MOLECULAR EPIDEMIOLOGY AND EVOLUTION OF INFLUENZA A(H1N1)pdm09 VIRUS IN KENYA

GEORGE GACHARA MAINA, BEd(Sc) KU, M.Sc. (KU)

P70/13284/2009

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY (MOLECULAR VIROLOGY) OF KENYATTA UNIVERSITY

NOVEMBER 2013

Declaration

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

George Gachara Maina

Department of Medical Laboratory Sciences,

Kenyatta University

Signature: _____ Date: _____

We confirm that the work reported in this thesis was carried out by the student under our supervision.

Wallace Bulimo, PhD

Chief Virologist, National Influenza Center,

US-Army Medical Research Unit Nairobi, Kenya.

Signature: _____ Date: _____

Michael F. Otieno, PhD

Associate Professor,

Department of Medical Laboratory Sciences,

Kenyatta University

Signature: _____ Date: _____

Dedication

To the late Drs Walter Onalo Ochieng and John Nzyoka Mbithi who were involved in conceptualizing of this study but departed before its conclusion.

Acknowledgements

I would like to express my heartfelt gratitude to Dr. Wallace Bulimo for his exquisite supervision and guidance, not only in my project but also regarding my attitude towards molecular virology and research in general. His felicitous encouragement to develop a scholarly mind has definitely shaped me and shall allow me to be prepared for future challenges. Dr Bulimo will always be a role model for me. I am grateful also to my second supervisor Prof. Michael Otieno. Prof took me under his wing when I lost my initial supervisor and provided me with professional guidance and continuous encouragement throughout my PhD study. I thank him for being so honest and patient with me. I am very pleased to know him and humbled to serve in the same department with him.

Special thanks to Samuel Symekher for his technical assistance. I thank him for sharing his knowledge and for the long periods we spent picking each other's brains. Sam pushed me to learn skills which I never thought I would be able to do on my own.

My thanks go to the National Influenza Laboratory staff all of whom contributed directly or indirectly to my work. These include Japheth Magana, Rachel Achilla, Janet Majanja, Benjamin Opot, Meshack Wadegu, Steve Omondi, Janet Nyambura, Finnley Osuna, Josphat Mwangi and Sylvanos Mukunzi. This team shared their excellent knowledge with me. Through various discussions with them, I had the opportunity to explore and broaden my views and knowledge.

Finally, I take this opportunity to express my heartfelt thanks to my dear family. My wife Lucy, kids Anthony, Abigael, Asha and Andrew - I love you so much!

To God be the glory forever and ever……..Amen

The Fog of War: Uncertainty

Where is the enemy?

What is his strength?

What counterattack?

The Fog of Epidemics: Uncertainty

Where is the microbe?

How many; how virulent; how communicable?

What counterattack?

Perceived Miscalculations

1975 Swine flu outbreak

Response too rapid

1981 HIV/AIDS occurrence

Response too slow

Richard Krausse, 1988

Table of contents

List of Tables

List of Figures

Abbreviations and Acronyms

CDC	Centers for Disease Control
cDNA	Complementary DNA
cRNA	Complementary RNA
CVR	Center for Virus Research
DNA	Deoxyribonucleic acid
HA	Haemagglutinin
KEMRI	Kenya Medical Research Institute
M	Matrix
mRNA	Messenger RNA
NA	Neuraminidase
NCBI	National Center for Biotechnology Information
NEP	Nuclear export protein
NIC	National Influenza Center
NP	Nucleoprotein
NS	Nonstructural protein
PCR	Polymerase chain reaction
PB1	Polymerase basic 1
PB2	Polymerase basic 2
RT-PCR	Reverse transcriptase polymerase chain reaction
rtRT-PCR	Real time reverse transcriptase PCR
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
S-OIV	Swine origin influenza virus
vRNA	Viral RNA
vRNP	Viral ribonucleoprotein
WHO	World Health Organization

Abstract

In April 2009, a novel influenza A virus was detected in Mexico and California. Thereafter, it spread globally to cause the first influenza pandemic of the 21st century. The virus was found to be antigenically unrelated to human seasonal influenza viruses but genetically related to viruses known to circulate in pigs. In view of its likely swine origin, it was initially referred to as 'swine-origin influenza virus' (S-OIV) A/H1N1, or pandemic influenza A (H1N1) 2009 virus but was later renamed by the World Health Organization as influenza A(H1N1)pdm09 virus. In its first year of circulation, the virus is estimated to have caused between 151,700 and 575,400 deaths globally. In Kenya, the virus was first detected in late June 2009 and in the next one year was the dominant virus in circulation. However, no laboratory confirmed deaths occurred in the country. The emergence and subsequent rapid global spread of this influenza virus provided a unique opportunity to observe the evolutionary dynamics of a pandemic influenza virus in Kenya, a tropical region where the virus circulates throughout the year. Understanding the evolution of influenza A(H1N1)pdm09 virus within the country is essential for studying global diversification, the emergence, spread and resistance of the viruses circulating in this region of the world, as well as determining the genetic relationships among the Kenyan strains and vaccine strains. The aim of this study therefore was to reconstruct the evolutionary dynamics of the A(H1N1)pdm09 influenza virus in Kenya during its first year of circulation. To accomplish this, the study undertook whole genome Sanger sequencing of 40 influenza A(H1N1)pdm09 virus isolates sampled nationwide during the pandemic period. To understand the evolutionary dynamics of the local A(H1N1)pdm09 viruses, the study employed the Bayesian evolutionary framework to analyze the resulting 320 individual gene sequences and the 40 complete genomes and compared them with sequences from two African countries, UK, USA and China isolated during the same period. The phylogenetic analyses showed that all of the Kenyan sequences sampled in the pandemic period grouped into at least four highly significant clusters and were interspersed with isolates from other countries. Two global clades (2 and 7) were identifiable within the first two weeks of the pandemic in Kenya, with clade 7 undergoing further diversification while clade 2 was not detected beyond the introductory foci. The time of the most recent common ancestor of the strains circulating in Kenya was estimated to be between April and June 2009, two months before the first laboratory confirmed case. High evolutionary rates and fast population growths was also observed. Progressive drift away from the vaccine strain was observed at both the nucleotide and amino acid level, with 2010 strains clustering separate from 2009 strains. A few unique clusters of amino acid changes were identified among all gene segments in the course of the pandemic, but no mutations previously associated with increased virulence were detected. The local strains were shown to be sensitive to neuraminidase inhibitors but resistant to adamantanes. Overall, results from this study indicate that two clades of influenza A(H1N1)pdm09 virus were introduced in Kenya and that the pandemic was sustained by multiple importations. They also indicate that clade 7 viruses dominated local transmission with an efficient community spread that was devoid of any spatial patterns but a progressive genetic drift was evident. In conclusion, adaptive evolution and viral migration seem to play a vital role in shaping the evolutionary dynamics of local A(H1N1)pdm09 viruses.

1.0: INTRODUCTION

1.1 Background of the study

Influenza is the paradigm of a viral disease in which continued evolution of the virus is of paramount importance for annual epidemics and occasional pandemics of the disease in humans (Webster *et al.,* 1992). Of the 3 types of influenza viruses (A, B, and C), only influenza A viruses are established in animals other than humans (Neumann and Kawaoka, 2006). Influenza A (Family *Orthomyxoviridae*, Genus *Influenzavirus A*) is currently the greatest pandemic disease threat to humankind (Gatherer, 2009). Its rivals for this title (HIV-1, Ebola, SARS, and pneumonic plague) have higher mortality if untreated, but either lack influenza's rapid inter-personal transmission (HIV-1) or its widespread seasonal distribution (Ebola, SARS, pneumonic plague).

Although influenza A viruses are best known for the annual epidemics and occasional pandemics in humans, wild birds form their natural reservoir (Olsen *et al.,* 2006; Webster *et al.,* 1992). Whereas only 3 influenza A virus subtypes circulated in humans in the past century, all known subtypes have been isolated from wild birds. Numerous subtypes have also been detected in poultry, generally causing severe disease, mild disease or no disease at all (de Wit and Fouchier, 2008). Several mammals including pigs and horses are also known to be naturally infected by influenza A viruses. Pigs are susceptible to all subtypes of avian influenza viruses in experimental settings (Horimoto and Kawaoka, 2001) and this has led to their label as a 'mixing vessel' allowing for two-way transmission of viruses between birds and humans. Occasionally,

a new subtype of influenza A virus is introduced into the human population from swine or birds that is able to spread efficiently from human to human, causing a pandemic. This process is called antigenic shift and may be the result of a reassortment event of avian and human influenza A viruses (Scholtissek *et al.*, 1978), or of adaptation of a fully avian virus to humans (Claas and Osterhaus, 1998; Tumpey *et al.*, 2005).

During the past century, three such pandemics occurred. The first of these is the so called 'Spanish influenza' which occurred in 1918 and which is considered as the greatest natural disaster of the 20th century. This pandemic occurred as a result of the introduction of an influenza A virus of the H1N1 subtype probably derived from an unidentified avian-like precursor virus which became adapted to mammals (Morens *et al.*, 2009). An estimated 50 million people died as a result (Patterson and Pyle, 1991). The second such pandemic occurred in 1957 due to the introduction of the H2N2 subtype causing the so called 'Asian influenza'. This pandemic virus emerged by reassortment between the viruses in humans at that time, and the virus from some aquatic avian reservoir killing an estimated 2 million people world wide (Webster, 1993). The last pandemic of the 20th century occurred in 1968 and was due to an influenza A H3N2 virus subtype derived from both human and avian viruses. This pandemic was relatively mild (de Wit and Fouchier, 2008).

In March and early April 2009, a new swine-origin influenza A (H1N1) virus (S-OIV) emerged in Mexico and the United States (CDC, 2009c). Since then, the virus (later renamed influenza A(H1N1)pdm09 virus) spread rapidly across

the world by human-to-human transmission. On the 11th of June 2009, the World Health Organization declared a global pandemic of 2009 H1N1 infection making it the first pandemic of the 21st century (WHO, 2009). By April 2010, there were 18,500 laboratory confirmed deaths with recent studies estimating a higher figure of 201,200 deaths (Viboud and Simonsen, 2012). The outbreak strain was identified as a swine origin influenza virus that resulted from a reassortment of two previously circulating strains. The new strain contains six segments (PB2, PB1, PA, HA, NP, and NS) that were similar to the ones previously found in triple-reassortant swine influenza viruses circulating in pigs in North America (Dawood *et al.,* 2009). The PB1 gene was itself seeded in swine from humans (Garten *et al.,* 2009) and two segments (NA and M) from the Eurasian lineage (Dawood *et al.,* 2009). The amino acid sequence divergence of the 2009 pandemic H1 from human seasonal influenza H1 is around 20–24% and for this reason, it is being considered as halfway towards a new serotype (Gatherer, 2009).

The severity during the early 2009 pandemic was estimated to be less than that seen in the 1918 influenza pandemic and comparable to that seen in the 1957 pandemic. However, there were concerns that due to the rapid evolutionary rate of influenza viruses, new variants could emerge leading to increased disease severity. Additionally, all the 20th century pandemics have occurred in successive waves.

1.2 Problem statement

Estimates of case fatality rates of influenza A(H1N1)pdm09 virus have only been available from high-income locations but vary three to nine times between studies. Using a probabilistic multiplier approach, the estimated numbers of deaths associated with pandemic influenza A H1N1 was 201 200 (range 105 700—395 600) for respiratory diseases and an additional 83 300 (46 000—179 900) for cardiovascular diseases in the first year of virus circulation (Dawood *et al.,* 2012). This approach suggests that the highest mortality occurred in Africa.

Complete influenza virus genomes provide critical and deeper understanding of influenza genetic structure and provide insight into effective control options. Sequencing of the whole influenza virus genome facilitates comparison and understanding of the evolutionary dynamics of circulating viruses and the prediction of potential evolution events that are likely to result in new strains (Greninger *et al.,* 2010). It also allows closer examination of the role played by all genes in influenza outbreaks and vaccine selection.

In Sub-Saharan Africa, there is a paucity of complete influenza genomes in the public domain. Most influenza genome sequencing in these countries has focused on the HA1 domain of the hemagglutinin gene, where mutations usually have the greatest effect on antigenic structure (Webster *et al.,* 1992). As such, it has not been possible to understand the contribution of Africa in the global ecology of influenza viruses. It is therefore clear that, complete understanding of influenza metapopulation dynamics is incomplete without

data from these regions (Viboud *et al.*, 2013). Complete influenza genome analysis of influenza A (H1N1)pdm09 viruses from Africa have been at best scanty. The genomic signatures and characteristics of local 2009 pandemic H1N1 viruses are therefore not well understood.

1.3 Justification of the study

The 2009 influenza pandemic caused by influenza A (H1N1)pdm09 virus presented a unique opportunity to integrate molecular and epidemiological analyses of global and regional spread patterns and epidemic growth dynamics. Such methods were instrumental in tracking how the pandemic unfolded across the globe and in confirming the Mexican origin of the most recent common viral ancestor in January–February 2009 (Lemey *et al.,* 2009). In Africa and Kenya in particular, there has been limited studies that has examined the molecular and epidemiological characteristics of influenza A (H1N1)pdm09 virus. This has impeded the complete understanding of this virus in the region.

Molecular study of Kenyan strains of influenza A (H1N1)pdm09 virus is anticipated therefore to give a preview of the direction of evolution of this virus. This is imperative so as to monitor the genetic make-up of this virus, to understand its adaptability and evolutionary dynamics in the country. This is essential for studying global diversification, the emergence, spread and drug resistance of influenza A (H1N1)pdm09 virus strains circulating in the country. It is also of importance in determining the genetic relationships

among Kenyan strains and vaccine strains included in the recommended influenza vaccines. This may provide insights into disease trajectory for future influenza pandemic planning and could contribute to improving public health in Kenya. At the same time drug susceptibility of circulating influenza viruses is of utmost necessity particularly in a pandemic situation.

1.4 Research questions

1. Are there genetic differences between the Kenyan influenza (H1N1)pdm09 viruses and others isolated elsewhere in the world?
2. What is the evolutionary rate and time of most recent ancestor of the local circulating influenza A(H1N1)pdm09 viruses?
3. Did genetic drift occur amongst the Kenyan influenza (H1N1)pdm09 viruses and if so, did it increase the virulence of pandemic H1N1 strains?
4. Were local pandemic strains sensitive to anti-influenza drugs that are in current use?

1.5 Objectives of the study

1.5.1 General objective

To describe the molecular epidemiology and evolution of the influenza A(H1N1)pdm09 virus during the pandemic period in Kenya.

1.5.2 Specific objectives

a) To isolate influenza A(H1N1)pdm09 viruses identified in the country and amplify their complete genomes using *in vitro* and molecular techniques.

b) To determine full genome nucleotide sequences and infer phylogenetic relationships of the influenza A(H1N1)pdm09 viruses isolated in Kenya during the pandemic period relative to vaccine/reference strains using bioinformatics tools.

c) To determine the evolutionary rate and ancestral time analyses of the Kenyan influenza A(H1N1)pdm09 viruses using Bayesian methods.

d) To describe the genetic drift occurring in the eight genomic segments of influenza A(H1N1)pdm09 viruses in Kenya during the pandemic by using bioinformatics tools..

e) To determine neuraminidase inhibitors sensitivity and adamantane sensitivity in the isolated strains by examining known genotypic markers.

1.6 Significance of the study

Understanding the evolution of influenza A(H1N1)pdm09 virus within the country is essential for studying global diversification, the emergence, spread and resistance of influenza A(H1N1)pdm09 strains circulating in this region of the world. The current study lays the groundwork for future influenza pandemic and epidemic studies in the country. It is also of importance in determining the genetic relationships among Kenyan strains

and vaccine strains included in the influenza vaccine recommended for the Southern Hemisphere.

2.0: LITERATURE REVIEW

2.1 Classification and nomenclature

The influenza viruses together with the Thogoto-like viruses (Thogoto, Batken and Dhori viruses) belong to the family *Orthomyxoviridae* (Cox, 2000). *Myxo* is the Greek word for mucus, which means that members of the family have a strong affinity for and posses an enzyme capable of removing chemical side chains from mucoproteins. These properties facilitate infection of cells of the mucous membranes in the respiratory tract (Metselaar and Simpson, 1982). The family *Orthomyxoviridae* is divided into five genera; *Influenzavirus* A, *Influenzavirus* B, *Influenzavirus* C, *Isavirus* and *Thogotovirus*, based on antigenic differences in two of the major structural proteins of the virus, the nucleoprotein (NP) and the matrix protein (M). Influenza A viruses are further classified into 16HA subtypes and 9NA subtypes based on the properties of their surface major membrane glycoproteins (Luke and Subbarao, 2006).

Two glycoproteins on the surface of the virus particle, the Neuraminidase or NA protein and the Haemagglutinin or HA (Haemagglutinin-esterase in influenza C) protein are involved in the interaction between virus and host cells. Influenza A viruses are classified into subtypes based on the antigenic differences between these two glycoproteins. Currently there are 16 (H1-H16) distinct HA subtypes and 9 (N1-N9) NA subtypes (Luke and Subbarao, 2006) though several new subtypes have recently been described (Tong *et al.*, 2013). Historically however, influenza A virus subtypes infecting humans have been limited to H1, H2, and H3 and in the past 100 years to N1 and N2. In recent

years, H5N1, H7N7 and H9N2 have sporadically infected humans and continue to pose a pandemic threat due to their possible interspecies transmission into humans. There is only one HA and one NA subtype that has been identified among type B influenza viruses (Nicholson *et al.,* 2003).

The World Health Organization (WHO) system of nomenclature includes the host of origin, geographical origin, strain number and year of isolation; then follows in parentheses the antigenic description of the haemagglutinin and the neuraminidase, e.g., A/Swine/Iowa/3/70(H1N1) was isolated from swine in Iowa in 1970. If from a human host, the origin is not given, e.g., A/Scotland/42/89(H3N2) was isolated from a human host in 1989 in Scotland (Nguyen-Van-Tam and Hampson, 2003).

2.2 Influenza A virion structure

Influenza A virions are spherical, 80-120nm in diameter, but may be filamentous, sometimes up to several micrometres in length as illustrated in Figure 1. Two distinct types of spikes (approximately 16nm in length), corresponding to the HA and NA molecules, reside on the surface of the virions. The HA spike appears rod shaped and protrudes from the envelope as a trimer (Wilson *et al.,* 1981); the NA spike is a mushroom shaped tetramer (Colman *et al.,* 1987). These two glycoproteins are anchored to the lipid enveloped derived from the plasma membrane of host cells by short sequences of hydrophobic amino acids (the transmembrane region).

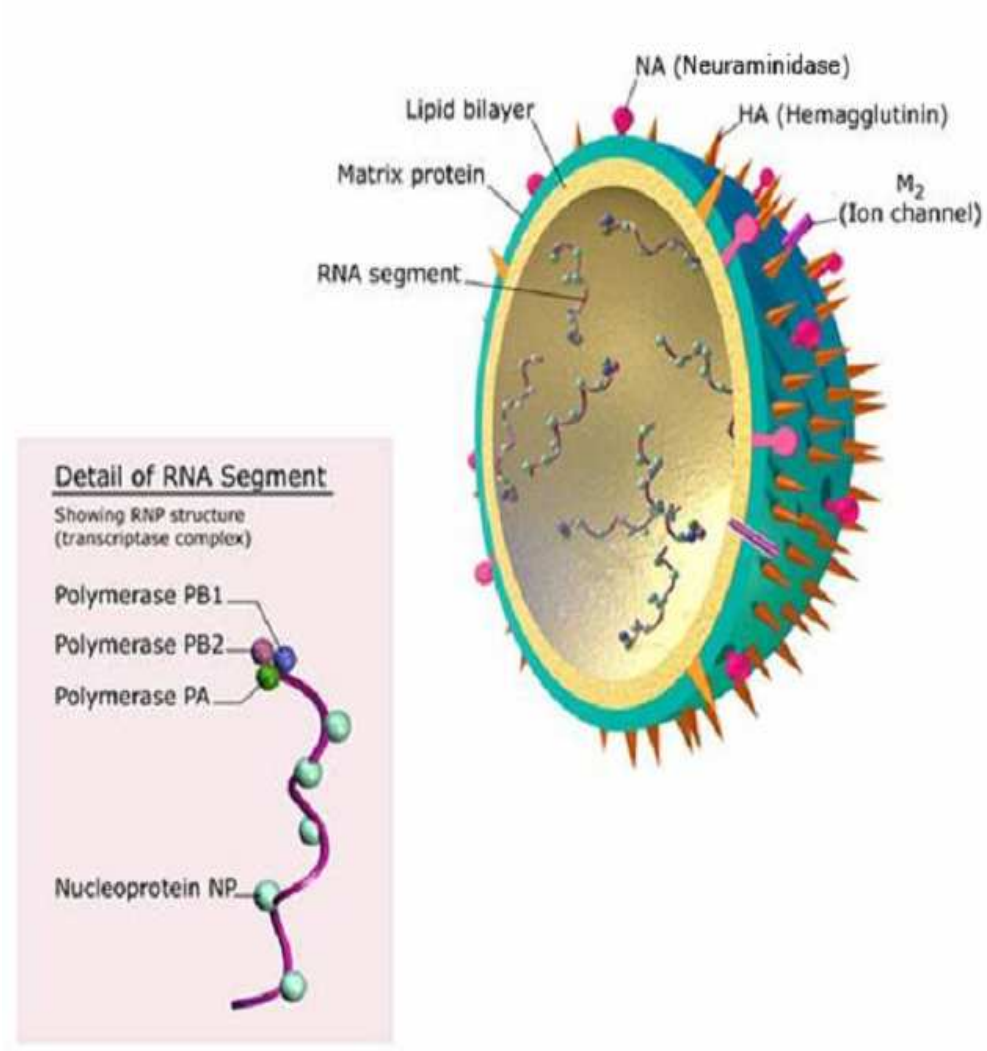


Figure 1: Influenza virion

Source: http://agrolink.moa.my/jph/epv/kertaskerja01.html

In addition to HA and NA, a limited number of M2 proteins are integrated into influenza A virions. The mRNA for M2 is transcribed from RNA segment 7 and is an intergral membrane protein whose membrane-spanning domain also serves as a signal for transport to the cell surface (Webster *et al.,* 1992). M2 proteins form tetramers and have H+ ion channel activity. When activated by the low pH in endosomes, they acidify the inside of the virion, facilitating its uncoating (Pinto *et al.,* 1992). M1 protein that lies within the envelope is also derived from the splicing of RNA segment 7. It is the most abundant protein in the influenza virus virion (Webster *et al.,* 1992). M1 is thought to play an important role in assembly and budding.

Eight segments of single stranded RNA molecules (negative sense, or complementary to mRNA) are contained within the viral envelope, in association with NP and three subunits of viral polymerase (PB1, PB2, and PA), which together form a ribonucleoprotein (RNP) complex that participates in RNA replication and transcription. NS2 and NS1 proteins are encoded in the RNA segment 8. NS2 protein is thought to play the role of export of RNP from the nucleus through interaction with M1 protein (Horimoto and Kawaoka, 2001). NS1 protein, the only nonstructural protein of influenza viruses, has multiple functions, including regulation of splicing and nuclear export of cellular mRNAs as well as stimulation of translation. Its major function seems to be to counteract the interferon activity of the host (Hale *et al.,* 2008). Also present in the virion is the viral RNA-dependent RNA polymerase PB1 and PB2 encoded by RNA segment 2 and 1 respectively. These two are important

for infectivity since the virion RNA is negative sense and therefore has to be transcribed to produce viral messenger RNA (mRNA). Nucleoprotein (NP) is encoded by RNA segment 5. It is transported into infected cell nucleus where it encapsidates viral RNA. In addition to its structural role, it is believed to play a role in the switching of viral RNA polymerase activity from mRNA synthesis to cRNA and vRNA synthesis (Webster *et al.,* 1992).

2.3 Genomic organization of Influenza A virus

The influenza A virus genome contains eight negative sense RNA segments (Palese, 1977). The viral mRNAs from segments 1 to 6 are monocistronic while viral mRNAs derived from segments 7 or 8 are spliced to form mRNAs coding for two proteins (Lamb and Krug, 2001). The sizes of the viral RNA segments and the proteins encoded are summarized in Table 1. Of these proteins, only the NS1 protein from segment 8 (NS segment) is a non-structural protein.

Table 1: Influenza A virus genome RNA segments

vRNA Segment	Encoded polypeptide	Abbreviation	vRNA length (bps)	mRNA length (bps)
1	Polymerase Basic 2	PB2	2341	2320
2	Polymerase Basic 1	PB1	2341	2320
3	Polymerase Acid	PA	2233	2211
4	Haemagglutinin	HA	1778	1757
5	Nucleoprotein	NP	1565	1540
6	Neuraminidase	NA	1413	1392
7	Matrix	MP	1027	1005 (M1) 315 (M2)
8	Nonstructural Protein	NS	890	868 (NS1) 395(NS2/NEP)

Table adapted from (Lamb and Krug, 2001)

2.3.1 Segment 1 – Basic polymerase protein 2 (PB2)

The first segment of influenza A viruses encodes a 2.3Kb protein that that forms part of the influenza viral polymerase subunit named PB2. Studies have shown that PB2, PB1, PA and NP form the minimum set of proteins required for viral transcription and replication (Honda *et al.,* 2002; Perales and Ortin, 1997). Like the other polymerase proteins, PB2 contains a nuclear localization signal (Mukaigawa and Nayak, 1991; Perales *et al.,* 1996) and is transported into the nucleus of infected cells for viral transcription and replication (Jones *et al.,* 1986). PB2 is an important protein for generating the cap structure for viral mRNAs. Studies of the influenza viral polymerase have demonstrated that the PB2 subunit is a cap binding protein (Braam *et al.,* 1983). The PB2 protein specifically plays a role in generating 5′-capped RNA fragments from cellular pre-mRNA molecules that are used as primers for viral transcription (Guilligay *et al.,* 2008).

Several studies indicate that PB2, PB1 and PA interact to form a polymerase complex for viral transcription and replication. Immunoprecipitation assays on influenza viral polymerase have shown that PB2 is associated with the PB1 subunit (Digard *et al.,* 1989). Specifically, analysis of deletion mutants of PB2 indicate that the amino-terminus of this protein is a binding site for PB1 (Toyoda *et al.,* 1996). Recently, functional analysis of PB2 protein has shown that this polymerase subunit contains a novel binding site for PB1 subunit and two regions for binding nucleoprotein (NP) with regulatory interactions potential (Poole *et al.,* 2004).

The PB2 gene is reported to play a key role in the adaptation of a virus that infects birds to one that infects humans. In particular, it has been reported that the amino acid at position 627 is critical for this adaptation (Van Hoeven *et al.,* 2009). Viruses that infect birds typically have a glutamic acid (E) in this position. On the other hand, influenza A viruses that are fully adapted to humans usually have a lysine (K) at this position. This same position in PB2 appears to be key in determining the lethality of flu viruses (Hatta *et al.,* 2001). In the 1918 pandemic and in H5N1 viruses, a lysine in this position was associated with a higher level of lethality. The presence of a lysine at position 627 appears to permit flu viruses to replicate in both the lungs and nose and thus spread more easily from person to person.

2.3.2 Segment 2 – Basic polymerase protein 1 (PB1)

The PB1 RNA polymerase subunit of influenza viruses is encoded by segment 2. Several lines of evidence have indicated that PB1, itself, is an RNA polymerase. The central location of the polymerase domain is predicted from observing the presence of conserved motifs characteristic of segmented negative-strand RNA-dependent polymerases (Poch *et al.,* 1989) and that mutations in these motifs abolished the polymerase activity.

Several studies have described the functional domains of PB1 involved in interaction with the other polymerase subunits. Immunoprecipitation studies of the influenza virus RNA polymerase indicate that PB1 contains independent

binding sites for PB2 and PA (Digard *et al.,* 1989) while deletion mutant analyses of PB1 suggest that the amino- and carboxyl-termini of PB1 are binding sites for the PA and PB2 polymerase subunits, respectively (Gonzalez *et al.,* 1996; Toyoda *et al.,* 1996). The nuclear localization signal of PB1 has been mapped to a region near the amino-terminus (Nath and Nayak, 1990). The PB1 subunit plays a key role in both the assembly of the three polymerase protein subunits and serves the catalytic function of RNA polymerization. It has been proposed that the catalytic specificity of PB1 subunit is modulated to the transcriptase by binding PB2 or the replicase by interaction with PA (Honda *et al.,* 2002).

2.3.3 Segment 3 – Acidic polymerase protein (PA)

The PA protein is encoded by segment 3 and is the smallest subunit of the influenza RNA polymerase complex. Like the other influenza viral polymerase subunits, it contains nuclear localization signals (Nieto *et al.,* 1994) required for transport into the nucleus (Jones *et al.,* 1986). The PA subunit has no significant homology to other proteins, and for a long time, its function was unclear (Perales and Ortin, 1997; Webster *et al.,* 1992). Various functions were proposed including helicase and ATP binding activities (de la Luna *et al.,* 1989). It has also been suggested to be a protease (Sanz-Ezquerro *et al.,* 1996) but this property is not related to any known viral function.

Recent studies however have shown that PA is separable by trypsinization into a large carboxy-terminal domain—the crystal structure of which has recently been reported (Obayashi *et al.,* 2008) and a small N-terminal domain, which contains residues important for protein stability, promoter binding, cap-binding and endonuclease activity of the polymerase complex (Hara *et al.,* 2006). The viral endonuclease activity which is critical for synthesizing viral messenger mRNA's have previously been thought to reside in the PB2 (Shi *et al.,* 1995) or PB1 (Li *et al.,* 2001) subunits. However, recent biochemical and structural studies have shown that the amino-terminal 209 residues of the PA subunit contain the endonuclease active site and not PB2 (Dias *et al.,* 2009). As such, during transcription, the PB2 subunit binds the 5′,7-methylguanosine cap of a host pre-mRNA molecule, which is subsequently cleaved 10–15 nucleotides downstream by the PA endonuclease. The resulting short capped RNA primer is used to initiate polymerization by the RNA-dependent RNA polymerase of the PB1 subunit using 5′- and 3′-bound vRNA as template, resulting in capped, polyadenylated, chimeric mRNA molecules that are exported to the cytoplasm for translation into viral proteins (Boivin *et al.,* 2010).

2.3.4 Segment 4 – Haemagglutinin (HA)

HA plays an essential role in the early stages of infection and is responsible for the virus binding to its receptor, sialic acid, which is present on the host cell surface and promotes fusion of viral and endosomal membranes and eventually facilitates viral entry into the host cell (Webster *et al.,* 1992).

During infection and vaccination, HA elicits neutralizing antibodies. It is therefore considered the most important target of antibody-mediated protection (Steinhauer and Wharton, 1998).

The HA is synthesized as a precursor polypeptide, HA0 (Lamb and Krug, 2001; Webster *et al.,* 1992). This precursor polypeptide is post-translationally cleaved into two disulphide-linked subunits, HA1 and HA2. The cleavage of the HA0 is a prerequisite for viral infectivity. This process liberates the "fusion peptide" at the amino-terminus of HA2 required for membrane fusion. In addition, this cleavage also allows the native HA molecule to undergo a conformational change, a process which is triggered by an acidic environment and is essential for membrane fusion (Skehel *et al.,* 1982). In general, the HA0 is believed to be cleaved by trypsin-like proteases extracellularly. However, the presence of multiple basic amino acid residues within the cleavage site allow the protein to be cleaved by intracellular proteases, e.g. furin (Horimoto *et al.,* 1994), which are ubiquitously expressed in most tissues. Hence, influenza viruses containing HA with multiple basic amino acids near the cleavage site are often highly infectious and can infect a wide range of cells. The generated HA1 surface subunit mediates the binding to cell surface sialic acid receptors and the HA2 transmembrane subunit that mediates membrane fusion between viral and endosomal membranes after endocytosis (Skehel and Wiley, 2000)

The specificity of the interaction of HA with sialic acid (SIA), the cellular receptor, largely explains the host range of influenza A viruses (Taubenberger and Kash, 2010). Thus, viruses that infect humans bind preferentially to SIA

linked to the penultimate galactose via an α 2–6 configuration, whereas avian viruses prefer binding to SIA with α 2–3 linkages (Connor *et al.,* 1994). The HA receptor binding site (RBS) is formed by three structural elements at the tip of the HA molecule, an α-helix composed of residues 190–198 (the 190-helix) and two loop structures formed by residues 133–138 (the 130-loop) and 220–229 (the 220-loop). Four conserved residues, comprising Tyr^{98}, Trp^{153}, His^{183} and Tyr^{195}, form the base of the RBS (Skehel and Wiley, 2000). The amino acid residues in the RBS that are critical for the recognition of either avian or human receptors have been well characterized (Connor *et al.,* 1994; Matrosovich *et al.,* 2000). For H1, glutamic acid and glycine residues at positions 190 and 225, respectively, result in binding to avian SIA receptors, whereas H1 proteins that carry aspartic acid residues at these positions interact with human SIA receptors. For H2 and H3, mutations of glutamine and glycine residues at positions 226 and 228 to leucine and serine, respectively, correlate with a shift from avian to human receptor specificity (Matrosovich *et al.,* 2000). The same mutations also allow binding of H5 to human SIA receptors (Stevens *et al.,* 2006).

