid alignments (Figure S3). The sequences from Naivasha had 77% - 98% amino acid identity in the *env* protein. These sequences from Naivasha exhibited 77% - 89% identity to sequences from Mombasa and 80% - 91% identity to sequences from Kisumu. The env amino acid sequence from a from an olive baboon (MG590126 Baboon5) sampled in Kisumu had 96% identity with the amino acid sequence from an AGM (MG590111 AGM5) also from Kisumu despite maintaining 80% - 89% identity to the rest of the sequences. This further supports epidemiological linkage of these two isolates as implied by phylogenetic analysis (Figure 12).

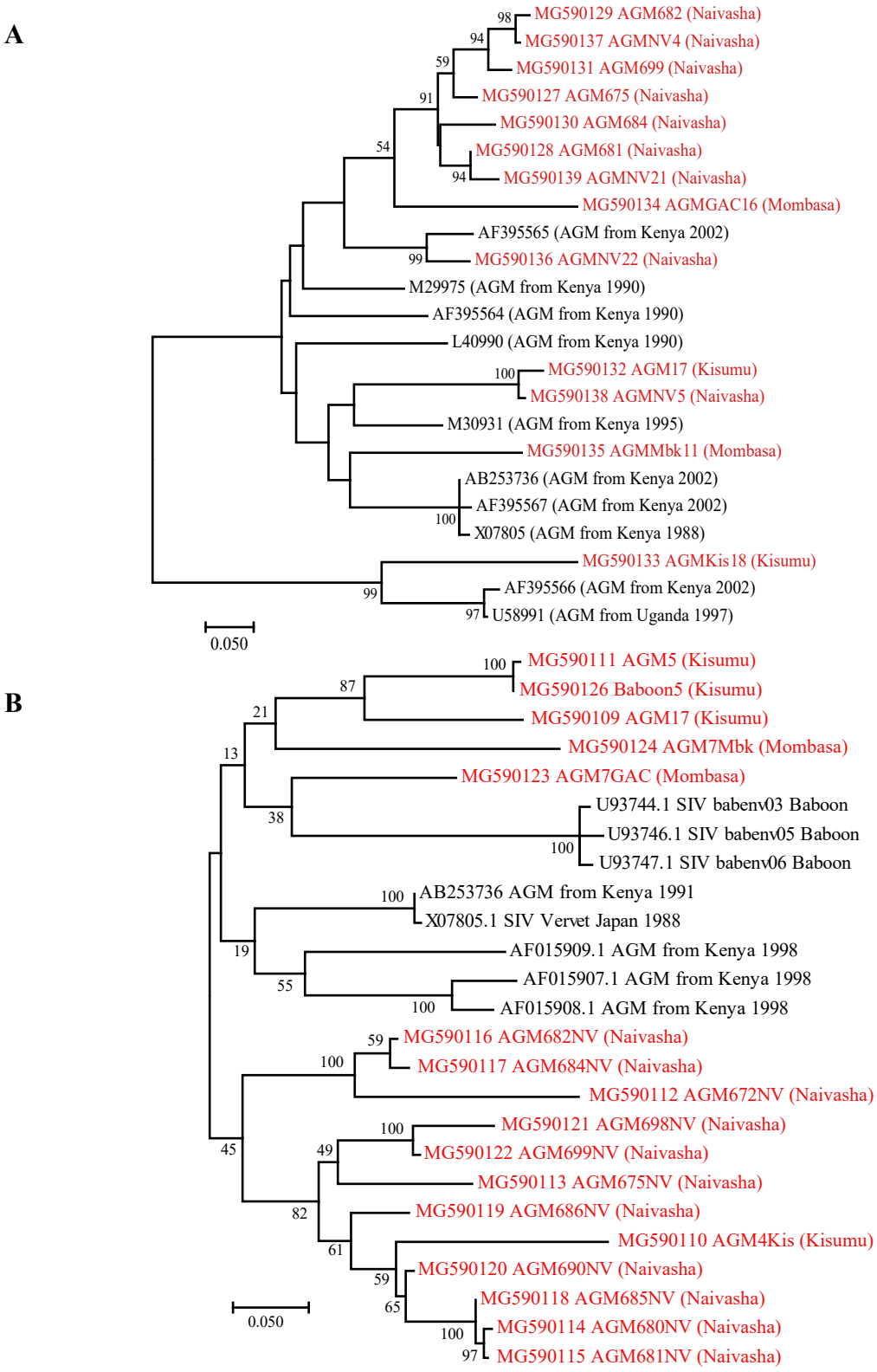


Figure 13: Maximum likelihood phylogenetic tree generated from partial *pol* gene (**A**) and *env* gene (**B**) illustrating genetic diversity of SIV from synanthropic (urban dwelling) AGMs from three urban centres in Kenya alongside previously characterized sequences from Kenya. Sequences from this study are highlighted in red while previously characterised Kenyan sequences retrieved from GenBank are shown in black. The percent bootstrap support values are shown at the nodes of the phylogenetic tree. The branch support evaluation was done using 1000 bootstrap replicates.

4.4.3 SIV *pol* and *env* protein diversity

Overall, protein sequence conservation was higher in the *pol* protein sequences compared to the *env* protein sequences (Figure S1). The identity of the amino acids for the partial *pol* protein ranged between 91 - 99% for the samples obtained from Naivasha. Samples obtained from Mombasa had an amino acid identity of 88%. Amino acid sequences from AGM Kis18 sampled in Kisumu, had 79% - 84% identity with other sequences from Kenya further supporting the genetic distant demonstrated by phylogenetic analysis (Figure 11). Similar to the pattern observed for nucleotide sequences, the sequences from Naivasha had 91 - 94% identity to sequences from Mombasa while also having 88 – 92% identity to sequences from Kisumu supporting lack of phylogeographical clustering. The overall amino acid identity observed for the *pol* protein suggests that codon usage was maintained despite the high nucleotide sequences diversity. Despite phylogenetic and amino acid diversity, the conserved domains and hypervariable regions (V3-V5) were conserved as expected in the *env* protein sequence (Figure 13 B). The identified V3 domain had 34 amino acids as reported previously while maintaining the well-conserved MAG motif that largely determines cellular tropism in SIV (Figure 13. C) (Carr *et al*., 2017). Eight of 12-conserved cysteine in the *env* region (Müller *et al*., 1993; Ma *et al*., 2013), were present. Conservation of major *env* functional domains including the CD4 binding sites, N-linked glycosylation sites and fusion peptide suggests protein function was maintained.

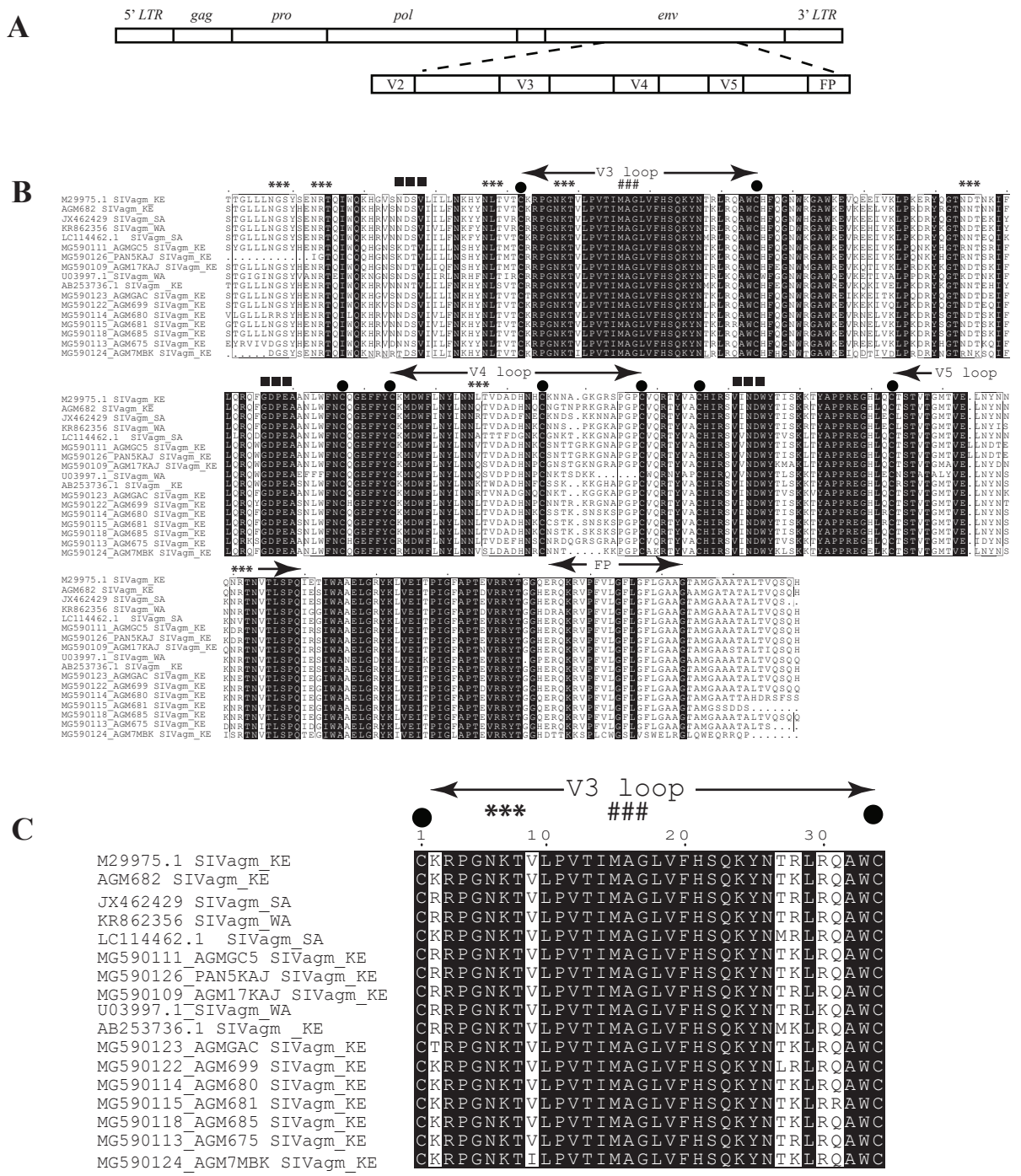


Figure 14. SIV genome architecture and partial *env* protein domains. A. A cartoon representing SIV genome architecture showing respective genes and offset is a representation of the V3-V5 and the fusion protein (FP) domain) domains of the partial *env* glycoprotein. **B.** *Env* protein amino acids alignment showing the V3 to V5 domains, the fusion protein, the potential N-linked glycosylation sites highlighted with asterisks (***). The conserved cysteine residues are highlighted with black dots while potential CD4 binding sites are indicated as black squares, the MAG motif is highlighted hash signs (###). **C.** A trimmed alignment of V3 hypervariable domain comprising 34 residues and showing the MAG motif as reported from previous studies.

4.5 Evolutionary selection

Analysis of evolutionary selection pressure demonstrated signatures of episodic and pervasive diversification on the *env* gene. IFEL and FEL approaches revealed codons 18 and 58 (Table 4) to be under purifying/negative selection (0.05 significance level). FUBAR revealed 14 sites under pervasive diversifying selection at posterior probability of >0.9 while MEME revealed 21 sites to be under episodic diversifying selection at a 0.05 level of significance. Overall, eight sites were identified as being under diversifying selection by FUBAR and MEME (Table 4).

Table 4: SIV *env* gene codons under evolutionary selection pressure. Analysis was undertaken using methods implemented on the Datamonkey web server (Delport *et al*., 2010).

Codon	**Diversifying selection**		**Negative selection**	
	FUBAR (β>α) 0.9)	**MEME (P<0.05)**	**IFEL (P<0.05)**	**FEL (P<0.05)**
	Pervasive	**Episodic**	**Negative**	**Negative**
2	0.930754	0.006223	-	-
18	-	-	0.010375	0.03724
58	-	-	0.020702	0.046966
70	0.926138	0.001725	-	-
124	0.939274	0.002007	-	-
161	0.926871	0.019448	-	-
170	0.965838	0.01065	-	-
296	0.927782	0.01812	-	-
297	0.960401	0.00871	-	-
301	0.927481	0.012432	-	-

4.6 Summary

In summary, this study sought to determine the prevalence and genetic diversity of simian immunodeficiency virus (SIV) in common free-ranging urban restricted non-human primates. The target NHP species included, AGMs and olive baboons captured in Mombasa, Naivasha and Kisumu urban centres. Molecular techniques by PCR-HRM, gel-electrophoresis and sequencing were used to screen DNA samples extracted from the sampled NHPs for SIV. The overall SIV prevalence was 32% (39/124) in AGMs and 3% (2/65). More adults (27 out of 72, 37.5%), than subadults (9 out 36, 25%) and infants (3 out of 16, 18.8 %) were positive for SIV. High-resolution melt analysis of the PCR amplicons revealed high diversity of SIV in the sampled animals. Despite demonstrating extensive diversity, HRM did not distinguish between different strains circulating in the sampled NHPs that led to sequencing to assist in characterisation of the detected SIVs. The nucleotide sequence plasticity observed from sequence comparison could be the reason why HRM did not resolve different SIV strains detected. Partial SIV *pol* and *env* genes were sequenced from this study. Comparison with characterised sequences from GenBank placed these sequences within the SIVagm lineage including the one isolated from an olive baboon. This also revealed nucleotide sequence plasticity upon sequence alignment. Phylogenetic analysis confirmed the extensive diversity that was demonstrated by HRM. There was no phylogeographic clustering according to the NHPs sampling location but rather all the sequences were homogenously distributed within the Kenyan clade of SIVs. Analysis of evolutionary selection pressure revealed episodic and pervasive diversifying selection within the env gene of SIVs indication continuous SIV evolution within its host.