2.3.5 Segment 5 – Nucleoprotein (NP)

NP protein which is an essential component for transcription and replication is encoded by segment 5 of influenza A viruses. NP is distinct among influenza A, B, and C viruses and is recognized as one of the type-specific viral antigens. It encodes a protein with approximately 500 amino acids; it plays an important

role in assembly and budding of influenza virus and has a putative role in host range (Ruigrok *et al.,* 2010; Snyder *et al.,* 1987). The primary function of NP is to form oligomers and bind with the viral RNA segments to form the nucleocapsid of a virus particle playing a pivotal role in the viral genome transcription machinery (Perales and Ortin, 1997).

Like the other influenza viral polymerase subunits, it contains nuclear localization signals and has been shown to be important for vRNA nuclear transport (Whittaker *et al.,* 1996). During the early stage of viral infection, the transport of incoming vRNPs from the viral particle into the nucleus is believed to be mediated by NP whereas, in the late infection stage, progeny vRNAs associated with NP, M1 and NS2 are exported to the cytoplasm for viral packaging (Whittaker *et al.,* 1996). To achieve its biological functions, recent studies have revealed that the NP is capable of interacting with various host proteins (Taubenberger and Kash, 2010; Wang *et al.,* 2009).

2.3.6 Segment 6 – Neuraminidase (NA)

The three-dimensional structure of the NA has revealed that the NA monomer is a homotetramer (Hausmann *et al.,* 1997). It consists of a box-shaped globular head, a thin stalk, a transmembrane domain and a cytoplasmic domain (Varghese and Colman, 1991). The NA is a surface glycoprotein and the glycosylation of the NA might be an important determinant (but not the sole determinant) of the neurovirulence of influenza viruses (Li *et al.,* 1993). It has

receptor-destroying activity to cleave the α–ketosidic linkage between a terminal sialic acid and an adjacent D-galactose or D-galactosamine residue (Colman, 1998).

The NA is a type II intergral membrane protein. It is a sialidase that possesses an enzymatically active domain which cleaves sialic acids from host cells, leading to the release of progeny virions and preventing the aggregation of virion during the budding process (Palese and Schulman, 1976). Strong binding to receptor analogs on mucins, cilia, and cellular epithelia would inhibit virus access to functional receptors on surface membrane of target cells. Therefore, NA is important for releasing viruses from the decoy receptor and plays an essential role of virus entry in the early stage of infection (Couceiro *et al.*, 1993; Matrosovich *et al.*, 2004). It is for this reason that NA is a favorable target for antiviral drugs aiming at reducing its sialidase enzymatic activity and thus release of progeny virions (Suzuki *et al.*, 2005).

2.3.7 Segment 7 – Matrix proteins (M1 and M2)

The segment 7 of influenza A virus encodes two proteins, M1 and M2, by differential splicing of mRNA. The M1 protein is the product of collinear transcript of the mRNA while differential splicing of the M gene RNA yields the M2 protein (Lamb and Choppin, 1981). M1 is recognized as another type specific antigen and constitutes the most abundant polypeptide in the virion which provides rigidity to the viral membrane and functions in the viral

assembly and budding during viral life cycle (Lamb and Choppin, 1981). It is found to interact with the viral surface glycoproteins and the RNP complex (Taubenberger and Kash, 2010).

The M2 is an integral membrane protein (Lamb and Choppin, 1981) which forms an ion channel on the viral envelope. The M2 ion channel is known to control the pH across the Golgi apparatus during viral maturation which is essential for virus uncoating in the virus infection process. In the endosome of infected cells, the ion channel activity of M2 allows acidification of the interior of the incoming viral particle to enable uncoating. The acidification of the viral particle is believed to be essential for viral replication, because it allows incoming vRNPs to dissociate from M1 proteins for nuclear import (Martin and Helenius, 1991). The ion channel activity of M2 is also reported to maintain a high pH in the Golgi vesicles so as to stabilize the native conformation of newly synthesized HA during the intracellular transport for viral assembly (Takeuchi and Lamb, 1994). The M2 protein is also a target for anti-influenza drugs called amantadines.

2.3.8 Segment 8 – Non-structural proteins (NS1 and NS2)

The eighth vRNA segment of the influenza A virus directs the synthesis of two mRNAs. The first of these encodes the non-structural (NS) protein, NS1, while the other is derived from splicing of the NS1 mRNA is translated into a protein that localizes to the cell nucleus and which was originally named NS2

(Nemeroff *et al.,* 1992) but which has now been renamed the nuclear export protein (NEP) (O'Neill *et al.,* 1998). NS1 is the only non structural protein of influenza virus and is found mainly in the host cell nucleus. It is a multifunctional protein involved in nuclear exportation of mRNA, posttranscriptional regulation, and inhibition of cellular interferon response (Hale *et al.,* 2008). Recently, NS1 protein was found to be responsible for the unusual severity of H5N1 diseases by inducing exaggerated proinflammatory cytokine responses (Cheung *et al.,* 2002).

NS2 protein exists in low amounts and binds to M1 protein in virion. The protein may promote the formation of a stable export complex of new viral RNP (Neumann *et al.,* 2000). In association with the matrix protein 1 (M1), it interacts with cellular export factor (CEF1) and mediates the nuclear export of viral ribonucleoprotein (vRNP) complexes by connecting the cellular export machinery with vRNPs (Neumann *et al.,* 2000).

2.4 Replication cycle of Influenza A viruses

2.4.1 Attachment, endocytosis and uncoating of influenza A virus

To infect host cells, the influenza viruses utilize their surface glycoprotein, HA, to bind to the sialic acid receptors which present on the host cell surface (Fig 2). The HA protein exists as precursor, the HA0, which comprises the HA1 and HA2 peptides. The receptor-binding site is found on the HA1 while the HA2 houses the fusion peptide (Webster *et al.,* 1992). The HA molecules of

influenza viruses from different species have different specificity for sialic acid receptors. Influenza viruses from humans recognize α-(2, 6) linked sialic acid receptors and viruses from avian species preferably bind to the-(2, 3) linked receptors. In pigs both types are present (Greenwood, 2003) and for this reason they are considered as “mixing vessels” for avian and human influenza viruses.

After attachment of the viral haemagglutinin to the specific receptors, the viruses enter cells by receptor-mediated endocytosis in clathrin-coated vesicles. The HA contains a fusion peptide which at a low pH, undergoes a conformational change which exposes the peptide. The low pH is achieved by the opening up of M2 ion channel leading to the influx of H+ ions. The acidic environment results in a conformational change of the HA in which the HA2 fusion peptide is exposed causing the fusion between viral and endosomal membranes, which facilitates the subsequent release of vRNP into the cytoplasm of host cells. The M1 protein is also dissociated from the vRNP under the acidic pH (Lennette, 1995).

2.4.2 Nuclear import of viral ribonucleoproteins in host cells

The replication and transcription of the viral genome takes place inside the nucleus of host cells (Webster *et al.*, 1992). It is therefore necessary to import the vRNP to the nucleus of infected cells to start the genome replication and transcription. The virus uses an active transport machinery to transport vRNP into the nucleus of host cells (Engelhardt and Fodor, 2006). Each of the subunit

of vRNP complex contains nuclear localization signals (NLS) which are recognized by importin α, also called karyopherin α (Watanabe *et al.,* 2010). Likewise, many host factors are reported to be involved in the active transport of vRNP into the nucleus. RanBP5 is reported to interact with the PB1 subunit and facilitate transport of vRNP (Deng *et al.,* 2006).

2.4.3 Viral RNA replication and transcription

Since the genome of the influenza A virus is organized in a negative sense manner, the viral RNAs have to be converted into positive sense before they can be transcribed into viral mRNAs for subsequent translation. Despite the fact that only 11 viral proteins are encoded, the influenza A viruses have evolved a sophisticated mechanism enabling itself to “hijack” host cell machinery for its viral RNA and mRNA synthesis (Nagata *et al.,* 2008).

The mechanism of viral RNA transcription is unique. The 5' cap from cellular mRNAs is cleaved by a viral endonuclease and used as a primer for transcription by the viral transcriptase. Six of eight RNA segments are transcribed into mRNAs in a monocistronic manner and translated into HA, NA, NP, PB1, PB2, and PA. By contrast, two RNA segments are each transcribed to two mRNAs by splicing. For both M and NS genes, these mRNAs are translated in different reading frames, generating M1 and M2 proteins and NS1 and NS2 proteins, respectively (Webster *et al.,* 1992).

The replication of viral RNA comprises two stages. Firstly, it involves the synthesis of full length viral RNAs which serve as template RNAs. These are positive sense complementary RNA (cRNA). Afterwards, exact copies of viral RNAs are generated by using the cRNAs as templates. This process is facilitated by the NP protein through interacting itself with the cRNAs during the production of nascent vRNA molecules (Engelhardt and Fodor, 2006).

2.4.4 Viral ribonucleoprotein export from nucleus to cytoplasm in host cells

The newly synthesized viral mRNAs are then exported to the cytoplasm of the infected cells in order for them to be translated into viral proteins. Since the viral mRNAs are incorporated with the polymerase complexes to form vRNPs, viral mRNAs are exported to the cytoplasm with the aid of the nuclear export machinery found on vRNPs (Ma *et al.*, 2001). The nuclear export of vRNPs is mediated through the cellular Crm1/Exportin pathway and two viral proteins, M1 and NS2 have been reported involving in the nuclear export of vRNPs (Watanabe *et al.*, 2010).

The M1 protein mediates the nuclear export by forming complex with vRNP in the nucleus of host cells. Subsequently, the NS2 protein, which is also named as nuclear export protein (NEP) for the presence of nuclear export signal (NES) in the protein, would associate with the M1-vRNP complex and the cellular Crm1 to mediate the export of vRNP to the cytoplasm (Ma *et al.*, 2001)

2.4.5 Virus assembly and budding

Once the vRNPs and viral proteins are synthesized, they will be localized around the basal membrane, and the viral progeny particles are ready to be packaged and bud off from the host cell. Generally, a viral core of vRNPs becomes encapsidated in a layer of M1 proteins and buds outward through the cell membrane, enclosing itself within a bubble of membrane as its own envelope, complete with the viral surface glycoproteins (Webster *et al.,* 1992). Interactions between M1 and the cytoplasmic domains of HA, NA or M2 have been proposed as signals for budding and NA activity of progeny virions releases them from the host cell (Webster *et al.*, 1992).

When the budding process is complete, HA still binds to the sialic acid on the cell surface until virus particles are released by the sialidase activity of NA protein which is anchored to the viral envelope by a transmembrane domain (Colman *et al.,* 1983; Varghese *et al.,* 1983). NA cleaves the terminal sialic acid residues from cell-surface glycoproteins and gangliosides to release progeny virus from the host cell. NA could also remove sialic acid residues from the virus envelope itself, preventing viral particles from self aggregation, in order to enhance the viral infectivity (Palese *et al.,* 1974).

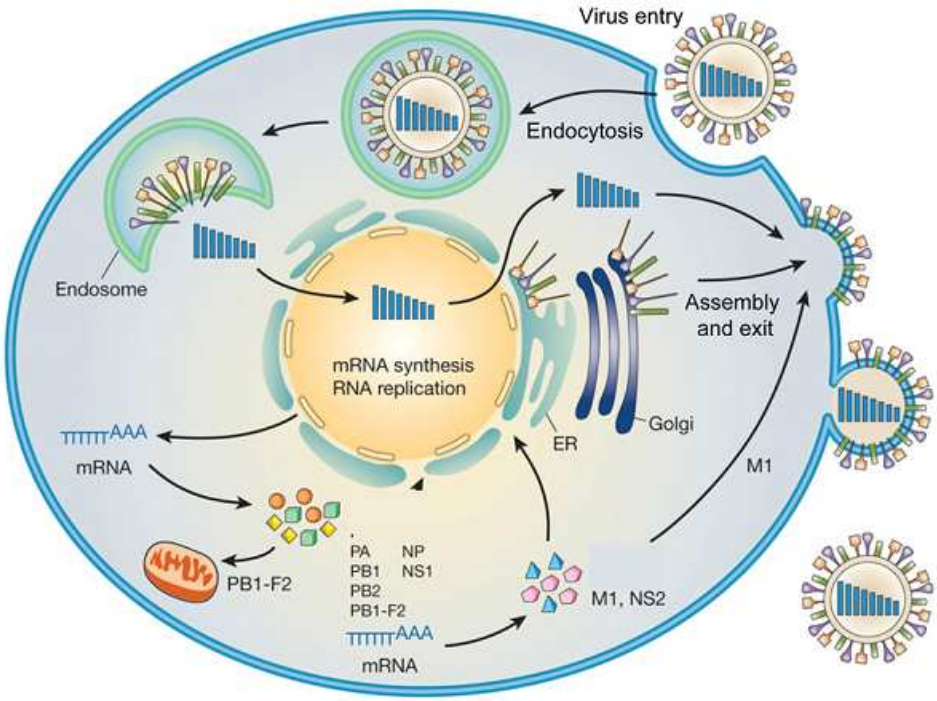


Figure 2: An illustration of influenza A virus replication.
(Neumann *et al.*, 2009)

2.5 Laboratory diagnosis of Influenza

Diagnosis of influenza is important in order to prescribe the correct antivirals, for immunoprophylaxis, and for epidemiological reasons. Successful diagnosis depends on the timing and quality of specimens collected. Clinical specimens should be taken early, preferably within 3 days of onset of symptoms since this is the time maximum virus shedding occurs. Nasopharyngeal or throat swabs combined or alone, nasopharyngeal aspirates and nasal washes are optimal specimens for virus culture and for the direct detection of viral antigens or nucleic acids. Bronchoalveolar lavage fluid and trancheal aspirate specimens may be considered if clinically warranted during serious respiratory infections.

The specimens must be placed in a suitable transport medium to stabilize the infectivity of the influenza viruses. Tryptose phosphate broth, veal infusion broth, commercially available cell culture medium or Hanks balanced salt solution (HBSS), sucrose phosphate buffer, all supplemented with 0.5% bovine serum albumin or 0.1% gelatin, antibiotics and anti-mycotics are appropriate transport media (Johnson, 1990). Virus infectivity is relatively well preserved if the specimens are stored at 4^0C in these liquids for up to 4 days. After collection, specimens should be transported to the laboratory without delay and preferably refrigerated at -70^0C at all times until processing for virus isolation or nucleic acid detection (Cox and Ziegler, 2003). Multiple freeze-thawing of samples should be avoided since they greatly reduce the virus yield.

2.5.1 Influenza virus isolation

Although many cells have been shown to support the replication of influenza viruses, primary monkey kidney cells (PMK), and continous cell lines such as Madin Darby Canine Kidney (MDCK) and Rhesus monkey kidney (LLC-MK2) cell lines are the most widely used. Of these, MDCK are the most commonly used (Bhatia and Ichhpujani, 1999). MDCK support the growth of influenza A, B, and C viruses when 1 to 2 μg of trypsin per ml is added to the maintenance medium to sustain multiple cycles of multiplication (Meguro *et al.*, 1979). For standard virus isolation, cell culture tubes or flasks are seeded with MDCK cells in growth medium containing 5 to 10% fetal bovine serum. The tubes are incubated at 37^0C, and before the cells are completely confluent,

growth medium is removed, the cell sheet is washed with HBSS or serum-free medium to remove nonspecific inhibitors that are present in serum and then serum-free maintenance medium is added. For inoculation, the medium is removed, 100 to 300µl of specimen added and allowed to adsorb for one hour at 37^0C. The inoculum is aspirated and replaced with maintenance medium containing trypsin. Majority of influenza viruses destroy the cell sheet within a few days after inoculation. However, in the absence of cytopathic effect (CPE), cultures should be tested by haemadsorption or by haemagglutination at 2 to 3 day intervals after 6 or 7 days (Cox and Ziegler, 2003).

Isolation in embryonated eggs has been in the earlier years the standard method of virus recovery from clinical specimens. Embryonated eggs are inoculated in the allantoic and amnionic cavities with clinical specimens and the fluid collected from these cavities is tested for viral agglutination after three-day incubation (Bhatia and Ichhpujani, 1999). It has been found that 10 to 11 day old embryonated eggs are optimal for recovery of influenza A and B viruses while 7 to 8 day old embryonated eggs are preferred for the recovery of influenza C viruses. After inoculation, the eggs are incubated at 33 to 34^0C for 2 to 3 days in case of influenza A and B viruses and approximately 5 days for influenza C virus. After one or two passages, most influenza A and B viruses will grow efficiently in the allantoic cavity but influenza C grows only in the amnionic cavity. The presence of influenza viruses in egg fluids is detected by the haemagglutination test (Cox and Ziegler, 2003).

2.5.2 Haemagglutination inhibition (HAI) test

Typing, subtyping, and further antigenic characterization of influenza viruses that have been isolated is usually done by the HAI test. Influenza viruses generally agglutinate human type O, guinea pig, or chicken and turkey erythrocytes. When these red blood cells are mixed with influenza virus in the appropriate ratio, the virus bridges the red blood cells resulting in agglutination of the cells (haemagglutination) and a change in their settling behaviour. Antibodies specific to the viral haemagglutinin interfere with the haemagglutinating activity and this is the basis of the HAI test which allows the identification and the differentiation of the variant strains that frequently appear (Young, 2009).

2.5.3 Fluorescent - antibody staining

Detection of infected, exfoliated epithelial cells in respiratory specimens is a rapid and sensitive method for laboratory diagnosis of respiratory virus infections. Epithelial cells are washed free of mucus, fixed to microscope slides, and stained with specific, well-characterized monoclonal antibodies. When specimens are collected properly and refrigerated at all times, epithelial cells are handled carefully, and the slides are read by an experienced microscopist with a good fluorescent microscope, the sensitivity of this technique is 80 to 90% of that for standard virus isolation (Cox and Ziegler, 2003).

2.5.4 Immunoassays

Several laboratories have developed sensitive and specific radioimmunoassays, enzyme immunoassays (EIAs), and fluoroimmunoassays for the detection of respiratory virus antigens in clinical samples. These assays can produce a result within a few hours, but these tests have limited sensitivity and specificity compared with standard virus isolation. Rapid EIAs for influenza A and B which produce result in 15 to 20 minutes are commercially available. These are useful in clinical laboratories of hospitals and nursing homes for early detection of infection to help patient management as well as detection of nosocomial influenza outbreaks (Cox and Ziegler, 2003). Presence of false positive and false negative results is the major draw backs in the use of rapid tests.

2.5.5 Molecular methods

Molecular methods are now being applied more widely in the diagnosis of influenza virus infection as well as characterization of influenza virus isolates. These methods are likely to supplant virus culture as the "gold standard" for virus detection because they are more sensitive and fast. The genetic analysis of a large number of influenza viruses isolated throughout the world each year provides information for a timely update of vaccine strains and of molecular reagents for diagnosis. For reverse transcriptase polymerase chain reaction (RT-PCR), viral nucleic acids are extracted from clinical specimens, allantoic fluid of embryonated hen eggs, or cell culture material. Complementary DNA

(cDNA) is synthesized by *in vitro* reverse transcription of viral RNA primed either by synthetic oligonucleotides matching known nucleotide sequences on the viral genes or by random hexamers. This cDNA is amplified with specific primers and DNA polymerase. Finally the amplified product is detected by agarose gel electrophoresis in conventional RT-PCR or fluorescence dye in real time PCR. The first report of the use of RT-PCR amplification of influenza RNA to diagnose influenza was published in 1991 (Zhang and Evans, 1991). The superiority of RT-PCR over virus culture has thereafter been reported in several studies (Boivin *et al.*, 2001; Zambon and Hayden, 2001).

2.5.6 Serological tests

Influenza virus infections can also be detected by measuring a fourfold or greater increase of specific antibody titers between serum samples collected at least 10 days apart during the acute and convalescent phases of illness. Thus, while serological methods are accurate, they rarely produce a result that would allow patient treatment with antivirals or timely prophylaxis of contacts. However, serodiagnosis is useful if virus detection and isolation fail and for surveillance and epidemiological studies. The complement fixation test, HAI, the neutralization test (NT), and EIA are the most widely used tests in serodiagnosis and seroepidemiological studies (Cox and Ziegler, 2003). Complement fixation detects type-specific antibodies to NP but is relatively insensitive in detecting diagnostic rises in titer between acute and convalescent phase sera. HAI and NT are type, subtype and strain specific.

2.6 Epidemiology

Influenza occurs throughout the world and affects every person in society. All influenza A virus subtypes naturally infect shorebirds and waterfowl. Influenza infections in these hosts are generally assymptomatic and are limited to the gastrointestinal and/or the respiratory tract. In addition, poultry, pigs, maritime mammals, horses, and occasionally other mammals maybe infected by certain subtypes of influenza A viruses (Cox and Ziegler, 2003). Interspecies transmission of influenza A viruses occurs and may result in severe human illness, as the 1997 outbreak in Hong Kong caused by an H5N1 virus of avian origin amply demonstrated. H7 and H9 subtypes from birds have also been isolated in humans. Influenza B viruses appear to naturally infect only humans and cause annual seasonal influenza infections and epidemics; however, influenza B virus infection of seals in Netherlands has been reported (Osterhaus *et al.*, 2000). Influenza C viruses infect humans and have also been isolated from pigs (Guo *et al.*, 1983). Since the re-emergence of influenza A (H1N1) in 1977, this and A (H3N2) subtype viruses have co-circulated in seasonal influenza illness with influenza B in varying proportions and levels. Molecular evidence exists for the occasional infection of humans with avian viruses containing H5, H7 and H9 genes (Gatherer, 2009).

2.7 Transmission and pathogenesis

Transmission of influenza in humans probably involves both respiratory infection by aerosols and droplets together with some contact transmission

from contaminated surfaces. Unfortunately, there are insufficient solid data to determine which of these is the more important and to provide guidance in protection of front-line medical staff, particularly regarding the required porosity of facemasks, during a pandemic. Nevertheless, influenza virus is relatively short-lived on most surfaces, and clear examples of aerosol transmission have been described in the literature. In addition, early experiments indicated greater infectivity by small particle aerosol than by nasal instillation of virus.

For many years, it was thought that for the efficient introduction of "novel" influenza viruses to humans, pigs were required to act as an intermediate host, or "mixing vessel". Since pigs are susceptible to both human and avian influenza viruses, the genomes of such viruses could reassort and/or adapt to this mammalian host, giving rise to new viruses with altered phenotypes that could more easily be transmitted to humans. Indeed, the pandemics of 1957 and 1968 were both caused by reassortant influenza viruses that may have been generated in pigs (Fouchier *et al.,* 2003). Moreover, several incidents of transmission of influenza viruses from pigs to humans have been described in the past decades. However, since 1996, the viruses H7N7, H5N1 and H9N2 have been transmitted directly from birds to humans but have apparently failed to spread from human to human in the human population. Such incidents are rare but demonstrate the potential of avian viruses to be transmitted directly to humans. It cannot be excluded that upon dual infection of humans with avian

and human influenza A viruses, man himself may act as a "mixing vessel" for the generation of novel human-avian reassortant viruses.

After transmission to pigs, horses, humans or birds, the method of spread of influenza within the population is mainly inhalation of respiratory droplets containing the virus. Less frequently, the virus is spread by person to person contact or contact with contaminated items. Pathogenicity of influenza viruses is multifaceted and may involve viral, host and environmental factors. Inhaled virus is deposited on the mucous membranes lining the respiratory tract or directly to the alveoli, the level depending on the size of the droplets inhaled. In the former state, it is exposed to mucoproteins containing sialic acid that can bind to the virus, thus blocking its attachment to the respiratory tract epithelial cells (Metselaar and Simpson, 1982). However, the action of neuraminidase allows the virus to break this bond. Specific local secretory IgA antibodies, if present from a previous infection, may neutralize the virus before attachment occurs, provided the antibody corresponds to the infecting virus type. If not prevented by one of these mechanisms, the virus attaches to the surface of respiratory epithelial cells and the intracellular replication cycle is initiated.

The major site of infection is the ciliated columnar epithelial cell. The first alteration is the disappearance of the elongated form of these cells which becomes round and swollen, the nucleus shrinks, becomes pyknotic and fragments. Vacuolization of the cytoplasm may occur. As the nucleus disintegrates, the cytoplasm shows inclusion bodies and the cilia are lost. Release of virus from the cell allows it to spread via the mucus blanket to other

areas of the respiratory tract. The cell damage initiates an acute inflammatory response with oedema and the attraction of phagocytic cells. The earliest response is the synthesis and release of interferons from the infected cells; these can diffuse to and protect both adjacent and more distant cells before the virus arrives. It appears that interferons released in this way cause many of the systemic features of the “flu like” syndrome that characterize the infection.

While viral components are absorbed and trigger the immune system, the virus itself is confined to the epithelium of the respiratory tract. Specific antibody will help to limit the extra cellular spread of the virus, while T cell responses are directed against the viral glycoproteins on the surface of infected cells leading to their destruction by cytotoxic T cells and also by antibody dependent cell cytotoxicity (Greenwood, 2003).

2.8 Antigenic variation

The surface antigens of influenza viruses undergo two types of variations necessitating the replacement of vaccine strains every several years. The first kind of variation occurs as a result of accumulation of point mutations in the surface antigens driven by the immune response; this is referred to as *antigenic drift* (Metselaar and Simpson, 1982). Point mutations, substitutions in one or a few of the individual amino acids making up the H protein, occur randomly as the virus is copied in infected cells. The immune response takes care of some variants in the haemagglutinin, but variants emerge that are not neutralized by

the circulating antibodies that infected hosts make in response to the original infection and so the virions go on accumulating point mutations. Antigenic (genetic) drifts are the main cause of the annual influenza epidemics. Over a period of time and because the haemagglutinins of the new variants remain antigenically distantly relative to the original representative of the subtype, immunity in the population to the antigen broadens generated by the new variants, eventually making the epidemic variants less and less able to cause severe disease.

The second kind of variation called *antigenic shift*, has only been observed in influenza A viruses. This has caused all the influenza pandemics of the 20th century, including the 1918"Spanish flu", the Asian flu (1957) and Hong Kong flu (1968). Antigenic shift is caused either by direct transmission of non-human influenza viruses to humans or the reassortment of genes from different influenza viruses that have infected a single cell. Theoretically, 256 different combinations of RNA are produced from the shuffling of the eight different genomic segments of the virus. Genomic reassortment is well documented both *in vitro* and *in vivo* under laboratory conditions (Webster and Laver, 1975). More importantly, mixed infections occur relatively frequently in nature and can lead to genetic reassortment. The result of this is the appearance and continued circulation in the human population of a new subtype of influenza A, of which either the haemagglutinin or the haemagglutinin and the neuraminidase differ completely or to a large extent in antigenic composition from the peplomers of the subtype that was circulating in the preceding years

(Metselaar and Simpson, 1982) as shown in Figure 3 below.

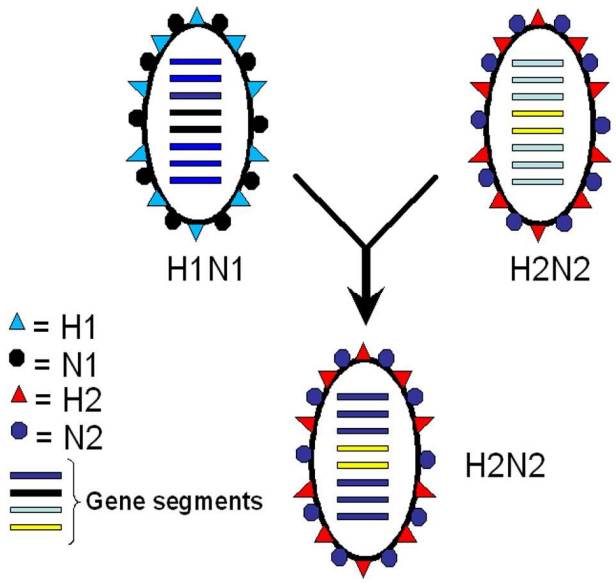


Figure 3: Mechanism of reassortment in influenza A viruses
(From: http://homepage.usask.ca/~vim458/advirol/SPCV/evolution/4.jpg)

2.9 Antivirals used in treatment and control of influenza

WHO estimates that seasonal influenza epidemics result in three to five million cases of severe illness and 250 000 to 500 000 deaths each year in the industrialized world alone (Stohr, 2003). Although vaccination remains the most important measure for reducing this sizeable public health burden, antiviral drugs have been welcomed as long awaited tools for treatment and prevention. Antiviral drugs are effective in treatment as well as prophylaxis, but in prophylaxis they must be administered daily throughout the period of risk of exposure. For this reason, they may be the only intervention available

during the early months of a pandemic when supplies of vaccines maybe severely limited (Monto, 2003). Currently, two classes of influenza antivirals are available to manage influenza. Each class inhibits a different step in the viral replication cycle. Type A but not type B influenza viruses contain an M2 protein. This protein is responsible for uncoating the viral nucleoprotein during replication. It functions as an ion channel, preventing exposure of the viral haemagglutinin to low intracellular pH to which it is sensitive. Amantadine and rimantidine inhibit this activity and they are termed 'M2 inhibitors'. These drugs have no effect on type B influenza viruses because influenza B viruses do not posses an M2 protein (Monto, 2003).

The viral neuraminidase of both type A and B viruses facilitates the release of virus from the infected cell after replication is complete and prevents the viral clumping before the next infectious cycle begins. Zanamivir (Relenza) and Oseltamivir (Tamiflu) are the “old” neuraminidase inhibitors (Nis) that have been widely used and continue to be used in influenza treatment (Barik, 2012). They have been specifically designed to interrupt the replication cycle by preventing virus release and allowing virus to clump. Zanamivir is administered by inhalation, while oseltamivir is taken orally. Both classes of influenza antivirals are approximately 70-90% efficacious when used as prophylaxis (Monto, 2003). Tamiflu has to be given within 48 hours of onset of clinical signs for maximum effect and is less effective if given thereafter. Recently, three new neuraminidase inhibitors have been developed and are in various stages of development. These include Laninamivir which has been

approved for use in Japan, Favipiravir and Peramivir. Both Favipiravir and Peramivir are in various stages of clinical trials (Barik, 2012).

2.9.1 Resistance to antiviral drugs

From the licensure of the neuraminidase inhibitors oseltamivir and zanamivir in 1999 until recently, resistance to these agents had remained at a low level, even in the countries responsible for most of their use worldwide (Monto, 2006). Therapeutic use of anti influenza drugs is frequently associated with emergence of drug resistant variants. Resistance to M2 inhibitors is generally conferred by emergence of single amino acid substitutions in the trans-membrane domain of the M2 protein at positions 26, 27, 30 or 31 (Hay, 1996). Common mutations associated with resistance include V27G, A/I27S/T, S31N and A30T. Resistance to neuraminidase inhibitors typically results from single amino acid changes in the NA protein active enzyme site that alters drug binding. Resistance to Oseltamivir is conferred by I117V, E119V, D198N, I222V, H274Y, R292K, N294S and I314V (N2 numbering) mutations. Double mutations with synergistic oseltamivir resistance phenotype have been noted as well. This includes the E119V+I222V double mutant, isolated from an immunocompromised child infected with H3N2 virus, and H247Y+I222V, from patients infected with influenza A(H1N1)pdm09 virus (Barik, 2012). Mutations related to zanamivir resistance include, V116A, R118K, E119G/A/D, Q136K, D151E, R152K, R224K, E276D, R292K and R371K (N2 numbering).