CHAPTER FIVE
DISCUSSION

This study reports on comprehensive molecular survey of SIV variants circulating in common monkeys, the AGMs and olive baboons inhabiting urban and peri-urban centres in Kenya. Apart from serologic survey (Ellis *et al*., 2004) and opportunistic sampling for SIV in captive animals (Otsyula *et al*., 1996), comprehensive molecular epidemiology studies of SIVs in free ranging monkey populations in Kenya have not been previously conducted to the best of our knowledge. We report an overall SIV prevalence of 32% amongst free-ranging urban-restricted AGMs and low SIV infection (3%) in peri-urban dwelling olive baboons in Kenya. Data from this study reveals higher overall SIV prevalence infecting AGMs in Kenya compared to previously documented 26% - 28% seroprevalence in captive and semi-captive AGMs (Ohta *et al*., 1988; Otsyula *et al*., 1996). We also report for the first time detection of SIVagm strain in olive baboons from Kenya. Previous studies have reported SIVagm strains from yellow baboons and chacma baboons from Tanzania and South Africa respectively (Kodama *et al*., 1989; Jin *et al*., 1994).

Although the overall SIV infection rate in AGMs from the urban centres was 33% in Mombasa, 24% in Kisumu and 45% in Naivasha, some sampling sites such as Mbaraki (Mombasa) and Kimana (Naivasha) had much higher infection rates of 48% and 46%, respectively (Table S1). The findings from these sites are comparable to previous molecular studies that reported 44% and 45% overall SIVagm prevalence in wild caught AGMs from Gambia and South Africa, respectively (Ma *et al*., 2013; 2014). Likewise, the AGMs sampled in Tudor (Mombasa) and Kajulu (Kisumu) had lower infection rates of 8% and 19%, respectively (Table S1). This was comparable to a study in Zambia that reported 19% SIVagm prevalence in AGMs (Carr *et al*., 2017). Generally, the prevalence according to sampling sites ranged from 8% - 48% in our study. This was similar to SIV incidence variations reported by Baileys and co-authors (2016) from AGMs from various sampling locations in South Africa. The variations in the prevalence exhibited from our study could be attributed to a smaller sample size as compared to the study from Ma and co-authors (2013). The study from Baileys *et al*. (2016) also had a smaller sample size per sampling site (ranging from 12 to 40 individuals) thus implying a higher prevalence when all the NHPs captured per site individuals turned out to be infected with SIV. Our findings provide further molecular evidence of high SIVagm prevalence in Kenya thus corroborating the consideration of AGMs as natural hosts of SIV in Africa. Despite not showing uniformity of SIV prevalence as suggested from other

studies, our findings corroborate the observation that SIVagm infections are high and common in AGM species across Africa.

Prevalence of SIV amongst different age groups was non-uniform with more adult than sub-adult and infant AGMs being SIV positive. The plausible reason for low infection rate amongst infants is the rare vertical transmission of SIV in AGMs as previously reported (Otsyula *et al*., 1996; Chahroudi *et al*., 2012). Ma and co-authors in 2014 demonstrated immunological suppression of SIV replication in infants thereby limiting infection. In addition, immunological adaptations in infants and juveniles has been shown to confer them with resistance to maternal virus transmission thus lowering SIV prevalence amongst this age-group (Pandrea *et al*., 2012). In contrast, sexual transmission is the main route of SIV transmission in adults and therefore enhanced exposure during mating leads to high prevalence among adults (Phillips-Conroy *et al*., 1994). The prevalence data from this study is comparable to previously documented seroprevalence exceeding 30% amongst extensively studied species including wild and captive AGMs in Africa (Santiago *et al*., 2005; Aghokeng *et al*., 2006; Ma *et al*., 2013).

Detection of SIVagm in an olive baboon from Kenya corroborates epidemiological findings from earlier studies from chacma baboons in Southern Africa and yellow baboons in Tanzania (Kodama *et al*., 1989; Jin *et al*., 1994; Carr *et al*., 2017). Elsewhere, wild caught chacma baboons also tested positive for SIV serologically, while sequencing confirmed the virus to be of SIVagm strain (van Rensburg *et al*., 1998) suggesting natural infection from sympatric AGMs. Our findings confirm the possibility of SIVagm establishing in new hosts. Given that a baboon-specific SIV has not been identified this far, it would be necessary to determine the extent of SIV transmission from other species into baboons. Characterisation of these viruses will not only be important in elucidating the determinants of cross-species transmission but also host adaptation to the virus. This complicates the lineage classification based on host. The infectious outcomes of SIVagm in baboons could offer insights into development of a baboon model for HIV therapeutic research.

To evaluate the diversity and attempt characterisation of SIV strains detected in these free ranging NHPs, PCR amplicons for SIV were subjected to HRM analysis. Despite its robustness, this technique could not explicitly distinguish SIV strains in this study based on their melt profiles. However, variation in melting profiles showed extensive diversity of SIV strains from the same sampling site (Figure 6). In studies where single nucleotide polymorphisms are critical, the robustness of HRM in analysing miniature differences between PCR generated DNA fragments has proved to be reliable (Hinsberger *et al*., 2019). To further

characterise SIV sequences amplified by PCR-HRM and sequenced, we conducted a BLAST search on NCBI's GenBank database. This gave SIVagm as topmost hits for all the sequences from this study. The identities were further confirmed by conducting sequence alignments and generation of identity matrices that placed our sequences within the SIVagm lineage (Figures 8 and 9).

Therefore, variation in HRM melt profiles suggests existence of SIV intra-lineage diversity. This was supported by the results from multiple sequence alignments and generation of percent identity matrices (Figure 8). From the alignments, SIV sequences showed nucleotide plasticity, which could probably contribute to the observed varying melt profiles for the same SIVagm lineage. Likewise, the percent identities from the matrices ranged from 70% to 97% (Figure 9). This shows that despite the sequences falling within the SIVagm lineage, there were imminent differences in their sequences, thus giving rise to varying melt profiles. Therefore, this observation presents a limitation in HRM characterisation of SIV that exhibits low nucleotide conservation on the sequence alignments. As reviewed by Tong & Giffard (2012), an increase in length of the amplicon size beyond 100bp becomes difficult to resolve by HRM for sequences with nucleotide plasticity such as SIV and HIV. Despite the HRM's shortfall, the technique stood out as a robust method in longitudinal evaluation of HIV diversity within an infected individual (Towler *et al*., 2010). In Towler's study, HIV diversity was documented in terms of variation in the melt curve width and not the different shapes/profiles of the melt curves. Progression of HIV infection correlated with virus diversification from founder strains, which correlated with increasing melt curve width. However, PCR-HRM is a cost-effective technique for monitoring prevalence of primate lentiviruses in urban-restricted NHPs rather than independent evaluation of SIV diversity/characterisation especially in virus isolates from different hosts. Despite its shortfalls in SIV characterisation, HRM is emerging as a sensitive and robust probe free technique for detecting infectious pathogens and differentiate genetic variants of different viruses (Lin *et al*., 2008; Steer *et al*., 2009; Yu *et al*., 2015; Villinger *et al*., 2017) including HIV (Towler *et al*., 2010).

Upon phylogenetic analysis, both *env* and *pol* sequences from this study clustered with previous isolates infecting captive AGMs from Kenya (Figure 11 and 12). However, these Kenyan isolates were distinct from those from infected monkeys in South Africa and West Africa. One sequence MG590133-AGMKis18 from Kisumu clustered with SIVagm Tan-1 isolated from Eastern Uganda (Soares *et al*., 1997), suggesting that the two had a common evolutionary history (Figure 11). This could be attributed to geographic proximities of their hosts' ranges as previously reported for simian T-lymphotrophic virus (STLV) infecting

mustached monkeys in Gabon (Liégeois *et al*., 2012). This observation agrees with the west-to-east SIV transmission model across existing AGM geographic ranges given the location of Kisumu relative to Uganda boarder (Wertheim & Worobey, 2007). However, there is need to have a more detailed genetic analysis including full length genome sequencing of the isolate to exclude the possibilities of it being a recombinant/mosaic strain. Based on the overall phylogenetic tree topologies, both *pol* and *env* trees, SIV isolates from West Africa appeared to be ancestral to the SIV strains from East Africa and South Africa. Despite extensive diversity demonstrated by PCR-HRM, phylogenetic analysis did not show phylogeographical clustering of SIV according to different locations in Kenya from which monkey populations were sampled. However, based on identification of closely related phylogenetic clusters with strong node bootstrap supports, epidemiologically linked SIV sequences could be inferred from NHPs sampled in the same population such as Naivasha suggesting high levels of local virus transmission (Figures 13). Bootstrap support values (100%) amongst SIV isolates from Naivasha imply continuous local virus transmission. In contrast, the interspersing of the isolates from Kisumu and Mombasa indicate that divergent strains can also occur in one location as previously reported in mustached monkeys (Liégeois *et al*., 2012) and chimpanzees (Van Heuverswyn *et al*., 2007). Within the cluster of SIV isolated from Kenya, there was no clear segregation according to the sampling towns amongst the newly sequenced *pol* and *env* (Figure 13) SIV gene fragments. This suggests and overall relatedness of these sequences as demonstrated by robust bootstrap values for the clusters on both phylogenetic trees.

For the first time, using molecular tools, we confirmed SIV infection in two olive baboons sampled in Kenya. The SIV sequence recovered from a baboon shared 97% similarity with an isolate from an AGM (MG590111-AGM5) from Kisumu (Figure 10). Phylogenetic analysis placed the baboon-derived sequence in a strongly supported cluster (100% bootstrap) with MG590111-AGM5 sampled in Kisumu (Figure 12). Upon translation and alignment of the nucleotide sequences, these two isolates had 98% amino acid identity (Table S3). Based on these results, the two SIV isolates were closely related and could be epidemiologically linked to a common ancestor within the SIVagm lineage. Detection of SIVagm in an olive baboon implies cross-species SIVagm transmission and diversification, increasing probabilities of establishing in this new baboon host. The most plausible explanation for cross-species transmission could be attributed to the aggressive behaviour of baboons (Jin *et al*., 1994; Carr *et al*., 2017), which strongly imply baboons could be infected from preying on AGMs (Kodama *et al*., 1989). Further, cross-species acquisition of SIVagm in Patas monkey resulted from physical contact involving aggression and bites amongst AGMs and Patas monkeys sharing a

habitat in Senegal (Bibollet-Ruche *et al*., 1996). Likewise, the behavior of hunting and preying on small monkeys resulted in SIVgsn and SIVrcm co-infection in chimpanzees whose outcome was recombination and emergence of SIVcpz in chimpanzees (Bailes *et al*., 2003). Elsewhere, physical aggression resulting in bites from NHPs has been reported as a cause of human zoonotic infection with simian foamy virus (SFV) in Gabon (Mouinga-Ondémé & Kazanji, 2013). All these scenarios provide the chance of exposure to body fluids, one of the modes of retroviruses transmission. This further substantiates the evidence of the complex evolution and cross-species transmission of primate lentiviruses in their natural habitat. The potential of free-ranging urban-restricted AGMs and baboons serving as sources of emerging retroviruses in human populations cannot be understated. Notably, SIVs from AGMs have successfully been cultured on immortalized human CD4 cells (Carr *et al*., 2017), suggesting the possibility of replicating in human cells.

There was an overall amino acid conservation in terms of percentage identity and sequence alignments within the *pol* protein (Table S2 and Figure S1) compared to the *env* protein (Table S3 and Figure S1) as previously reported in northern talapoins (*Miopithecus ogouensis*) (Liegeois *et al*., 2006). However, the amino acid's percent identity (81%-99%) was higher than the nucleotide sequences (66% - 99%) for *pol* gene and (77% - 98%) amino acid identity compared to (62% - 99%) in nucleotide sequence similarity for *env* gene. Equally, percentage amino acid similarities were higher suggesting prevalence of silent mutation. Despite the reduction of the amino acid identity/similarity range compared to the nucleotide sequences suggesting its utility to improve lineage classification, its reliability is still complex. For instance, up to 31% amino acid differences within SIVagm sequences have been reported suggesting high variability in sequences from within the AGM species (van Rensburg *et al*., 1998). In contrast, 84% - 94% amino acid similarity between SIV sequences from red-capped mangabeys (SIVrcm) and agile mangabeys (SIVagi) were reported suggesting close relationship of these two sequences from two different host species (Ahuka-Mundeke *et al*., 2010). Elsewhere, genetic distance assessment of the entire SIVmus proteome revealed identity differences for the isolates from this lineage within the *env* proteome (Aghokeng *et al*., 2007). These findings suggest that classification of SIV into lineages is complex and requires holistic combination of various factors including identity of the host species this gets complex in case of cross-species transmission and recombination. Generally, the *pol* protein conservation enables accurate assignment of newly sequenced SIVs into their respective lineages. Whereas diversity in the *env* domains are associated with subtle but important structural and hence functional changes (Leonard *et al*., 1990).

Despite *env* amino acid diversity, V3-V5 hypervariable domains revealed a similar structure to a reference SIVagmVER-2 (GenBank accession U04004.1) as previously described (Ma *et al*., 2013), suggesting that these sequences could be translated into functional proteins. Further, the expected N-linked glycosylation sites, the cysteine residues and the potential CD4 binding sites were conserved in the env proteins suggesting maintained biological functions despite observed nucleotide sequence diversity (Figure 13). These results suggests existence of synonymous and non-synonymous nucleotide substitutions. Analysis of the synonymous and non-synonymous substitution rates (dS and dN) for the env protein revealed average dS = 0.0309, dN = 0.0957, dS/dN = 0.3149 and pS/pN = 0.3286 (Table S4). The observed dS/dN ratio (0.3149, which is < 1) illustrates the presence of purifying selection against deleterious mutations and hence maintain a viable virus population. Non-synonymous mutations induced by selection pressure affects protein function thereby influence virus fitness. The SIVs from the targeted animals in this study exhibited extensive genetic diversity, indicating that they thrive in a suitable niche for viral evolutionary changes within natural hosts (Simmonds *et al*., 2018). When selection favours mutations within amino acids critical for host-virus interactions, the virus fitness in a new host is favoured (Shackelton *et al*., 2005). Therefore, the diversity and selection patterns of SIV in AGMs could be a recipe for emergence of novel strains that could cross species barriers. The risk of SIV transmission into different host species including human cannot be ignored in the wake of high SIV prevalence in free-ranging urban restricted AGMs with evidence of SIVagm infection in an olive baboon because of cross-species transmission.