Initial testing of the influenza A(H1N1)pdm09 virus found it susceptible to neuraminidase inhibitors (CDC, 2009b). However, sporadic cases of resistance to the neuraminidase inhibitors are increasingly being reported some with novel mutations besides the common H275Y mutation (CDC, 2009b; Hurt *et al.,* 2009). The pandemic virus contains the M2 inhibitors resistance-conferring change S31N in the M2 protein and is thus resistant to the M2 inhibitor drugs (Dawood *et al.,* 2009; Garten *et al.,* 2009).

2.10 Influenza A(H1N1)pdm09 virus

Influenza A(H1N1)pdm09 virus has been described as a triple reassortant virus. The virus contains a combination of gene segments that previously have not been reported in swine or human influenza viruses anywhere in the world (Garten *et al.,* 2009). The NA and M gene segments are from the Eurasian swine genetic lineage. Viruses with NA and M gene segments in this lineage were originally derived from a wholly avian influenza virus and thought to have entered the Eurasian swine population in 1979 (Pensaert *et al.,* 1981), continue to circulate throughout Eurasia (Maldonado *et al.,* 2006), and have not been previously reported outside Eurasia. The HA, NP, and NS gene segments are in the classical swine lineage. Viruses that seeded this lineage are thought to have entered swine around 1918 (Shope, 1931) and subsequently circulated in classical swine viruses and triple reassortant swine viruses (Olsen, 2002).

The PB2 and PA gene segments are in the swine triple reassortant lineage. Viruses that seeded this lineage, originally of avian origin, entered swine in North America around 1998 (Zhou *et al.,* 1999). Finally, the PB1 gene segment is in the swine triple reassortant lineage. This lineage of PB1 was seeded in swine from humans at the time of the North American swine triple reassortment events (Zhou *et al.,* 1999) and was itself seeded from birds around 1968 (Kawaoka *et al.*, 1989). This is shown in Fig 4.

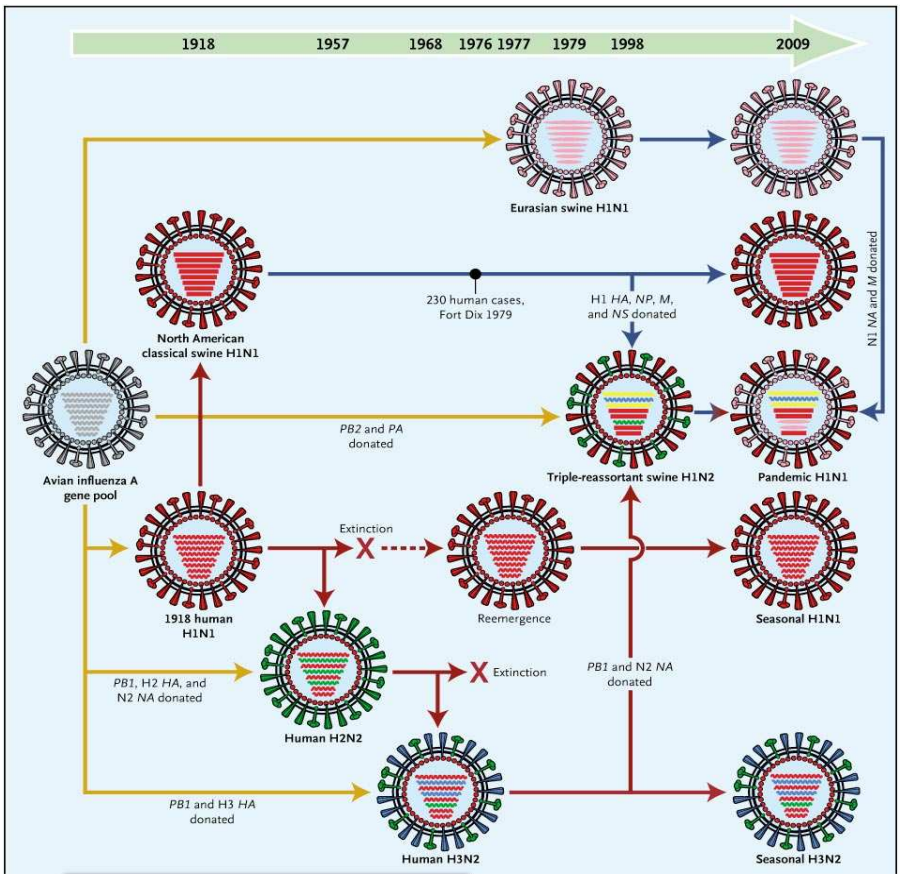


Figure 4: An illustration of the origin of influenza A(H1N1)pdm09 virus
Yellow arrows reflect exportation of one or more genes from the avian influenza A virus gene pool. The dashed red arrow indicates a period without circulation. Solid red arrows indicate the evolutionary paths of human influenza virus lineages; solid blue arrows, of swine influenza virus lineages; and the blue-to-red arrow, of a swine-origin human influenza virus (Morens *et al.*, 2009).

Clinically, infection with influenza A(H1N1)pdm09 virus appears indistinguishable in severity and symptoms from seasonal H1N1 or H3N2 influenza, with some notable exceptions (Dawood *et al.*, 2009; Shinde *et al.*, 2009). First, a greater percentage of patients have complained of having gastrointestinal symptoms than what is usually reported for seasonal influenza (Dawood *et al.*, 2009). Correspondingly, viral replication of S-OIV within the intestinal tracts of inoculated ferrets has been documented (Maines *et al.*, 2009). A second striking feature is the high rate of infection and severe disease in the younger population, rather than in the elderly, as normally seen in seasonal influenza. A putative explanation for this observation is the presence of cross-reactive antibodies generated during previous exposures to 1918-like viruses, which circulated between 1918 and 1943, or H1N1 influenza vaccination in 1976, both of which were shown to confer protection to influenza A(H1N1)pdm09 virus in mice (Manicassamy *et al.*, 2010). Structural data demonstrate that the virus lacks certain glycosylation sites in the HA head region that have been gradually acquired by human H1 HAs to mask neutralizing epitopes (Xu *et al.*, 2010), making the virus antigenically similar to viruses that circulated after the 1918 pandemic. These data suggest that HAs in swine viruses can remain antigenically frozen in time because of the lack of selection pressure in this host and can re-emerge in the future (Tscherne and Garcia-Sastre, 2011).

3.0: MATERIALS AND METHODS

3.1 Study site and design

This study was carried out at the National Influenza Centre (NIC) laboratories located at the Centre for Virus research (CVR) within the Kenya Medical Research Institute (KEMRI). It was a cross sectional laboratory based retrospective study involving molecular characterization of influenza H1N1 pandemic viruses obtained from isolates derived from the 2009 influenza pandemic surveillance around the country. However, in order to minimize founder effects, the isolates from this study were filtered such that not more than two isolates per week from a specific region were included.

3.2 Ethical considerations and sample origin

The specimens used in this study were from the pandemic outbreak surveillance of influenza at the NIC. The Kenyan NIC is part of the WHO Global Influenza Surveillance and Response System that is tasked with conducting national influenza surveillance to monitor circulating viruses by performing preliminary analysis (WHO, 2013). The samples for this study were therefore collected as part of the global influenza pandemic response, routine diagnostic treatment by hospital physicians and were also provided by sentinel hospitals and medical facilities around the country from the sites shown in Figure 5. However, ethical clearance to carry out this particular protocol (Protocol #1753) was sought and approval granted by both the

Scientific Steering Committee and Ethical Review Committee of KEMRI (Appendix 1). For the sake of this study, samples collected in July 2009 are considered 'early' pandemic samples; those collected between August and October 2009 are termed 'peak' pandemic samples while those collected between November 2009 and August 2010 when the pandemic was declared over are considered 'late' pandemic samples.

Figure 5: Map of Kenya showing the sites where study samples originated from.

3.3 Collection of clinical samples

Samples were collected from patients who met the WHO suspected case definition by the Ministry of Public Health and Sanitation/KEMRI/CDC surveillance team. A suspected case of influenza A(H1N1)pdm09 virus infection was defined as a person with acute febrile respiratory illness with onset:

1. Within 7 days of close contact with a person who was a confirmed case of influenza A(H1N1)pdm09 virus infection, or
2. Within 7 days of travel to community internationally where there were one or more confirmed influenza A(H1N1)pdm09 cases, or
3. Resided in a community where there was one or more confirmed influenza A(H1N1)pdm09 cases.

Duplicate nasopharyngeal samples were collected using a flocked Dacron swab and placed in cryovials containing 1 ml Virus Transport Medium (VTM), kept at 4^0C and stored in a liquid nitrogen dry shipper within 8 hrs of collection. All samples were transported immediately from the sites to the National Influenza Center (NIC) laboratory located within KEMRI, Nairobi Kenya, maintaining the cold chain throughout.

3.4 RNA extraction

The virus RNA was extracted from the clinical samples and isolates using the QIAamp® Viral RNA extraction kit (Qiagen, Germany) following the

manufacturer's protocol. Briefly, 100µl of the virus was added to 500µl of lysis buffer per tube and allowed to incubate at room temperature for 10 min to allow for the lysis. A 500µl aliquot of ethanol was added and pulse vortex performed for 15s to give a homogeneous solution. A 630µl volume of the lysed solution was applied to the spin columns and centrifuged at 6000 xg for 1 minute and column placed in a clean collection tube. Then 500µl of Buffer AW1 (wash buffer 1) was added to the spin column and centrifuged at 12000 X g for 1min in a Eppendorff 5415R centrifuge (Eppendorf AG, Barkhausenweg, Hamburg, Germany) and the column placed in a clean collection tube. Then the column was washed with 500µl of Buffer AW2 (wash buffer 2) and centrifuged at 13,000 rpm for 3 min in Eppendorf 5415R centrifuge. The spin column was placed in a 1.5 ml micro centrifuge tube and 60µl of Buffer AVE (elution buffer) added to the column and allowed to incubate at room temperature for 1 min. The column was centrifuged at 8,000 rpm in the Eppendorf centrifuge at room temperature for 1 minute and the filtrate (RNA) stored at -80^0C.

3.5 Real time RT-PCR

Real-time RT-PCR (rRT-PCR) amplification and detection was performed on an ABI 7500 (Applied Biosystems, CA, USA) using the AgPath-ID one-step RT-PCR kit (Applied Biosystems, Foster City, CA, USA). The 25µl reaction volume for each sample contained 5 µl of extracted RNA, 12·5 µl of AgPath Kit 2X buffer, 1 µl of AgPath 25X enzyme mix, 5 pmol of Taqman probe, 10

pmol of each of the forward and reverse primers, and 6 µl of RNase-free water. Each RNA sample was tested for four sets of genes namely: matrix protein gene segment for the identification of influenza A viruses, nucleoprotein gene segment for the identification of swine A influenza viruses, haemagglutinin gene segment for the subtyping of swine A H1 viruses and lastly RNP gene segment was amplified to test the RNA extraction procedure. Reverse transcription was achieved at 50^0C for 30 min and 95^0C for 15 min. PCR was achieved after 45 cycles of 94^0C for 15s and 55^0C for 30s. Cycle threshold $\leq$40 was interpreted as positive in the early and peak phase of the pandemic but this was adjusted to $\leq$35 in the late phase.

3.6 Virus Isolation

All samples that were influenza A(H1N1)pdm09 positive by rRT-PCR were inoculated in Madin-Darby canine kidney (MDCK) cell in tissue culture tubes (Nunclon, Denmark). Influenza virus isolations were performed in MDCK cells followed by hemagglutination inhibition assay (HAI) using guinea pig red blood cells and reference antiserum in accordance with World Health Organization (WHO) protocols (WHO, 2002). Briefly, 100 µl of each sample was added onto 70-90% confluent MDCK cells in flat sided tubes after pre-treatment with L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK) trypsin in order to facilitate virus entry into the cells and allowed to adsorb for 45 min. The cells were then overlaid with 1ml of Gibco® Dulbecco's Modified Eagles Medium (Invitrogen Corporation, NY, USA). Inoculated cells were

incubated at 33°C with 5% CO_2 and observed daily for 10 days for visual cytopathic effect (CPE) using an Olympus CKP inverted microscope (Olympus Corporation, Tokyo, Japan). When CPE was observed, the supernatant fluid was collected and the hemagglutination titer measured as described by the WHO (WHO, 2002). Positive isolates were confirmed using 1% guinea pig red blood cells as described previously (Ziegler *et al.,* 1995) with post-infection ferret antisera obtained from CDC, Atlanta, Georgia.

3.7 Conventional RT-PCR for genomic amplification

The genes of the influenza A(H1N1)pdm09 viruses that were isolated were sequenced to study their molecular characteristics. Conventional PCR to amplify each of the 8 gene segments of the virus was done using M13 tagged primers (Primer sequences in Appendix 2). Complete NS, M and NP genes were amplified as one fragment while NA and HA were amplified as two overlapping fragments. The larger genes namely PA, PB1, and PB2 were amplified as four overlapping fragments. The RT-PCR was performed using Superscript III One-Step RT-PCR system (Invitrogen Corporation, NY, USA). The reaction mix was prepared by mixing 12.5µl of the 2x reaction mix, 0.5 µl of the forward primer (20 µM), 0.5 µl the reverse primer (20 µM), 1.0 µl Superscript III RT/Platinum Taq mix and this mixture was then topped using 7.5 µl of distilled water to make a total of 22 µl. 3 µl of the RNA template was then added. Thermocycling was carried out on a 9700 ABI Thermal Cycler (Applied Biosystems, CA, USA) using the following cycling conditions; 1

cycle of reverse transcription at 50^0C for 30 min followed by an initial denaturation of 94^0C for 2min. This was followed by 35 cycles of; denaturation at 94^0C for 30 s, annealing at 55^0C for 30 s and strand extension at 68^0C for 1 min. Lastly, the reaction mixture was incubated at 68^0C for 1min to allow for extension of recessed ends of the amplicons. A final pause was set at 7^0 C.

3.8 Agarose Gel electrophoresis and visualization of the amplicons

A concentration of 1.5% agarose was prepared in 1x TBE buffer. The solution was mixed by swirling gently and then heating in a microwave until all the agarose was completely dissolved. The gel was then left to cool for a few min at room temperature and ethidium bromide was added to a final concentration of 0.5µg/ml. The gel was then poured onto an electrophoretic tank containing combs and left to set for 30 min at room temperature. The combs were then carefully removed. 3µl of the PCR samples were mixed with the 2µl of the blue orange gel loading dye (Invitrogen, NY, USA) and then loaded onto the wells. A 100bp or 1kb DNA ladder marker (Invitrogen, NY, USA) was loaded on the first lane of each of the wells. The tank was then connected onto a PowerPac Universal Power Supply (Bio-Rad, CA, USA) and run at 150 volts for about 30-45 min. The gel was then visualized and the gel photo photographed using the AlphaImager gel documentation system (Alpha Innotech, CA, USA).

3.9 Clean-up of PCR products using Exosap-IT

PCR product clean-up was conducted to remove any unconsumed dNTPs and primers remaining as they interfere with sequencing. PCR tubes containing 10 µl of the PCR products to be purified were pulse-spun. 3 µl of the ExoSap-IT enzyme (U. S Biological, Swampscott, MA, USA) was added to each of the PCR tubes followed by a brief vortex using Vortex Genie 2 (USA Scientific, FL, USA) for 10s. They were then pulse spun for 30s. The PCR tubes were then placed into a 9700 ABI Thermal Cycler (Applied Biosystems, CA, USA) and incubated for 30 min at 37^0C. This was followed by inactivation of the ExoSap-IT enzyme by incubation for fifteen minutes at 80^0 C before storing the product at 4^0C.

3.10 Cycle sequencing of the purified PCR products.

PCR amplicons incorporating fluorescent-labeled dideoxy-chain terminators were synthesized using an ABI BigDye Terminator version 3.1 cycle sequencing Kit (Applied Biosystems, Forster City, USA). The reaction mixture for both the forward and reverse reactions were prepared by adding 2µl of BigDye to 2µl BigDye 5X buffer followed by addition of 1µl (4µM) of the M13R/F primers and 3µl of distilled water to make a total volume of 8µl. This reaction mixture was loaded into each annotated well on the 96-well plate followed by addition of 2µl of the purified PCR product. The plates were covered with a sealing mat, vortexed briefly and pulse-spun. The PCR was then run with the cycling conditions as follows: 1 cycle of initial denaturation

at 95^0C for 5 min, followed by 30 Cycles of denaturation at 95^0C for 15 s, annealing at 45^0C for 30 s and strand extension at 68^0C for 2 min and 30 s. This was followed by a final incubation at 68^0C for 3 min to allow for extension of recessed ends of the amplicons before storing the product at 4^0C.

3.11 Purification of the cycle sequencing products using sephadex spin columns.

Dry sephadex G-50 medium powder (Sigma - Aldrich Co Ltd, St. Louis, MO, USA) was loaded into unused wells of 96-well Column Loader (Millipore, MA, USA). The 96 well Multi-Screen®-HV Plate (Millipore Ireland B.V., County Cork, Ireland) was then placed upside down on top of the column loader, and both the Multi-Screen Plate and the column loader were held together and inverted. The top of the Column loader was tapped to release the resin. 300µl of Milli-Q water was then added to each well containing sephadex to swell the resin and the setup allowed to incubate at room temperature for 3 hours. The MultiScreen HV plate was then placed on top of a standard 96-well microplate and spun at 910× g for 5 min on an Eppendorf 5810R bench top centrifuge (Eppendorf, Hamburg, Germany) using a two-tray rotor to remove excess water from the columns. The excess water was then discarded. The sequencing products were carefully added to the centre of each Sephadex well. The 96-well plate with the excess water was replaced with a new 96-well microplate (USA Scientific, FL, USA), and centrifugation done at 910× g for 5 min ensuring that approximately 10µl of product came through

the column. Then, 10µl of Hi-Di™ Formamide (Forster City, CA, USA) was added to ensure the sequencing fragments were maintained as single strands and to hydrate any dry (empty) wells to avoid destroying the capillaries.

3.12 Genetic analyzer procedure

The purified PCR products and the Hi-Di in the 96 well plate were placed into 24-capillaries ABI 3500 XL Genetic Analyzer (Applied Biosystems). These were left to run and nucleotide sequences were obtained using the sequence analysis software (Applied Biosystems).

3.13 Contig assembly

To generate contiguous nucleotide sequences (contigs) from the reverse and forward sequence runs for each amplified segment from the Genetic Analyzer, the sequences were uploaded into the contig assembly program (CAP) of DNA Baser Sequence Assembler v3 (Heracle BioSoft SRL Romania, http://www.DnaBaser) and run under default parameters. The generated contig was then visually inspected against individual chromatograms and any ambiguities corrected.

3.14 Similarity searches

To determine whether the obtained nucleotide sequences were similar to influenza A(H1N1)pdm09 sequences deposited in genomic databases, a similarity search against sequences in the influenza virus resource (Bao *et al.,* 2008) and GenBank (Benson *et al.,* 2006) databases was performed using the basic local alignment search tool (BLAST) (Altschul *et al.,* 1997) with the default parameters of the program.

3.15 Multiple sequence alignment

The nucleotide sequences were put in a multiple sequence file in a fasta format before uploading them for alignment using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) V3.8 (Edgar, 2004). To achieve this, the program defaults parameters were used according to the user guide. The output file was in the aligned fasta format. The same procedures were followed for generating protein multiple sequence alignments.

3.16 Preliminary phylogenetic reconstruction

The phylogenetic reconstruction was carried out utilizing the Bayesian method of tree inference using the MrBayes program v3.1.2 (Ronquist and Huelsenbeck, 2003). The nucleotide fasta data of individual gene segments were converted into the Nexus file format using the Concatenator program

(Pina-Martins and Paulo, 2008). To allow for comparison, representative nucleotide sequences of pandemic viruses isolated in England, China, Africa and America during the study period were downloaded from the NCBI H1N1 Flu Influenza Resource database (Bao *et al.,* 2008) and included in the nexus file. The gene segments of the vaccine strain A/California/7/2009 were similarly downloaded and incorporated in the analysis. For the complete genome analyses, the aligned coding region of each gene was concatenated in the following sequence; PB2, PB1, PA, HA, NP, NA, M and NS using the SequenceMatrix software (Gaurav *et al.,* 2011) to produce a nexus file. In addition, the generalized time reversible (GTR) model parameters and priors were incorporated into the nexus file for execution of all analyses in MrBayes.

The data in nexus format was then executed in MrBayes. After completion of the analysis, the probabilities values were summarized and the trees summarized to generate a consensus tree. The tree was converted into a graphic and visualized using FigTree version 1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/).

3.17 Evolutionary analysis.

Evolutionary rates, molecular clock phylogenies and other evolutionary parameters were estimated from the full length concatenated genome of Kenyan influenza A(H1N1)pdm09 viruses and individual gene segments using the relaxed-clock Bayesian Markov chain Monte Carlo (MCMC) method as

implemented in Bayesian Evolutionary Analysis Sampling Trees (BEAST) v1.7.4 (Drummond and Rambaut, 2007). This approach has been validated extensively in the context of human influenza (Rambaut *et al.,* 2008). The method allows variable nucleotide substitution rates among lineages and also incorporates phylogenetic uncertainty by sampling phylogenies and parameter estimates in proportion to their posterior probability (Drummond and Rambaut, 2007).

For the complete genome analyses, the aligned coding region of each gene was concatenated in the following sequence; PB2, PB1, PA, HA, NP, NA, M and NS using the SequenceMatrix software (Gaurav *et al.,* 2011). The resulting nexus file was imported into BEAUti (Bayesian Evolutionary Analysis Utility) programe which is incorporated in BEAST package. BEAUti serves to set the evolutionary model and options for the MCMC analysis and also generate an XML file of the full length concatenated genome for use in the BEAST software. The XML file was then used in the BEAST analyses. For the individual gene segments, the above procedure was followed but without the concatenation step.

Bayesian MCMC analysis for each gene sequence was run for 30 million states and 50 million states for the concatenated genome with sampling every 10,000 states to ensure adequate sample size of all analysis parameters including the posterior, prior, nucleotide substitution rates, and likelihoods. Convergence was assessed on the basis of the effective sampling size (ESS) after a 10% burn-in using Tracer software version 1.5 (Rambaut and Drummond, 2009).

Only ESS's of ≥ 200 were accepted. Statistical support for specific clades was obtained by calculating the posterior probability of each monophyletic clade. As coalescent priors, the four simple parametric demographic models (constant population size, and exponential, expansion and logistic population growth) and a piecewise-constant model, the Bayesian skyline plot (BSP) under a relaxed (uncorrelated log-normal) clock were compared. Uncertainty in the estimates was indicated by 95% highest posterior density (95% HPD) intervals, and the best-fitting models were selected using a Bayes factor (BF with using marginal likelihoods) implemented in BEAST.

The resulting tree with the maximum product of posterior probabilities (maximum clade credibility tree) for analyzing the MCMC data set was annotated by TreeAnotator incorporated in the BEAST package. All trees generated were rooted using the prototype vaccine strain A/California/07/2009 and visualized using FigTree version 1.3.1. In order to identify Kenya-specific transmission clusters, two criteria were used; (i) the cluster must be significantly supported, with a phylogenetic posterior probability of ≥90% and (ii) the cluster must contain more than two isolates (an exception was made for one cluster due to its uniqueness). Clusters were named to reflect their placement within global clades 1 to 7, defined previously (Nelson *et al.*, 2009); hence, KENA-GC2 indicates Kenyan cluster A which falls within global clade 2.

3.18 Analysis of genetic drift and antiviral drug associated mutations

Analysis of antigenic drift was elucidated by analyzing and comparing changes in the amino acid signature patterns of each gene segment between the Kenyan viruses and the vaccine strain. To accomplish this, the translated protein alignments of the coding region of each gene (Kenyan samples and the vaccine strain) were entered in Molecular Evolutionary Genetics Analysis software (MEGA) version 5 (Tamura *et al.,* 2011) and analyzed. The variable amino acid sites were then highlighted and exported to Microsoft Excel spreadsheet.

In order to analyze amino acid changes associated with resistance to antiviral drugs, a multiple sequence alignment of the matrix 2 proteins was analyzed for adamantine sensitivity while the Neuraminidase protein was analyzed for neuramindase inhibitor sensitivity. This analysis included the A/California/07/2009 virus whose sensitivity to these drugs in known as a reference strain and known mutations associated with resistance examined. For neuramindase inhibitor sensitivity, all the sequences were examined for oseltamivir - (I117V, E119V, D198N, I222V, H274Y, R292K, N294S and I314V) and zanamivir-related mutations (V116A, R118K, E119G/A/D, Q136K, D151E, R152K, R224K, E276D, R292K and R371K) while in the M2, the S31N mutation conferring resistance to adamantanes was examined in the local isolates.

4.0: RESULTS

4.1 Viral isolation, amplification and sequencing

In total, 369 out of 2,519 samples (14.6%) collected during the study period (July 2009 and August 2010) were positive for influenza A(H1N1)pdm09 using real time RT-PCR. Out of these samples, 55 were successfully cultured after the first passage. These were filtered to 40 isolates collected in all the provinces in the country and based on the criteria described in section 3.1. The information regarding age, sex, nationality, date of collection and collection site for these isolates can be found in Appendix 3. A figure showing the RT-PCR amplified gene segments is shown in Figure 6. The 40 complete genome sequences from this study have been deposited in the Global Initiative on Sharing All Influenza Data (GISAID) (http://platform.gisaid.org/) and Genbank (Benson *et al.*, 2006) databases. The sequences together with their accession numbers can be found in Appendix 5.

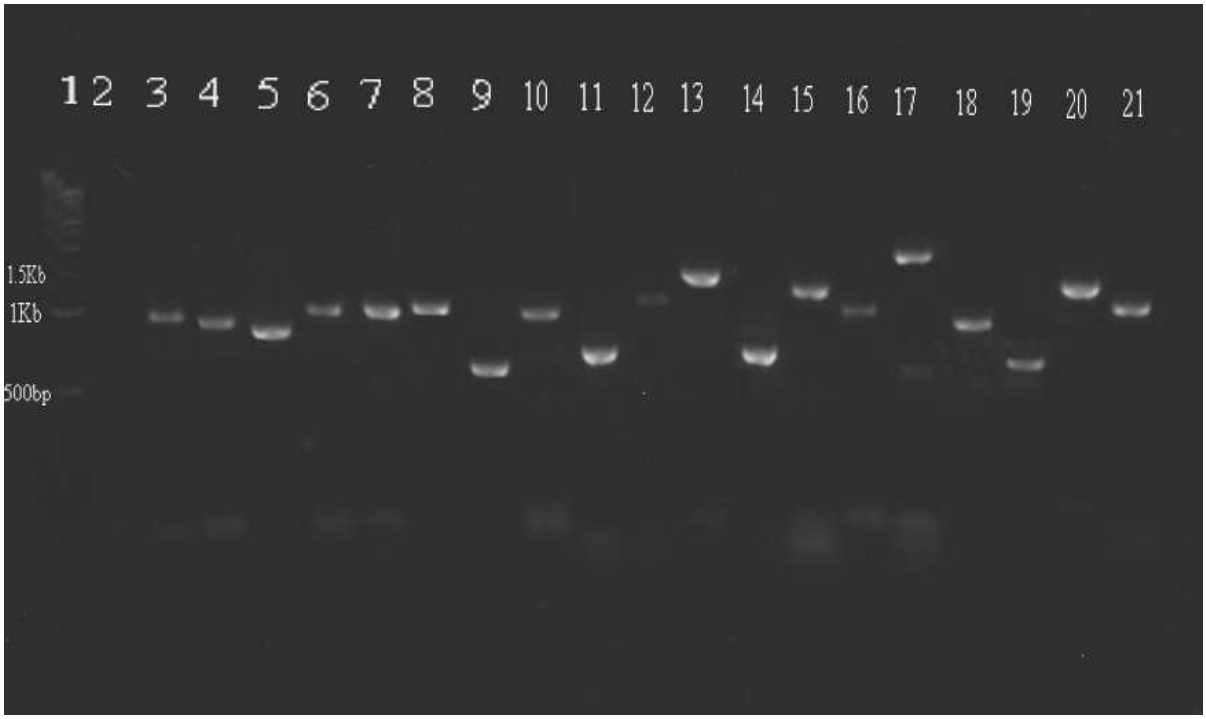


Figure 6: Gel photo showing the PCR amplification of the eight influenza gene segments

Lane 1, 1kb ladder marker., Lane 2, negative control., Lane 3, PB2 fragment 1, Lane 4, PB2 fragment 2., Lane 5, PB2 fragment 3., Lane 6 PB2 fragment 4., Lane 7, PB1 fragment 1., Lane 8, PB1 fragment 2., Lane 9, PB1 fragment 3., Lane 10, PB1 fragment 4., Lane 11, PA fragment 1., Lane 12, PA fragment 2., Lane 13, PA fragment 3., Lane 14, PA fragment 4., Lane 15, HA fragment 1., Lane 16, HA fragment 2, Lane 17, NP segment., Lane 18, NA fragment 1., Lane 19, NA fragment 2., Lane 20, M segment., Lane 21, NS segment.

4.2 Phylogenetic characterization

Phylogenetic characterization of the nucleotide sequences of isolated local influenza A(H1N1)pdm09 viruses was carried out to compare local viruses with globally circulating viruses and also to estimate evolutionary relationships between them. The phylogenetic tree of each of the eight genes and of the concatenated genome showed similar topologies (Figs 7-15). All the trees were

rooted using the nucleotide sequences of the prototype vaccine strain A/California/07/2009. The phylogeny of each of the eight genes was relatively shallow, with large polytomies observed. The maximum depth of the tree (from root to furthest tip) being 5-9 nodes indicating close evolutionary relationships and subsequent lack of phylogenetic resolution.

The phylogenetic analysis showed that Kenyan viruses were interspersed with sequences from other countries. It was observed that the four earliest Kenyan isolates grouped together in two clusters among all the genes (except NS). It was further observed that there was co-circulation of at least four distinct clusters (pp≥70%) in majority of the genes during the pandemic. These clusters were distributed in such a way that, there were more clusters co-circulating during the peak phase of the pandemic than in the early and late phases. Further, the trees do not show any evidence of spatial evolution. However, temporal evolution was evident during the late phase of the pandemic and isolates during this period clustered together in a well supported monophyletic clade in all the genes (except NP).

4.2.1 PB2

The phylogeny of the PB2 nucleotide sequences shows that local isolates collected in the early and peak phases of the pandemic are dispersed within the phylogenetic tree with no clear patterns (Fig. 7). However, majority of the isolates from the late phase clustered together in a well supported cluster. At the same time, it was observed that local isolates in these early and peak phases are interspersed with PB2 sequences from other countries isolated in the same period. Some local isolates including A/Kisumu/61/2009 and A/Malindi/238/2009 were found to cluster together with isolates from the United Kingdom that were collected between April/May 2009 and October 2009 respectively. The tree does not indicate any spatial evolutionary patterns except in the late pandemic phase.

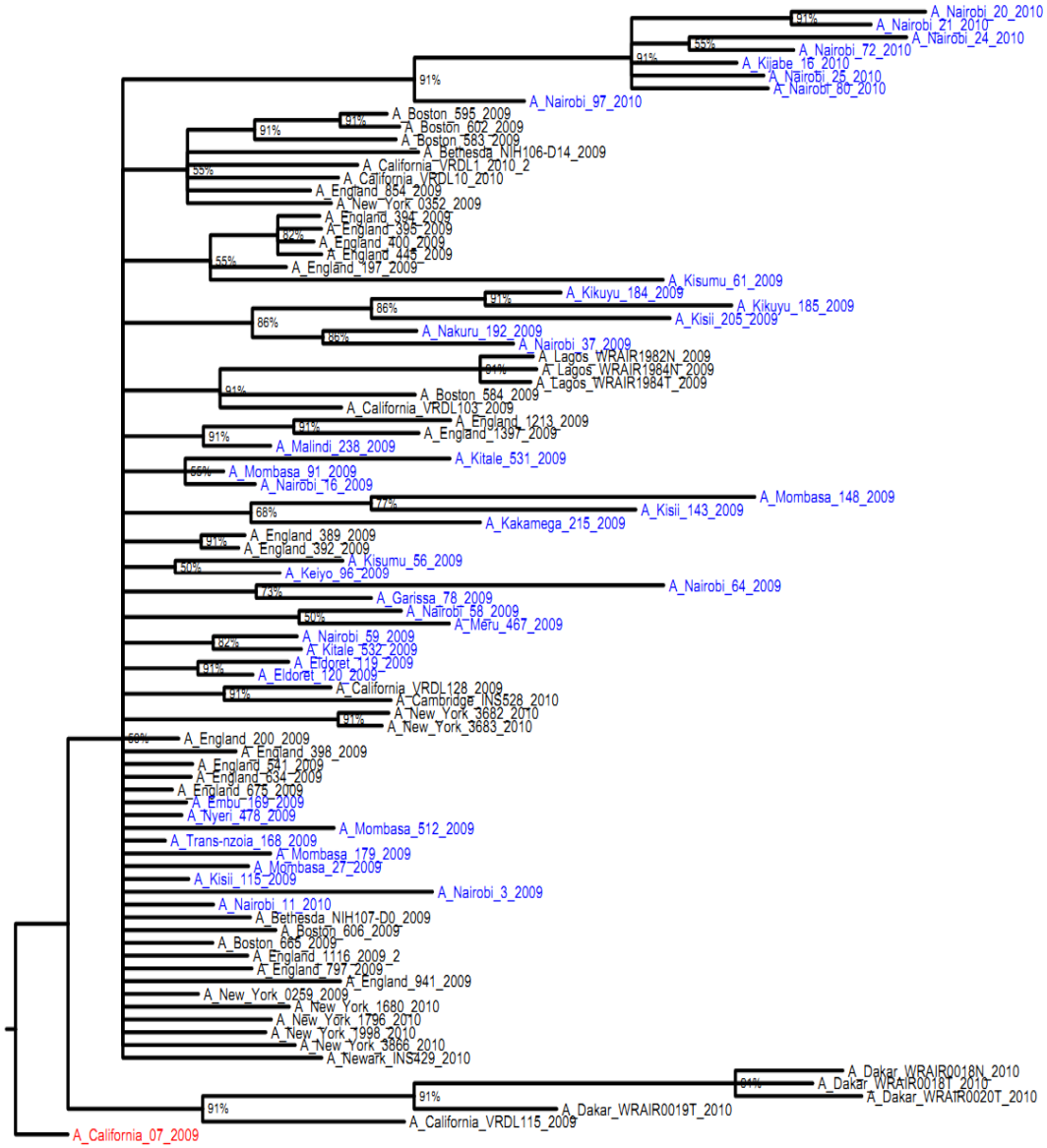


Figure 7: Phylogenetic tree of PB2 gene segment of Kenyan samples.

Phylogeny was inferred with representative PB2 sequences from Africa, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities.

4.2.2 PB1

The phylogenetic tree of the PB1 reveals a similar topology to that of the PB2 tree (Fig. 8). The PB1 sequences of early and peak phase isolates are largely distributed throughout the tree with no discernable temporal or spatial patterns while the late phase isolates group together to form a well supported cluster. Majority of the local isolates were noted to form clusters that were interspersed with global isolates. The late phase isolates were shown to cluster together with one isolate from the United Kingdom (A/England/941/2009) which was collected in December 2009. None of the PB1 sequences from Kenyan isolates clustered with the PB1 sequences from the isolates from other African countries that are currently in the public databases.

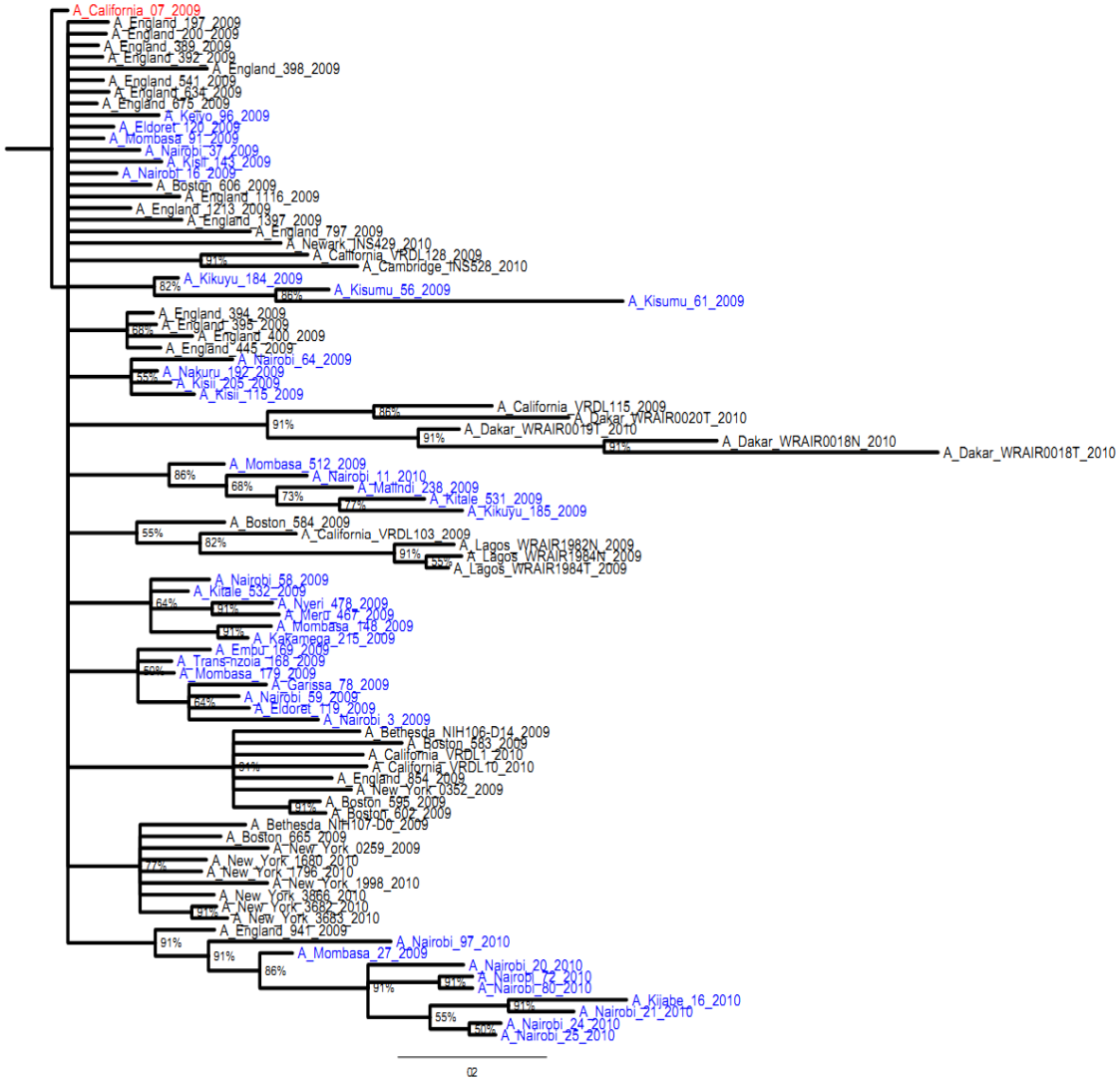


Figure 8: Phylogenetic tree of PB1 gene segment of Kenyan samples.

Phylogeny was inferred with representative PB1 sequences from Africa, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities

4.2.3 PA

The phylogeny of the PA nucleotide sequences does not show any discernable spatial and temporal patterns in the early and peak pandemic phases (Fig. 9). Like the PB2 and PB1 phylogenetic trees, the PA nucleotide sequences of local isolates are interspersed with PA sequences of isolates from other countries in the world collected at the same period. The tree showed that two of the early isolates from Kenya namely A/Kisumu/56/2009 and A/Kisumu/61/2009 clustered with several isolates from the United Kingdom that were sampled between April and June 2009. The late pandemic phase isolates also clustered together.

Figure 9: Phylogenetic tree of PA gene segment of Kenyan samples.

Phylogeny was inferred with representative PA sequences from Africa, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities

4.2.4 HA

The phylogenetic tree of the HA nucleotide sequences showed the maximum depth by having 9 nodes from the root (Fig. 10). The tree can easily be divided into two clades, one containing two of the earliest Kenyan isolates namely A/Kisumu/56/2009 and A/Kisumu/61/2009 while the second containing all the other isolates. The two isolates from Kisumu collected in July 2009 cluster closely with isolates from the United Kingdom collected between April and June 2009. Another local isolate A/Mombasa/27/2009 collected in October 2009 clustered with A/England/634/2009 which was collected in July 2009. As in previous phylogenetic trees, the local HA sequences are interspersed with sequences from other influenza viruses isolated globally. The tree also showed no clear spatial evolutionary patterns but isolates from the late pandemic phase cluster together indicating temporal evolution in this phase.

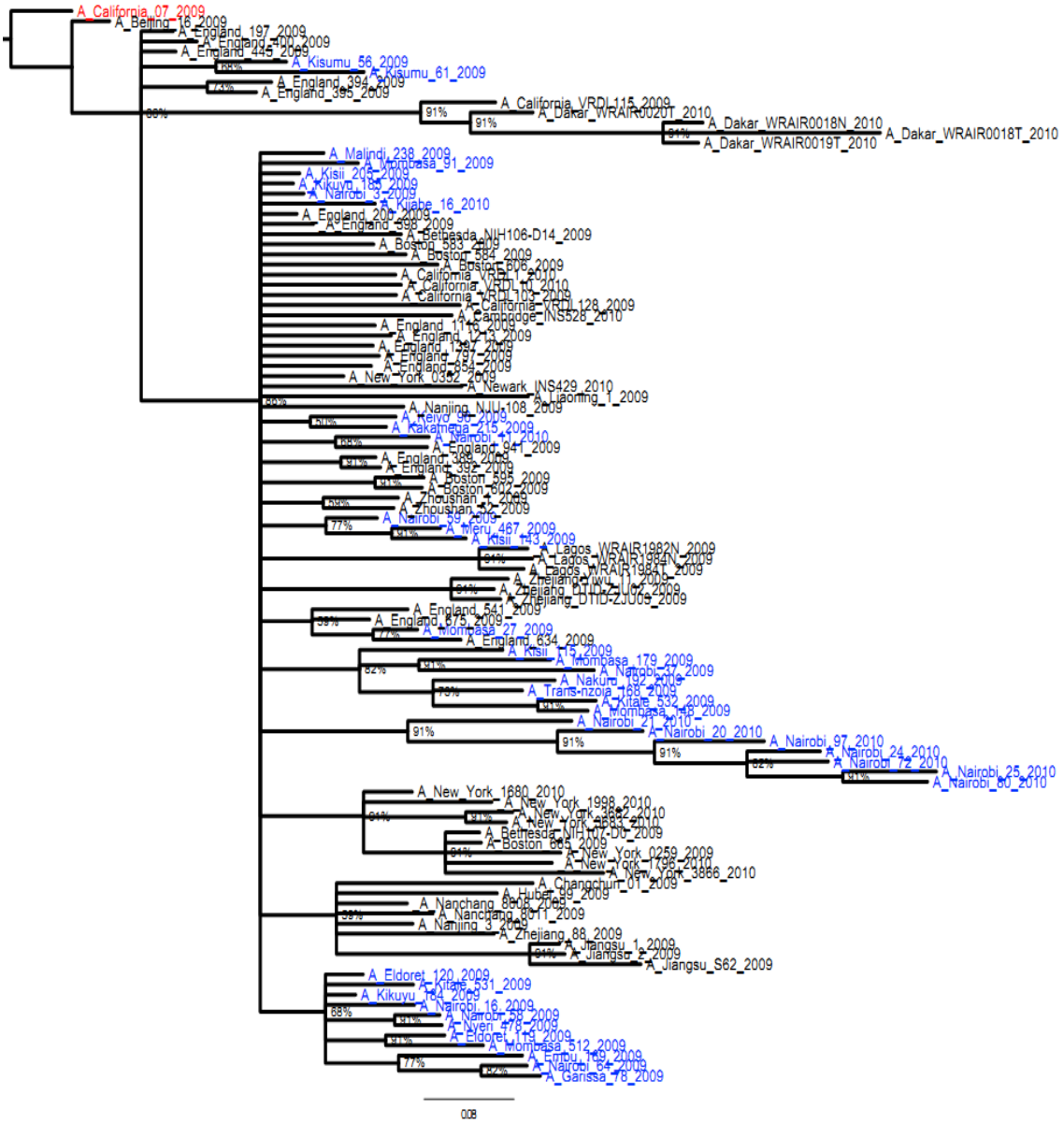


Figure 10: Phylogenetic tree of HA gene segment of Kenyan samples.

Phylogeny was inferred with representative HA sequences from Africa, Asia, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities

4.2.5 NP

The NP phylogenetic tree can also be divided into two clades, one comprising of the two earliest Kenyan isolates A/Kisumu/56/2009 and A/Kisumu/61/2009 while the second comprising all the other isolates (Fig. 11). The two Kisumu isolates were shown to cluster together with isolates from the United Kingdom collected between April and June 2009. All the other local isolates did not show any spatial or temporal patterns. The NP nucleotide sequences of the local isolates were shown to be interspersed with NP sequences from other countries.

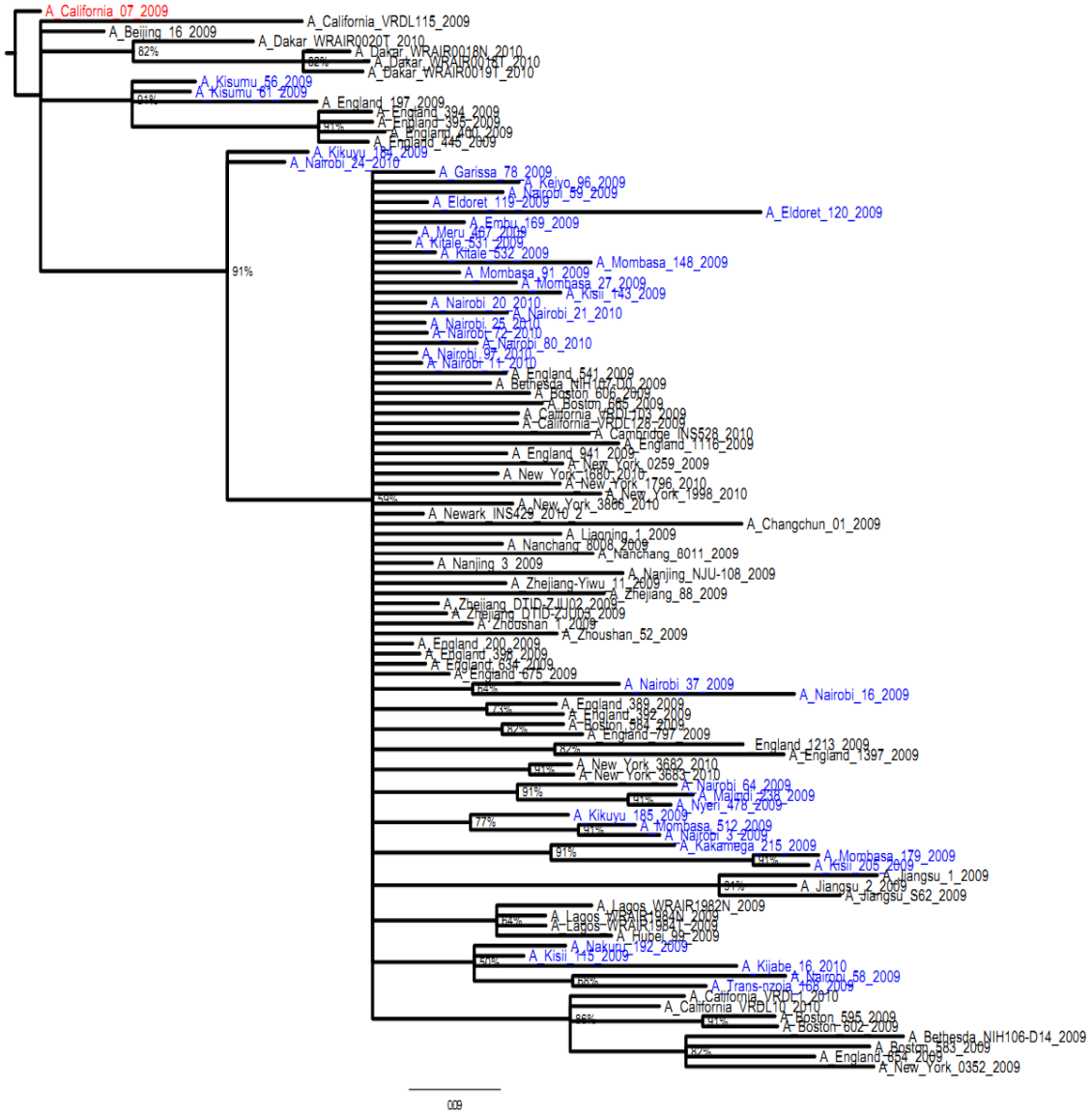


Figure 11: Phylogenetic tree of NP gene segment of Kenyan samples

Phylogeny was inferred with representative NP sequences from Africa, Asia, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities.

4.2.6 NA

The NA phylogenetic tree showed that the local isolates clustered into two distinct clades (Fig. 12). One clade comprising two of the earliest Kenyan isolates A/Kisumu/56/2009 and A/Kisumu/61/2009 while the other comprising all the other isolates. As with the NP and HA sequences, the two Kisumu samples cluster together samples from the United Kingdom that were collected in June 2009. The Kenyan isolates are also shown to be interspersed with isolates from other countries globally. Majority of the 2010 isolates (55%) clustered together while all the other isolates were dispersed throughout the phylogenetic tree.

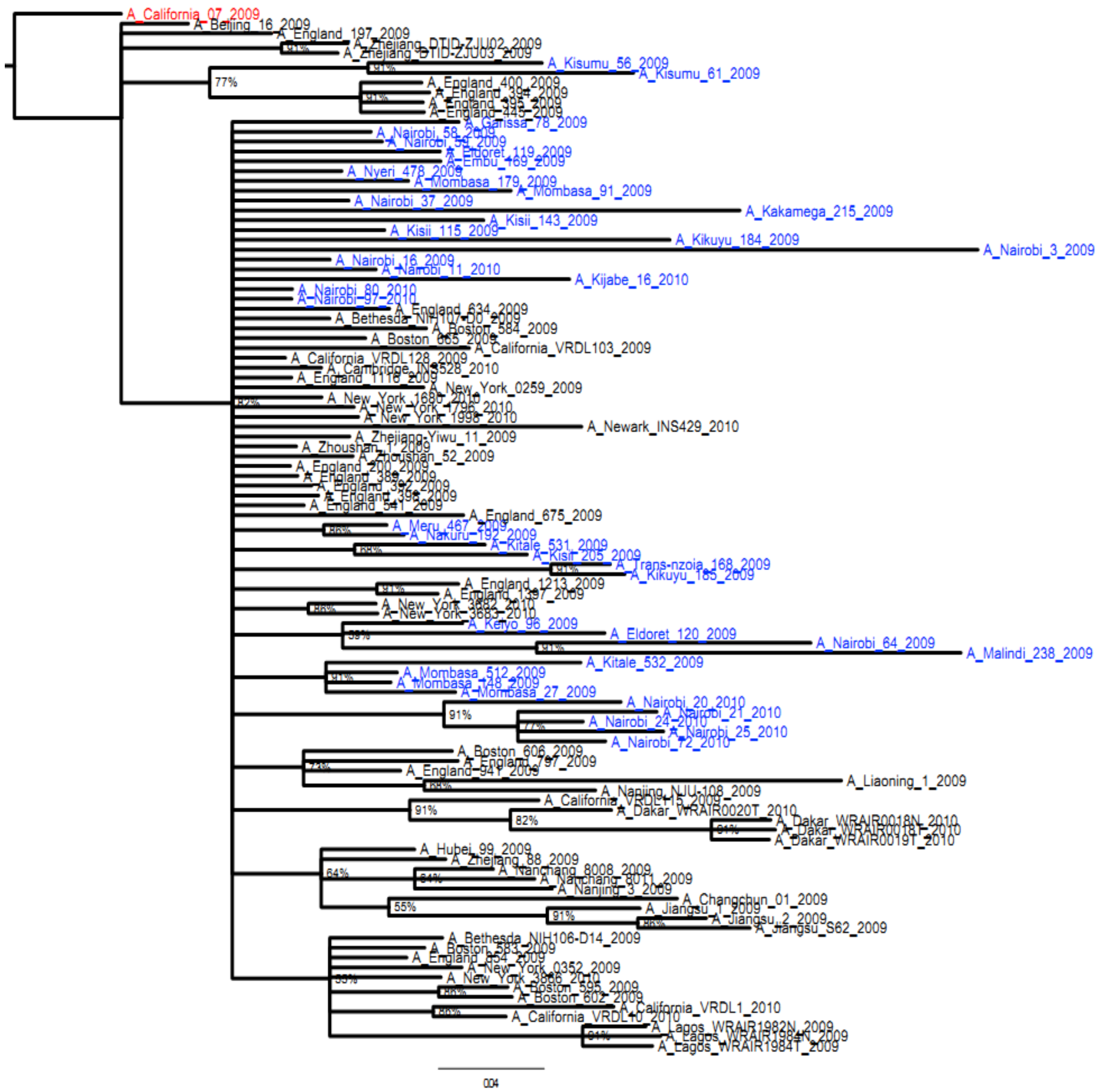


Figure 12: Phylogenetic tree of NA gene segment of Kenyan samples.

Phylogeny was inferred with representative NA sequences from Africa, Asia, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities

4.2.7 M

The phylogeny of the M nucleotide sequences of local isolates was shown to be made up of two distinct clades (Fig. 13). The first comprising two of the earliest Kenyan isolates A/Kisumu/56/2009 and A/Kisumu/61/2009 and the second comprising of all the other isolates. The 2010 local isolates formed a well supported cluster which also had the maximum depth of 9 nodes from the root showing greater diversification in comparison with the other isolates. The local M nucleotide sequences were also interspersed with global sequences in the phylogenetic tree though this was not as explicit as in the other nucleotide sequences.

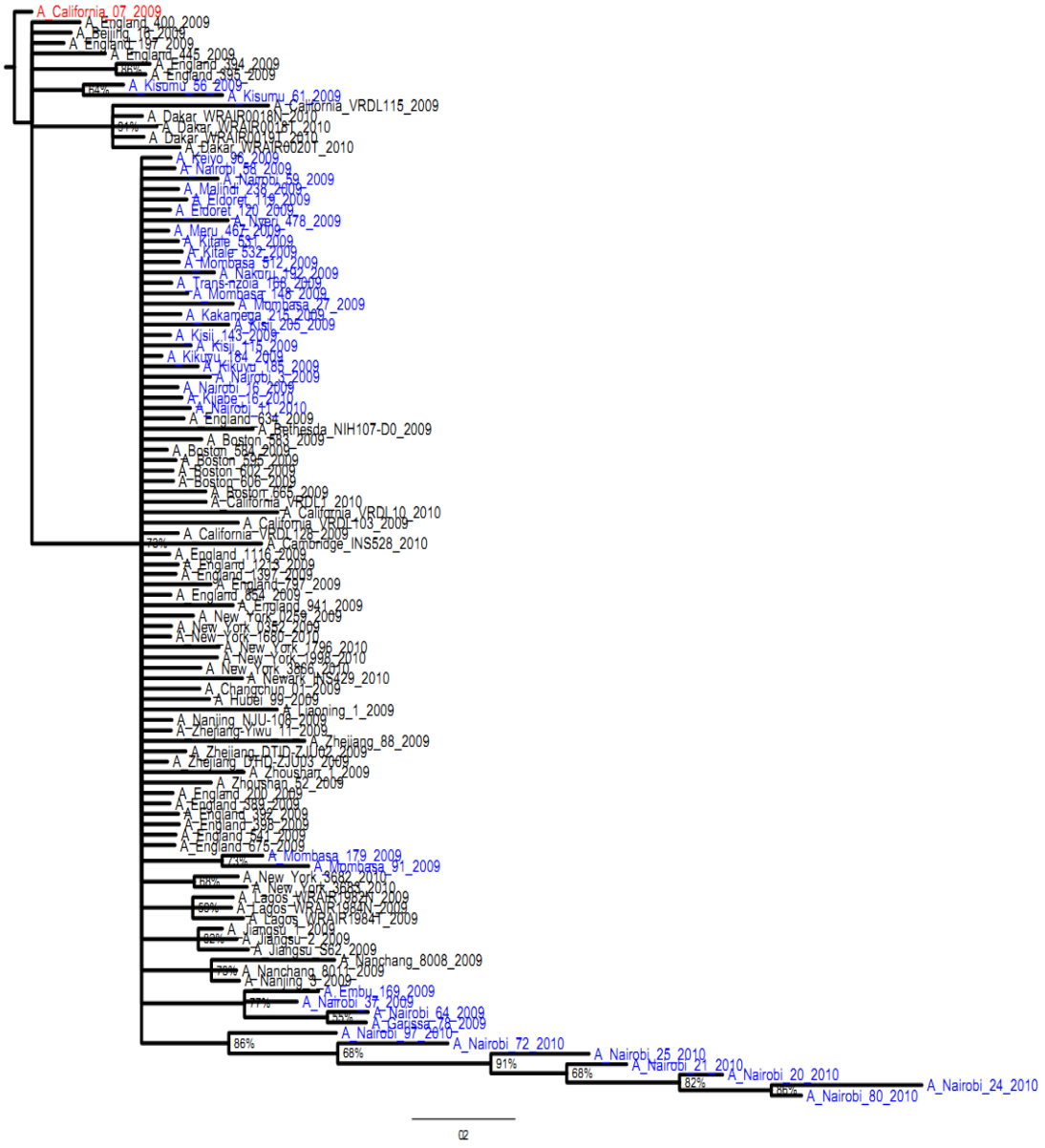


Figure 13: Phylogenetic tree of M gene segment of Kenyan samples.

Phylogeny was inferred with representative M sequences from Africa, Asia, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities

4.2.8 NS

The phylogenetic tree of the NS nucleotide sequences of local isolates shows that they are interspersed with NS sequences from isolates from other countries (Fig 14). One cluster comprised of local isolates mostly from the peak phase also incorporates two isolates from the United Kingdom and China. The two foreign isolates namely A/England/941/2009 and A/Liaoning/1/2009 were collected in December and October 2009 respectively. Isolates from the early and peak phases of the pandemic do not show any clear clustering patterns. However, majority (67%) of the 2010 isolates cluster together to form a well supported cluster. None of the Kenyan isolates were found to cluster with isolates from other African countries.

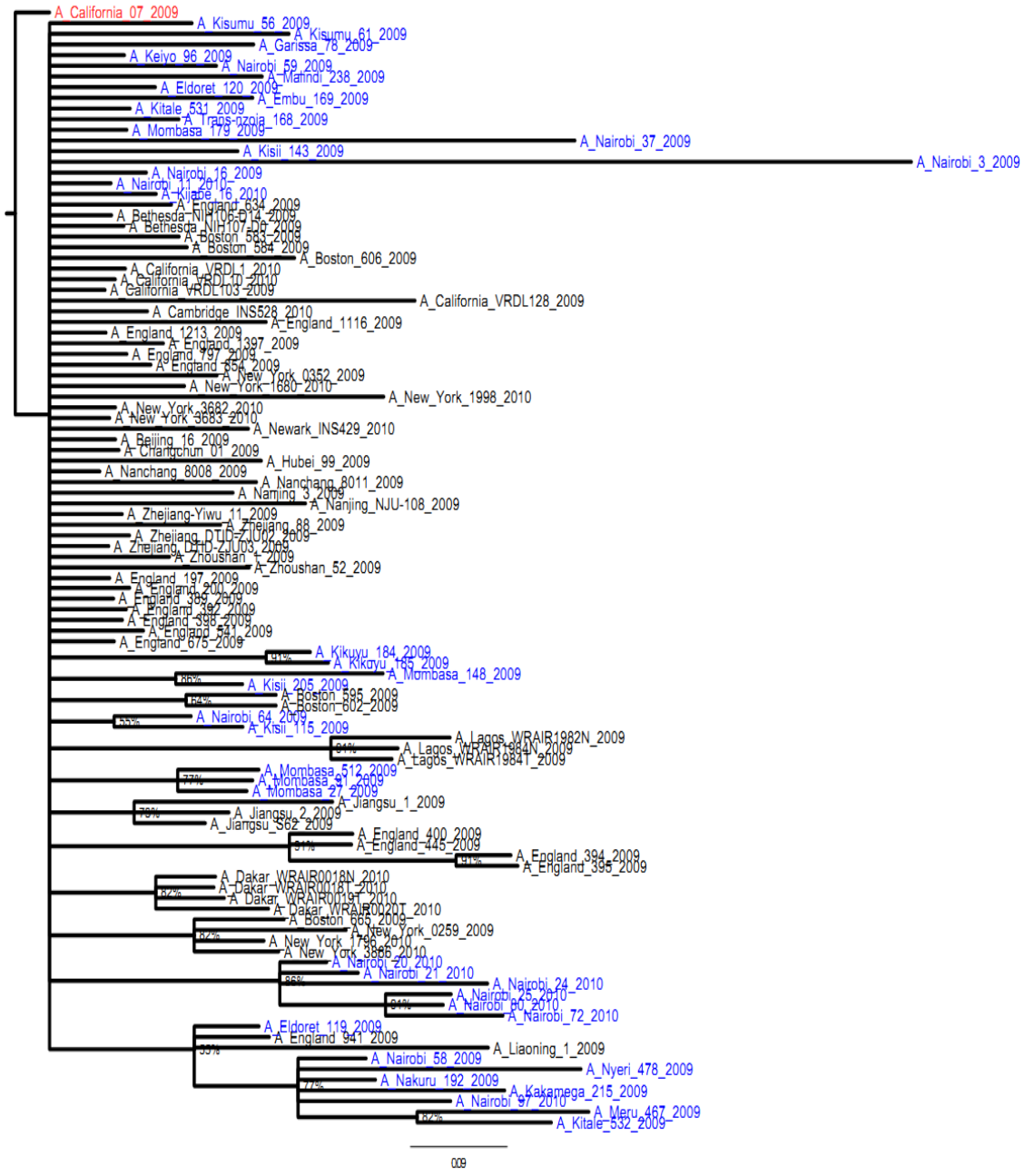


Figure 14: Phylogenetic tree of NS gene segment of Kenyan samples.

Phylogeny was inferred with representative NS sequences from Africa, Asia, Europe and America (black) using MrBayes software (version 3.1.2). The tree was rooted with the influenza A(H1N1)pdm09 reference strain A/California/07/2009 (red). Kenyan samples are shown in blue**.** The numbers on the internal nodes indicate posterior probabilities

4.2.9 Concatenated genome

The phylogenetic tree of the concatenated genome which also included viral genomes isolated elsewhere whose clade identities are known is shown in Fig. 15. It shows that 2(40%) of the earliest isolates from Kenya clustered with global clade II viruses while the other 3(60%) clustered with clade VII viruses. Throughout the pandemic, clade VII viruses were noted to dominate local transmission and no clade II viruses were observed after the initial introductions in the country (Fig 15). Of the seven clades of pandemic H1N1 virus that has been shown to circulate globally, the Kenyan isolates belong to only two different clades. The other globally circulating influenza A(H1N1)pdm09 clades namely I, III, IV, V and V1 were not detected in the Kenyan population. Like in the individual gene segments, the concatenated genomes of the Kenyan isolates were interspersed with genomes of viruses from other countries. Majority of the Kenyan isolates also clustered often with isolates from the United Kingdom than with isolates from other countries. Majority (89%) of the Kenyan 2010 isolates were also shown to cluster together.