The codon-oriented FUBAR and MEME methods revealed evidence of diversifying (positive) selection within the SIV *env* genes sequenced in this study. These signatures of evolutionary selection pressure implies ongoing viral molecular adaptation. The sequenced gene fragment corresponded to the surface (SU) fragment that is accessible to the immune system thus playing a role in host-virus interaction. This interaction with the host elicits immunological pressure on the *env* gene resulting in enhanced evolution, sequence diversity and increased fitness (Lynch *et al*., 2009). Under diversifying or positive selection, more fit variants from a virus quasispecies (a virus population comprising related sequences arising from high mutation rates during viral RNA replication) are selected over variants with deleterious mutations thereby ensuring viable virus population are able to infect new cells. Despite resulting in non-viable individuals, existence of virus quasispecies creates a pool of beneficial mutations that enhance adaptation to new environments and host immunological pressure during infection. This sequence diversification and evolutionary selection in a natural

host is a typical characteristic of potentially zoonotic viruses as opposed to viruses that diversify in a dead-end host thus limiting onward transmission (Vanderford *et al*., 2006). Located within the SIV *env* glycoprotein is a V3 domain, which is the most conserved of all the five hypervariable regions. V3 is crucial to viral entry by being a critical determinant of the chemokine receptor's specificity. Therefore, amino acid sequence changes that affect overall charge will have an impact on co-receptor usage during the host cell invasion, thereby influencing adaptation to a new host (Demma *et al*., 2005; Van Heuverswyn *et al*., 2007). This phenomenon underscores the diversification of SIV as a potential precursor for pathogen spill over.

Our results demonstrate extensive SIVagm diversity in free-ranging AGMs from Kenya, besides complementing the high prevalence findings from other African countries. This study also documents, for the first time in Kenya, SIVagm infection in olive baboons, suggesting cross-species SIVagm transmission. The SIVagm virus is known to be species-specific to the genus *Chlorocebus* (Allan *et al*., 1991). Further, the occurrence of highly related viruses infecting AGMs from Naivasha indicates identifiable local transmission within the population. Immense immunological pressure is being exerted on the *env* protein leading to concentration of episodic diversifying selection in this region and probably increasing chances of establishing in this new host, the baboon. Further sampling of AGMs and baboons is necessary to reveal the historical connection of SIVs infecting monkeys in East Africa and identify the extent of cross-species transmission into baboons in the wild and the associated outcomes of the infection in the new primate host. This could be important for the establishment of a baboon model for HIV vaccine research. With reports of continued transmission of simian retroviruses, including SIV, STLV, and SFV to humans, our findings underscore the importance of characterising diverse SIV strains in East African primate populations to identify those that can potentially cross species barriers and could establish in humans. Finally, our findings demonstrate the urgency of developing local surveillance programmes to identify the extent of retrovirus genetic diversity, including possible cross-species transmission events across major human-NHP interfaces with high transmission risks, and adequate SIV control strategies at human-wildlife interfaces.

CHAPTER SIX
CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

i. This study demonstrates SIV prevalence above 30% in AGMs which is comparable to previous studies that have shown prevalence greater than 40% in wild AGMs in Africa, implying that urban-restricted AGMs also exhibit high levels of SIV infection. Further, this study detected SIVagm strain in olive baboons that commonly shared a habitat with AGMs in Mombasa and Kisumu suggesting possibility of cross-species transmission.

ii. There is extensive genetic diversity of SIV infecting free-ranging NHPs found in urban and peri-urban centres in Kenya as illustrated by both phylogenetics and by amino acid sequence similarity.

iii. Phylogenetic analysis of SIV sequences from this study did not show evidence of phylogeographical clustering according to the origin of the host NHPs. This implies the strains from Kenya could be having a common ancestor including the strain detected in olive baboons. Signatures of diversification detected on the SIV partial *env* sequences implies continuous virus evolution in their hosts that probably generated the strain with the ability to establish in a new host.

6.2 Recommendations

The following are the recommendations drawn from this study:

i. There is potential of cross-species transmission of SIV amongst sympatric NHP species. Cross-species transmission increases the risk of co-infection with different strains, which provides a probable chance for virus recombination which can generate unique SIV strains with unknown zoonotic potential.

ii. With increased frequency of human-wildlife contacts across different human-wildlife interfaces in Kenya, the potential of emergence of novel zoonotic human retroviruses with a significant impact on human health as evidenced by HIV/AIDS pandemic cannot be overlooked. This calls for elucidation of the zoonotic risks posed by SIV circulating in NHP species across various human-NHP interfaces in terms of infection outcomes in a non-natural hosts such as olive baboons. Such studies will help describe the overall picture of the emergence of zoonotic pathogens with a primate origin such as HIV/SIV and inform strategies for mitigating unforeseen pandemics.

iii. There is also a need for community sensitisation and engagement with policy makers about the zoonosis risks involved during aggressive human encounters with NHPs.

REFERENCES

Aghokeng, A. F., Ayouba, A., Mpoudi-Ngole, E., Loul, S., Liegeois, F., Delaporte, E., & Peeters, M. (2010). Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions. *Infection, Genetics and Evolution, 10*(3), 386-396.

Aghokeng, A. F., Bailes, E., Loul, S., Courgnaud, V., Mpoudi-Ngolle, E., Sharp, P. M., Delaporte, E., & Peeters, M. (2007). Full-length sequence analysis of SIVmus in wild populations of mustached monkeys (*Cercopithecus cephus*) from Cameroon provides evidence for two co-circulating SIVmus lineages. *Virology, 360*(2), 407-418.

Aghokeng, A. F., Liu, W., Bibollet-Ruche, F., Loul, S., Mpoudi-Ngole, E., Laurent, C., Mwenda, J. M., Langat, D. K., Chege, G. K., & McClure, H. M. (2006). Widely varying SIV prevalence rates in naturally infected primate species from Cameroon. *Virology, 345*(1), 174-189.

Ahuka-Mundeke, S., Liegeois, F., Ayouba, A., Foupouapouognini, Y., Nerrienet, E., Delaporte, E., & Peeters, M. (2010). Full-length genome sequence of a simian immunodeficiency virus (SIV) infecting a captive agile mangabey (*Cercocebus agilis*) is closely related to SIVrcm infecting wild red-capped mangabeys (*Cercocebus torquatus*) in Cameroon. *The Journal of General Virology, 91*(Pt 12), 2959.

Akinyi, M. Y., Tung, J., Jeneby, M., Patel, N. B., Altmann, J., & Alberts, S. C. (2013). Role of grooming in reducing tick load in wild baboons (*Papio cynocephalus*). *Animal Behaviour, 85*(3), 559-568.

Allan, J. S., Short, M., Taylor, M., Su, S., Hirsch, V., Johnson, P., Shaw, G., & Hahn, B. (1991). Species-specific diversity among simian immunodeficiency viruses from African green monkeys. *Journal of Virology, 65*(6), 2816-2828.

Apetrei, C., Kaur, A., Lerche, N. W., Metzger, M., Pandrea, I., Hardcastle, J., Falkenstein, S., Bohm, R., Koehler, J., & Traina-Dorge, V. (2005). Molecular epidemiology of simian immunodeficiency virus SIVsm in US primate centers unravels the origin of SIVmac and SIVstm. *Journal of Virology, 79*(14), 8991-9005.

Asante-Appiah, E., & Skalka, A. M. (1997). Molecular mechanisms in retrovirus DNA integration. *Antiviral Research, 36*(3), 139-156.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403-410.

Ayouba, A., Akoua-Koffi, C., Calvignac-Spencer, S., Esteban, A., Locatelli, S., Li, H., Li, Y., Hahn, B. H., Delaporte, E., & Leendertz, F. H. (2013). Evidence for continuing cross-

species transmission of SIVsmm to humans: characterization of a new HIV-2 lineage in rural Côte d'Ivoire. *AIDS (London, England), 27*(15).

Ayouba, A., Njouom, R., Chia, J. E., Ahuka-Mundeke, S., Kfutwah, A., Ngole, E. M., Nerrienet, E., Delaporte, E., & Peeters, M. (2015). Molecular characterization of a new mosaic Simian Immunodeficiency Virus in a naturally infected tantalus monkey (*Chlorocebus tantalus*) from Cameroon: A challenge to the virus–host co-evolution of SIVagm in African green monkeys. *Infection, Genetics and Evolution, 30*, 65-73.

Bailes, E., Gao, F., Bibollet-Ruche, F., Courgnaud, V., Peeters, M., Marx, P. A., Hahn, B. H., & Sharp, P. M. (2003). Hybrid origin of SIV in chimpanzees. *Science, 300*(5626), 1713-1713.

Bailey, A. L., Lauck, M., Ghai, R. R., Nelson, C. W., Heimbruch, K., Hughes, A. L., Goldberg, T. L., Kuhn, J. H., Jasinska, A. J., & Freimer, N. B. (2016). Arteriviruses, pegiviruses, and lentiviruses are common among wild African monkeys. *Journal of Virology, 90*(15), 6724-6737.Barouch, D. H. (2008). Challenges in the development of an HIV-1 vaccine. *Nature, 455*(7213), 613.

Bengis, R., Leighton, F., Fischer, J., Artois, M., Morner, T., & Tate, C. (2004). The role of wildlife in emerging and re-emerging zoonoses. *Revue Scientifique et Technique-office International des Epizooties, 23*(2), 497-512.

Bibollet-Ruche, F., Galat-Luong, A., Cuny, G., Sarni-Manchado, P., Galat, G., Durand, J.-P., Pourrut, X., & Veas, F. (1996). Simian immunodeficiency virus infection in a patas monkey (*Erythrocebus patas*): evidence for cross-species transmission from African green monkeys (*Cercopithecus aethiops sabaeus*) in the wild. *Journal of General Virology, 77*(4), 773-781.

Bibollet-Ruche, F., Bailes, E., Gao, F., Pourrut, X., Barlow, K. L., Clewley, J. P., Mwenda, J. M., Langat, D. K., Chege, G. K., & McClure, H. M. (2004). New simian immunodeficiency virus infecting De Brazza's monkeys (*Cercopithecus neglectus*): evidence for a cercopithecus monkey virus clade. *Journal of Virology, 78*(14), 7748-7762.

Carr, M., Kawaguchi, A., Sasaki, M., Gonzalez, G., Ito, K., Thomas, Y., Hang'ombe, B. M., Mweene, A. S., Zhao, G., & Wang, D. (2017). Isolation of a simian immunodeficiency virus from a malbrouck (*Chlorocebus cynosuros*). *Archives of Virology, 162*(2), 543-548.

Chahroudi, A., Bosinger, S. E., Vanderford, T. H., Paiardini, M., & Silvestri, G. (2012). Natural SIV hosts: showing AIDS the door. *Science, 335*(6073), 1188-1193.

Chahroudi, A., Meeker, T., Lawson, B., Ratcliffe, S., Else, J., & Silvestri, G. (2011). Mother-to-infant transmission of simian immunodeficiency virus is rare in sooty mangabeys and is associated with low viremia. *Journal of Virology, 85*(12), 5757-5763.

Chahroudi, A., Permar, S., & Pandrea, I. (2014). Chapter 13 - SIV Transmission in natural hosts. In A. A. Ansari & G. Silvestri (Eds.), Natural hosts of SIV (pp. 257-268). Amsterdam: Elsevier.

Chakrabarti, L., Hurtrel, M., Maire, M., Vazeux, R., Dormont, D., Montagnier, L., & Hurtrel, B. (1991). Early viral replication in the brain of SIV-infected rhesus monkeys. *The American Journal of Pathology, 139*(6), 1273.

Charon, J., Grigg, M. J., Eden, J.-S., Piera, K. A., William, T., Rose, K., Davenport, M. P., Anstey, N. M., & Holmes, E. C. (2019). Matryoshka RNA virus 1: a novel RNA virus associated with Plasmodium parasites in human malaria. *BioRxiv*, 624403

Chen, Z., Telfier, P., Gettie, A., Reed, P., Zhang, L., Ho, D. D., & Marx, P. A. (1996). Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. *Journal of Virology, 70*(6), 3617-3627.

Chretien, J.-P., Anyamba, A., Small, J., Britch, S., Sanchez, J. L., Halbach, A. C., Tucker, C., & Linthicum, K. J. (2015). Global climate anomalies and potential infectious disease risks: 2014-2015. *PLoS Currents, 7*.

Chua, K. B., & Gubler, D. J. (2013). Perspectives of public health laboratories in emerging infectious diseases. *Emerging Microbes and Infections, 2*(1), 1-6.

Coffin, J. M. (1992). Structure and classification of retroviruses. In J. A. Levy (Ed.), *The Retroviridae* (pp. 19-49). Boston, MA: Springer.

Courgnaud, V., Salemi, M., Pourrut, X., Mpoudi-Ngole, E., Abela, B., Auzel, P., Bibollet-Ruche, F., Hahn, B., Vandamme, A.-M., & Delaporte, E. (2002). Characterization of a novel simian immunodeficiency virus with a vpu gene from greater spot-nosed monkeys (*Cercopithecus nictitans*) provides new insights into simian/human immunodeficiency virus phylogeny. *Journal of Virology, 76*(16), 8298-8309.

D'arc, M., Ayouba, A., Esteban, A., Learn, G. H., Boué, V., Liegeois, F., Etienne, L., Tagg, N., Leendertz, F. H., & Boesch, C. (2015). Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proceedings of the National Academy of Sciences, 112*(11), E1343-E1352.

Dazza, M. C., Ekwalanga, M., Nende, M., Shamamba, K. B., Bitshi, P., Paraskevis, D., & Saragosti, S. (2005). Characterization of a novel vpu-harboring simian immunodeficiency virus from a Dent's Mona monkey (*Cercopithecus mona denti*). *Journal of Virology, 79*(13), 8560-8571.

de Jong, Y. A., & Butynski, T. M. (2010). Three sykes's monkey *Cercopithecus mitis*× vervet monkey *Chlorocebus pygerythrus* hybrids in Kenya. *Primate Conservation, 2010*(25), 43-57.