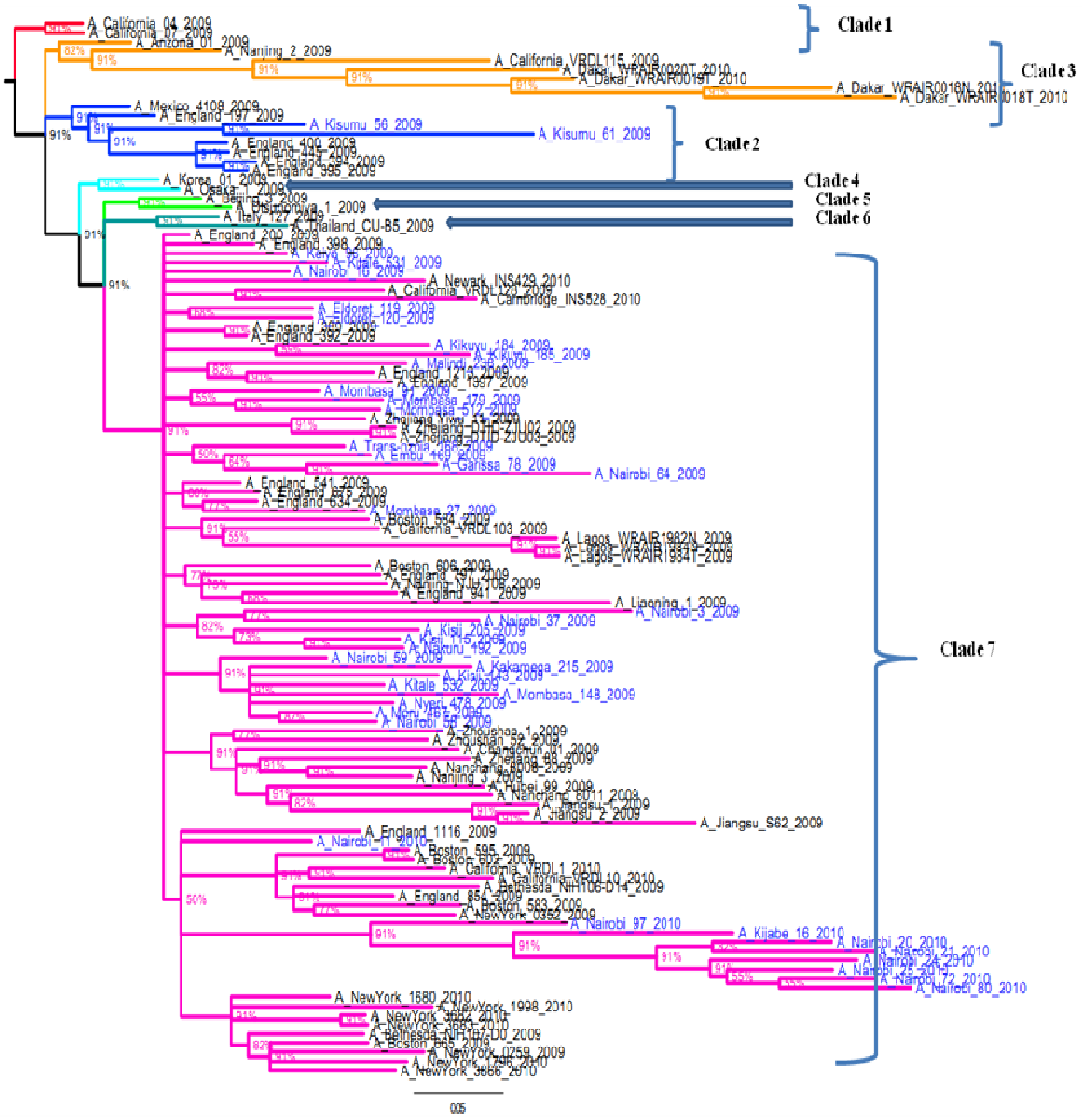


Figure 15: Bayesian phylogenetic tree of the concatenated influenza A(H1N1)pdm09 genome of Kenyan and selected global isolates.

The 40 Kenyan isolates characterised in this study are highlighted in blue. The global influenza A(H1N1)pdm09 clades are indicated and also their branches are shown using different colours.

4.3 Evolutionary analysis

4.3.1 Circulating Kenyan clusters and their divergence times

The study sought to identify the Kenyan-specific circulating clusters and their divergence times based on the complete genome. The resulting temporal phylogeny of the complete genome of Kenyan influenza A(H1N1)pdm09 viruses is shown in Figure 16. This phylogeny demonstrates the existence of multiple clusters during the early and peak pandemic phase. Isolates were considered to belong to one cluster if the branch supporting that cluster showed a Bayesian posterior probability (pp) of ≥ 85% among other criteria described previously. The time of the most recent common ancestor (tMRCA) of each cluster, corresponding to the date at which they emerged was also estimated. The study identified six such Kenyan circulating clusters each containing 2 to 8 Kenyan isolates.

The viruses isolated during the early phase of the pandemic clustered in two different clusters namely KENA-GC2 and KENB-GC7 which were both well supported with posterior probabilities of ≥95%. KENA-GC2 which was made up of only two samples did not spread beyond the early pandemic period. However, KENE-GC7 continued to circulate into the peak pandemic phase. During the peak phase of the pandemic, four clusters namely KENB-GC7, KENC-GC7, KEND-GC7 and KENE-GC7 were co-circulating. KENF-GC7 cluster which was well supported (pp=100%) dominated in the late phase of the pandemic.

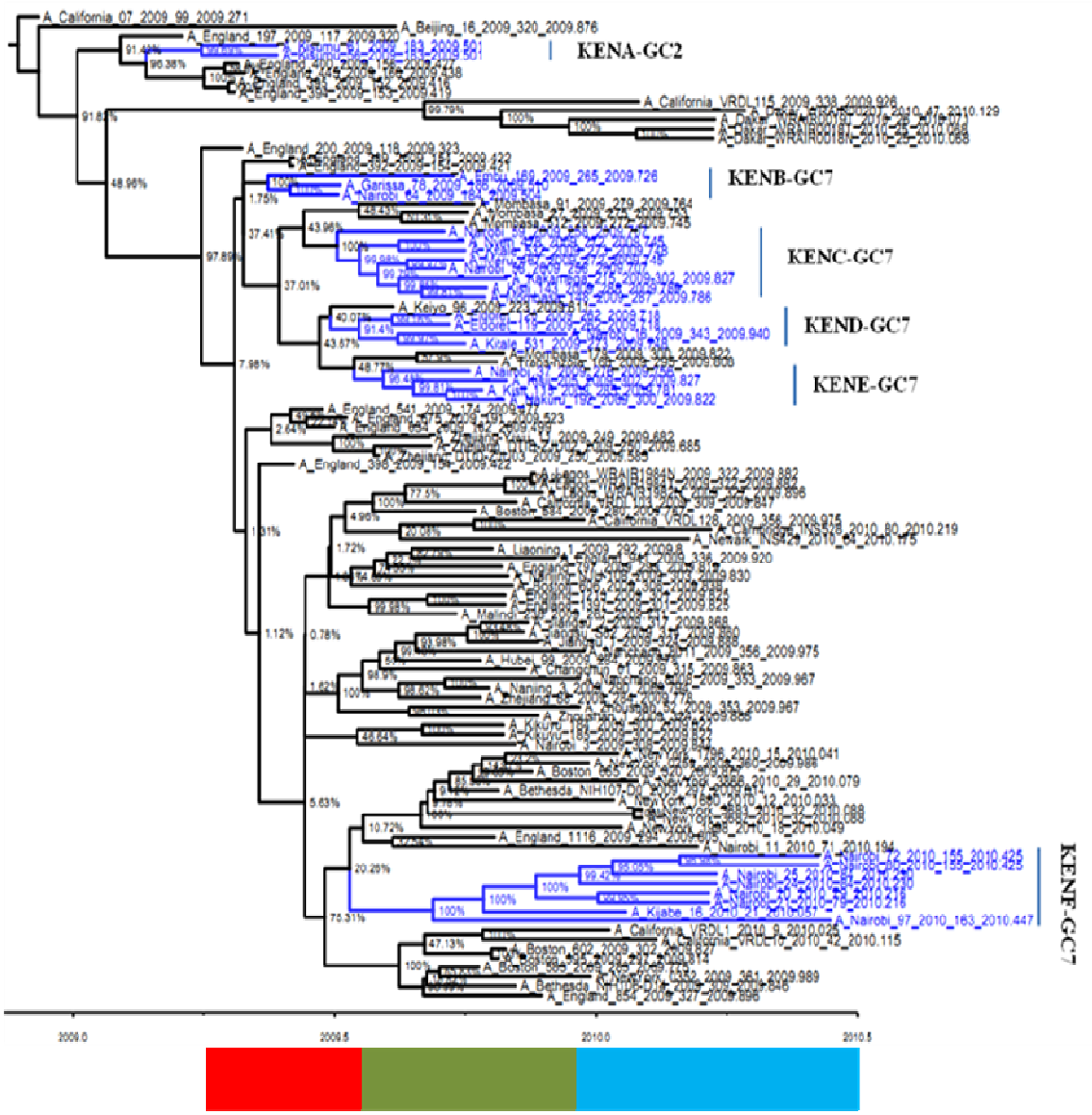


Figure 16: Simplified time scaled maximum clade credibility tree showing distribution of circulating Kenyan clusters of influenza A(H1N1)pdm09.

Identified Kenyan clusters are shown in blue and named to reflect their placement within global clades. The numbers on the internal nodes indicate posterior probabilities. Red colour shows the early phase, green the peak phase and blue the late pandemic phase

The mean time of the most recent common ancestor (tMRCA) of the Kenyan influenza A(H1N1)pdm09 viruses (corresponding to the tree root) was estimated to be 28th February 2009, with a 95% credibility interval of between October 2008 and May 2009. The tMRCAs and the most probable months in the calendar time scale of the main clusters are shown in Table 2. KENA-GC2 which is one of the two earliest clusters in Kenya was shown to emerge on 19th April 2009 around the same time that the virus was first detected in the United Kingdom. The other earliest Kenyan cluster was shown to emerge in June 2009 around the same time the pandemic was peaking in the UK.

Table 2: Divergence times and detection ranges of Kenyan clusters

Cluster*	N° of isolates	TMRCA	TMRCA	
			Low 95% HPD	High 95% HPD
KENA-GC2	2	19th April 2009	27th Dec 2008	6th July 2009
KENB-GC7	3	4th June 2009	16th April 2009	29th June 2009
KENC-GC7	8	8th August 2009	19th May 2009	2nd October 2009
KEND-GC7	4	7th August 2009	10th June 2009	17th Sept 2009
KENE-GC7	4	4th August 2009	10th June 2009	2nd October 2009
KENF-GC7	8	2ndOctober 2009	20th June 2009	10th Dec 2009
root	40	28th Feb 2009	8th Oct 2008	19th May 2009

*The Kenyan-specific circulating clusters are named to reflect their placement within global influenza A(H1N1)pdm09 clades

4.3.2 Substitution rate estimations and population dynamics

The study estimated population dynamics of local influenza A(H1N1)pdm09 virus throughout the pandemic period using the Bayesian Skyline Plot (BSP). The Bayesian skyline plot of the concatenated genome of the Kenyan viruses (Fig 17) indicated an exponential growth in the population size which occurred only during the first three months (July-September 2009) of the pandemic in the country. Thereafter, a slight decline was observed followed by a constant population demography which was maintained till the end of the study period.

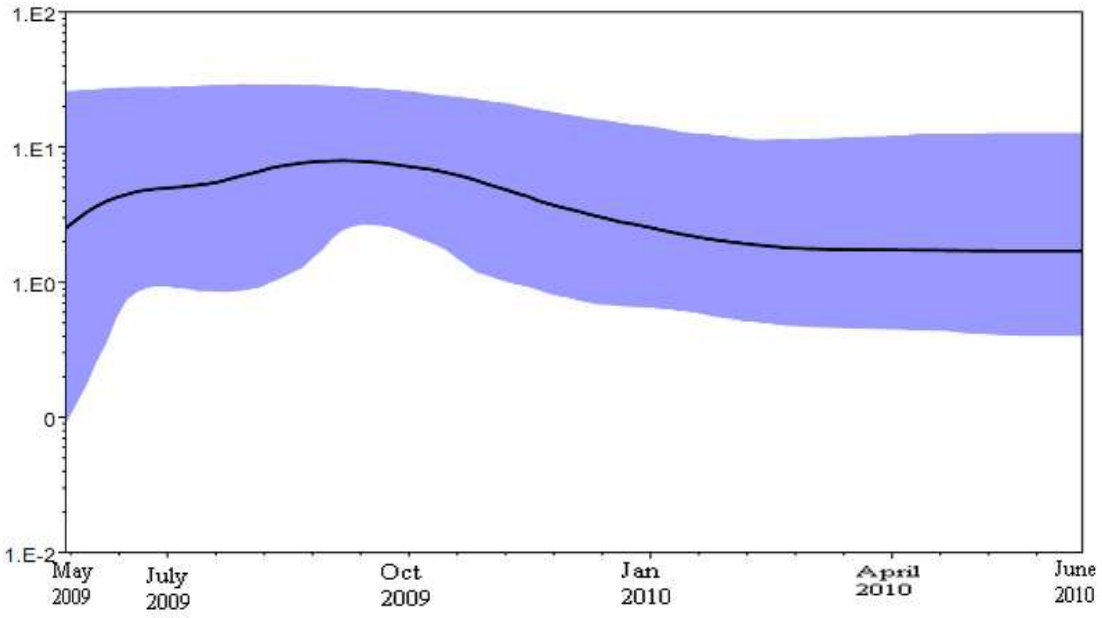


Figure 17: Bayesian skyline plot for the influenza A(H1N1)pdm09 isolates under the uncorrelated lognormal clock.

X-axis indicates time and Y-axis indicates number of effective infections at time t (Ne(t)). The thick solid line represents the median value, and the blue area the 95% HPD of the Ne(t) estimates.

The complete genome of the Kenyan influenza A(H1N1)pdm09 viruses had an estimated rate of nucleotide substitution of 4.9 X 10^{-3} substitutions/site/year (95% confidence interval [CI], 2.6 X 10^{-3} – 7.2 X 10^{-3}) as shown in Table 3 below. Among the gene segments, the rate of nucleotide substitution was lowest in the Nucleoprotein gene at 0.8 X 10^{-3} substitutions/site/year and highest in the Matrix gene at 9.8 X 10^{-3} substitutions/site/year.

Table 3: Mean nucleotide substitution rates for each of the Kenyan influenza A(H1N1)pdm09 virus genomic segments

Gene	**Mean evolutionary rate X 10^{-3} (substitutions per site per year)***
PB2	4.01 (1.47, 6.45)
PB1	3.89 (1.29, 7.22)
PA	1.86 (3.03, 6.19)
HA	5.58 (2.75, 9.28)
NP	0.80 (0.4, 2.04)
NA	4.07 (1.47, 7.73)
M	9.88 (5.58, 14.5)
NS	5.22 (1.64, 9.17)
Genome[#]	4.96 (2.65, 7.21)

*The numbers in brackets indicate the 95% confidence interval estimates.

[#] is the evolutionary rate for the concatenated genome.

4.4 Genetic drift and amino acid changes

Genetic analysis of influenza A(H1N1)pdm09 viruses indicated that the amino acid sequences of all genes of the Kenyan samples were largely similar (sequence similarity of 99%-100%) to that of the pandemic prototype and vaccine strain, A/California/07/2009. The largest numbers of amino acid substitutions were observed during the peak period of the pandemic in all genes except in the Matrix gene.

4.4.1 PB2

The analysis of the 40 PB2 sequences from this study in comparison with the vaccine strain is shown in Table 4. It indicates that local isolates had a 99%-100% similarity with the prototype and vaccine strain. In the early pandemic phase, a K526R substitution was seen among two of the earliest isolates in Kenya. An examination of 3,462 global A(H1N1)09pdm PB2 sequences deposited in Genebank showed that this substitution was prevalent between April and May 2009. This substitution was rarely detected globally after August 2009. All other local isolates had K526. No Amino acid substitutions were fixed in the early and peak phases of the pandemic. The most common amino acid substitutions seen during the peak of the pandemic in local isolates were R175K, P679S and S684C. In the late phase of the pandemic, the following substitutions generally became fixed; R299K, V344M, I354L and N456S. V344M and I354L and R299K and N456S seem to be parallel mutations appearing late in the pandemic.

Table 4: Amino acid sequence variations of PB2 protein of influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus isolate	PB2 amino acid position																								
	127	153	157	175	299	344	354	412	456	480	490	526	533	539	551	561	679	684	686	691	710	735	741	747
A/California/07/2009	**H**	**D**	**K**	**R**	**R**	**V**	**I**	**K**	**N**	**V**	**S**	**K**	**S**	**I**	**Q**	**K**	**P**	**S**	**V**	**L**	**I**	**M**	**S**	**S**
A/Kisumu/56/2009	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisumu/61/2009	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	T
A/Nairobi/64/2009	.	.	.	.	.	.	.	.	.	L	.	.	.	T	.	.	.	.	.	.	.	K	T	T
A/Garissa/78/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	N	.	.	.
A/Keiyo/96/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/58/2009	.	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/59/2009	.	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Malindi/238/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
A/Eldoret/119/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Eldoret/120/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Embu/169/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nyeri/478/2009	.	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Meru/467/2009	.	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.
A/Kitale/531/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	N	.	.	.
A/Kitale/532/2009	.	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/512/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nakuru/192/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	C	G	.	.	.	.	.
A/TransNzoia/168/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/179/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/148/2009	.	.	.	K	.	.	.	.	.	L	T	.	T	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/91/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/27/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/37/2009	.	E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	C	G	M	.	.	.	.
A/Kakamega/215/2009	.	.	R	K	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.
A/Kisii/205/2009	.	.	.	.	.	.	.	N	.	.	.	.	.	F	H	.	S	C	.	.	.	.	.	.
A/Kisii/143/2009	.	E	.	K	.	.	.	.	.	.	T	.	T	.	.	.	.	.	G	M	.	.	.	.
A/Kisii/115/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kikuyu/184/2009	.	.	.	.	.	.	.	N	.	.	.	.	.	F	.	.	S	C	.	.	.	.	.	.
A/Kikuyu/185/2009	.	.	R	.	.	.	.	.	.	.	.	.	.	F	.	.	S	C	.	.	.	.	.	.
A/Nairobi/3/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	G	M	.	.	.	.
A/Nairobi/16/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/11/2010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kijabe/16/2010	.	.	.	.	K	M	L	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/20/2010	Q	.	.	.	K	M	L	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/21/2010	Q	.	.	.	K	M	L	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/24/2010	.	.	.	.	K	M	L	.	S	.	.	.	.	.	.	N	.	.	.	.	.	.	.	.
A/Nairobi/25/2010	.	.	.	.	K	M	L	.	S	.	.	.	.	.	.	.	.	.	.	.	.	L	.	.
A/Nairobi/72/2010	.	.	.	.	K	M	L	.	S	.	.	.	.	.	.	N	.	.	.	.	.	.	.	.
A/Nairobi/80/2010	.	.	.	.	K	M	L	.	S	.	.	.	.	.	E	.	.	.	.	.	.	.	.	.
A/Nairobi/97/2010	.	.	.	.	.	M	L	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.4.2 PB1

A comparison of the PB1 protein of the study samples and the vaccine strain is shown in Table 5. The PB1 protein of the study samples exhibited an amino acid similarity of 98.4-100% to that of the vaccine strain. Two of the five earliest isolates from Kenya had T257A, V587A, S712C and A717G amino acid substitutions in comparison with A/California/07/2009. Of these, V587A has been observed in majority of clade 2 viruses isolated globally while T257A has been observed in some Asian isolates between June 2009 and May 2010. The substitutions S712C and A717G found in the early pandemic samples were unique to Kenyan samples. On the other hand, two of the other three early pandemic phase isolates had an E739K substitution which has been observed in some clade 7 viruses globally.

There were no amino acid substitutions that became fixed in this protein during the pandemic. The amino acid substitutions Q268R, A717G and E739K were observed in all the phases of the pandemic in the Kenyan isolates. Q569H, S703N, R707T and V709D were observed in the peak and late phases of the pandemic only while S712C was observed during the early and peak phases of the pandemic in Kenya. All these mutations except Q269R and E739K were unique to the Kenyan samples and were not observed elsewhere.

Table 5: Amino acid sequence variations of PB1 protein of influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus isolate	PB1 amino acid position											
	257	268	289	569	587	678	703	707	709	712	717	739
A/California/07/2009	**T**	**Q**	**M**	**Q**	**V**	**S**	**S**	**R**	**V**	**S**	**A**	**E**
A/Kisumu/56/2009	A	.	.	.	A	.	.	.	.	C	G	.
A/Kisumu/61/2009	A	.	.	.	A	.	.	.	.	C	G	.
A/Nairobi/64/2009	.	.	.	.	.	.	.	.	.	.	.	K
A/Garissa/78/2009	.	R	.	.	.	.	.	.	.	.	.	K
A/Keiyo/96/2009	.	R	K	.	.	.	.	.	.	.	.	.
A/Nairobi/58/2009	.	.	.	.	.	.	N	.	.	.	.	.
A/Nairobi/59/2009	.	R	.	.	.	.	.	.	.	.	.	.
A/Malindi/238/2009	.	.	.	.	.	.	N	T	D	.	G	.
A/Eldoret/119/2009	.	R	.	.	.	.	.	.	.	.	.	.
A/Eldoret/120/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Embu/169/2009	.	.	K	.	.	.	.	.	.	.	G	.
A/Nyeri/478/2009	.	.	.	H	.	.	.	.	.	.	.	.
A/Meru/467/2009	.	.	.	H	.	.	N	.	.	.	.	.
A/Kitale/531/2009	.	.	.	.	.	.	N	T	D	C	G	.
A/Kitale/532/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/512/2009	.	.	.	.	.	.	.	T	D	.	.	.
A/Nakuru/192/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Trans-Nzoia/168/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/179/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/148/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/91/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/27/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/37/2009	A	.	.	.	.	.	.	.	.	.	.	.
A/Kakamega/215/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisii/205/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisii/143/2009	.	.	.	.	.	.	.	.	.	.	.	K
A/Kisii/115/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kikuyu/184/2009	.	.	.	.	.	.	.	.	.	C	G	.
A/Kikuyu/185/2009	.	.	.	.	.	T	.	T	D	C	G	.
A/Nairobi/3/2009	.	R	.	.	.	.	.	.	.	.	.	.
A/Nairobi/16/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/11/2010	.	.	.	.	.	.	N	T	D	.	.	.
A/Kijabe/16/2010	.	.	.	.	.	T	.	.	.	.	G	.
A/Nairobi/20/2010	.	.	.	.	.	.	.	.	D	.	.	K
A/Nairobi/21/2010	.	R	.	.	.	.	.	.	.	.	.	.
A/Nairobi/24/2010	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/25/2010	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/72/2010	.	.	.	H	.	.	.	.	.	.	.	.
A/Nairobi/80/2010	.	.	.	H	.	.	.	.	.	.	.	.
A/Nairobi/97/2010	.	.	.	H	.	.	.	.	.	.	.	.

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.4.3 PA

The PA protein of the local isolates had 99.4-100% similarity with the prototype strain. Amino acid differences between the PA of the study samples and the vaccine strain is shown in Table 6. Two of the earliest isolates had an M581L substitution which is a clade 2 defining mutation. This mutation was not observed in any other Kenyan isolate during the pandemic. All Kenyan isolates shared a common signature in the P224S substitution. This substitution has been found worldwide in majority of non clade 1 influenza A(H1N1)pdm09 viruses. Three substitutions namely L425F, L589F and I690M arose during the peak phase of the pandemic but these replacements did not persist into the late phase of the pandemic. In the late phase of the pandemic, substitutions N321K and I330V which appeared as parallel mutations became dominant.

Table 6: Amino acid sequence variations of PA protein influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus isolate	PA amino acid position							
	5	224	321	330	425	581	589	690
A/California/07/2009	**V**	**P**	**N**	**I**	**L**	**M**	**L**	**I**
A/Kisumu/56/2009	.	S	.	.	.	L	.	.
A/Kisumu/61/2009	L	S	.	.	.	L	.	.
A/Nairobi/64/2009	.	S	.	.	.	.	.	.
A/Garissa/78/2009	.	S	.	.	.	.	.	.
A/Keiyo/96/2009	.	S	.	.	.	.	.	.
A/Nairobi/58/2009	L	S	.	.	.	.	.	.
A/Nairobi/59/2009	.	S	.	.	.	.	.	.
A/Malindi/238/2009	.	S	.	.	.	.	.	.
A/Eldoret/119/2009	.	S	.	.	.	.	.	.
A/Eldoret/120/2009	.	S	.	.	.	.	.	.
A/Embu/169/2009	.	S	.	.	.	.	.	.
A/Meru/467/2009	.	S	.	.	.	.	.	.
A/Nyeri/478/2009	.	S	.	.	.	.	F	M
A/Kitale/531/2009	.	S	.	.	.	.	.	.
A/Kitale/532/2009	.	S	.	.	.	.	F	M
A/Nakuru/192/2009	.	S	.	.	F	.	.	.
A/Trans-Nzoia/168/2009	.	S	.	.	.	.	.	M
A/Mombasa/179/2009	.	S	.	.	.	.	.	M
A/Nairobi/37/2009	L	S	.	.	F	.	.	.
A/Mombasa/91/2009	.	S	.	.	.	.	.	.
A/Kisii/115/2009	.	S	.	.	F	.	.	.
A/Kisii/143/2009	.	S	.	.	.	.	.	.
A/Mombasa/148/2009	.	S	.	.	.	.	.	.
A/Mombasa/27/2009	.	S	.	.	.	.	.	.
A/Kikuyu/184/2009	.	S	.	.	.	.	.	M
A/Kikuyu/185/2009	.	S	.	.	.	.	.	.
A/Kisii/205/2009	.	S	.	.	F	.	.	.
A/Kakamega/215/2009	.	S	.	.	.	.	.	.
A/Mombasa/512/2009	.	S	.	.	.	.	F	M
A/Nairobi/16/2009	.	S	.	.	.	.	.	M
A/Nairobi/3/2009	.	S	.	.	.	.	.	.
A/Nairobi/20/2010	.	S	K	V	.	.	.	.
A/Nairobi/21/2010	.	S	K	V	.	.	.	.
A/Nairobi/24/2010	.	S	K	V	.	.	F	.
A/Nairobi/25/2010	.	S	K	V	.	.	F	.
A/Nairobi/72/2010	.	S	K	V	.	.	F	.
A/Nairobi/80/2010	.	S	K	V	.	.	.	.
A/Nairobi/97/2010	.	S	K	.	.	.	.	.
A/Nairobi/11/2010	L	S	.	.	.	.	.	.
A/Kijabe/16/2010	.	S	K	V	.	.	.	.

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.4.4 HA

The HA protein of local samples was compared with that of the vaccine strain (Table 7). The Kenyan samples had an HA amino acid similarity of 98.6-99.5% in comparison to the vaccine strain . All the Kenyan samples had 100S and 338V both of which are non clade 1 amino acid signatures. At the same time, all Kenyan samples except two of the five earliest isolates had an S220T substitution. These three substitutions at positions 100, 220 and 338 are clade 7 defining. Furthermore, 100S and 338V were present at the beginning of the spread of the virus in Kenya and were maintained throughout the study period. Amino acid position 220 is located within antigenic site Ca1, while the substitution at position 338 is outside the major defined HA antigenic epitopes**.**

Several mutations were observed in varying proportions during the peak of the pandemic but these did not persist into the late phase. These include F12Y, H54P, A65S, W77R, N146D, D239E and F309I. Amino acid position 239 lies within antigenic site Ca2 of the HA1 protein. In the late phase of the pandemic, the frequency of D114N, S202T, E391K and S468N substitutions which had not been observed earlier increased. D114N and 202T seem to be parallel mutations. The amino acid position 202 is located within the HA1 antigenic site Sb.

Table 7: Amino acid sequence variations of HA protein of influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus isolate	HA Amino acid position																
	12	21	54	65	77	100	114	146	202	220	239	281	309	314	338	391	468
A/California/07/2009	**F**	**C**	**H**	**A**	**W**	**P**	**D**	**N**	**S**	**S**	**D**	**G**	**F**	**P**	**I**	**E**	**S**
A/Kisumu/56/2009	.	.	P	.	.	S	.	.	.	.	.	.	.	.	V	.	.
A/Kisumu/61/2009	.	.	.	.	.	S	.	.	.	.	.	.	.	.	V	.	.
A/Nairobi/64/2009	.	.	P	S	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Garissa/78/2009	.	.	P	S	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Keiyo/96/2009	.	.	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Nairobi/58/2009	.	.	P	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Nairobi/59/2009	.	W	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Malindi/238/2009	.	.	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Eldoret/119/2009	.	.	P	.	.	S	.	.	.	T	.	.	I	.	V	.	.
A/Eldoret/120/2009	.	.	P	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Embu/169/2009	.	.	P	S	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Nyeri/478/2009	.	.	P	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Meru/467/2009	.	W	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Kitale/531/2009	Y	.	P	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Kitale/532/2009	Y	.	.	S	.	S	.	D	.	T	.	.	I	.	V	.	.
A/Mombasa/512/2009	.	.	P	.	.	S	.	.	.	T	.	.	I	.	V	.	.
A/Nakuru/192/2009	Y	.	P	S	.	S	.	D	.	T	.	.	.	.	V	.	.
A/Transnzoia/168/2009	Y	.	.	S	.	S	.	D	.	T	.	.	.	.	V	.	.
A/Mombasa/179/2009	.	.	.	S	R	S	.	D	.	T	.	.	.	.	V	.	.
A/Mombasa/148/2009	Y	.	.	S	.	S	.	D	.	T	.	.	I	.	V	.	.
A/Mombasa/91/2009	.	.	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Mombasa/27/2009	.	.	.	.	.	S	.	.	.	T	E	.	.	.	V	.	.
A/Nairobi/37/2009	.	.	.	S	R	S	.	D	.	T	E	.	.	.	V	.	.
A/Kakamega/215/2009	.	.	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Kisii/205/2009	.	.	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Kisii/143/2009	.	W	.	.	R	S	.	.	.	T	.	.	.	.	V	.	.
A/Kisii/115/2009	.	.	.	S	.	S	.	D	.	T	.	.	.	R	V	.	.
A/Kikuyu/184/2009	.	.	P	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Kikuyu/185/2009	.	.	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Nairobi/3/2009	.	.	.	.	.	S	.	.	.	T	.	.	.	.	V	.	.
A/Nairobi/16/2009	.	.	P	.	.	S	.	.	.	T	.	.	.	R	V	.	.
A/Kijabe/16/2010	.	W	.	.	.	S	.	.	.	T	.	.	.	.	V	K	.
A/Nairobi/11/2010	.	.	.	.	.	S	.	.	.	T	.	.	.	R	V	K	.
A/Nairobi/20/2010	.	.	.	.	.	S	N	.	T	T	.	.	.	.	V	K	N
A/Nairobi/21/2010	.	.	.	.	.	S	.	.	.	T	.	A	.	.	V	K	N
A/Nairobi/24/2010	.	.	.	.	.	S	N	.	T	T	.	.	.	.	V	K	N
A/Nairobi/25/2010	.	.	.	.	.	S	N	.	T	T	.	A	.	.	V	K	N
A/Nairobi/72/2010	.	.	.	.	.	S	N	.	T	T	.	.	C	.	V	K	N
A/Nairobi/80/2010	C	.	.	.	.	S	N	.	T	T	.	.	.	.	V	K	N
A/Nairobi/97/2010	.	.	.	.	.	S	N	.	T	T	.	.	.	.	V	K	N

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.4.5 NP

Amino acid alignment of the NP of Kenyan isolates against the vaccine strain was performed, and the results are summarized in Table 8. The amino acid differences of the NP of Kenyan isolates were Y10F, E18D, I41N, V100I, Q149H, T373I, L381I, R416G, A423S, M426K, A428P, V444F, E454K and M481L. Among these amino acid differences, T373I observed in two of the earliest five isolates from Kenya is of significance since it is a clade II defining mutation. At the same time, the V100I which was found in 38 (95%) of all the Kenyan isolates is a clade IV-VII defining mutation. Besides V100I, no other mutation became fixed in this protein in the Kenyan isolates. Overall, the Kenyan isolates shared a 98.8-99.6% amino acid homology with the vaccine strain.

Table 8: Amino acid sequence variations of NP protein of influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus Isolate	NP Amino acid position													
	10	18	41	100	149	373	381	416	423	426	428	444	454	481
A/California/07/2009	**Y**	**E**	**I**	**V**	**Q**	**T**	**L**	**R**	**A**	**M**	**A**	**V**	**E**	**M**
A/Kisumu/56/2009	.	.	.	.	.	I	.	.	.	.	.	.	.	.
A/Kisumu/61/2009	.	.	.	.	.	I	.	.	.	.	.	.	.	.
A/Nairobi/64/2009	.	D	.	I	.	.	.	.	.	.	.	.	.	L
A/Garissa/78/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Keiyo/96/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/58/2009	.	.	.	I	.	.	.	G	.	K	.	.	.	L
A/Nairobi/59/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Malindi/238/2009	F	D	.	I	.	.	.	.	.	.	.	.	.	L
A/Eldoret/119/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Eldoret/120/2009	.	.	.	I	.	.	I	.	.	.	.	.	.	.
A/Trans-nzoia/168/2009	.	.	.	I	.	.	.	G	.	.	.	.	.	.
A/Embu/169/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nakuru/192/2009	.	.	.	I	.	.	I	.	.	.	.	.	.	L
A/Nyeri/478/2009	F	D	.	I	.	.	.	.	.	.	.	.	.	L
A/Meru/467/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Kitale/531/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Kitale/532/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	L
A/Mombasa/512/2009	.	.	.	I	.	.	.	.	S	.	.	F	.	.
A/Mombasa/179/2009	.	.	.	I	.	.	.	.	S	.	P	.	K	.
A/Mombasa/148/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Mombasa/91/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Mombasa/27/2009	.	.	.	I	.	.	.	.	.	K	.	.	.	.
A/Nairobi/37/2009	.	.	.	I	H	.	.	.	.	.	.	.	.	.
A/Kakamega/215/2009	.	.	.	I	.	.	.	.	.	.	.	.	K	.
A/Kisii/205/2009	.	.	.	I	.	.	.	.	S	.	P	.	K	.
A/Kisii/143/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Kisii/115/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	L
A/Kikuyu/184/2009	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Kikuyu/185/2009	.	.	.	I	.	.	.	.	.	.	.	F	K	.
A/Nairobi/3/2009	.	.	.	I	.	.	.	.	S	.	.	F	.	.
A/Nairobi/16/2009	.	.	N	I	H	.	.	.	.	.	.	.	.	.
A/Kijabe/16/2010	.	.	N	I	.	.	.	.	.	.	.	.	.	L
A/Nairobi/20/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/21/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/24/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/25/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/72/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/80/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/97/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.
A/Nairobi/11/2010	.	.	.	I	.	.	.	.	.	.	.	.	.	.

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.4.6 NA

The number of amino acid differences in NA between the Kenyan isolates and the vaccine strain (A/California/07/2007) ranged from 2 to 5. This is shown in Table 9. The frequency of isolates with cumulative numbers of amino acid differences (equal to and greater than 4) was generally higher in the late phase of the pandemic than during the earlier periods. The NA protein showed very few mutations in the transmembrane and linker region but the catalytic neuraminidase domain harbored certain mutations which have become fixed over the study period. Both V106I and N248D amino acid substitutions occurred early in the pandemic and persisted throughout the pandemic. The two together with V100I in the NP and S220T in the HA are clade VII defining mutations. The I396T substitution which appeared in several isolates during the pandemic peak phase was not observed during the late pandemic phase. G11V and N369K mutations were seen in increasing prevalence in the late pandemic phase and were the only ones that emerged during this phase.

Table 9: Amino acid sequence variations of NA protein of influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus isolate	NA amino acid position													
	8	11	25	106	248	262	277	329	369	386	396	460	461	463
A/California/07/2009	**I**	**G**	**Q**	**V**	**N**	**K**	**E**	**N**	**N**	**N**	**I**	**G**	**A**	**L**
A/Kisumu/56/2009	.	.	.	.	.	R	.	I	.	.	.	.	.	.
A/Kisumu/61/2009	.	.	.	.	.	.	.	I	.	.	.	.	.	.
A/Nairobi/64/2009	.	.	L	I	D	.	Q	.	.	.	.	V	.	.
A/Garissa/78/2009	.	.	.	I	D	R	.	.	.	.	.	.	.	.
A/Keiyo/96/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Nairobi/58/2009	.	.	.	I	D	.	.	.	.	.	T	.	.	.
A/Nairobi/59/2009	.	.	.	I	D	.	.	.	.	.	T	.	.	.
A/Malindi/238/2009	L	.	L	I	D	.	.	.	.	.	.	.	.	.
A/Eldoret/119/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	F
A/Eldoret/120/2009	L	.	.	I	D	.	Q	.	.	.	.	.	.	.
A/Embu/169/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	W
A/Nyeri/478/2009	.	.	.	I	D	.	.	.	.	.	T	.	.	.
A/Meru/467/2009	.	.	.	I	D	.	.	.	.	.	T	.	.	F
A/Kitale/531/2009	.	.	.	I	D	.	.	.	.	.	.	R	.	.
A/Kitale/532/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Mombasa/512/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Nakuru/192/2009	.	.	.	I	D	.	.	.	.	.	T	.	.	F
A/Trans-Nzoia/168/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Mombasa/179/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Mombasa/148/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Mombasa/91/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Mombasa/27/2009	.	.	.	I	D	.	.	.	.	.	.	.	P	.
A/Nairobi/37/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Kakamega/215/2009	.	.	.	I	D	.	.	.	.	.	T	.	.	.
A/Kisii/205/2009	.	.	.	I	D	.	.	.	.	D	.	R	.	.
A/Kisii/143/2009	.	.	.	I	D	.	.	.	.	.	T	.	.	.
A/Kisii/115/2009	.	.	.	I	D	.	.	.	.	D	.	.	.	F
A/Kikuyu/184/2009	V	.	.	I	D	.	.	.	.	.	.	.	P	.
A/Kikuyu/185/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Nairobi/3/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Nairobi/16/2009	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Nairobi/11/2010	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Kijabe/16/2010	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Nairobi/20/2010	.	.	.	I	D	.	.	.	K	.	.	.	.	.
A/Nairobi/21/2010	.	V	.	I	D	.	.	.	K	.	.	.	.	.
A/Nairobi/24/2010	.	V	.	I	D	.	.	.	K	.	.	.	.	.
A/Nairobi/25/2010	.	V	.	I	D	.	.	.	K	.	.	.	.	.
A/Nairobi/72/2010	.	V	.	I	D	.	.	.	K	.	.	.	.	.
A/Nairobi/80/2010	.	.	.	I	D	.	.	.	.	.	.	.	.	.
A/Nairobi/97/2010	.	.	.	I	D	.	.	.	.	.	.	.	.	.

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.4.7 M

The two Matrix proteins were found to be highly conserved with amino acid homology of 98-100% with the vaccine strain. The M1 protein was generally more conserved (92-100%) than the M2 protein (77-100%). The M1 protein had amino acid substitutions at 7 sites while the shorter M2 protein had amino acid substitutions at 10 sites. These results are shown in Table 10. Isolates from the late phase of the pandemic generally had more cumulative number of amino acid differences to the vaccine strain than isolates from the early or peak phases of the pandemic.

In the late phase of the pandemic, the substitution V80I in the M1 protein was the most common among the 2010 Kenyan samples. The other substitutions in this protein (R78P, Q191K, Q241H and M244R) were mainly transient and did not become fixed. Amino acids at positions 64 and 91 in the M2 located in the extracellular domain were most mutable.

Table 10: Amino acid sequence variations of M proteins of influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus Isolate	M1 Amino acid position							M2 Amino acid position								
	43	78	80	191	233	241	244	12	15	44	53	58	64	72	82	91	96
A/California/07/2009	**M**	**R**	**V**	**Q**	**L**	**Q**	**M**	**R**	**W**	**D**	**R**	**G**	**S**	**M**	**S**	**F**	**L**
A/Kisumu/56/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisumu/61/2009	L	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
A/Nairobi/64/2009	L	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
A/Garissa/78/2009	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
A/Keiyo/96/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/58/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/59/2009	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
A/Malindi/238/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Eldoret/119/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Eldoret/120/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Embu/169/2009	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
A/Nyeri/478/2009	.	P	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.
A/Meru/467/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kitale/531/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kitale/532/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/512/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nakuru/192/2009	.	.	.	.	I	.	.	.	.	.	.	.	.	.	.	.	.
A/Trans-nzoia/168/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/179/2009	.	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.
A/Mombasa/148/2009	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.
A/Mombasa/91/2009	.	.	.	.	I	.	.	.	.	.	G	.	C	.	.	.	.
A/Mombasa/27/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	L	.
A/Nairobi/37/2009	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
A/Kakamega/215/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisii/205/2009	.	.	.	.	.	H	.	T	.	.	.	.	.	.	.	.	.
A/Kisii/143/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisii/115/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kikuyu/184/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kikuyu/185/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/3/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	L	.
A/Nairobi/16/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kijabe/16/2010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/11/2010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/20/2010	.	.	I	K	.	.	.	.	.	H	G	V	C	L	T	L	.
A/Nairobi/21/2010	.	P	I	.	.	H	.	T	.	H	G	V	C	L	T	L	.
A/Nairobi/24/2010	.	.	I	K	.	.	R	.	G	.	.	V	C	L	.	.	.
A/Nairobi/25/2010	.	P	I	.	.	.	.	.	.	.	G	V	C	.	T	L	V
A/Nairobi/80/2010	.	.	I	K	.	.	R	.	G	H	G	V	C	L	T	L	.
A/Nairobi/72/2010	.	.	I	.	.	H	.	T	.	.	.	.	.	L	T	L	.
A/Nairobi/97/2010	.	.	I	K	.	.	R	.	G	.	.	.	.	.	.	L	V

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.4.8 NS

A comparison of the NS1 and NEP proteins of the Kenyan isolates from this study and the vaccine strain revealed that there were substitutions at 6 and 4 amino acid positions in the NS1 and NS2/NEP proteins, respectively (Table 11). The NS1 protein was found to be more conserved (98-100%) than the NS2/NEP (94-100%) protein. The only fixed amino acid substitution occurred at position 123 in the NS1. This position is located within the effector domain of this protein. This change occurred as a result of a non-synonymous substitution at the 367–369 codon from ATC to GTC causing an isoleucine to valine substitution. The only isolates lacking this mutation A/Kisumu/56/2009 and A/Kisumu/61/2009 represent two of the five earliest isolates introduced in Kenya. The amino acid at position 77 of the NEP was the most mutable resulting in R77M and R77K mutants. Besides the I123V mutation, there were no other spatial or temporal specific mutations that were observed in this protein.

Table 11: Amino acid sequence variations of NS proteins influenza A(H1N1)pdm09 strains from Kenya compared with the vaccine strain

Virus isolate	NS1 amino acid position						NEP amino acid position			
	55	123	157	162	185	205	77	108	114	120
A/California/07/2009	**E**	**I**	**V**	**P**	**L**	**N**	**R**	**E**	**R**	**L**
A/Kisumu/56/2009	.	.	.	.	.	.	.	.	.	R
A/Kisumu/61/2009	.	.	.	.	.	.	K	.	.	.
A/Nairobi/64/2009	.	V	.	.	.	.	M	.	.	.
A/Garissa/78/2009	.	V	.	.	.	.	K	.	G	.
A/Keiyo/96/2009	.	V	.	.	.	.	.	.	.	.
A/Nairobi/58/2009	.	V	.	.	.	.	.	.	.	.
A/Nairobi/59/2009	.	V	.	.	.	.	.	.	.	.
A/Malindi/238/2009	.	V	.	.	.	.	.	.	G	.
A/Eldoret/119/2009	.	V	.	.	.	.	.	.	.	.
A/Eldoret/120/2009	.	V	.	.	.	.	K	.	.	.
A/Embu/169/2009	.	V	.	L	.	.	.	.	.	.
A/Nyeri/478/2009	.	V	.	.	.	.	K	.	.	R
A/Meru/467/2009	.	V	.	.	.	.	M	.	.	.
A/Kitale/531/2009	.	V	.	.	.	.	.	.	.	.
A/Kitale/532/2009	.	V	.	.	.	.	M	.	.	.
A/Mombasa/512/2009	.	V	.	.	.	D	.	.	.	.
A/Nakuru/192/2009	.	V	.	.	.	.	.	.	.	.
A/Trans-nzoia/168/2009	.	V	.	.	.	.	.	.	.	R
A/Mombasa/179/2009	.	V	.	.	.	.	.	.	.	.
A/Mombasa/148/2009	.	V	.	L	F	.	.	D	.	.
A/Mombasa/91/2009	.	V	.	.	.	D	.	.	.	.
A/Mombasa/27/2009	.	V	.	.	.	D	.	.	.	.
A/Nairobi/37/2009	.	V	I	.	.	.	.	.	.	.
A/Kakamega/215/2009	.	V	.	.	.	.	.	.	.	.
A/Kisii/205/2009	.	V	.	L	F	.	.	.	.	.
A/Kisii/143/2009	.	V	.	.	.	.	.	.	.	.
A/Kisii/115/2009	.	V	.	.	.	.	M	.	.	.
A/Kikuyu/184/2009	K	V	.	.	.	.	.	D	.	.
A/Kikuyu/185/2009	K	V	.	.	.	.	.	D	.	.
A/Nairobi/3/2009	.	V	I	.	.	.	M	.	.	.
A/Nairobi/16/2009	.	V	.	.	.	.	.	.	.	.
A/Nairobi/11/2010	.	V	.	.	.	.	.	.	.	.
A/Kijabe/16/2010	.	V	.	.	.	.	.	.	.	.
A/Nairobi/20/2010	.	V	.	.	.	.	.	.	.	.
A/Nairobi/21/2010	.	V	.	.	.	.	.	.	.	.
A/Nairobi/24/2010	.	V	.	.	.	.	.	.	.	.
A/Nairobi/25/2010	.	V	.	.	.	.	.	.	.	.
A/Nairobi/80/2010	.	V	.	.	.	.	.	.	.	.
A/Nairobi/72/2010	.	V	.	.	.	.	.	.	.	.
A/Nairobi/97/2010	.	V	.	.	.	.	.	.	.	.

Early pandemic samples are shown in red, peak pandemic samples in black and late pandemic samples in blue. The vaccine strain is shown in bold.

4.5 Analysis of antiviral drug-associated mutations

The susceptibility of the Kenyan isolates to anti-influenza drugs in current use was determined.

4.5.1 Sensitivity to neuraminidase inhibitors

To determine the presence of genetic markers associated with susceptibility to neuramindase inhibitors, a multiple sequence alignment of the NA protein of Kenyan viruses and two NA proteins of a resistant and sensitive influenza A(H1N1)pdm09 virus was conducted. A multiple protein sequence alignment of the Kenyan isolates from this study is shown in Figure 18. The alignment showed that amino acids that have been shown confer resistance to these drugs such as I117V, E119V, D198N, I222V, H274Y, R292K, N294S, I314V, V116A, R118K, E119G/A/D, Q136K, D151E, R152K, R224K, E276D, R292K and R371K [N2 numbering]) were found to be conserved among all the isolates except in two isolates that had an E277Q (N1 numbering) substitution. All the local isolates were therefore found to be sensitive to neuraminidase inhibitors.

Figure 18: A Multiple sequence alignment of the Neuraminidase protein amino acid positions 270-300.

The alignment includes NA protein of A/California/07/2009 (highlighted in red) which is sensitive to Oseltamivir and A/Sydney/DD3-48/2010 (highlighted in blue) which is resistant to Oseltamivir. Amino acid position 275 associated with Oseltamivir resistance is shown in red. The Kenyan isolates are not highlighted and they lack the Oseltamivir resistance conferring mutation H275Y as shown. Other neuraminidase inhibitor resistance conferring mutations are also absent.

4.5.2 Sensitivity to Adamantanes

To determine the presence of genetic markers associated with susceptibility to Adamantane drugs, a multiple sequence alignment of the M2 protein of Kenyan viruses and two M2 proteins of a resistant and sensitive virus was conducted (Figure 19). Since there is currently no adamantine sensitive influenza A(H1N1)pdm09 virus, the M2 protein of an influenza A/H3N2 virus (A/New-York/32/2003) which is sensitive to adamantine drugs was used. Analysis of sensitivity to adamantanes showed that all the local isolates were resistant to these drugs. All the local virus strains had the amino acid 31N in their M2 protein which confers them with resistance to all M2 inhibitors. Amino acid positions 27 and 30 also involved in resistance were also conserved in the Kenyan viruses.

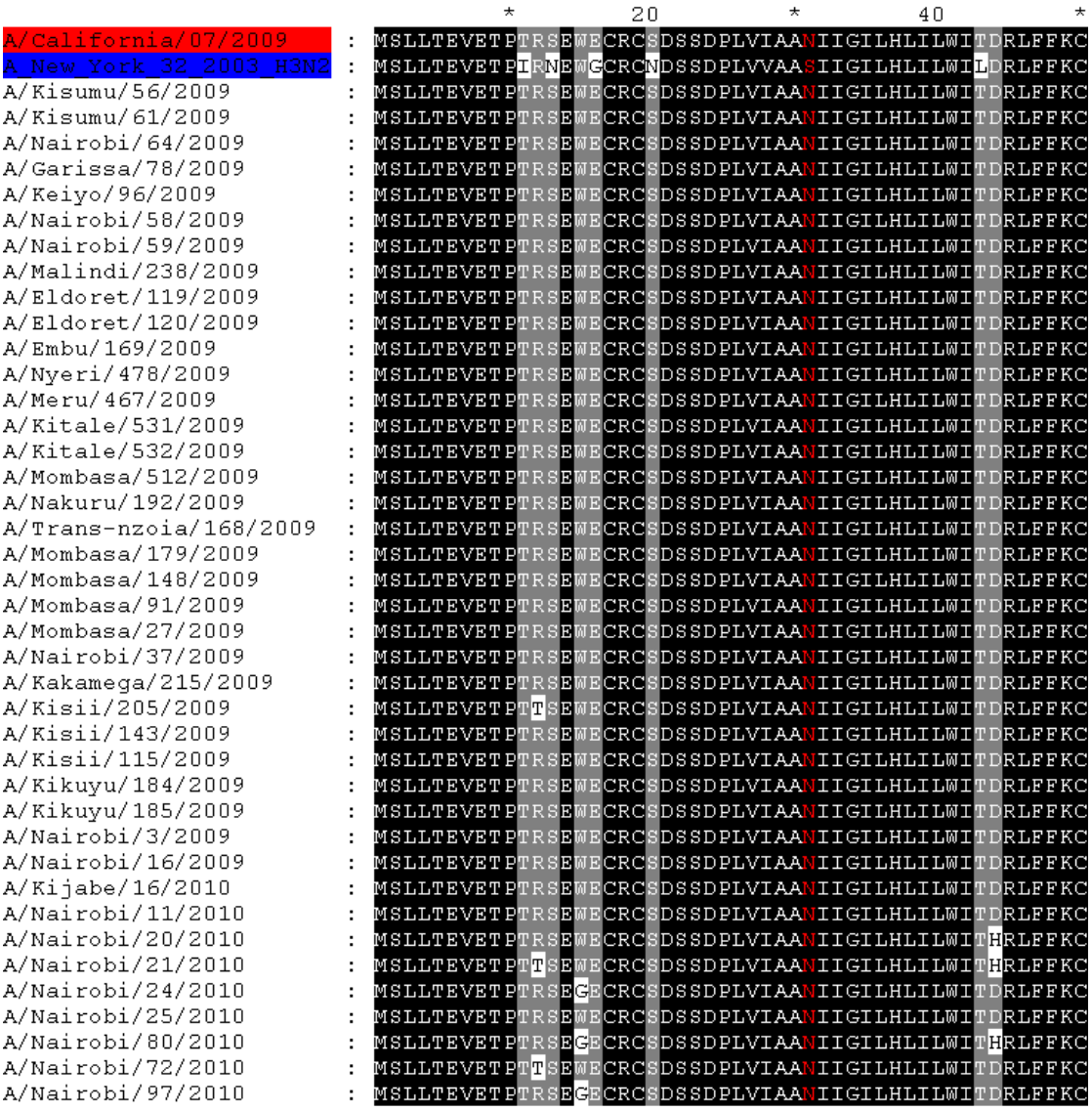


Figure 19: Multiple sequence alignment of the M2 protein amino acid positions 1-50.

The alignment includes M2 protein of A/California/07/2009 (highlighted in red) which is adamantane resistant and A/New-York/32/2003 [H3N2] (highlighted in blue) which is adamantane sensitive. Amino acid position 31 associated with adamantine resistance is shown in red. The Kenyan isolates are not highlighted and they possess the adamantane resistance conferring mutation S31N.

5.0: DISCUSSION

Influenza A has the distinction of being an old disease, a recurring disease, and an ‘emerging’ disease thus presenting constant challenge to influenza scientists. This prompted the great influenza scientist Edwin Kilbourne to note that (the) "student of influenza is constantly looking back over his shoulder and asking 'what happened?' in the hope that understanding of past events will alert him to the catastrophes of the future" (Kilbourne, 1973). The current study attempts to accomplish this by employing a sophisticated Bayesian evolutionary framework to understand the molecular evolution of influenza A(H1N1)09pdm virus in Kenya. Molecular phylogenies can reveal many aspects of the transmission, epidemiology, and evolution of rapidly evolving pathogens (Pybus and Rambaut, 2009). Analysis of influenza virus genomes during the emergence of influenza A(H1N1)09pdm virus, causing the first influenza pandemic in 40 years therefore provides a unique opportunity to track the transmission dynamics of a new influenza virus in an immunologically naïve population (Baillie *et al.*, 2012).

The application of complete virus genome sequencing and analysis has already provided detailed insights into seasonal influenza virus infections (Ghedin *et al.*, 2005; Rambaut *et al.*, 2008). This analysis facilitates comparison and understanding of the evolutionary dynamics of circulating viruses and the prediction of potential evolution events that are likely to result in new strains (Greninger *et al.*, 2010). It also allows closer examination of the importance of

other genes in influenza outbreaks and vaccine selection. However, despite the availability of complete influenza genome data from several developed countries, African countries still provide incomplete genomes into the public domain. Most influenza sequencing in these countries has focused on the HA1 domain of the haemagglutinin gene where mutations have the greatest effect on the antigenic structure (Barr *et al.*, 2010). In fact, by January 2013, there were only eight complete genomes of influenza A(H1N1)pdm09 virus from Africa that were publicly available in the Influenza Virus Resource database (Bao *et al.*, 2008)

This current study uses whole genome sequences combined with sophisticated Bayesian evolutionary framework for the molecular characterisation and reconstruction of the phylodynamics of the A(H1N1)09pdm influenza virus in Kenya during the pandemic period. In order to describe the Kenyan pandemic situation in the setting of the widespread diffusion of infection throughout the world, the study analysed full genomes of 40 newly characterised local virus isolates and a series of reference isolates from different countries globally retrieved from public databases. First, the individual gene segments and the concatenated genomes were analyzed to infer phylogenetic relationships. This was followed by a reconstruction of the Kenyan-specific transmission dynamics of the novel virus in the country. Lastly, the study analyzed amino acid polymorphisms in each of the virus proteins to determine the trends in genetic drift and antiviral sensitivity.

5.1 Phylogenetic relationships of Kenyan influenza A(H1N1)pdm09 viruses

Phylogenetic trees for each of the gene segments of the Kenyan and selected global influenza A(H1N1)pdm09 viruses were constructed using Bayesian inference and are shown in Figs 7-15. The topologies of these trees are generally similar. Even when the Neighbor Joining method was used, the observed topologies did not change. There was observed in most of the genes a clustering together of the 2010 samples which was not observed in samples from the early and peak phases of the pandemic. This is indicative of circulation of a distinct lineage of viruses during this phase. Such temporal evolution has been observed in several studies elsewhere (Mullick *et al.,* 2011; Venter *et al.,* 2012). The topology also did not show any spatial clustering in the local viruses. An earlier study however had shown global spatial diversification patterns (Nelson *et al.,* 2009). This discordance may be attributed to the shorter sampling period used in the early studies and the fewer number of sequences available.

An examination of the phylogeny of individual gene segments also showed that the Kenyan isolates sampled during the pandemic season were interspersed with sequences sampled from other countries. This is probably indicative of multiple independent introductions of the virus into the country as opposed to a single introduction (or a few introductions) followed by clonal expansion (*in situ* evolution). This observation is in agreement with recent studies aiming to understand short term evolution of influenza A viruses (Nelson *et al.,* 2006;

Rambaut *et al.,* 2008). These studies suggest that the evolution of influenza A viruses is maintained by continuous importation of viral lineages of the same subtype during the time course of the influenza season. The resultant effect of this viral immigration is to produce genetic diversity of influenza A viruses within season thus providing an abundant raw material for reassortment. This is important to ensure evasion of host immunity by the virus leading to continuous transmission.

Extensive phylogenetic analysis based on concatenated whole genome sequences of representative influenza A(H1N1)pdm09 viruses sampled globally from different geographical regions revealed seven distinct clades 1-7 (Nelson *et al.,* 2009). The vaccine strain A/California/07/2009 belongs to clade 1. Clade 2 is represented by A(H1N1)pdm09 viruses isolated from California, Canada, Netherlands, and United States. Clade 3 is represented by viruses isolated from England, Russia, China, and the United States. Clade 4 is represented by viruses isolated from two East Asian countries, Korea and Japan. Clade 5 is represented by viruses isolated from Canada, China, Japan, the United States (mainly Wisconsin isolates) along with India. Clade 6 is represented by viruses isolated from China, Japan with new additions from Taiwan, Thailand, India and United States. The clade 7, which is the largest clade is represented by viruses isolated from Japan, Mexico, China, Asia and several states of the USA (Sharma *et al.,* 2013).

Using global isolates whose clade identities are known, the phylogenetic analysis was able to show that the initial introductions of the influenza

A(H1N1)09pdm virus in the country belonged to two different global clades namely 2 and 7 (Fig. 15). Besides differences in geographical distribution, these two clades differ genetically in the mutations they harbour relative to the prototype A/California/7/09 strain. Clade 2 viruses have the M581L and T373I mutations in the PA and NP proteins respectively while clade 7 viruses lack these mutations but possess the mutations S220T in their HA, V100I in their NP, V106I and N248D in their NA and I123V in their NS1 (Nelson *et al.,* 2009). In this study, clade 2 viruses were not isolated beyond the initial introduction cases. The absence of clade 2 viruses beyond the introduction cases is worth noting and this may mean that this particular clade did not acquire adequate fitness for sustenance.

On the other hand, clade 7 viruses were isolated throughout the pandemic period in Kenya. This dominance of circulation of clade 7 viruses throughout the pandemic period and the disappearance of all other clades have been observed in other studies in United Kingdom, Argentina, India and United States of America (Baillie *et al.,* 2012; Barrero *et al.,* 2011; Mullick *et al.,* 2011; Nelson *et al.,* 2011). While clade 7 viruses have dominated globally, they have also been the most diverse and their genetic diversity may have arisen from reassortment among clades considering that they have co-circulated with all the other clades. This predominance of clade 7 viruses is hypothesized to be due to its fitness and adaptability and more efficient human transmissibility (Potdar *et al.,* 2010). However, this requires further

investigation because currently there is no proof that clade 7 viruses are fitter or more adapted to humans than other clades.

Why were clades 1, 3, 4, 5 and 6 not detected in the local influenza A(H1N1)09pdm virus transmission? This is more puzzling considering that the introduction of influenza A(H1N1)09pdm virus into the country was from the UK where it was observed that all these clades except clade 4 circulated (Baillie *et al.,* 2012). While clade 1 viruses did not circulate globally beyond May 2009, it would have been expected that clades 3, 5 and 6 which were circulating in the UK at the same time that clade 2 and 7 viruses were imported into Kenya would also have been imported. It should however be noted that this finding is not unique to Kenya. Studies in several countries observed circulation of only one or two clades in their local transmission (Barrero *et al.,* 2011; Goni *et al.,* 2012; Lycett *et al.,* 2012). Perhaps the earlier clades were not fit enough to allow sustained human to human transmission thus restricting their spread.

5.2 Evolutionary rate and ancestral time analyses of local A(H1N1)pdm09 viruses

Using sequence data to infer population dynamics is playing an increasing role in the analysis of outbreaks such as the influenza A(H1N1)09pdm virus. With the growth of faster and more reliable sequencing technologies (and consequently the availability of genetic data), there have been a number of

statistical and computational innovations to analyse this proliferation of sequence data (de Silva *et al.,* 2012). One successful application has been the analysis of epidemic trends using the so-called phylodynamic methods that use pathogen sequences to infer pathogen diversity, and the changing number of infected individuals as well as more subtle effects on pathogen selection and population structure (Grenfell *et al.,* 2004).

The pathogen diversity of viruses like influenza A which produces acute infections with partial cross-immunity is a result of an interplay between genotypic diversity and host immunity. Studies on evolutionary genetics of viral emergence of RNA viruses indicate that the most important driver of their evolution (and thus their genetic diversity) is their capacity for mutation (Holmes, 2010). The vast majority of estimates of mutation rates in RNA viruses are in the range of 0.1 to 1.0 mutations per genome, per replication. These rates are several orders of magnitude higher than those in most DNA-based organisms (Duffy *et al.,* 2008). Such remarkable error rates are largely due to replication using a low-fidelity RNA-dependent RNA polymerase (Holland *et al.,* 1982). High mutation rates, coupled with rapid replication, are also the basis for the high rates of nucleotide substitution (fixation) recorded in RNA viruses (Holland *et al.,* 1982; Holmes, 2010; Jenkins *et al.,* 2002). Mean substitution rates in this case are usually in the realm of 10^{-3} to 10^{-4} nucleotide substitutions per site per year (subs/site/year) and hence some six orders of magnitude higher than those seen in eukaryotes (Jenkins *et al.,* 2002). Consequently, evolutionary rates of this magnitude are a major reason

clinically important traits, such as drug resistance, escape from vaccine coverage, and host range expansion, appear so readily in some RNA viruses. They are therefore the main drivers of antigenic variation in influenza viruses.

In order to reconstruct the population dynamics of the 2009 H1N1 pandemic in Kenya on a calendar time scale, the study first estimated the evolutionary rate of the Kenyan isolates using the relaxed (uncorrelated lognormal) clock implemented in the Bayesian framework. Using this framework, the complete genome of the local viruses had an estimated rate of nucleotide substitution of 4.9×10^{-3} substitutions/site/year. In 2009 Smith analyzed evolutionary rates during the early phase of the pandemic and found that the rate for the complete genome to be 3.66×10^{-3} with a range of $0.61\text{-}6.58 \times 10^{-3}$ (Smith *et al.*, 2009). Therefore, the observed rate of substitution for the complete genome for the Kenyan viruses is similar to that estimated during the early phase of the pandemic. The rate of substitutions for the individual gene segments of the Kenyan viruses in this study ranged from 0.8×10^{-3} substitutions/site/year in the NP gene to 9.8×10^{-3} substitutions/site/year in the M gene (Table 3). These rates observed are generally higher than previous estimations performed during the initial spread of these pandemic viruses in all the genes except in the NP. In studies carried out early in the pandemic, the evolutionary rate for individual gene segments was estimated at $2.55\text{–}3.67 \times 10^{-3}$ substitutions/site/year (Smith *et al.*, 2009). However, it needs to be noted that the elapsed time since the emergence of this virus is too short (12 months of sampling from July 2009-

period to June 2010). Thus this short period precludes confident prediction of evolutionary rate trends for this virus.

The surface expressed gene segments (HA and NA), the transmembrane expressed segment (M) and the NS gene segment had the highest rates of evolution compared to internally expressed segments (PB2, PB1, PA, and NP). Thus, the study identified a major evolutionary rate difference between those segments expressed on the surface of the virion (high rates) and those that are expressed inside the virion (low rates). This observation has also been shown previously among seasonal influenza viruses (Rambaut *et al.,* 2008).