Delport, W., Poon, A. F., Frost, S. D., & Kosakovsky Pond, S. L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics, 26*(19), 2455-2457.

Demma, L. J., Logsdon Jr, J. M., Vanderford, T. H., Feinberg, M. B., & Staprans, S. I. (2005). SIVsm quasispecies adaptation to a new simian host. *PLoS Pathogens, 1*(1), e3.

Desrosiers, R. C. (1990). The simian immunodeficiency viruses. *Annual Review of Immunology, 8*(1), 557-578.

Dwyer-Lindgren, L., Cork, M. A., Sligar, A., Steuben, K. M., Wilson, K. F., Provost, N. R., Mayala, B. K., VanderHeide, J. D., Collison, M. L., & Hall, J. B. (2019). Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature, 570*(7760), 189.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32*(5), 1792-1797.

Edinger, A. L., Blanpain, C., Kunstman, K. J., Wolinsky, S. M., Parmentier, M., & Doms, R. W. (1999). Functional dissection of CCR5 coreceptor function through the use of CD4-independent simian immunodeficiency virus strains. *Journal of Virology, 73*(5), 4062-4073.

Ellis, B. R., Munene, E., Elliott, D., Robinson, J., Otsyula, M. G., & Michael, S. F. (2004). Seroprevalence of simian immunodeficiency virus in wild and captive born Sykes' monkeys (*Cercopithecus mitis*) in Kenya. *Retrovirology, 1*(1), 34.

Engelman, A., & Cherepanov, P. (2012). The structural biology of HIV-1: mechanistic and therapeutic insights. *Nature Reviews Microbiology, 10*(4), 279-290.

Engering, A., Hogerwerf, L., & Slingenbergh, J. (2013). Pathogen–host–environment interplay and disease emergence. *Emerging Microbes and Infections, 2*(2), e5.

Faust, T. B., Binning, J. M., Gross, J. D., & Frankel, A. D. (2017). Making sense of multifunctional proteins: human immunodeficiency virus type 1 accessory and regulatory proteins and connections to transcription. *Annual Review of Virology, 4*, 241-260.

Foley, B. T., Leitner, T., Paraskevis, D., & Peeters, M. (2016). Primate immunodeficiency virus classification and nomenclature. *Infection, Genetics and Evolution, 46*, 150-158.

Friant, S., Paige, S. B., & Goldberg, T. L. (2015). Drivers of bushmeat hunting and perceptions of zoonoses in Nigerian hunting communities. *PLoS Neglected Tropical Diseases, 9*(5), e0003792.

Fu, X., Fang, B., Liu, Y., Cai, M., Jun, J., Ma, J., Bu, D., Wang, L., Zhou, P., & Wang, H. (2018). Newly emerged porcine enteric alphacoronavirus in southern China: Identification, origin and evolutionary history analysis. *Infection, Genetics and Evolution, 62*, 179-187.

Groves, C. (2016). Primates (Taxonomy). In M. Bezanson, K. C. MacKinnon, E. Riley, K. C. J. Campbell, & A. E. (. Nekaris, A. F. Di Fiore, S. Ross, L. E. Jones-Engel, B. Thierry, R. W. Sussman, C. Sanz, J. Loudon, S. Elton and A. Fuentes) (Eds.), *The International Encyclopedia of Primatology*. Hoboken, NJ: John Wiley & Sons.

Galat-Luong, A., Bibollet-Ruche, F., Pourrut, X., Durand, J., Sarni, P., Pichon, G., & Galat, G. (1994). Social organisation and SIV sero-epidemiology of a patas monkey population in Senegal. *Folia Primatologica, 63*(4), 226-228.

Gallo, R. C., & Montagnier, L. (2003). The discovery of HIV as the cause of AIDS. *New England Journal of Medicine, 349*(24), 2283-2285.

Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M., & Shaw, G. M. (1999). Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*, *397*(6718), 436.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*(5), 696-704.

Haesevelde, M. M. V., Peeters, M., Jannes, G., Janssens, W., Guido, V. D. G., Sharp, P. M., & Saman, E. (1996). Sequence analysis of a highly divergent HIV-1-related lentivirus isolated from a wild captured chimpanzee. *Virology, 221*(2), 346-350.

Hinsberger, A., Theulier Saint Germain, S., Guerrero, P., Blachère-López, C., López-Ferber, M., & Bayle, S. (2019). A Combination of Real-Time PCR and High-Resolution Melting Analysis to Detect and Identify CpGV Genotypes Involved in Type I Resistance. *Viruses*, *11*(8), 723.

Hirsch, V. M., Dapolito, G., Goeken, R., & Campbell, B. J. (1995). Phylogeny and natural history of the primate lentiviruses, SIV and HIV. *Current Opinion in Genetics & Development, 5*(6), 798-806.

Hirsch, V. M., McGann, C., Dapolito, G., Goldstein, S., Ogen-Odoi, A., Biryawaho, B., Lakwo, T., & Johnson, P. R. (1993). Identification of a new subgroup of SIVagm in tantalus monkeys. *Virology, 197*(1), 426-430.

Howard, C. R., & Fletcher, N. F. (2012). Emerging virus diseases: can we ever expect the unexpected? *Emerging Microbes & Infections, 1*(1), 1-9.

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., & Abdelalim, A. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet, 392*(10159), 1789-1858.

Jeneby, M. M., Suleman, M., & Gichuki, C. (2002). Sero-epizootiologic survey of *Trypanosoma brucei* in Kenyan nonhuman primates. *Journal of Zoo and Wildlife Medicine, 33*(4), 337-342.

Jeneby, M. M., Suleman, M. A., Akinyi, M., Ozwara, H., Kariuki, T., & Carlsson, H.-E. (2011). Prevalence of *Babesia microti* in free-ranging baboons and African green monkeys. *Journal of Parasitology, 97*(1), 63-68.

Johnson, P., Fomsgaard, A., Allan, J., Gravell, M., London, W., Olmsted, R., & Hirsch, V. (1990). Simian immunodeficiency viruses from African green monkeys display unusual genetic diversity. *Journal of Virology, 64*(3), 1086-1092.

Johnson, P., Gravel, M., Allan, J., Goldstein, S., Olmsted, R. A., Dapolito, G., McGann, C., London, W., Purcell, R. H., & Hirsch, V. M. (1989). Genetic diversity among simian immunodeficiency virus isolates from African green monkeys. *Journal of Medical Primatology, 18*(3-4), 271-277.

Jin, M. J., Rogers, J., Phillips-Conroy, J. E., Allan, J. S., Desrosiers, R. C., Shaw, G. M., Sharp, P. M., & Hahn, B. H. (1994). Infection of a yellow baboon with simian immunodeficiency virus from African green monkeys: evidence for cross-species transmission in the wild. *Journal of Virology, 68*(12), 8454-8460.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature, 451*(7181), 990.

Kagira, J., Kivai, S., Ndeereh, D., Ngotho, M., Akinyi, M., Mwangi, D. K., & Maloba, F. (2017). Socioeconomic and health implications of human-wildlife interactions in Nthongoni, Eastern Kenya. *African Journal of Wildlife Research, 46*(2), 87-103.

Kapoor, A., Kumar, A., Simmonds, P., Bhuva, N., Chauhan, L. S., Lee, B., Jin, Z., Morse, S. S., Shaz, B., & Burbelo, P. D. (2015). Virome analysis of transfusion recipients reveals

a novel human virus that shares genomic features with hepaciviruses and pegiviruses. *MBio, 6*(5), e01466-01415.

Kapusinszky, B., Mulvaney, U., Jasinska, A. J., Deng, X., Freimer, N., & Delwart, E. (2015). Local virus extinctions following a host population bottleneck. *Journal of Virology*, 89(16), 8152-8161.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., & Duran, C. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics, 28*(12), 1647-1649.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., & Mitchell, C. E. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature, 468*(7324), 647.

Kimwele, C., Karisa, B., Stokes, M., Junga, J., Hanotte, O., Skilton, R. A., & McElroy, D. (2012). DNA species surveillance: Monitoring bushmeat poaching and trading in Kenya using partial cytochrome b gene. *African Journal of Biotechnology, 11*(78), 14276-14286.

Kirmaier, A., Wu, F., Newman, R. M., Hall, L. R., Morgan, J. S., O'Connor, S., Marx, P. A., Meythaler, M., Goldstein, S., & Buckler-White, A. (2010). TRIM5 suppresses cross-species transmission of a primate immunodeficiency virus and selects for emergence of resistant variants in the new species. *PLoS Biology, 8*(8), e1000462.

Klatt, N. R., Silvestri, G., & Hirsch, V. (2012). Nonpathogenic simian immunodeficiency virus infections. *Cold Spring Harbor Perspectives in Medicine, 2*(1), a007153.

Kodama, T., Silva, D. P., Daniel, M. D., Phillips-Conroy, J. E., Jolly, C. J., Rogers, J., & Desrosiers, R. C. (1989). Prevalence of antibodies to SIV in baboons in their native habitat. *AIDS Research and Human Retroviruses, 5*(3), 337-343.

Krief, S., Escalante, A. A., Pacheco, M. A., Mugisha, L., André, C., Halbwax, M., Fischer, A., Krief, J.-M., Kasenene, J. M., & Crandfield, M. (2010). On the diversity of malaria parasites in African apes and the origin of *Plasmodium falciparum* from Bonobos. *PLoS Pathogens, 6*(2), e1000765.

Lauck, M., Switzer, W. M., Sibley, S. D., Hyeroba, D., Tumukunde, A., Weny, G., Taylor, B., Shankar, A., Ting, N., & Chapman, C. A. (2013). Discovery and full genome characterization of two highly divergent simian immunodeficiency viruses infecting black-and-white colobus monkeys (*Colobus guereza*) in Kibale National Park, Uganda. *Retrovirology, 10*(1), 107.

Lefort, V., Longueville, J. E., & Gascuel, O. (2017). SMS: smart model selection in PhyML. *Molecular Biology and Evolution, 34*(9), 2422-2424.

Leonard, C. K., Spellman, M. W., Riddle, L., Harris, R. J., Thomas, J. N., & Gregory, T. (1990). Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *Journal of Biological Chemistry, 265*(18), 10373-10382.

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research, 44*(W1), W242-W245.

Liégeois, F., Boué, V., Mouacha, F., Butel, C., Ondo, B. M., Pourrut, X., Leroy, E., Peeters, M., & Rouet, F. (2012). New STLV-3 strains and a divergent SIVmus strain identified in non-human primate bushmeat in Gabon. *Retrovirology, 9* (1), 28.

Liégeois, F., Courgnaud, V., Switzer, W. M., Murphy, H. W., Loul, S., Aghokeng, A., Pourrut, X., Mpoudi-Ngole, E., Delaporte, E., & Peeters, M. (2006). Molecular characterization of a novel simian immunodeficiency virus lineage (SIVtal) from northern talapoins (*Miopithecus ogouensis*). *Virology, 349*(1), 55-65.

Lin, J.-H., Tseng, C.-P., Chen, Y.-J., Lin, C.-Y., Chang, S.-S., Wu, H.-S., & Cheng, J.-C. (2008). Rapid differentiation of influenza A virus subtypes and genetic screening for virus variants by high-resolution melting analysis. *Journal of Clinical Microbiology, 46*(3), 1090-1097.

Liovat, A.-S., Jacquelin, B., Ploquin, M. J., Barré-Sinoussi, F., & Muller-Trutwin, M. C. (2009). African non-human primates infected by SIV-why don't they get sick? Lessons from studies on the early phase of non-pathogenic SIV infection. *Current HIV Research, 7*(1), 39-50.

Liu, W., Li, Y., Learn, G. H., Rudicell, R. S., Robertson, J. D., Keele, B. F., Ndjango, J.-B. N., Sanz, C. M., Morgan, D. B., & Locatelli, S. (2010). Origin of the human malaria parasite Plasmodium falciparum in gorillas. *Nature, 467*(7314), 420-425.

Locatelli, S., Liegeois, F., Lafay, B., Roeder, A. D., Bruford, M. W., Formenty, P., Noë, R., Delaporte, E., & Peeters, M. (2008). Prevalence and genetic diversity of simian immunodeficiency virus infection in wild-living red colobus monkeys (*Piliocolobus badius badius*) from the Taï forest, Côte d'Ivoire: SIVwrc in wild-living western red colobus monkeys. *Infection, Genetics and Evolution, 8*(1), 1-14.

Locatelli, S., & Peeters, M. (2012). Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. *AIDS, 26*(6), 659-673.

Luciw, P. A., & Leung, N. J. (1992). Mechanisms of retrovirus replication. In J. A. Levy (Ed.), *The retroviridae* (pp. 159-298). Boston, MA: Springer.

Lynch, R. M., Shen, T., Gnanakaran, S., & Derdeyn, C. A. (2009). Appreciating HIV type 1 diversity: subtype differences in Env. *AIDS Research and Human Retroviruses, 25*(3), 237-248.

Ma, D., Jasinska, A., Kristoff, J., Grobler, J. P., Turner, T., Jung, Y., Schmitt, C., Raehtz, K., Feyertag, F., & Sosa, N. M. (2013). SIVagm infection in wild African green monkeys from South Africa: epidemiology, natural history, and evolutionary considerations. *PLoS Pathogens, 9*(1), e1003011.

Ma, D., Jasinska, A. J., Feyertag, F., Wijewardana, V., Kristoff, J., He, T., Raehtz, K., Schmitt, C. A., Jung, Y., & Cramer, J. D. (2014). Factors associated with siman immunodeficiency virus transmission in a natural African nonhuman primate host in the wild. *Journal of Virology, 88*(10), 5687-5705.

Mouinga-Ondémé, A., & Kazanji, M. (2013). Simian foamy virus in non-human primates and cross-species transmission to humans in Gabon: an emerging zoonotic disease in central Africa? *Viruses, 5*(6), 1536-1552.

Morens, D. M., & Fauci, A. S. (2013). Emerging infectious diseases: threats to human health and global stability. *PLoS Pathogens, 9*(7), e1003467.

Morens, D. M., Folkers, G. K., & Fauci, A. S. (2004). The challenge of emerging and re-emerging infectious diseases. *Nature, 430*(6996), 242.

Morse, S. S. (2001). Factors in the emergence of infectious diseases. In *Plagues and Politics* (pp. 8-26): Springer.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S. L., & Scheffler, K. (2013). FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Molecular Biology and Evolution, 30*(5), 1196-1205.

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Pond, S. L. K. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics, 8*(7), e1002764.

Mutemi, M., Njenga, M., Lamond, G., Kuria, A., Öborn, I., Muriuki, J., & Sinclair, F. L. (2017). Using local knowledge to understand challenges and opportunities for

enhancing agricultural productivity in Western Kenya. *Sustainable Intensification in Smallholder Agriculture*, 177.

Mwangi, D. K., Akinyi, M., Maloba, F., Ngotho, M., Kagira, J., Ndeereh, D., & Kivai, S. (2016). Socioeconomic and health implications of human-wildlife interactions in Nthongoni, Eastern Kenya. *African Journal of Wildlife Research, 46*(2), 87-103.

Newman, T. K., Jolly, C. J., & Rogers, J. (2004). Mitochondrial phylogeny and systematics of baboons (*Papio*). *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists, 124*(1), 17-27.

Nyakarahuka, L., Ojwang, J., Tumusiime, A., Balinandi, S., Whitmer, S., Kyazze, S., Kasozi, S., Wetaka, M., Makumbi, I., & Dahlke, M. (2017). Isolated case of Marburg virus disease, Kampala, Uganda, 2014. *Emerging Infectious Diseases, 23*(6), 1001.

Ndembi, N., Kaptue, L., & Ido, E. (2009). Exposure to SIVmnd-2 in southern Cameroon: public health implications. *AIDS Reviews, 11*(3), 135-139.

Neil, S. J., Zang, T., & Bieniasz, P. D. (2008). Tetherin inhibits retrovirus release and is antagonised by HIV-1 Vpu. *Nature, 451*(7177), 425.

Otsyula, M., Yee, J., Jennings, M., Suleman, M., Gettie, A., Tarara, R., Isahakia, M., Marx, P., & Lerche, N. (1996). Prevalence of antibodies against simian immunodeficiency virus (SIV) and simian T-lymphotropic virus (STLV) in a colony of non-human primates in Kenya, East Africa. *Annals of Tropical Medicine & Parasitology, 90*(1), 65-70.

Pandrea, I., Parrish, N. F., Raehtz, K., Gaufin, T., Barbian, H. J., Ma, D., Kristoff, J., Gautam, R., Zhong, F., & Haret-Richter, G. S. (2012). Mucosal simian immunodeficiency virus transmission in African green monkeys: susceptibility to infection is proportional to target cell availability at mucosal sites. *Journal of Virology, 86*(8), 4158-4168.

Pandrea, I., Silvestri, G., & Apetrei, C. (2009). AIDS in African nonhuman primate hosts of SIVs: a new paradigm of SIV infection. *Current HIV Research, 7*(1), 57-72.

Paules, C. I., Eisinger, R. W., Marston, H. D., & Fauci, A. S. (2017). What recent history has taught us about responding to emerging infectious disease threats. *Annals of Internal Medicine, 167*(11), 805-811.

Peeters, M., Courgnaud, V., Abela, B., Auzel, P., Pourrut, X., Bibollet-Ruche, F., Loul, S., Liegeois, F., Butel, C., & Koulagna, D. (2002). Risk to human health from a plethora of simian immunodeficiency viruses in primate bushmeat. *Emerging Infectious Diseases, 8*(5), 451.

Peeters, M., D'Arc, M., & Delaporte, E. (2014). The origin and diversity of human retroviruses. *AIDS Reviews, 16*(1), 23.

Peeters, M., & Delaporte, E. (2012). Simian retroviruses in African apes. *Clinical Microbiology and Infection, 18*(6), 514-520.

Peeters, M., Jung, M., & Ayouba, A. (2013). The origin and molecular epidemiology of HIV. *Expert Review of Anti-infective therapy, 11*(9), 885-896.

Petersen, E., Petrosillo, N., Koopmans, M., Beeching, N., Di Caro, A., Gkrania-Klotsas, E., Kantele, A., Kohlmann, R., Lim, P. L., & Markotic, A. (2018). Emerging infections - an increasingly important topic: review by the Emerging Infections Task Force. *Clinical Microbiology and Infection, 24*(4), 369-375.

Petropoulos, C. (1997). Retroviral taxonomy, protein structures, sequences, and genetic maps. *Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*, 757-805.

Petrovas, C., Yamamoto, T., Gerner, M. Y., Boswell, K. L., Wloka, K., Smith, E. C., Ambrozak, D. R., Sandler, N. G., Timmer, K. J., & Sun, X. (2012). CD4 T follicular helper cell dynamics during SIV infection. *The Journal of Clinical Investigation, 122*(9), 3281-3294.

Phillips-Conroy, J., Jolly, C., Petros, B., Allan, J., & Desrosiers, R. C. (1994). Sexual transmission of SIVagm in wild grivet monkeys. *Journal of Medical Primatology, 23*(1), 1-7.

Pigott, D. M., Golding, N., Mylne, A., Huang, Z., Weiss, D. J., Brady, O. J., Kraemer, M. U., & Hay, S. I. (2015). Mapping the zoonotic niche of Marburg virus disease in Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 109*(6), 366-378.

Pond, S., & Frost, S. (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution, 22*(5), 1208-1222.

Pond, S., Frost, S. D., Grossman, Z., Gravenor, M. B., Richman, D. D., & Brown, A. J. L. (2006). Adaptation to different human populations by HIV-1 revealed by codon-based analyses. *PLoS Computational Biology, 2*(6), e62.

Prabakaran, P., Dimitrov, A. S., Fouts, T. R., & Dimitrov, D. S. (2007). Structure and function of the HIV envelope glycoprotein as entry mediator, vaccine immunogen, and target for inhibitors. *Advances in Pharmacology, 55*, 33-97.

Rhyan, J., & Spraker, T. (2010). Emergence of diseases from wildlife reservoirs. *Veterinary Pathology, 47*(1), 34-39.

Richman, D. D., Margolis, D. M., Delaney, M., Greene, W. C., Hazuda, D., & Pomerantz, R. J. (2009). The challenge of finding a cure for HIV infection. *Science, 323*(5919), 1304-1307.

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research, 42*(W1), W320-W324.

Roesch, F. (2018). SAMHD1: mechanisms of regulation and viral evasion. *Virologie, 22*(4), 215-225.

Saéz, A. M., Weiss, S., Nowak, K., Lapeyre, V., Zimmermann, F., Düx, A., Kühl, H. S., Kaba, M., Regnaut, S., & Merkel, K. (2015). Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Molecular Medicine, 7*(1), 17-23.

Santiago, M. L., Bibollet-Ruche, F., Bailes, E., Kamenya, S., Muller, M. N., Lukasik, M., Pusey, A. E., Collins, D. A., Wrangham, R. W., & Goodall, J. (2003). Amplification of a complete simian immunodeficiency virus genome from fecal RNA of a wild chimpanzee. *Journal of Virology, 77*(3), 2233-2242.

Santiago, M. L., Range, F., Keele, B. F., Li, Y., Bailes, E., Bibollet-Ruche, F., Fruteau, C., Noë, R., Peeters, M., & Brookfield, J. F. (2005). Simian immunodeficiency virus infection in free-ranging sooty mangabeys (*Cercocebus atys atys*) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *Journal of Virology, 79*(19), 12515-12527.

Santoro, M. M., & Perno, C. F. (2013). HIV-1 genetic variability and clinical implications. *ISRN Microbiology, 2013*.

Seki, S., & Matano, T. (2012). CTL escape and viral fitness in HIV/SIV infection. *Frontiers in Microbiology, 2*, 267.

Shackelton, L. A., Parrish, C. R., Truyen, U., & Holmes, E. C. (2005). High rate of viral evolution associated with the emergence of carnivore parvovirus. *Proceedings of the National Academy of Sciences, 102*(2), 379-384.

Shah, S., Alexaki, A., Pirrone, V., Dahiya, S., Nonnemacher, M. R., & Wigdahl, B. (2014). Functional properties of the HIV-1 long terminal repeat containing single-nucleotide polymorphisms in Sp site III and CCAAT/enhancer binding protein site I. *Virology Journal, 11*(1), 92.

Sharp, P. M., & Hahn, B. H. (2010). The evolution of HIV-1 and the origin of AIDS. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1552), 2487-2494.

Sharp, P. M., & Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harbor Perspectives in Medicine, 1*(1), a006841.

Sharp, P. M., Shaw, G. M., & Hahn, B. H. (2005). Simian immunodeficiency virus infection of chimpanzees. *Journal of Virology, 79*(7), 3891-3902.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., & Söding, J. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology, 7*(1).

Silvestri, G., Paiardini, M., Pandrea, I., Lederman, M. M., & Sodora, D. L. (2007). Understanding the benign nature of SIV infection in natural hosts. *The Journal of Clinical Investigation, 117*(11), 3148-3154.

Simmonds, P., Aiewsakun, P., & Katzourakis, A. (2018). Prisoners of war-host adaptation and its constraints on virus evolution. *Nature Reviews Microbiology*, 1.

Simonin, Y., Sillam, O., Carles, M. J., Gutierrez, S., Gil, P., Constant, O., Martin, M. F., Girard, G., Van de Perre, P., & Salinas, S. (2018). Human Usutu virus infection with atypical neurologic presentation, Montpellier, France, 2016. *Emerging Infectious Diseases, 24*(5), 875.

Smith, D. M., Richman, D. D., & Little, S. J. (2005). HIV superinfection. *The Journal of Infectious Diseases, 192*(3), 438-444.

Smith, K., Machalaba, C., Seifman, R., Feferholtz, Y., & Karesh, W. (2019). Infectious disease and economics: The case for considering multi-sectoral impacts. *One Health*, 100080.

Soares, M. A., Robertson, D. L., Hui, H., Allan, J. S., Shaw, G. M., & Hahn, B. H. (1997). A Full-length and replication-competent proviral clone of SIVagm from tantalus monkeys. *Virology, 228*(2), 394-399.

Sodora, D. L., Allan, J. S., Apetrei, C., Brenchley, J. M., Douek, D. C., Else, J. G., Estes, J. D., Hahn, B. H., Hirsch, V. M., & Kaur, A. (2009). Toward an AIDS vaccine: lessons from natural simian immunodeficiency virus infections of African nonhuman primate hosts. *Nature Medicine, 15*(8), 861.

Steer, P. A., Kirkpatrick, N. C., O'Rourke, D., & Noormohammadi, A. H. (2009). Classification of fowl adenovirus serotypes by use of high-resolution melting-curve analysis of the hexon gene region. *Journal of Clinical Microbiology, 47*(2), 311-321.

Steve, A.-M., Ahidjo, A., Placide, M.-K., Caroline, F., Mukulumanya, M., Simon-Pierre, N.-K., Octavie, L.-M., Valentin, M.-A., Jean-Jacques, M.-T., & Eric, D. (2017). High prevalence and a wide genetic diversity of simian retroviruses in non-human primate bushmeat in rural areas of the Democratic Republic of Congo. *EcoHealth, 14*(1), 100-114.

Strøm, G. E., Tellevik, M. G., Hanevik, K., Langeland, N., & Blomberg, B. (2014). Comparison of four methods for extracting DNA from dried blood on filter paper for

PCR targeting the mitochondrial Plasmodium genome. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 108*(8), 488-494.

Takehisa, J., Kraus, M. H., Ayouba, A., Bailes, E., Van Heuverswyn, F., Decker, J. M., Li, Y., Rudicell, R. S., Learn, G. H., & Neel, C. (2009). Origin and biology of simian immunodeficiency virus in wild-living western gorillas. *Journal of Virology, 83*(4), 1635-1648.

Tong, S. Y., & Giffard, P. M. (2012). Microbiological applications of high-resolution melting analysis. *Journal of Clinical Microbiology, 50*(11), 3418-3421

Towler, W. I., James, M. M., Ray, S. C., Wang, L., Donnell, D., Mwatha, A., Guay, L., Nakabiito, C., Musoke, P., & Jackson, J. B. (2010). Analysis of HIV diversity using a high-resolution melting assay. *AIDS Research and Human Retroviruses, 26*(8), 913-918.

Tsetsarkin, K. A., Chen, R., & Weaver, S. C. (2016). Interspecies transmission and chikungunya virus emergence. *Current Opinion in Virology, 16*, 143-150.

UNAIDS. (2019). Global HIV & AIDS statistics - 2019 fact sheet.

Van Heuverswyn, F., Li, Y., Bailes, E., Neel, C., Lafay, B., Keele, B. F., Shaw, K. S., Takehisa, J., Kraus, M. H., & Loul, S. (2007). Genetic diversity and phylogeographic clustering of SIVcpz*Ptt* in wild chimpanzees in Cameroon. *Virology, 368*(1), 155-171.

van Rensburg, E. J., Engelbrecht, S., Mwenda, J., Laten, J. D., Robson, B. A., Stander, T., & Chege, G. K. (1998). Simian immunodeficiency viruses (SIVs) from eastern and southern Africa: detection of a SIVagm variant from a chacma baboon. *Journal of General Virology, 79*(7), 1809-1814.

Vanderford, T. H., Demma, L. J., Feinberg, M. B., Staprans, S. I., & Logsdon Jr, J. M. (2006). Adaptation of a diverse simian immunodeficiency virus population to a new host is revealed through a systematic approach to identify amino acid sites under selection. *Molecular Biology and Evolution, 24*(3), 660-669.

Villinger, J., Mbaya, M. K., Ouso, D., Kipanga, P. N., Lutomiah, J., & Masiga, D. K. (2017). Arbovirus and insect-specific virus discovery in Kenya by novel six genera multiplex high-resolution melting analysis. *Molecular Ecology Resources, 17*(3), 466-480.