What was the cause of this phenomenon in Kenyan viruses whereby the surface expressed segments had higher substitution rates while the internally expressed segments had lower substitution rates? Genes that code for surface proteins (HA and NA) may be subject to strong selection pressure by neutralizing antibodies of host immune systems (Webster *et al.,* 1992). These two genes are therefore subject to strong positive selection to allow for antigenic variation and thus immune escape of the virus accounting for their high evolutionary rate. The extracellular domain of the M2 ion channel is a key target in the development of a 'universal' influenza vaccine (Gerhard *et al.,* 2006), whereas the M1 protein is the major component of the viral capsid and a potential target of cellular immune responses (Berkhoff *et al.,* 2005). Likewise, NS1 encodes a pleiotropic, nonstructural protein that down-regulates double-stranded-RNA-induced antiviral responses (Krug *et al.,* 2003). Therefore, immune-mediated selection may also explain the increased rates in M and NS gene segments.

In contrast, the three segments that constitute the viral polymerase (PB2, PB1, PA), as well as the RNA-binding NP, were shown to exhibit lower evolutionary rates revealing stronger selective constraints against amino acid variation. Previous studies have identified that the activity of influenza polymerase is highly host- and cell type-specific and that compatibility among viral ribonucleoprotein (vRNP) complex proteins is necessary for optimal genome replication and transcription, and can be restrictive in reassortant influenza viruses (Boivin *et al.,* 2010; Gabriel *et al.,* 2008; Naffakh *et al.,* 2008). As a consequence, the polymerase genes are therefore not expected to undergo significant host-specific adaptive evolution.

To understand the population dynamics of this pandemic in the country, the study attempted to infer demographic changes during the pandemic from sequence data using the coalescent Bayesian Skyline Plot (BSP) using full genomic sequences. In common with all RNA viruses, influenza A viruses have a high mutation rate and short generation time such that genetic changes encode epidemiological information. The combination of these properties allow use of both sequence data and temporal information obtained early on in the pandemic to compute infection dynamics even while the pandemic is still progressing (de Silva *et al.,* 2012). Inferring demographic history during a pandemic using BSP involves estimating the genealogy and inferring the effective population size at different points along the genealogical timescale. The effective population size reflects the number of individuals that contribute offspring to the descendent generation.

Analysis of the BSP showed that the growth of the effective number of infections was exponential only during the first four months (July to September 2009) of the pandemic (Figure 17). Thereafter, constant population demography was maintained till the end of the study period. The exponentially growing effective population size suggests that the viral diversity was increasing before the peak of the pandemic (September 2009) in the country. These findings are consistent with surveillance data reported in the FluNet (http://www.who.int/influenza/gisrs_laboratory/flunet/en/), a global tool for influenza virological surveillance. FluNet surveillance data on Kenya shows that the number of infections with the pandemic virus peaked between September and October 2009 and thereafter declined. It has been hypothesized that the pandemic strain generated enough herd immunity through immunization by exposure in the first four months thereby suppressing its transmission. However, this immunity wanes with time allowing the virus to re-emerge. It is also possible for viruses to re-emerge if they have mutated enough to escape immunity generated against the 'parent' strain (Majanja *et al.*, 2013). Again this is consistent with FluNet's surveillance data that shows that the pandemic virus re-emerged in November 2010 and dominated for several months thereafter.

To identify possible temporal or spatial discrete lineages in influenza A(H1N1)09pdm viruses circulating in Kenya, the study conducted a maximum clade credibility tree analysis (i.e. the tree sampled from the Bayesian MCMC with the highest product of individual clade

probabilities). This aimed at seeking molecular insights into the trajectory of A/H1N1pdm following its introduction into the country. The chronological phylogeny for the Kenyan viruses revealed that six distinct A/H1N1pdm viral clusters co-circulated throughout the pandemic period. The tMRCA estimate of the tree root dating back to February 2009 is in line with the majority of the previous estimations, which place the origin of the pandemic H1N1 strain to January 2009 with intervals of credibility between late 2008 and March 2009 (Fraser *et al.,* 2009; Shiino *et al.*, 2010; Smith *et al.*, 2009). This pre-dates the first identification of A(H1N1)09pdm infections in Kenya and suggests that, the earliest Kenyan influenza A(H1N1)09pdm viruses were circulating elsewhere several months before their introduction and detection in Kenya. This observation has also been made in Italy and UK (Baillie *et al.,* 2012; Zehender *et al.,* 2012).

Using contact tracing methods, it was shown that influenza A(H1N1)09pdm virus was introduced in Kenya at three points; the first 2 points were in Western Kenya while the third was in Nairobi and that all these initial introductions were from the UK. (CDC, 2009a). Using molecular data, this study shows that the two introductions in Western Kenya were due to both clade 2 and clade 7 viruses while the introductions in Nairobi was due to a clade 7 virus. In the UK, the first laboratory confirmed cases were detected in April 2009 and the first wave of infections peaked in June - July 2009 (Baillie *et al.,* 2012). In agreement with this observation, the tMRCA's of the earliest clusters identified in Kenya was estimated to be between April and June 2009.

It is worth noting that the tMRCA of the Kenyan clade 2 viruses overlap with that of the UK clade 2 viruses (Baillie *et al.,* 2012). At the same time, both local clade 2 and 7 two viruses are closely associated with UK isolates as shown in the time scaled MCC tree shown in Figure 16 and not with isolates from USA, Asia or Africa. Molecular data presented in this study therefore confirms previous classical epidemiological work while at the same time shedding light on the genetic characteristics of the viruses introduced into the country.

The tMRCA's of the four clusters that co-circulated in the country during the peak pandemic phase was dated between June and August 2009. These clusters are also more closely associated with UK isolates than isolates from other regions. Furthermore, the tMRCA's of these clusters coincide and overlap with those of clusters identified in the UK (Baillie *et al.,* 2012). Similar results were obtained with cluster KENE-GC7 which dominated during the late phase of the pandemic in Kenya. The tMRCA for this cluster was estimated at October 2009. One UK isolate A/England/1116/2009 sampled on 21st October 2009 was found to be closely related with this cluster. Further, the fact that during this period the second wave of pandemic H1N1 infections in UK peaked is noteworthy. Taken together, this information strongly suggests that not only was the pandemic virus introduced into the country from the UK, but it was also sustained via multiple importations from the same country.

An interesting observation seen in this study is that though the introduction of influenza A(H1N1)09pdm virus occurred in June 2009, the pandemic peaked

in September/October. The reasons for this apparent lag between the introductions into the country and the peak of infections are not exactly known. However, what is clear is that the pandemic peaked in the country in September immediately after schools opened for their third term. This suggests that school closure for the August holiday and the movements associated with school holidays could have played a role in the intense circulation of the virus in the country. Similar observations have been made in the United States of America where elevated outpatient visits for influenza-like illness occurred an average of 14 days after schools opened in the fall of 2009 (Chao *et al.,* 2010). Studies have also previously shown that school activities have been found to be significantly correlated with influenza transmission rates (Cauchemez *et al.,* 2008; Monto *et al.,* 1985).

No clear signal of the geographical spread of A(H1N1)09pdm influenza virus was detected in this study. The study observed that all the circulating clusters were geographically spread and were not restricted to one sampling location. This may be due to the small sample size utilized. At the same time, this finding may be indicative of efficient community spread of these clusters throughout the country. Genetic drift however was visible over time, with the 2010 strains clustering separately from 2009 strains and further from the Mexico strain and the A/California/7/2009 vaccine strain forming a well supported clade. A recent study carried out in South Africa also showed rapid spread of circulating clusters throughout the country and also a progressive genetic drift among 2010 isolates (Venter *et al.,* 2012).

5.3 Genetic drift in the proteins of local A(H1N1)09pdm viruses

The most salient feature of influenza evolution in humans is its antigenic drift. This process is characterized by structural changes in the virus's B-cell epitopes and ultimately results in the ability of the virus to evade immune recognition and thereby reinfect previously infected hosts (Yuan and Koelle, 2013). Occasionally the degree of antigenic drift is sufficient that a very large proportion of the population is susceptible and severe epidemics occur in a proportion of the population that has become immunologically susceptible to the drift variants. In a pandemic situation if this was to happen, it would lead to repeated waves of infection some of which may be more virulent than others. It is for this reason that the study undertook surveillance of amino acid changes in local viruses over the pandemic period in all the genes.

In the PB2 protein of influenza A viruses, E627K and D701N mutations are associated with increased virulence and adaptation of avian influenza viruses to the human host (Herfst *et al.,* 2010). However, examination of the PB2 amino acid residues of the Kenyan A(H1N1)09pdm influenza viruses showed that these two mutations were absent during the study period. The polymerase complex of influenza A(H1N1)09pdm virus was derived from triple-reassortant swine viruses, the PB2 gene of which was of avian origin and entered pigs around 1998 (Garten *et al.,* 2009). The fact that this avian-origin PB2 gene did not significantly benefit from the substitutions E627K and D701N (locally and globally) which allow for enhanced replication in mammals suggests that other mutations may have compensated for their absence.

The K526R substitution in the PB2 protein was observed in the local isolates in only two of the earliest Kenyan isolates. An analysis of over 1,000 global PB2 protein sequences deposited in the Influenza Virus Resource database reveals that this substitution was only observed early in the pandemic. The absence of K526R substitution beyond the early phase suggests that K526 offers transmission or replication advantage over R526 in humans. This study hypothesizes that the acquisition of K526 could compensate for the absence of the E627K and D701N mutations associated with mammalian adaptation of avian influenza viruses. Using minireplicon and animal studies, a recent study on the mammalian adaptation of avian H5N1 viruses found that PB2 possessing 591R supports efficient viral replication in mammals, so that there is no strong selective pressure to acquire the mammalian-type amino acids at position 627 and 701 (Yamada *et al.,* 2010). These findings therefore support the hypothesis that a mutation in the C terminus of PB2 such as K526R could offer replication advantage in influenza A(H1N1)09pdm viruses.

Interestingly, most of the substitutions observed in the PB2 during the peak phase of the pandemic in Kenya were not observed elsewhere in the world except K157R. None of these mutations also became fixed in the Kenyan isolates. As would be expected, the increased transmission during this phase coupled with the error-prone nature of influenza RNA polymerase created some variants that were not able to transmit effectively. This could account for the lack of any fixed amino acid mutations during this period. In the late phase of the pandemic, four dominant mutations were observed in the Kenyan

isolates. These included R299K, V344M, I354L and N456S. An examination of global PB2 protein sequences from influenza A(H1N1)09pdm viruses isolated elsewhere indicates that these four mutations were observed increasingly from late 2009 suggesting that they offered a selective advantage. It is therefore expected that these mutations will increase in subsequent seasons particularly as they tended to occur as parallel mutations implying that they are beneficial and characteristic of instances of convergent evolution (Bull *et al.*, 1997).

PB1 is the core subunit for assembly of the virus RNA polymerase as the N-terminal tip of PB1 binds to the C terminus of PA while the C terminus of PB1 binds to the N terminus of PB2 (Toyoda *et al.*, 1996). In the current study, no fixed amino acid substitutions were observed in the PB1 during the pandemic amongst the Kenyan isolates. The main observation in the PB1 protein of viruses from this study was that most of the sites (7/12) harboring amino acid substitutions are located towards the C-terminus of the protein which associate with the PB2 protein by binding to the N-terminal end of the PB2 protein during replication. Mutations in each of the subunits comprising the influenza polymerase impact on the efficiency of viral replication by affecting RNA polymerase activity. It is therefore interesting to note that no mutations were observed in the N-terminal of PB2 in the local isolates implying that the higher rate of mutations seen in the C-terminal of PB1 may have been compensatory to allow for optimal RNA polymerase activity. This is more so as mutations at

the PB1-PB2 interface have been shown to inhibit RNA synthesis (Sugiyama *et al.,* 2009).

Like the PB2, the influenza A(H1N1)09pdm virus PA is also avian-like. It contains 7 avian virus unique residues (28P, 55D, 57R, 65S, 100V, 312K, and 552T), and only 3 human virus residues (356R, 382D, and 409N). This is despite the fact that the gene was maintained in swine viruses for over a decade (Chen and Shih, 2009). The main observation in this study regarding genetic drift in the PA gene segment of local isolates was that all local isolates that circulated beyond the introduction cases had 224S and 581L changes. The study suggests that these two mutations were crucial in efficient replication of influenza A(H1N1)09pdm viruses. This is due to the fact that amino acid 224 has been shown to significantly affect viral replication in H5N1 (Song *et al.,* 2011) while position 581 is located in the C terminal domain of the PA which bind to PB1 for complex formation and nuclear transport (Yuan *et al.,* 2009). Late in the pandemic, N321K and I330V also located in the C terminal domain were increasingly seen to dominate in the Kenyan isolates. These two mutations have also been reported in Thailand and Taiwan from 2010 (Makkoch *et al.,* 2012; Yang *et al.,* 2011). Their distribution suggests that they confer a selective advantage to the fitness of the virus and are thus expected to dominate in successive seasons.

The influenza HA is involved in two major functions: recognition of target cells, by binding to their sialic acid-containing receptors, and fusion of the viral and the endosomal membranes succeeding endocytosis (Chiu *et al.,* 1997).

Mutations in this protein therefore have implications in receptor specificity, host range and pathogenicity. Mapping of the HA mutations observed in the Kenyan isolates on to the known antigenic sites (Caton *et al.,* 1982) and receptor binding sites (Skehel and Wiley, 2000) was carried out. Of the 14 amino acid substitutions observed in the HA protein amongst the Kenyan isolates, 11 (78.5%) occurred in the HA1 domain while only 3 (21.5%) occurred in the HA2. Since HA1 is the most important in viral pathogenesis in initiating infection, the protein is a primary target for antibodies to neutralize the virus and is therefore more variable than HA2 (Skehel and Wiley, 2000). Thus these findings about amino acid changes in the HA are in agreement with our understanding of influenza biology. Another important observation noted was that 3 of the 11 sites (i.e. at positions 202, 220 and 239 [H1 numbering]) that had amino acid changes in the HA1 are located within known antigenic sites. Of these, 2 (sites 202 and 239) are located at antigenic site Ca while the other (site 220) is located at site Sb (Caton *et al.,* 1982). These results indicate a dominance of mutations within antigenic site Ca over mutations in other sites. Several studies on both A(H1N1)09pdm virus and seasonal influenza A/H1N1 viruses have suggested that sites Sb and Ca are of greatest significance in the immune escape and thus evolution of H1 viruses (Ding *et al.,* 2010; Raymond *et al.,* 1986).

Regarding amino acid changes in the receptor binding site, only amino acid position 239 and 202 showed substitutions in two isolates. At position 239, two local isolates possessed 239E while all the remaining had 239D. At this

position among influenza A(H1N1)09pdm viruses, 239D is the wild type amino acid. Studies have shown that mutations at this position alter influenza virus receptor binding specificity (Abed *et al.,* 2011). During the 2009 pandemic, substitution at this position involving the change from D to G/E was shown to be associated with severe clinical outcome including fatalities in several studies (Galiano *et al.,* 2011; Kilander *et al.,* 2010; Mak *et al.,* 2010). Considering that this substitution was not observed on other local isolates, it is possible that it did not offer a selective advantage to the virus. Late in the pandemic, S202T also located at the RBS emerged and seemed to dominate. Since this position is also an antigenic site, the study speculates that this mutation may offer selective advantage probably by favouring transmissibility of this new virus.

In the late phase of the pandemic, the emergence of parallel mutations D114N, S202T, E391K and S468N in the HA protein was evident. The hemagglutinin position 391 corresponds to the HA2-subunit position 47. E391K mutation has been shown to occur simultaneously with D114N (Maurer-Stroh *et al.,* 2010). This observation was also shown in this study. This (E391K) mutation has been shown to drastically alter the oligomerization interface of HA1 and HA2 in a region that undergoes structural changes required for membrane fusion (Maurer-Stroh *et al.,* 2010). Furthermore, this region was recently identified as a highly conserved epitope recognized by antibodies that neutralize the closely related 1918 H1N1 virus by blocking the structural changes associated with membrane fusion (Ekiert *et al.,* 2009). It is possible that the substitution from S

to N at position 468, a residue also located in the HA2 domain, may cause structural changes associated with membrane fusion. We hypothesize that these parallel mutations offer selective advantage and that they are likely to dominate in the HA protein of isolates in consecutive seasons.

The NP assembles with the three subunits of the polymerase into a ribonucleoprotein complex (RNP) which controls transcription and replication. In this complex, it regulates the balance between transcription and replication during the virus cycle (Portela and Digard, 2002). The most dominant amino acid change in this protein of local isolates was found to be the V100I substitution. Globally, in the pre-epidemic period only 10% of influenza A(H1N1)09pdm viruses had this valine to isoleucine change at position NP-100, whereas about 57%, 80% and 93% of the virus isolates collected in the early, middle and late periods possessed this change, respectively (Pan *et al.*, 2010). Strikingly, the 1918 H1N1pdm virus also had the residue isoleucine at position NP-100, while other pandemic viruses and human IAVs display valine at this position. The local isolates were also noted to possess 373T except two of the earliest isolates that possessed 373I. In light of these observations, this study hypothesizes that the acquisition of 100I and 373T in the NP was necessary for adaptation of influenza A(H1N1)09pdm virus in humans and may have played a role in its increased transmissibility or replication.

The influenza NA removes sialic acid from virus and cellular glycoproteins to facilitate virus progeny release and spread of infection to new cells. In this protein among local influenza A(H1N1)09pdm virus isolates, two dominant

mutations were observed in all isolates except in two of the earliest isolates. First, the avian-like residue, valine at NA-106 mutated to the human-like residue isoleucine, which was present in the 1918 and 1977 H1N1 pandemic viruses. Secondly, the non-charged residue, asparagine at NA-248, mutated to a negatively charged residue, aspartic acid, which was only presented in the 1977 H1N1 pandemic viruses. These two mutations, V106I and N248D, are non clade 1 defining mutations suggesting that no clade 1 viruses circulated in the country. Both these amino acid substitutions are located in previously defined or predicted B-cell antigenic regions. In addition, position 248 is in the proximity of the catalytic NA pocket (Maurer-Stroh *et al.,* 2009). These results reveal that the novel substitutions found in influenza A(H1N1)09pdm viruses isolated in Kenya are located on the surface of the protein (since B-cell epitopes are located on the surface) and may not interfere with the active site. This is in agreement with previous studies (Kao *et al.,* 2012; Maurer-Stroh *et al.,* 2009).

Another mutation observed in the NA of local isolates is the N369K mutation, which was only observed during the late phase of the pandemic in Kenya. Studies have shown that this mutation observed in local isolates, if acquired together with V241I and N386S, could potentially facilitate accommodation of the H275Y substitution without loss of fitness (Hurt *et al.,* 2012). This accommodation would lead to neuraminidase inhibitor resistant influenza viruses that are easily transmissible among humans. As a result, this mutation should be closely monitored in the successive seasons. Mutations associated

with resistance to neuraminidase inhibitors located in the NA protein are discussed under section 5.4.

The M protein gene of influenza A viruses encodes two proteins, M1 and M2, derived by splicing of mRNA. The Ml internal protein is a major component of the virus particle with an essential role in virus assembly and budding while the M2 has a proton-selective ion channel activity and is involved in virus assembly (Garcia-Robles *et al.,* 2005). It has been indicated previously that M1 is an interior virion protein (matrix) that lies just beneath the viral envelope (Garcia-Robles *et al.,* 2005). On the other hand, the M2 is an integral membrane protein inserted into the viral envelope and projects from the surface of the virus as tetramers (Lamb *et al.,* 1985). Consequently, infection of influenza A virus is known to induce the host's immune response to M2 protein, especially to the extracellular domain (Gerhard *et al.*, 1997).

An examination of the amino acid polymorphisms seen in the M1 and M2 proteins encoded by the M gene showed that there were more amino acid substitutions in the M2 (10/97) than in the M1 (7/252) protein of local influenza A(H1N1)09pdm viruses. As a result, the M1 protein was more conserved than the M2. The host's immune response seems to have exerted stronger selective pressure on the M2 than that on the M1 protein, accounting for the increased number of amino acid substitutions in this protein in a bid to evade immune pressure. Studies have shown that even minor mutations in the M1 protein such as the substitution of Lys with Asn at position 102 or 104 may cause restriction of virus replication (Furuse *et al.,* 2009; Liu and Ye, 2002).

This could explain why M1 is under strong negative pressure to avoid creation of defective virions. It also could explain why the selective pressure on M1 is smaller than that on M2. It was however observed that isolates obtained during the late phase of the pandemic acquired more amino acid substitutions (3/7 in M1 and 5/10 in the M2). This indicates a progressive genetic drift in this gene segment. Mutations associated with resistance to Adamantanes which are normally located in the M2 protein are discussed under section 5.4 below.

The NS gene segment of the influenza A virus encodes two proteins; the non-structural protein 1 (NS1) and the nuclear export protein (NEP). NS1 is essential to inhibition of the host immune response as it suppresses the host type 1 interferon (IFN) (Hale *et al.,* 2008). NEP on the other hand in association with M1, mediates the nuclear export of viral ribonucleoprotein (vRNP) complexes (Neumann *et al.,* 2000). Deletions in the NS1 protein have been associated with increased influenza A virulence. Highly pathogenic H5N1 viruses isolated from waterfowl, poultry, and humans in Southeast Asia since 2000 possess a 15-nucleotide deletion from position 263–277 in the NS gene (Guan *et al.,* 2004; Li *et al.,* 2004) resulting in enhanced virulence in both chicken and mice (Long *et al.,* 2008). Studies have also identified single mutations in the NS1 gene (S42P, D92E and V149A) as well as multiple mutations in the PDZ ligand domain that increase influenza pathogenicity (Jiao *et al.,* 2008; Li *et al.,* 2006; Seo *et al.,* 2002). On the other hand, mutations in the nuclear export signal ($_{12}$ILMRMSKMQL$_{21}$) in the NEP have been shown to

affect virulence and replication of influenza viruses (Iwatsuki-Horimoto *et al.*, 2004).

In the current study, the known virulence markers in the NS1 including S42P, D92E and V149A mutations and the PDZ ligand domain are missing in the local isolates. A unique dominant mutation I123V occurred in the NS1 protein during the early pandemic phase in the local isolates. None of the influenza A viruses collected globally during the pre-epidemic period carried this mutation, while 29%, 40% and 78% of the influenza A(H1N1)09pdm viruses collected in the early, middle and late periods possessed this mutation, respectively (Pan *et al.*, 2010). This mutation has not been observed in other viruses that caused previous pandemics. Since NS1 is a virulence factor responsible for suppressing antiviral interferon (IFN) induction during viral replication, and since this mutation is located in the effector domain of NS1, it is hypothesized here that this mutation had a beneficial role in replication.

In the NEP, the nuclear export sequence was conserved in all the local isolates. The R77M and R77K substitution was the most common seen in the NEP during the early and peak phases of the pandemic. Of these two, the R77M substitution may have been more destabilizing to the NEP as it led to the replacement of a polar amino acid with a non polar one. The substitution at position 77 was noted to be adjacent to Trp 78 which mediates the binding of NEP and M1 to allow nuclear export of viral ribonucleoproteins (vRNPs) replicated in the nucleus to the cytoplasm (Garcia-Robles *et al.*, 2005). A recent study has shown that mutations at this position may not affect the ability

of NEP to regulate viral RNA levels (Robb *et al.,* 2009). However, this mutation was not observed beyond December 2009 in the local isolates which may indicate that it did not confer a selective advantage to the virus.

In this study, the two proteins encoded by the NS gene also exhibited a different number of amino acid substitutions. Generally, the NS1 protein was more conserved than the NEP. This observation has been made previously (Qu *et al.,* 2011). Perhaps this may be attributed to the fact that, functioning of NEP requires it to bind to M1, meaning that any mutation in the NEP could require compensatory mutations in the M1. On the other hand, this may not be required in the NS1 as it does not bind to other influenza virus proteins.

5.4 Antiviral sensitivity of local influenza A(H1N1)09pdm viruses

Two classes of antiviral drugs are currently in use for the treatment of influenza virus infections. These are the neuraminidase inhibitors and M2 inhibitors (Adamantanes). These drugs inhibit influenza virus replication by binding to the NA and M2 proteins respectively. While these two classes are generally effective, the emergence of drug resistant is always a constant threat. This is much so considering the error prone nature of influenza polymerases. Monitoring the susceptibility of circulating influenza viruses is therefore of utmost necessity particularly in a pandemic situation. The study therefore determined the susceptibility of the Kenyan isolates by examining for known molecular markers of resistance in the NA and M2 proteins.

Resistance to M2 inhibitors can be achieved by a single substitution of any of the amino acid residues located at positions 26, 27, 30, 31, or 34 of the transmembrane domain of the M2 protein (Boivin *et al.,* 2002). In the vast majority of cases, the basis for this resistance is a single serine to asparagine amino acid replacement (S31N) in the matrix M2 protein (Wang *et al.,* 1993). This mutation weakens (M2) transmembrane helical packing and thereby disrupts the drug binding pocket (Pielak *et al.,* 2009). None of the Kenyan isolates harboured mutations at positions 26, 27, 30, or 34. However, all the Kenyan isolates during the pandemic period had the S31N mutation and were thus resistant to M2 inhibitors. This resistance has also been found in all currently circulating influenza A(H1N1)09pdm viruses isolated globally (Neumann *et al.,* 2009).

Resistance to neuraminidase inhibitors is conferred by any of the following mutations in the NA; I117V, E119V, D198N, I222V, H274Y, R292K, N294S, I314V, V116A, R118K, E119G/A/D, Q136K, D151E, R152K, R224K, E276D, R292K and R371K [N2 numbering] (Colman *et al.,* 1983) These molecular markers associated with resistance to neuraminidase inhibitors are located in the active site of the NA protein at different positions depending on the virus subtype, thus altering its sensitivity to inhibition (Hayden *et al.,* 2005). The study examined the local isolates for oseltamivir and zanamivir-related mutations. Majority of the local isolates did not carry the resistance conferring mutations indicating their sensitivity to neuraminidase inhibitors. However, two isolates possessed an E277Q (E276 in N2 numbering) mutation. It has

been shown previously that E277 is an oseltamivir-interacting residue (Li *et al.*, 2012). Thus, the substitution of Glutamic acid with aspartic acid changes this interaction resulting in Oseltamivir resistant viruses. However, in the local isolates, though this mutation replaces an acidic amino acid residue with one that is basic, the study speculates that the negative charge of Glutamine under physiological conditions may alter the NA active site which in the wild-type is positively charged at physiological conditions. This may thus alter the binding of Oseltamivir to the NA active site. Phenotypic assays should therefore be carried out to determine the effect of this mutation on Oseltamivir activity.

5.5 Mea culpa

Although genetic analysis is important in informing our understanding of influenza virus evolution and antigenic changes, the lack of corresponding hemagglutination inhibition data of the Kenyan viruses is a limitation in this study. As such, it is not possible to correlate the observed mutations in the local isolates with antigenic drift. Also, the analysis in this study only involved the coding region of the gene segments. Although the non-coding regions are considered to be conserved, mutations that affect viral replication may occur and this information may have been lost in the current study. Additionally, since this study did not allow for follow up of the patients, it is not possible to associate any of the observed mutations with virulence. However, the strength of this study lies in the fact that it is among the first of its kind in Africa that

employs whole genome sequencing and phylodynamics in an attempt to understand the evolutionary dynamics of this novel virus in the region.

6.0: CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

a) On the basis of observations in this study, the study concludes that influenza A(H1N1)pdm09 virus was introduced into the country and sustained as a result of multiple importations by travellers coming from the UK.

b) Two clades of influenza A(H1N1)pdm09 virus namely clade 2 and 7 were introduced into the country. Of these two, clade 2 died out and did not circulate beyond the introduction sites while clade 7 disseminated and was transmitted throughout the country.

c) The surface expressed gene segments exhibited the highest rates of evolution while the internally expressed segments expressed lower evolutionary rates.

d) The effective number of infections and hence the number of infected individuals increased exponentially in the first three months of the presence of the virus in the country and thereafter declined.

e) The earliest clusters of influenza A(H1N1)pdm09 virus in the country are estimated to have emerged elsewhere between April – June 2009. This is 2-3 months before the first laboratory confirmed case was detected and coincides with the time of the most recent common ancestor (tMRCA) of specific clusters in the UK.

f) Majority of circulating clusters of A(H1N1)pdm09 virus were geographically spread and did not exhibit any observable spatial patterns indicating efficient community spread.

g) None of the previously described virulence markers were acquired by the local A(H1N1)pdm09 virus during its circulation in the pandemic period.

h) While many of the mutations observed in the gene segments were transient, demonstrating how the viral genome has shaped dynamically, the study clearly showed that a progressive genetic drift away from the A/California/7/2009 vaccine strain is apparent in the PB2, PB1, PA, HA, NA, M and NS gene segments particularly in samples isolated in 2010.

i) All the local influenza A(H1N1)pdm09 virus isolates during the pandemic period were found to be sensitive to neuraminidase inhibitors but resistant to adamantanes.

6.2 Recommendations

6.2.1 Application of research findings

i) The research findings indicate that the effective number of infections increased exponentially in the first three months. Due to this, vaccination or any other public health measure (such as prophylactic antiviral therapy) should be instituted as soon as the pandemic emerges so as to slow down the rate of new infections.

ii) The evolution of A(H1N1)pdm09 virus in the country was observed to be largely via progressive genetic drift. This may not have been clear if only one gene was sequenced. This finding therefore underscores the need for intensified surveillance applying whole genome data to monitor the diversity of influenza viruses.

iii) The study found out that local influenza A(H1N1)pdm09 viruses were sensitive to neuraminidase inhibitors but resistant to adamantanes. Neuraminidase inhibitors therefore offer the best option in the treatment of influenza A(H1N1)pdm09 virus infections in the country.

6.2.2 Recommendations for further research

a) Studies should be carried out to compare the transmissibility and replication kinetics of local clade 2 and clade 7 viruses' in-order to understand the local and global dominance of clade 7 viruses.

b) Antigenic cartography based on haemagglutination inhibition assays should be performed in-order to determine the effect of the genetic drift observed in local isolates on antigenic drift.

c) Phenotypic *in vitro* antiviral resistance assays should be performed to investigate the role of mutations observed in the NA protein of local isolates on their susceptibility to Oseltamivir and Zanamivir.

d) Reverse genetics studies should be carried out to understand the role of the mutations observed in the local isolates on their replication, virulence and transmissibility in animal models.

REFERENCES

Abed, Y., Pizzorno, A., Hamelin, M.E., Leung, A., Joubert, P., Couture, C., Kobasa, D., Boivin, G., 2011, The 2009 pandemic H1N1 D222G hemagglutinin mutation alters receptor specificity and increases virulence in mice but not in ferrets. J Infect Dis 204, 1008-1016.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25, 3389-3402.

Baillie, G.J., Galiano, M., Agapow, P.M., Myers, R., Chiam, R., Gall, A., Palser, A.L., Watson, S.J., Hedge, J., Underwood, A., Platt, S., McLean, E., Pebody, R.G., Rambaut, A., Green, J., Daniels, R., Pybus, O.G., Kellam, P., Zambon, M., 2012, Evolutionary dynamics of local pandemic H1N1/2009 influenza virus lineages revealed by whole-genome analysis. J Virol 86, 11-18.

Bao, Y., Bolotov, P., Dernovoy, D., Kiryutin, B., Zaslavsky, L., Tatusova, T., Ostell, J., Lipman, D., 2008, The influenza virus resource at the National Center for Biotechnology Information. J Virol 82, 596-601.

Barik, S., 2012, New treatments for influenza. BMC Med 10, 104.

Barr, I.G., McCauley, J., Cox, N., Daniels, R., Engelhardt, O.G., Fukuda, K., Grohmann, G., Hay, A., Kelso, A., Klimov, A., Odagiri, T., Smith, D., Russell, C., Tashiro, M., Webby, R., Wood, J., Ye, Z., Zhang, W., 2010, Epidemiological, antigenic and genetic characteristics of seasonal influenza A(H1N1), A(H3N2) and B influenza viruses: basis for the WHO recommendation on the composition of influenza vaccines for use in the 2009-2010 Northern Hemisphere season. Vaccine 28, 1156-1167.