Wang, H., Naghavi, M., Allen, C., Barber, R. M., Bhutta, Z. A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., & Coates, M. M. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1459-1544.

Weaver, S. C., Charlier, C., Vasilakis, N., & Lecuit, M. (2018). Zika, chikungunya, and other emerging vector-borne viral diseases. *Annual Review of Medicine, 69*, 395-408.

Wertheim, J. O., & Worobey, M. (2007). A challenge to the ancient origin of SIVagm based on African green monkey mitochondrial genomes. *PLoS Pathogens, 3*(7), e95.

WHO. (2004). Joint consultation report on emerging zoonotic diseases (joint consultation No. WHO/CDS/CPE/ZFK/2004.9). Geneva, Switzerland.

WHO. (2016). World health statistics 2016: monitoring health for the SDGs sustainable development goals: *World Health Organization*. Geneva, Switzerland.

Wolfe, N., Switzer, W. M., Carr, J. K., Bhullar, V. B., Shanmugam, V., Tamoufe, U., Prosser, A. T., Torimiro, J. N., Wright, A., & Mpoudi-Ngole, E. (2004). Naturally acquired simian retrovirus infections in central African hunters. *The Lancet, 363*(9413), 932-937.

Wolfe, N. D., Dunavan, C. P., & Diamond, J. (2007). Origins of major human infectious diseases. *Nature, 447*(7142), 279.

Wolfe, N. D., Escalante, A. A., Karesh, W. B., Kilbourn, A., Spielman, A., & Lal, A. A. (1998). Wild primate populations in emerging infectious disease research: the missing link? *Emerging Infectious Diseases, 4*(2), 149-158. doi:10.3201/eid0402.980202

Woolhouse, M., Scott, F., Hudson, Z., Howey, R., & Chase-Topping, M. (2012). Human viruses: discovery and emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1604), 2864-2871.

Woolhouse, M. E., Howey, R., Gaunt, E., Reilly, L., Chase-Topping, M., & Savill, N. (2008). Temporal trends in the discovery of human viruses. *Proceedings of the Royal Society B: Biological Sciences, 275*(1647), 2111-2115.

Yu, H.-Q., Cai, X.-Q., Lin, Z.-X., Li, X.-L., Yue, Q.-Y., Li, R., & Zhu, X.-Q. (2015). Rapid and specific detection of porcine parvovirus using real-time PCR and High Resolution Melting (HRM) analysis. *BMC Veterinary Research, 11*(1), 46.

APPENDICES

Appendix 1: Supplementary tables

Table S1: Prevalence of SIV infection in AGMs and olive baboons according to species and sampling sites from Mombasa, Naivasha and Kisumu as detected by PCR.

Sampling Town	Sampling Sites	Infected (%) SIV positive/Total sampled NHPs	
		AGMs	**Baboons**
Mombasa			
	Mbaraki	11/23 (48%)	
	Tudor	1/12 (8%)	-
	GAC	5/16 (31%)	-
	Mazeras		1/55 (1%)
	Subtotal	**17/51 (33.3%)**	1/55 (1%)
Kisumu			
	Kajulu	4/21 (19%)	1/10 (10%)
	Golf Course	1/5 (20%)	-
	Impala	7/25 (28%)	-
	Subtotal	**12/51 (23.5%)**	1/10 (10%)
Naivasha			
	Kimana	10/22 (46%)	
Total		**39/124 (32%)**	**2/65 (3.1%)**

Table S2. Percent SIV *pol* amino acid identity matrix generated by multiple sequence alignment using clustal O implemented on the webserver at the European Bioinformatics Institute (EMBL-EBI) (Sievers *et al*., 2011).

	AF395565_SIVagm	M29975.1_SIVagm	AF395564_SIVagm	L40990_SIVagm	AB253736-SIVagmTYO-1	AF395567_SIVagmTYO1	X07805_SIV-AGM	U58991	MG590131_AGM699	MG590128_AGM681	MG590139_AGMNV21	MG590129_AGM682	MG590137_AGMNV4	MG590127_AGM675	MG590130_AGM684	MG590136_AGMNV22	MG590133_AGMKis18	MG590132_AGMKis17	MG590135_AGMMbk11	MG590134_AGMGAC16
AF395565_SIVagm	100	95	95	88	91	90	90	81	89	90	90	90	90	90	91	97	85	85	91	90
M29975.1_SIVagm	95	100	95	91	92	91	92	82	91	93	92	92	92	92	94	97	83	88	93	92
AF395564_SIVagm	95	95	100	92	94	93	93	82	92	94	94	94	94	94	95	94	82	89	91	93
L40990_SIVagm	88	91	92	100	92	91	92	79	90	92	91	92	92	92	93	90	81	88	88	93
AB253736-SIVagmTYO-1	91	92	94	92	100	99	99	82	89	91	90	90	90	90	92	93	80	91	91	94
AF395567_SIVagmTYO1	90	91	93	91	99	100	98	81	88	90	88	89	89	89	90	92	79	90	90	92
X07805_SIV-AGM	90	92	93	92	99	98	100	81	88	90	89	90	90	90	91	92	79	90	90	93
U58991	81	82	82	79	82	81	81	100	81	82	83	82	82	82	81	82	93	82	83	81
MG590131_AGM699	89	91	92	90	89	88	88	81	100	95	95	97	97	97	96	90	82	88	90	91
MG590128_AGM681	90	93	94	92	91	90	90	82	95	100	99	97	97	98	98	93	82	88	90	94
MG590139_AGMNV21	90	92	94	91	90	88	89	83	95	99	100	97	97	98	97	92	82	88	88	92
MG590129_AGM682	90	92	94	92	90	89	90	82	97	97	97	100	100	99	97	91	83	89	89	92
MG590137_AGMNV4	90	92	94	92	90	89	90	82	97	97	97	100	100	99	97	91	83	89	89	92
MG590127_AGM675	90	92	94	92	90	89	90	82	97	98	98	99	99	100	97	92	82	90	89	93
MG590130_AGM684	91	94	95	93	92	90	91	81	96	98	97	97	97	97	100	93	82	87	90	94
MG590136_AGMNV22	97	97	94	90	93	92	92	82	90	93	92	91	91	92	93	100	84	88	92	93
MG590133_AGMKis18	85	83	82	81	80	79	79	93	82	82	82	83	83	82	82	84	100	79	82	80
MG590132_AGMKis17	85	88	89	88	91	90	90	82	88	88	88	89	89	90	87	88	79	100	88	88
MG590135_AGMMbk11	91	93	91	88	91	90	90	83	90	90	88	89	89	89	90	92	82	88	100	88
MG590134_AGMGAC16	90	92	93	93	94	92	93	81	91	94	92	92	92	93	94	93	80	88	88	100

Table S3. Percent SIV *env* amino acid identity matrix generated by multiple sequence alignment using clustal O implemented on the webserver at the European Bioinformatics Institute (EMBL-EBI) (Sievers *et al* ., 2011).

	AB253736_SIVagmTYO-1	AF015909	U93746_babenv05	U93744_babenv03	X07805_SIVagm	MG590109_AGM17Kis	MG590110_AGM4Kis	MG590126_AGM5Kis	MG590112_AGM672	MG590113_AGM675	MG590114_AGM680	MG590115_AGM681	MG590116_AGM682	MG590117_AGM684	MG590118_AGM685	MG590119_AGM686	MG590120_AGM690	MG590121_AGM698	MG590122_AGM699	MG590123_AGM7GAC	MG590124_AGM7Mbk	MG590126_Baboon5
AB253736_SIVagmTYO-1	100	84	78	84	99	86	88	86	81	83	88	89	88	86	89	88	88	82	89	86	82	86
AF015909	84	100	73	79	82	81	83	82	80	80	81	81	84	84	82	84	84	76	85	81	79	83
U93746_babenv05	78	73	100	90	77	78	79	77	73	75	80	79	80	78	80	79	80	74	79	77	75	78
U93744_babenv03	84	79	90	100	83	83	84	83	78	81	85	85	85	84	86	84	85	80	84	83	82	84
X07805_SIVagm	99	82	77	83	100	85	87	85	79	82	87	88	87	85	88	87	87	81	88	85	81	85
MG590109_AGM17Kis	86	81	78	83	85	100	89	88	81	82	84	85	87	86	84	84	87	81	86	84	82	89
MG590110_AGM4Kis	88	83	79	84	87	89	100	88	82	85	88	89	89	87	89	90	94	84	91	86	83	87
MG590126_AGM5Kis	86	82	77	83	85	88	88	100	80	84	86	87	89	87	87	87	88	80	87	84	82	96
MG590112_AGM672	81	80	73	78	79	81	82	80	100	80	82	82	86	87	82	80	81	77	82	80	77	82
MG590113_AGM675	83	80	75	81	82	82	85	84	80	100	87	87	87	87	88	86	87	80	88	85	83	85
MG590114_AGM680	88	81	80	85	87	84	88	86	82	87	100	98	90	89	98	89	92	83	90	88	85	87
MG590115_AGM681	89	81	79	85	88	85	89	87	82	87	98	100	90	89	98	89	92	83	90	89	85	87
MG590116_AGM682	88	84	80	85	87	87	89	89	86	87	90	90	100	97	91	89	90	83	91	88	86	91
MG590117_AGM684	86	84	78	84	85	86	87	87	87	87	89	89	97	100	90	88	89	83	89	87	85	90
MG590118_AGM685	89	82	80	86	88	84	89	87	82	88	98	98	91	90	100	90	93	84	91	88	85	87
MG590119_AGM686	88	84	79	84	87	84	90	87	80	86	89	89	89	88	90	100	93	85	92	87	84	87
MG590120_AGM690	88	84	80	85	87	87	94	88	81	87	92	92	90	89	93	93	100	85	92	87	86	87
MG590121_AGM698	82	76	74	80	81	81	84	80	77	80	83	83	83	83	84	85	85	100	86	81	80	80
MG590122_AGM699	89	85	79	84	88	86	91	87	82	88	90	90	91	89	91	92	92	86	100	88	86	86
MG590123_AGM7GAC	86	81	77	83	85	84	86	84	80	85	88	89	88	87	88	87	87	81	88	100	83	84
MG590124_AGM7Mbk	82	79	75	82	81	82	83	82	77	83	85	85	86	85	85	84	86	80	86	83	100	83
MG590126_Baboon5	86	83	78	84	85	89	87	96	82	85	87	87	91	90	87	87	87	80	86	84	83	100

Table S4. Analysis of the synonymous and non-synonymous substitution rates (dS and dN) for the *env* protein. Analysis was done on codon alignments of the env sequences and the run executed on hyphy platform via data monkey webserver (Delport *et al*., 2010).

Sequence	Compared against	SD	SN	S	N	pS	pN	dS	dN	dS/dN	pS/pN
MG590123 AGMGAC7 (Mombasa)	**MG590111 AGM5 (Kisumu)**	3.67	60.33	181.83	715.17	0.02	0.08	0.02	0.09	0.23	0.24
MG590123 AGMGAC7 (Mombasa)	**MG590132 AGM17(Kisumu)**	7.17	71.83	180.50	716.50	0.04	0.10	0.04	0.11	0.38	0.40
MG590123 AGMGAC7 (Mombasa)	**MG590115 AGM681 (Naivasha)**	6.17	64.83	175.00	698.00	0.04	0.09	0.04	0.10	0.36	0.38
MG590123 AGMGAC7 (Mombasa)	**MG590121 AGM698 (Naivasha)**	5.17	71.83	181.00	716.00	0.03	0.10	0.03	0.11	0.27	0.28
MG590123 AGMGAC7 (Mombasa)	**MG590113 AGM675 (Naivasha)**	202.83	488.17	186.00	693.00	1.09	0.70	nan	2.10	nan	1.55
MG590123 AGMGAC7 (Mombasa)	**MG590114 AGM680 (Naivasha)**	7.67	79.33	180.67	716.33	0.04	0.11	0.04	0.12	0.36	0.38
MG590123 AGMGAC7 (Mombasa)	**MG590116 AGM682 (Naivasha)**	4.50	44.50	182.00	715.00	0.02	0.06	0.03	0.07	0.39	0.40
MG590123 AGMGAC7 (Mombasa)	**MG590117 AGM684 (Naivasha)**	11.50	103.50	169.83	685.17	0.07	0.15	0.07	0.17	0.42	0.45
MG590123 AGMGAC7 (Mombasa)	**MG590118 AGM685 (Naivasha)**	3.33	49.67	180.67	716.33	0.02	0.07	0.02	0.07	0.26	0.27
MG590123 AGMGAC7 (Mombasa)	**MG590119 AGM686 (Naivasha)**	3.83	54.17	181.33	715.67	0.02	0.08	0.02	0.08	0.27	0.28
MG590123 AGMGAC7 (Mombasa)	**MG590120 AGM690 (Naivasha)**	4.17	53.83	182.00	715.00	0.02	0.08	0.02	0.08	0.29	0.30
MG590123 AGMGAC7 (Mombasa)	**MG590122 AGM699 (Naivasha)**	3.67	43.33	181.67	715.33	0.02	0.06	0.02	0.06	0.32	0.33
MG590123 AGMGAC7 (Mombasa)	**MG590126 Baboon5 (Kisumu)**	3.67	61.33	175.67	685.33	0.02	0.09	0.02	0.10	0.22	0.23

Averages of the first sequence (MG590123 AGMGAC7, Mombasa) compared to others:

dS = 0.0309 dN = 0.0957 dS/dN = 0.3149 pS/pN = 0.3286

Table S5: Sample identification for NHPs that were positive for SIV. The accession numbers for respective sequences submitted to NCBI's GenBank database.

Sample ID	GenBank Accession	SIV gene	Sample ID	GenBank Accession	SIV gene
AGMKIS17KAJ	MG590108	*env*	BABOON5	MG590126	*env*
AGMKIS4KAJ	MG590110	*env*	AGMNVS675	MG590127	*pol*
AGMKIS5GC	MG590111	*env*	AGMNVS681	MG590128	*pol*
AGMNVS672	MG590112	*env*	AGMNVS682	MG590129	*pol*
AGMNVS675	MG590113	*env*	AGMNVS684	MG590130	*pol*
AGMNVS680	MG590114	*env*	AGMNVS699	MG590131	*pol*
AGMNVS681	MG590115	*env*	AGMKIS17	MG590132	*pol*
AGMNVS682	MG590116	*env*	AGMKIS18	MG590133	*pol*
AGMNVS684	MG590117	*env*	AGM16GAC	MG590134	*pol*
AGMNVS685	MG590118	*env*	AGM11MBK	MG590135	*pol*
AGMNVS686	MG590119	*env*	AGMNVS22	MG590136	*pol*
AGMNVS690	MG590120	*env*	AGMNVS4	MG590137	*pol*
AGMNVS698	MG590121	*env*	AGMNVS5	MG590138	*pol*
AGMNVS699	MG590122	*env*	AGMNVS21	MG590139	*pol*
AGM17GAC	MG590123	*env*	AGM6GAC	MG757169	*pol*
AGM7MBK	MG590124	*env*	AGM12NVS	MG757170	*pol*
AGMNVS679	MG590125	*env*	AGM13NVS	MG757171	*pol*

Appendix 2: Supplementary figures

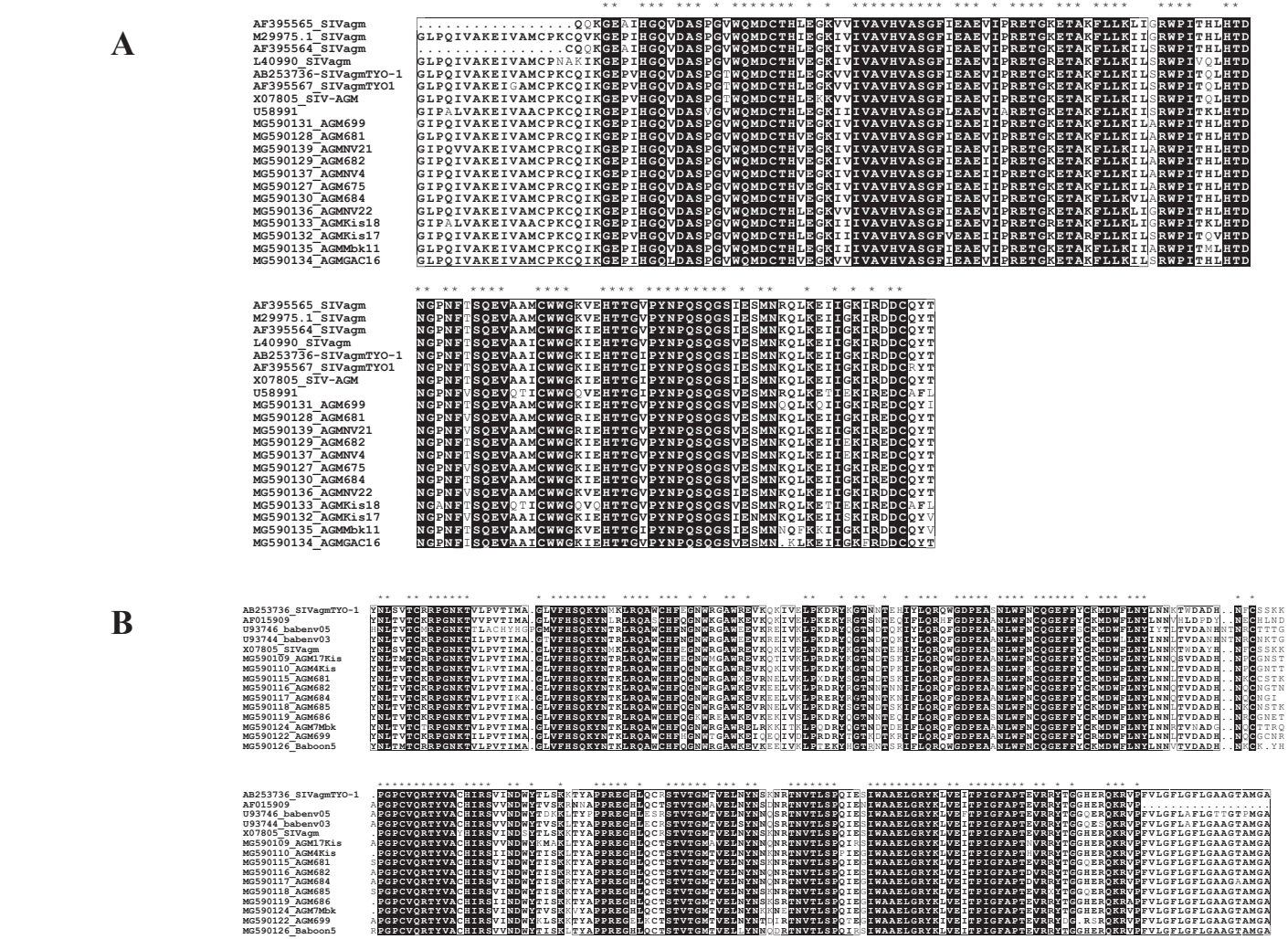


Figure S1: Multiple sequence alignments for the *pol* protein amino acid residues **(A)** and *env* protein amino acid residues **(B)** generated from the partial *pol* and *env* gene fragments amplified in this study alongside SIVagm sequences from GenBank. The conserved residues are highlighted with a dark background and the column highlighted with an asterisks (*). The conservative amino acid replacements can be observed for the non-conserved residues indicating maintenance of the normal protein function.

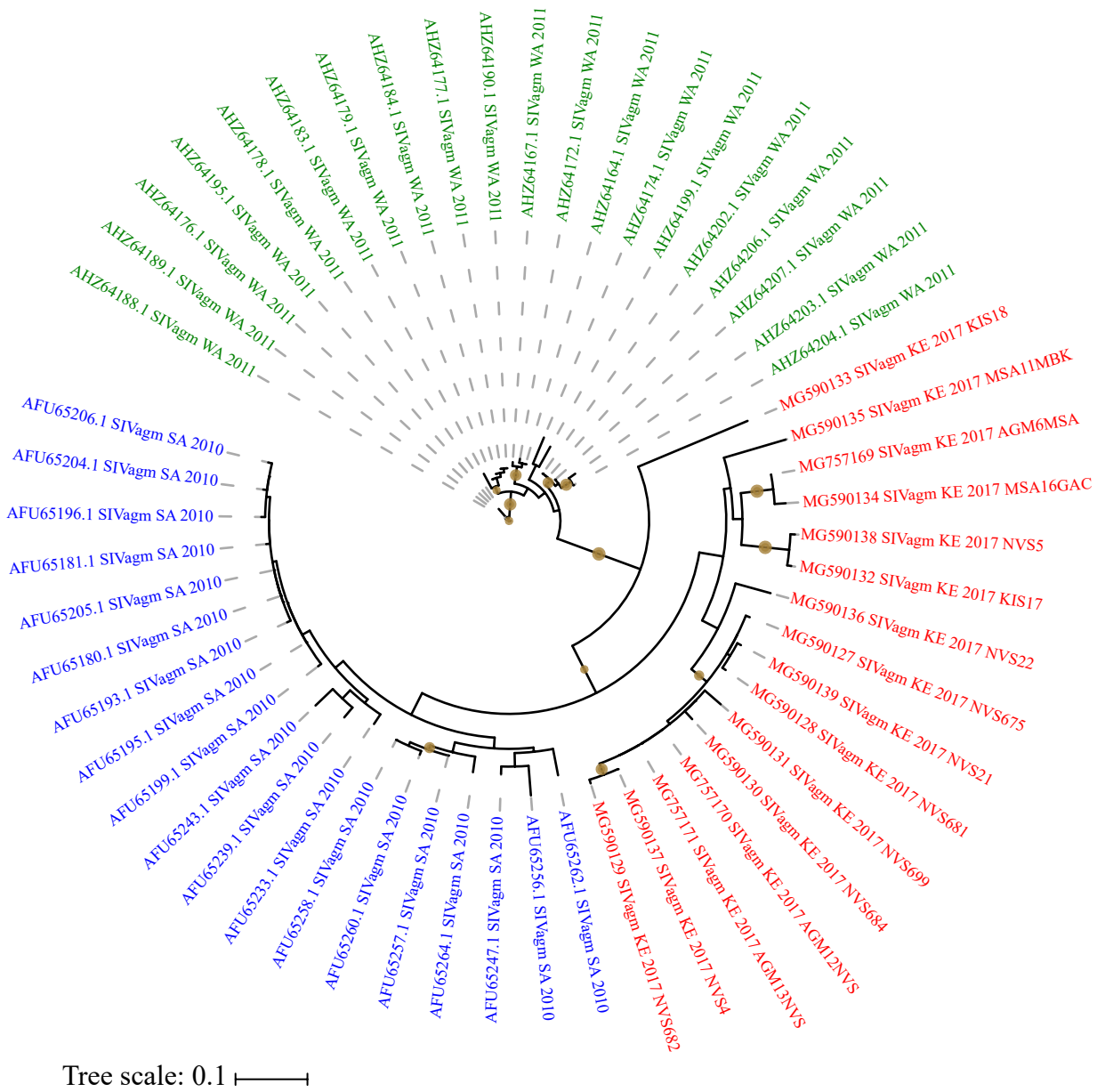


Figure S2: Maximum likelihood tree generated from partial *pol* protein amino acids illustrating genetic diversity of SIV from synanthropic (urban dwelling) AGMs from three urban centres in Kenya. Sequences from Kenya are highlighted in red, South Africa in blue and West Africa in green. Solid brown circles represent bootstrap support values of >70 while the bar represents the number of nucleotide substitutions (0.1) per site. Sequences generated in this study are named according to the monkey species (AGM), the animal number and sampling site.

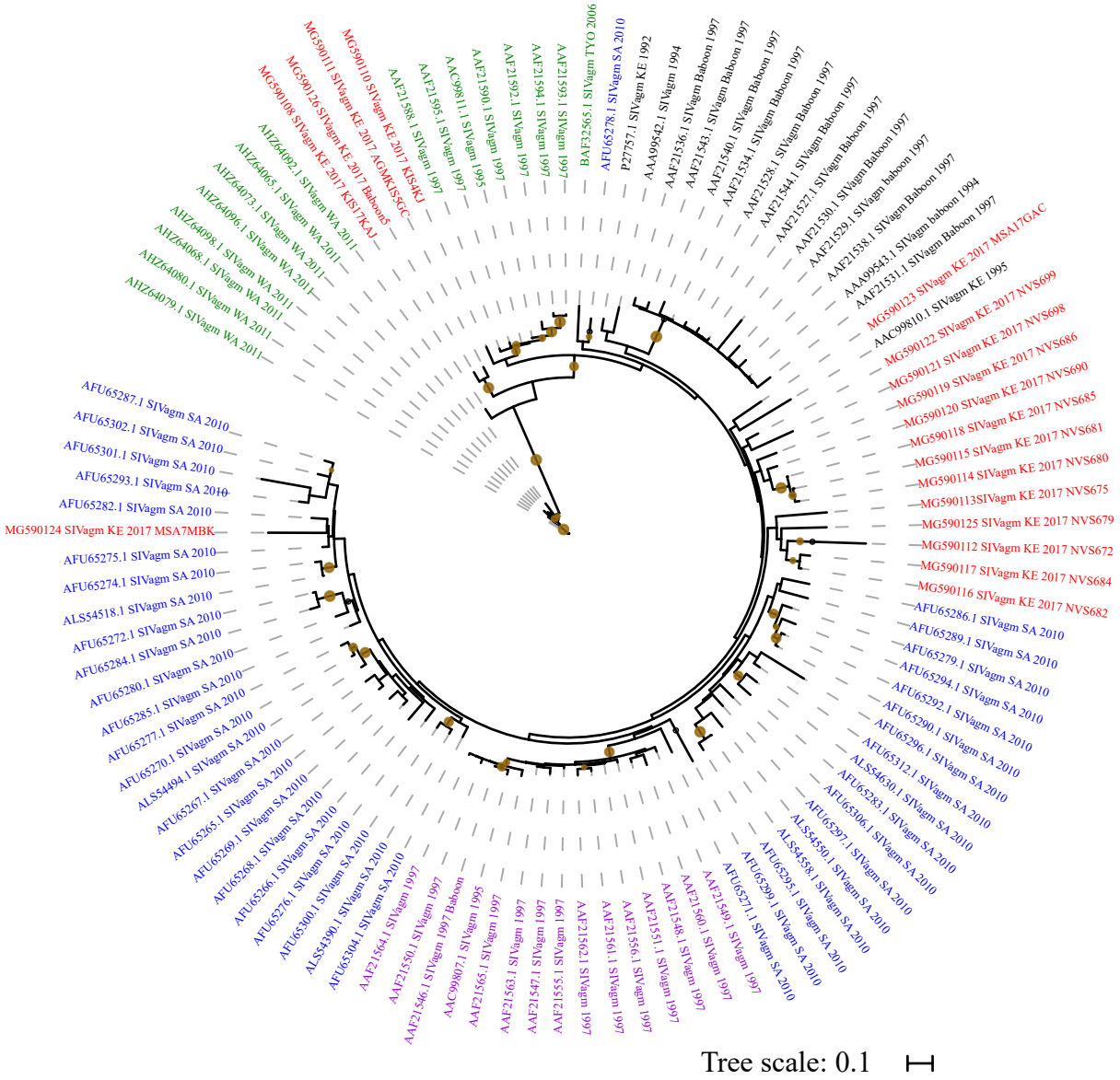


Figure S3: Maximum likelihood tree generated from partial *env* protein amino acids illustrating genetic diversity of SIV from synanthropic (urban dwelling) AGMs from three urban centres in Kenya. Sequences from Kenya are highlighted in red, South Africa in blue and West Africa in green. Sequences in black and those highlighted in purple were deposited in GenBank from captive AGMs and infected baboons in the US that were positive for SIVagm. Solid brown circles represent bootstrap support values of >70 while the bar represents the number of nucleotide substitutions (0.1) per site. Sequences generated in this study are named according to the monkey species (AGM), the animal number and sampling site.