Barrero, P.R., Viegas, M., Valinotto, L.E., Mistchenko, A.S., 2011, Genetic and phylogenetic analyses of influenza A H1N1pdm virus in Buenos Aires, Argentina. J Virol 85, 1058-1066.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2006, GenBank. Nucleic Acids Res 34, D16-20.

Berkhoff, E.G., de Wit, E., Geelhoed-Mieras, M.M., Boon, A.C., Symons, J., Fouchier, R.A., Osterhaus, A.D., Rimmelzwaan, G.F., 2005, Functional constraints of influenza A virus epitopes limit escape from cytotoxic T lymphocytes. J Virol 79, 11239-11246.

Bhatia, R., Ichhpujani, R.L., 1999, Essentials of Microbiology 2Edition. Jaypee Brothers, New Delhi.

Boivin, G., Goyette, N., Bernatchez, H., 2002, Prolonged excretion of amantadine-resistant influenza a virus quasi species after cessation of antiviral therapy in an immunocompromised patient. Clin Infect Dis 34, E23-25.

Boivin, G., Hardy, I., Kress, A., 2001, Evaluation of a rapid optical immunoassay for influenza viruses (FLU OIA test) in comparison with cell culture and reverse transcription-PCR. J Clin Microbiol 39, 730-732.

Boivin, S., Cusack, S., Ruigrok, R.W., Hart, D.J., 2010, Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. J Biol Chem 285, 28411-28417.

Braam, J., Ulmanen, I., Krug, R.M., 1983, Molecular model of a eucaryotic transcription complex: functions and movements of influenza P proteins during capped RNA-primed transcription. Cell 34, 609-618.

Bull, J.J., Badgett, M.R., Wichman, H.A., Huelsenbeck, J.P., Hillis, D.M., Gulati, A., Ho, C., Molineux, I.J., 1997, Exceptional convergent evolution in a virus. Genetics 147, 1497-1507.

Caton, A.J., Brownlee, G.G., Yewdell, J.W., Gerhard, W., 1982, The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31, 417-427.

Cauchemez, S., Valleron, A.J., Boelle, P.Y., Flahault, A., Ferguson, N.M., 2008, Estimating the impact of school closure on influenza transmission from Sentinel data. Nature 452, 750-754.

CDC, 2009a, Introduction and transmission of 2009 pandemic influenza A (H1N1) Virus--Kenya, June-July 2009. MMWR Morb Mortal Wkly Rep 58, 1143-1146.

CDC, 2009b, Oseltamivir-resistant 2009 pandemic influenza A (H1N1) virus infection in two summer campers receiving prophylaxis--North Carolina. Morbidity and Mortality Weekly Report 58, 969-972.

CDC, 2009c, Swine influenza A (H1N1) infection in two children--Southern California, March-April 2009. MMWR Morb Mortal Wkly Rep 58, 400-402.

Chao, D.L., Halloran, M.E., Longini, I.M., Jr., 2010, School opening dates predict pandemic influenza A(H1N1) outbreaks in the United States. J Infect Dis 202, 877-880.

Chen, G.W., Shih, S.R., 2009, Genomic signatures of influenza A pandemic (H1N1) 2009 virus. Emerg Infect Dis 15, 1897-1903.

Cheung, C.Y., Poon, L.L., Lau, A.S., Luk, W., Lau, Y.L., Shortridge, K.F., Gordon, S., Guan, Y., Peiris, J.S., 2002, Induction of proinflammatory cytokines in human macrophages by influenza A (H5N1) viruses: a mechanism for the unusual severity of human disease? Lancet 360, 1831-1837.

Chiu, W., Burnett, R.M., Garcea, R.L., 1997, Structural biology of viruses. Oxford university press New York.

Claas, E.C., Osterhaus, A.D., 1998, New clues to the emergence of flu pandemics. Nat Med 4, 1122-1123.

Colman, P.M., 1998, Structure and function of the neuraminidase, In: Nicholson, K.G., Webster, R. G., Hay, A. J. (Ed.) Textbook of influenza. Blackwell Science Ltd, pp. 65-73.

Colman, P.M., Laver, W.G., Varghese, J.N., Baker, A.T., Tulloch, P.A., Air, G.M., Webster, R.G., 1987, Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. Nature 326, 358-363.

Colman, P.M., Varghese, J.N., Laver, W.G., 1983, Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature 303, 41-44.

Connor, R.J., Kawaoka, Y., Webster, R.G., Paulson, J.C., 1994, Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. Virology 205, 17-23.

Couceiro, J.N., Paulson, J.C., Baum, L.G., 1993, Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. Virus Res 29, 155-165.

Cox, N.J., Fuller, F., Keverin, N., Klenk, H.D., Lamb, R.A., Mahy, B.W., McCauley, J., Nakamura, K., Palese, P. Webster, R.G, 2000, Orthomyxoviridae, In: M.H.V van Regenmortal, C.M.F., E.B Caustens, M.K Estes, S.M Lemon, J Maniloff, M.A Mayo, D.J McGeoch, C.R Pringle, R.B Wickner (Ed.) Virus taxonomy (Seventh report and the international committee on the taxonomy of viruses). . Academic Press, Inc, California.

Cox, N.J., Ziegler, T., 2003, Influenza viruses, In: Murray, P.R.B., E, J; Jorgensen, J,H; Pfaller, M,A; Yolken, R,H (Ed.) Manual of clinical microbiology. ASM Press, Washington.

Dawood, F.S., Iuliano, A.D., Reed, C., Meltzer, M.I., Shay, D.K., Cheng, P.-Y., Bandaranayake, D., Breiman, R.F., Brooks, W.A., Buchy, P., Feikin, D.R., Fowler, K.B., Gordon, A., Hien, N.T., Horby, P., Huang, Q.S., Katz, M.A., Krishnan, A., Lal, R., Montgomery, J.M., Mølbak, K., Pebody, R.,

Presanis, A.M., Razuri, H., Steens, A., Tinoco, Y.O., Wallinga, J., Yu, H., Vong, S., Bresee, J., Widdowson, M.-A., 2012, Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. The Lancet Infectious Diseases 12, 687-695.

Dawood, F.S., Jain, S., Finelli, L., Shaw, M.W., Lindstrom, S., Garten, R.J., Gubareva, L.V., Xu, X., Bridges, C.B., Uyeki, T.M., 2009, Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med 360, 2605-2615.

de la Luna, S., Martinez, C., Ortin, J., 1989, Molecular cloning and sequencing of influenza virus A/Victoria/3/75 polymerase genes: sequence evolution and prediction of possible functional domains. Virus Res 13, 143-155.

de Silva, E., Ferguson, N.M., Fraser, C., 2012, Inferring pandemic growth rates from sequence data. Journal of The Royal Society Interface.

de Wit, E., Fouchier, R.A., 2008, Emerging influenza. J Clin Virol 41, 1-6.

Deng, T., Engelhardt, O.G., Thomas, B., Akoulitchev, A.V., Brownlee, G.G., Fodor, E., 2006, Role of ran binding protein 5 in nuclear import and assembly of the influenza virus RNA polymerase complex. J Virol 80, 11911-11919.

Dias, A., Bouvier, D., Crepin, T., McCarthy, A.A., Hart, D.J., Baudin, F., Cusack, S., Ruigrok, R.W., 2009, The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature 458, 914-918.

Digard, P., Blok, V.C., Inglis, S.C., 1989, Complex formation between influenza virus polymerase proteins expressed in Xenopus oocytes. Virology 171, 162-169.

Ding, X., Jiang, L., Ke, C., Yang, Z., Lei, C., Cao, K., Xu, J., Xu, L., Yang, X., Zhang, Y., Huang, P., Huang, W., Zhu, X., He, Z., Liu, L., Li, J., Yuan, J., Wu, J., Tang, X., Li, M., 2010, Amino acid sequence analysis and identification of mutations under positive selection in hemagglutinin of 2009 influenza A (H1N1) isolates. Virus Genes 41, 329-340.

Drummond, A.J., Rambaut, A., 2007, BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7, 214.

Duffy, S., Shackelton, L.A., Holmes, E.C., 2008, Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 9, 267-276.

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

Ekiert, D.C., Bhabha, G., Elsliger, M.A., Friesen, R.H., Jongeneelen, M., Throsby, M., Goudsmit, J., Wilson, I.A., 2009, Antibody recognition of a highly conserved influenza virus epitope. Science 324, 246-251.

Engelhardt, O.G., Fodor, E., 2006, Functional association between viral and cellular transcription during influenza virus infection. Rev Med Virol 16, 329-345.

Fouchier, R.A., Osterhaus, A.D., Brown, I.H., 2003, Animal influenza virus surveillance. Vaccine 21, 1754-1757.

Fraser, C., Donnelly, C.A., Cauchemez, S., Hanage, W.P., Van Kerkhove, M.D., Hollingsworth, T.D., Griffin, J., Baggaley, R.F., Jenkins, H.E., Lyons, E.J., Jombart, T., Hinsley, W.R., Grassly, N.C., Balloux, F., Ghani, A.C., Ferguson, N.M., Rambaut, A., Pybus, O.G., Lopez-Gatell, H., Alpuche-Aranda, C.M., Chapela, I.B., Zavala, E.P., Guevara, D.M., Checchi, F., Garcia, E., Hugonnet, S., Roth, C., 2009, Pandemic potential of a strain of influenza A (H1N1): early findings. Science 324, 1557-1561.

Furuse, Y., Suzuki, A., Kamigaki, T., Oshitani, H., 2009, Evolution of the M gene of the influenza A virus in different host species: large-scale sequence analysis. Virol J 6, 67.

Gabriel, G., Herwig, A., Klenk, H.D., 2008, Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. PLoS Pathog 4, e11.

Galiano, M., Agapow, P.M., Thompson, C., Platt, S., Underwood, A., Ellis, J., Myers, R., Green, J., Zambon, M., 2011, Evolutionary pathways of the pandemic influenza A (H1N1) 2009 in the UK. PLoS ONE 6, e23779.

Garcia-Robles, I., Akarsu, H., Muller, C.W., Ruigrok, R.W., Baudin, F., 2005, Interaction of influenza virus proteins with nucleosomes. Virology 332, 329-336.

Garten, R.J., Davis, C.T., Russell, C.A., Shu, B., Lindstrom, S., Balish, A., Sessions, W.M., Xu, X., Skepner, E., Deyde, V., Okomo-Adhiambo, M., Gubareva, L., Barnes, J., Smith, C.B., Emery, S.L., Hillman, M.J., Rivailler, P., Smagala, J., de Graaf, M., Burke, D.F., Fouchier, R.A., Pappas, C., Alpuche-Aranda, C.M., Lopez-Gatell, H., Olivera, H., Lopez, I., Myers, C.A., Faix, D., Blair, P.J., Yu, C., Keene, K.M., Dotson, P.D., Jr., Boxrud, D., Sambol, A.R., Abid, S.H., St George, K., Bannerman, T., Moore, A.L., Stringer, D.J., Blevins, P., Demmler-Harrison, G.J., Ginsberg, M., Kriner, P., Waterman, S., Smole, S., Guevara, H.F., Belongia, E.A., Clark, P.A., Beatrice, S.T., Donis, R., Katz, J., Finelli, L., Bridges, C.B., Shaw, M., Jernigan, D.B., Uyeki, T.M., Smith, D.J., Klimov, A.I., Cox, N.J., 2009, Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science 325, 197-201.

Gatherer, D., 2009, The 2009 H1N1 influenza outbreak in its historical context. J Clin Virol 45, 174-178.

Gaurav, V., David, J.L., Rudolf, M., 2011, SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171-180.

Gerhard, W., Mozdzanowska, K., Furchner, M., Washko, G., Maiese, K., 1997, Role of the B-cell response in recovery of mice from primary influenza virus infection. Immunol Rev 159, 95-103.

Gerhard, W., Mozdzanowska, K., Zharikova, D., 2006, Prospects for universal influenza virus vaccine. Emerg Infect Dis 12, 569-574.

Ghedin, E., Sengamalay, N.A., Shumway, M., Zaborsky, J., Feldblyum, T., Subbu, V., Spiro, D.J., Sitz, J., Koo, H., Bolotov, P., Dernovoy, D., Tatusova, T., Bao, Y., St George, K., Taylor, J., Lipman, D.J., Fraser, C.M., Taubenberger, J.K., Salzberg, S.L., 2005, Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. Nature 437, 1162-1166.

Goni, N., Moratorio, G., Coppola, L., Ramas, V., Comas, V., Sonora, M., Chiparelli, H., Cristina, J., 2012, Bayesian coalescent analysis of pandemic H1N1 influenza A virus circulating in the South American region. Virus Res 170, 91-101.

Gonzalez, S., Zurcher, T., Ortin, J., 1996, Identification of two separate domains in the influenza virus PB1 protein involved in the interaction with the PB2 and PA subunits: a model for the viral RNA polymerase structure. Nucleic Acids Res 24, 4456-4463.

Greenwood, D., 2003, Medical microbiology : a guide to microbial infections ; pathogenesis, immunity, laboratory diagnosis and control. Livingstone, Edinburgh [u.a.].

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L.N., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004, Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. Science 303, 327-332.

Greninger, A.L., Chen, E.C., Sittler, T., Scheinerman, A., Roubinian, N., Yu, G., Kim, E., Pillai, D.R., Guyard, C., Mazzulli, T., Isa, P., Arias, C.F., Hackett, J., Schochetman, G., Miller, S., Tang, P., Chiu, C.Y., 2010, A metagenomic analysis of pandemic influenza A (2009 H1N1) infection in patients from North America. PLoS ONE 5, e13381.

Guan, Y., Poon, L.L., Cheung, C.Y., Ellis, T.M., Lim, W., Lipatov, A.S., Chan, K.H., Sturm-Ramirez, K.M., Cheung, C.L., Leung, Y.H., Yuen, K.Y.,

Webster, R.G., Peiris, J.S., 2004, H5N1 influenza: a protean pandemic threat. Proc Natl Acad Sci U S A 101, 8156-8161.

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R.W., Ortin, J., Hart, D.J., Cusack, S., 2008, The structural basis for cap binding by influenza virus polymerase subunit PB2. Nat Struct Mol Biol 15, 500-506.

Guo, Y.J., Jin, F.G., Wang, P., Wang, M., Zhu, J.M., 1983, Isolation of influenza C virus from pigs and experimental infection of pigs with influenza C virus. J Gen Virol 64 (Pt 1), 177-182.

Hale, B.G., Randall, R.E., Ortin, J., Jackson, D., 2008, The multifunctional NS1 protein of influenza A viruses. J Gen Virol 89, 2359-2376.

Hara, K., Schmidt, F.I., Crow, M., Brownlee, G.G., 2006, Amino acid residues in the N-terminal region of the PA subunit of influenza A virus RNA polymerase play a critical role in protein stability, endonuclease activity, cap binding, and virion RNA promoter binding. J Virol 80, 7789-7798.

Hatta, M., Gao, P., Halfmann, P., Kawaoka, Y., 2001, Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science 293, 1840-1842.

Hausmann, J., Kretzschmar, E., Garten, W., Klenk, H.D., 1997, Biosynthesis, intracellular transport and enzymatic activity of an avian influenza A virus neuraminidase: role of unpaired cysteines and individual oligosaccharides. J Gen Virol 78 (Pt 12), 3233-3245.

Hay, A., 1996, Amantadine and Rimantadine - Mechanisms, In: Richman, D.D. (Ed.) Antiviral drug resistance. Wiley, Chichester; New York.

Hayden, F., Klimov, A., Tashiro, M., Hay, A., Monto, A., McKimm-Breschkin, J., Macken, C., Hampson, A., Webster, R.G., Amyard, M., Zambon, M., 2005, Neuraminidase inhibitor susceptibility network position statement: antiviral resistance in influenza A/H5N1 viruses. Antivir Ther 10, 873-877.

Herfst, S., Chutinimitkul, S., Ye, J., de Wit, E., Munster, V.J., Schrauwen, E.J., Bestebroer, T.M., Jonges, M., Meijer, A., Koopmans, M., Rimmelzwaan, G.F., Osterhaus, A.D., Perez, D.R., Fouchier, R.A., 2010, Introduction of virulence markers in PB2 of pandemic swine-origin influenza virus does not result in enhanced virulence or transmission. J Virol 84, 3752-3758.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., VandePol, S., 1982, Rapid evolution of RNA genomes. Science 215, 1577-1585.

Holmes, E.C., 2010, Evolution in health and medicine Sackler colloquium: The comparative genomics of viral emergence. Proc Natl Acad Sci U S A 107 Suppl 1, 1742-1746.

Honda, A., Mizumoto, K., Ishihama, A., 2002, Minimum molecular architectures for transcription and replication of the influenza virus. Proc Natl Acad Sci U S A 99, 13166-13171.

Horimoto, T., Kawaoka, Y., 2001, Pandemic threat posed by avian influenza A viruses. Clin Microbiol Rev 14, 129-149.

Horimoto, T., Nakayama, K., Smeekens, S.P., Kawaoka, Y., 1994, Proprotein-processing endoproteases PC6 and furin both activate hemagglutinin of virulent avian influenza viruses. J Virol 68, 6074-6078.

Hurt, A.C., Hardie, K., Wilson, N.J., Deng, Y.M., Osbourn, M., Leang, S.K., Lee, R.T., Iannello, P., Gehrig, N., Shaw, R., Wark, P., Caldwell, N., Givney, R.C., Xue, L., Maurer-Stroh, S., Dwyer, D.E., Wang, B., Smith, D.W., Levy, A., Booy, R., Dixit, R., Merritt, T., Kelso, A., Dalton, C., Durrheim, D., Barr, I.G., 2012, Characteristics of a widespread community cluster of H275Y oseltamivir-resistant A(H1N1)pdm09 influenza in Australia. J Infect Dis 206, 148-157.

Hurt, A.C., Holien, J.K., Parker, M., Kelso, A., Barr, I.G., 2009, Zanamivir-resistant influenza viruses with a novel neuraminidase mutation. J Virol 83, 10366-10373.

Iwatsuki-Horimoto, K., Horimoto, T., Fujii, Y., Kawaoka, Y., 2004, Generation of influenza A virus NS2 (NEP) mutants with an altered nuclear export signal sequence. J Virol 78, 10149-10155.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C., 2002, Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J Mol Evol 54, 156-165.

Jiao, P., Tian, G., Li, Y., Deng, G., Jiang, Y., Liu, C., Liu, W., Bu, Z., Kawaoka, Y., Chen, H., 2008, A single-amino-acid substitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice. J Virol 82, 1146-1154.

Johnson, F.B., 1990, Transport of viral specimens. Clin Microbiol Rev 3, 120-131.

Jones, I.M., Reay, P.A., Philpott, K.L., 1986, Nuclear location of all three influenza polymerase proteins and a nuclear signal in polymerase PB2. EMBO J 5, 2371-2376.

Kao, C.L., Chan, T.C., Tsai, C.H., Chu, K.Y., Chuang, S.F., Lee, C.C., Li, Z.R., Wu, K.W., Chang, L.Y., Shen, Y.H., Huang, L.M., Lee, P.I., Yang,

C., Compans, R., Rouse, B.T., King, C.C., 2012, Emerged HA and NA mutants of the pandemic influenza H1N1 viruses with increasing epidemiological significance in Taipei and Kaohsiung, Taiwan, 2009-10. PLoS ONE 7, e31162.

Kawaoka, Y., Krauss, S., Webster, R.G., 1989, Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. J Virol 63, 4603-4608.

Kilander, A., Rykkvin, R., Dudman, S.G., Hungnes, O., 2010, Observed association between the HA1 mutation D222G in the 2009 pandemic influenza A(H1N1) virus and severe clinical outcome, Norway 2009-2010. Euro Surveill 15.

Kilbourne, E.D., 1973, The molecular epidemiology of influenza. J Infect Dis 127, 478-487.

Krug, R.M., Yuan, W., Noah, D.L., Latham, A.G., 2003, Intracellular warfare between human influenza viruses and human cells: the roles of the viral NS1 protein. Virology 309, 181-189.

Lamb, R.A., Choppin, P.W., 1981, Identification of a second protein (M2) encoded by RNA segment 7 of influenza virus. Virology 112, 729-737.

Lamb, R.A., Krug, R.M., 2001, Orthomyxoviridae: the viruses and their replication., In: Knipe, D.M., Howley, P. M. (Ed.) Fields Virology. Lippincott Williams and Wilkins, Philadelphia, pp. 1487-1531.

Lamb, R.A., Zebedee, S.L., Richardson, C.D., 1985, Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. Cell 40, 627-633.

Lemey, P., Suchard, M., Rambaut, A., 2009, Reconstructing the initial global spread of a human influenza pandemic: A Bayesian spatial-temporal model for the global spread of H1N1pdm. PLoS Curr 1, RRN1031.

Lennette, E.H., 1995, Diagnostic procedures for viral, rickettsial, and chlamydial infections. American Public Health Association, Washington, DC.

Li, K.S., Guan, Y., Wang, J., Smith, G.J., Xu, K.M., Duan, L., Rahardjo, A.P., Puthavathana, P., Buranathai, C., Nguyen, T.D., Estoepangestie, A.T., Chaisingh, A., Auewarakul, P., Long, H.T., Hanh, N.T., Webby, R.J., Poon, L.L., Chen, H., Shortridge, K.F., Yuen, K.Y., Webster, R.G., Peiris, J.S., 2004, Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. Nature 430, 209-213.

Li, L., Li, Y., Zhang, L., Hou, T., 2012, Theoretical Studies on the Susceptibility of Oseltamivir against Variants of 2009 A/H1N1 Influenza Neuraminidase. J Chem Inf Model 52, 2715-2729.

Li, M.L., Rao, P., Krug, R.M., 2001, The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. EMBO J 20, 2078-2086.

Li, S., Schulman, J., Itamura, S., Palese, P., 1993, Glycosylation of neuraminidase determines the neurovirulence of influenza A/WSN/33 virus. J Virol 67, 6667-6673.

Li, Z., Jiang, Y., Jiao, P., Wang, A., Zhao, F., Tian, G., Wang, X., Yu, K., Bu, Z., Chen, H., 2006, The NS1 gene contributes to the virulence of H5N1 avian influenza viruses. J Virol 80, 11115-11123.

Liu, T., Ye, Z., 2002, Restriction of viral replication by mutation of the influenza virus matrix protein. J Virol 76, 13055-13061.

Long, J.X., Peng, D.X., Liu, Y.L., Wu, Y.T., Liu, X.F., 2008, Virulence of H5N1 avian influenza virus enhanced by a 15-nucleotide deletion in the viral nonstructural gene. Virus Genes 36, 471-478.

Luke, C.J., Subbarao, K., 2006, Vaccines for pandemic influenza. Emerg Infect Dis 12, 66-72.

Lycett, S., McLeish, N.J., Robertson, C., Carman, W., Baillie, G., McMenamin, J., Rambaut, A., Simmonds, P., Woolhouse, M., Leigh Brown, A.J., 2012, Origin and fate of A/H1N1 influenza in Scotland during 2009. J Gen Virol 93, 1253-1260.

Ma, K., Roy, A.M., Whittaker, G.R., 2001, Nuclear export of influenza virus ribonucleoproteins: identification of an export intermediate at the nuclear periphery. Virology 282, 215-220.

Maines, T.R., Jayaraman, A., Belser, J.A., Wadford, D.A., Pappas, C., Zeng, H., Gustin, K.M., Pearce, M.B., Viswanathan, K., Shriver, Z.H., Raman, R., Cox, N.J., Sasisekharan, R., Katz, J.M., Tumpey, T.M., 2009, Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. Science 325, 484-487.

Majanja, J., Njoroge, R.N., Achilla, R., Wurapa, E.K., Wadegu, M., Mukunzi, S., Mwangi, J., Njiri, J., Gachara, G., Bulimo, W., 2013, Impact of Influenza A(H1N1)pdm09 Virus on Circulation Dynamics of Seasonal Influenza Strains in Kenya. Am J Trop Med Hyg.

Mak, G.C., Au, K.W., Tai, L.S., Chuang, K.C., Cheng, K.C., Shiu, T.C., Lim, W., 2010, Association of D222G substitution in haemagglutinin of 2009 pandemic influenza A (H1N1) with severe disease. Euro Surveill 15.

Makkoch, J., Suwannakarn, K., Payungporn, S., Prachayangprecha, S., Cheiocharnsin, T., Linsuwanon, P., Theamboonlers, A., Poovorawan, Y., 2012, Whole genome characterization, phylogenetic and genome signature analysis of human pandemic H1N1 virus in Thailand, 2009-2012. PLoS ONE 7, e51275.

Maldonado, J., Van Reeth, K., Riera, P., Sitja, M., Saubi, N., Espuna, E., Artigas, C., 2006, Evidence of the concurrent circulation of H1N2, H1N1 and H3N2 influenza A viruses in densely populated pig areas in Spain. Vet J 172, 377-381.

Manicassamy, B., Medina, R.A., Hai, R., Tsibane, T., Stertz, S., Nistal-Villan, E., Palese, P., Basler, C.F., Garcia-Sastre, A., 2010, Protection of mice against lethal challenge with 2009 H1N1 influenza A virus by 1918-like and classical swine H1N1 based vaccines. PLoS Pathog 6, e1000745.

Martin, K., Helenius, A., 1991, Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import. Cell 67, 117-130.

Matrosovich, M., Tuzikov, A., Bovin, N., Gambaryan, A., Klimov, A., Castrucci, M.R., Donatelli, I., Kawaoka, Y., 2000, Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. J Virol 74, 8502-8512.

Matrosovich, M.N., Matrosovich, T.Y., Gray, T., Roberts, N.A., Klenk, H.D., 2004, Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. J Virol 78, 12665-12667.

Maurer-Stroh, S., Lee, R.T., Eisenhaber, F., Cui, L., Phuah, S.P., Lin, R.T., 2010, A new common mutation in the hemagglutinin of the 2009 (H1N1) influenza A virus. PLoS Curr 2, RRN1162.

Maurer-Stroh, S., Ma, J., Lee, R.T., Sirota, F.L., Eisenhaber, F., 2009, Mapping the sequence mutations of the 2009 H1N1 influenza A virus neuraminidase relative to drug and antibody binding sites. Biol Direct 4.

Meguro, H., Bryant, J.D., Torrence, A.E., Wright, P.F., 1979, Canine kidney cell line for isolation of respiratory viruses. J Clin Microbiol 9, 175-179.

Metselaar, D., Simpson, D.I.H., 1982, Practical virology for medical students and practitioners in tropical countries. Oxford University Press, Oxford ; New York.

Monto, A.S., 2003, The role of antivirals in the control of influenza. Vaccine 21, 1796-1800.

Monto, A.S., 2006, Vaccines and antiviral drugs in pandemic preparedness. Emerg Infect Dis 12, 55-60.

Monto, A.S., Koopman, J.S., Longini, I.M., Jr., 1985, Tecumseh study of illness. XIII. Influenza infection and disease, 1976-1981. Am J Epidemiol 121, 811-822.

Morens, D.M., Taubenberger, J.K., Fauci, A.S., 2009, The persistent legacy of the 1918 influenza virus. N Engl J Med 361, 225-229.

Mukaigawa, J., Nayak, D.P., 1991, Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. J Virol 65, 245-253.

Mullick, J., Cherian, S.S., Potdar, V.A., Chadha, M.S., Mishra, A.C., 2011, Evolutionary dynamics of the influenza A pandemic (H1N1) 2009 virus with emphasis on Indian isolates: evidence for adaptive evolution in the HA gene. Infect Genet Evol 11, 997-1005.

Naffakh, N., Tomoiu, A., Rameix-Welti, M.A., van der Werf, S., 2008, Host restriction of avian influenza viruses at the level of the ribonucleoproteins. Annu Rev Microbiol 62, 403-424.

Nagata, K., Kawaguchi, A., Naito, T., 2008, Host factors for replication and transcription of the influenza virus genome. Rev Med Virol 18, 247-260.

Nath, S.T., Nayak, D.P., 1990, Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1 N1). Mol Cell Biol 10, 4139-4145.

Nelson, M., Spiro, D., Wentworth, D., Beck, E., Fan, J., Ghedin, E., Halpin, R., Bera, J., Hine, E., Proudfoot, K., Stockwell, T., Lin, X., Griesemer, S., Kumar, S., Bose, M., Viboud, C., Holmes, E., Henrickson, K., 2009, The early diversification of influenza A/H1N1pdm. PLoS Curr 1, RRN1126.

Nelson, M.I., Simonsen, L., Viboud, C., Miller, M.A., Taylor, J., George, K.S., Griesemer, S.B., Ghedin, E., Sengamalay, N.A., Spiro, D.J., Volkov, I., Grenfell, B.T., Lipman, D.J., Taubenberger, J.K., Holmes, E.C., 2006, Stochastic processes are key determinants of short-term evolution in influenza a virus. PLoS Pathog 2, e125.

Nelson, M.I., Tan, Y., Ghedin, E., Wentworth, D.E., St George, K., Edelman, L., Beck, E.T., Fan, J., Lam, T.T., Kumar, S., Spiro, D.J., Simonsen, L., Viboud, C., Holmes, E.C., Henrickson, K.J., Musser, J.M., 2011, Phylogeography of the spring and fall waves of the H1N1/09 pandemic influenza virus in the United States. J Virol 85, 828-834.

Nemeroff, M.E., Utans, U., Kramer, A., Krug, R.M., 1992, Identification of cis-acting intron and exon regions in influenza virus NS1 mRNA that

inhibit splicing and cause the formation of aberrantly sedimenting presplicing complexes. Mol Cell Biol 12, 962-970.

Neumann, G., Hughes, M.T., Kawaoka, Y., 2000, Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. EMBO J 19, 6751-6758.

Neumann, G., Kawaoka, Y., 2006, Host range restriction and pathogenicity in the context of influenza pandemic. Emerg Infect Dis 12, 881-886.

Neumann, G., Noda, T., Kawaoka, Y., 2009, Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature 459, 931-939.

Nguyen-Van-Tam, J.S., Hampson, A.W., 2003, The epidemiology and clinical impact of pandemic influenza. Vaccine 21, 1762-1768.

Nicholson, K.G., Wood, J.M., Zambon, M., 2003, Influenza. Lancet 362, 1733-1745.

Nieto, A., de la Luna, S., Barcena, J., Portela, A., Ortin, J., 1994, Complex structure of the nuclear translocation signal of influenza virus polymerase PA subunit. J Gen Virol 75 (Pt 1), 29-36.

O'Neill, R.E., Talon, J., Palese, P., 1998, The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. EMBO 17, 288-296.

Obayashi, E., Yoshida, H., Kawai, F., Shibayama, N., Kawaguchi, A., Nagata, K., Tame, J.R., Park, S.Y., 2008, The structural basis for an essential subunit interaction in influenza virus RNA polymerase. Nature 454, 1127-1131.

Olsen, B., Munster, V.J., Wallensten, A., Waldenstrom, J., Osterhaus, A.D., Fouchier, R.A., 2006, Global patterns of influenza a virus in wild birds. Science 312, 384-388.

Olsen, C.W., 2002, The emergence of novel swine influenza viruses in North America. Virus Res 85, 199-210.

Osterhaus, A.D., Rimmelzwaan, G.F., Martina, B.E., Bestebroer, T.M., Fouchier, R.A., 2000, Influenza B virus in seals. Science 288, 1051-1053.

Palese, P., 1977, The genes of influenza virus. Cell 10, 1-10.

Palese, P., Schulman, J.L., 1976, Mapping of the influenza virus genome: identification of the hemagglutinin and the neuraminidase genes. Proc Natl Acad Sci U S A 73, 2142-2146.

Palese, P., Tobita, K., Ueda, M., Compans, R.W., 1974, Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. Virology 61, 397-410.

Pan, C., Cheung, B., Tan, S., Li, C., Li, L., Liu, S., Jiang, S., 2010, Genomic signature and mutation trend analysis of pandemic (H1N1) 2009 influenza A virus. PLoS ONE 5, e9549.

Patterson, K.D., Pyle, G.F., 1991, The geography and mortality of the 1918 influenza pandemic. Bull Hist Med 65, 4-21.

Pensaert, M., Ottis, K., Vandeputte, J., Kaplan, M.M., Bachmann, P.A., 1981, Evidence for the natural transmission of influenza A virus from wild ducts to swine and its potential importance for man. Bull World Health