Appendix 3: Publication

Received: 11 October 2019 | Revised: 6 December 2019 | Accepted: 19 January 2020

DOI: 10.1111/jmp.12461

ORIGINAL ARTICLE

WILEY

Broad diversity of simian immunodeficiency virus infecting *Chlorocebus* species (African green monkey) and evidence of cross-species infection in *Papio anubis* (olive baboon) in Kenya

Richard Nyamota[1] | Vincent Owino[1,2] | Edwin Kimathi Murungi[3] | Jandouwe Villinger[2] | Moses Otiende[4] | Daniel Masiga[2] | John Thuita[5] | Isaac Lekolool[4] | Maamun Jeneby[2,6]

[1]Department of Biochemistry and Molecular Biology, Egerton University, Egerton, Kenya

[2]International Centre of Insect Physiology and Ecology (*icipe*), Nairobi, Kenya

[3]Department of Biochemistry, Kisii University, Kisii, Kenya

[4]Kenya Wildlife Services, Nairobi, Kenya

[5]Biotechnology Research Institute, Kenya Agricultural and Livestock Research Organization (BioRI-KALRO), Kikuyu, Kenya

[6]Department of Tropical and Infectious Diseases, Institute of Primate Research, Karen, Kenya

Correspondence
Maamun Jeneby, International Centre of Insect Physiology and Ecology (*icipe*), P.O Box 30772-00100, Nairobi, Kenya.
Email: majeneby@primateresearch.org

Funding information
This study received financial support from the Consortium for National Health Research (CNHR), Kenya (Project number RCDG-041/2012) and the World Federation of Scientists (WFS) through the National Scholarship Programme. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The views expressed herein do not necessarily reflect the official opinion of the donors.

Abstract

Background: Simian immunodeficiency virus (SIV) naturally infects African non-human primates (NHPs) and poses a threat of transmission to humans through hunting and consumption of monkeys as bushmeat. This study investigated the as of yet unknown molecular diversity of SIV in free-ranging *Chlorocebus* species (African green monkeys—AGMs) and *Papio anubis* (olive baboons) within Mombasa, Kisumu and Naivasha urban centres in Kenya.

Methods: We collected blood samples from 124 AGMs and 65 olive baboons in situ, and detected SIV by high-resolution melting analysis and sequencing of PCR products.

Results: Simian immunodeficiency virus prevalence was 32% in AGMs and 3% in baboons. High-resolution melting (HRM) analysis demonstrated distinct melt profiles illustrating virus diversity confirmed by phylogenetic analysis.

Conclusions: There is persistent evolutionary diversification of SIVagm strains in its natural host, AGMs and cross-species infection to olive baboons is occurring. Further study is required to establish pathogenesis of the diverse SIVagm variants and baboon immunological responses.

KEYWORDS

African green monkeys, evolution, high-resolution melting analysis, olive baboon, SIVagm, strain diversity

1 | INTRODUCTION

African non-human primates (NHPs) are reservoirs for diverse zoonotic retroviruses of human health importance including simian immunodeficiency virus (SIV), genus *Lentivirus* and simian foamy virus (SFV)[1] besides being incidental hosts for Ebola[2] and Marburg virus.[3] Diverse strains of SIV infect over 40 species of African NHPs.[4] Although not pathogenic to its natural host, the NHPs, the virus continuously poses a threat of transmission to humans because of hunting and consuming NHPs as bushmeat.[5,6] It has been shown that human immunodeficiency virus (HIV) originated from cross-species transmission of SIVcpz strain from chimpanzees (*Pan troglodytes*) and SIVsmm from sooty mangabeys (*Cercocebus atys*), resulting in zoonotic infections that respectively produced human immunodeficiency virus type 1 (HIV-1) and HIV-2 pandemic.[7,8] Despite the high prevalence of natural SIV infections in African NHPs species,[6,9,10] comprehensive surveys and molecular characterisation of lentiviruses infecting East African monkeys remain limited.

microbiologically documented without bacteremia (MDCD) 2-Clinically documented (CD) 3-Posible infection (PI): no MD no CD.

Results: 134 infectious episodes were diagnosed during 39 episodes of neutropenia in 33 patients, mean age 58(DS ± 18) years, female 78(58%), mean duration of neutropenia 30(DS ± 15) days, median ANC at diagnosis of infectious episode 41 cel/mm^3 (range, 0-962). We found 3,2(DS ± 1,7) infectious attacks by episode of neutropenia.

Among the 134, MD 58(43%): BUO 9(7%), BSI 30(22%), MDCD 19(29%); CD 58(43%) and PI 18(13%). Most common sites of infection: Lung 33(25%), intestinal tract 30(22%), skin and mucosa 20(15%); 39 bacteremias ocurred with 43 isolations: gram negative (GN) 21(49%), gram positive (GP) 18(42%), fungemia 3(7%), others 1(2%). K pneumoniae was the most common GN and coagulase-negative staphylococci between GP. The prevalence of multi drug resistance microorganisms (MDRM) in bacteremia was: 8(38%) in GN, and 6(33%) in GP. Mortality rate was 12%. Duration of neutropenia, median ANC, site of infection, and isolation of MDRM were not associated with mortality. On univariate and multivariate analyses the variables found to be significantly associated with mortality were the presence of bacteremia (OR 8: IC 2,4-28; p = 0,001) and age ≥ 60 years (OR 12: IC 2,56-61,7; p = 0,002)

Conclusion: We found a high rate of bacteremia between infectious episodes, as well as high percentage of MDRM isolations. An age ≥ 60 years and the presence of bacteremia were associated with more probability of death.

https://doi.org/10.1016/j.ijid.2018.04.3617

Final Abstract Number: Fri_Station 02.1
Session: Moderated ePoster Presentations: HIV and Tuberculosis
Date: Friday, March 2, 2018
Time: 12:45-13:45
Room: San Telmo

Type: Electronic/Moderated Poster Presentation

Extensive diversity of SIVagm infecting synanthropic African green monkeys and Olive Baboons in Kenya

R. Nyamota [1,*], O. Vincent [2], M. Edwin [2], O. Moses [3], V. Jandouwe [4], J. Maamun [5]

[1] *Egerton University, Nairobi, Nairobi, Kenya*
[2] *Egerton University, Nairobi, Kenya*
[3] *Kenya Wildlife Service, Nairobi, Kenya*
[4] *International Centre for Insect Physiology and Ecology, Nairobi, Kenya*
[5] *Institute of Primate Research, Nairobi, Kenya*

Background: Diverse strains of simian immunodeficiency virus (SIV) infect over 40 species of African nonhuman primates (NHP) and continuously pose potential threat of transmission to humans. However, strain diversity amongst common free-ranging *Chlorocebus aethiops* (African green monkey-AGMs) and *Papio anubis* (Olive baboon) in Kenya remains unclear. For the first time, we investigated epidemiology and diversity of SIV in AGMs and Baboon from major Kenyan urban centres.

Methods & Materials: A total of 126 AGMs and 65 olive baboons from Mombasa, Kisumu and Naivasha were sampled once *in situ* and released. Blood samples were spotted on whatman FTA cards for DNA/RNA extraction. Total DNA from dried blood on FTA (DB-FTA) was subjected to PCR followed by high resolution melting (HRM) analysis using consensus primers targeting a 650-bp fragment in partial *pol* gene and a 900-bp fragment of *env* gene encompassing hypervariable V3–V5 regions.

Results: PCR-HRM analysis illustrated amplification of SIV fragments giving an overall infection rate of 30.95% (39/126) in AGMs and 3.07% (2/65) in Olive baboons. Subsequent sequence identification confirmed SIV infection in AGM and for the first time in olive baboon. Phylogenetic analysis of *pol* and *env* genes revealed extensive genetic diversity among newly generated SIVagm sequences within groups sympatric NHPs and within a geographical location. Signatures of pervasive and episodic diversifying selection were also detected on the *env* gene indicating continuous SIV diversification.

Conclusion: This molecular evidence of SIVagm in olive baboons illustrates continuous simian-to-simian SIV transmission in the wild and can be linked to potential risks of transmission to humans through consumption of monkey bushmeat. In addition, this study shows that DB-FTA specimens and PCR-HRM analysis can be used as a cost-effective alternative sampling method for the surveillance and monitoring of SIV and other retrovirus of public health importance.

https://doi.org/10.1016/j.ijid.2018.04.3618

Final Abstract Number: Fri_Station 02.2
Session: Moderated ePoster Presentations: HIV and Tuberculosis
Date: Friday, March 2, 2018
Time: 12:45-13:45
Room: San Telmo

Type: Electronic/Moderated Poster Presentation

Functionality of CD8$^+$ T-cells in subjects under cART: implications on cure strategies

J. Salido [1,*], M.J. Ruiz [1], C. Trifone [1], M.I. Figueroa [1], M.P. Caruso [3], M.M. Gherardi [3], O. Sued [1], S. Horacio [3], L. Natalia [1], Y. Ghiglione [3], G. turk [3]

[1] *Instituto de Investigaciones Biomédicas en Retrovirus y SIDA, Ciudad Autónoma de Buenos Aires, Argentina*
[2] *Fundacion Huesped, CABA, Argentina*
[3] *Instituto de Investigaciones Biomédicas en Retrovirus y SIDA, CABA, Argentina*
[4] *Fundacion Huesped, CABA, Ciudad Autónoma de Buenos Aires, Argentina*

Background: Reaching HIV cure will largely depend on the capacity of HIV-specific memory CD8$^+$ T-cells (CD8TC) to eliminate the viral reservoir. However, CD8TC response is limited in subjects on cART. Here, we aimed to investigate the phenotype and function of in vitro expanded CD8TCs in HIV$^+$ subjects and the impact of ART initiation timing on these parameters.

Methods & Materials: PBMCs from 28 HIV$^+$ subjects on cART for 1 year were obtained. Twelve initiated treatment during chronic infection (Delayed Treatment, DT) and 16 within four months post-infection (Early Treatment, ET). PBMCs were stimulated with peptides spanning Nef and Gag plus IL-2 during 14 days. ELISPOT (pre and post-expansion) and Flow Cytometry (FC, post-expansion) were performed to assess expanded CD8TC function (CD107a/b, IFN-g, IL-2, MIP-1b and TNF-a) and phenotype (CD45RO, CCR7, CD95 and PD1). Data was analyzed using non-parametric statistics.

Results: Magnitude of ELISPOT responses increased after expansion by 10^3 times ($p < 0.002$), in both groups, being this effect more pronounced in CD8TCs, compared to CD4TCs ($p < 0.0001$), as confirmed by FC. Cells showed higher avidity after stimulation (evidenced by greater spot sizes, $p < 0.002$). DT subjects displayed a broader response to HIV than ET, after expansion. ET group had a significantly higher proportion of monofunctional

KENYA WILDLIFE SERVICE

ISO 9001:2008 Certified

Our REF:KWS/VET/3003.24

3rd March 2020

The Director
Graduate School,
Egerton University
P.O Box 536-20115,
Egerton.

Dear Prof Gitaka,

REF: MSC PROJECT FOR RICHARD NYAMOTA STUDENT No. SM14/23502/14

Kenya Wildlife Service (KWS) is a state corporation mandated to conserve and manage wildlife in Kenya on behalf of the Kenyan people. One of the key responsibility of KWS is to maintain viable and healthy wildlife population. Through its veterinary department KWS regularly collects wildlife samples during scheduled veterinary interventions, translocations and routine medical health checks in collaboration with other government institutions including IPR under existing material transfer agreements (MTA).

Richard Nyamota in his study entitled ***Prevalence and genetic diversity of simian immune deficiency virus infecting free ranging non-human primates in Kenyan Urban centres*** was part of KWS routine medical checks and did not require any external permits. Further all molecular analysis related to this work were conducted at the Forensic and Genetics laboratory at Kenya Wildlife Services under my direct supervision in collaboration with experts from IPR and ICIPE.

Yours Sincerely

Moses Y. Otiende PhD.
Manager Forensics and Genetics Laboratory,
Kenya Wildlife Services.

KENYA WILDLIFE SERVICE
FORENSIC & GENETICS LABORATORY
P. O. Box 40241 - 00100
NAIROBI

P.O Box 40241-00100, Nairobi, Kenya.
Tel: +254-020-2379407/8/9-15. Mobile: +254-735 663 421, +254-726 610 508/9.
Email: kws@kws.go.ke Website:www.kws.go.ke

Institute of Primate Research
Address: P. O. Box 24481-00502 Karen Nairobi Kenya | Tel: +254 02 2606235 | Fax: +254 02 2606231
URL: www.primateresearch.org | Email: directoripr@primateresearch.org

NATIONAL MUSEUMS OF KENYA
WHERE HERITAGE LIVES ON

INSTITUTIONAL REVIEW COMMITTEE (IRC)

FINAL PROPOSAL APPROVAL FORM

Our ref: **ISERC/04/18**

Dear **Dr. Mercy Akinyi,**

It is my pleasure to inform you that your proposal entitled **"Dynamics of co-infection in wild primates"** has been reviewed by the Institutional Review Committee (IRC) at a meeting of 17th April 2018. The proposal was reviewed on the scientific merit and ethical considerations on the use of animals for research purposes.

The committee is guided by the Institutional guidelines as well as International regulations, including those of WHO, NIH, PVEN and Helsinki Convention on the humane treatment of animals for scientific purposes and GLP.

This proposal has been approved and you are bound by the IPR Intellectual Property Policy.

Signed:.....Chairman IRC: DR. NJALLA JILLANI

INSTITUTE OF PRIMATE RESEARCH
INSTITUTIONAL REVIEW COMMITTEE
P. O. Box 24481-00502 KAREN
NAIROBI - KENYA
APPROVED 23 APRIL 2018

Signed:..... Secretary IRC: For DR. LUCY OCHOLA.

Date: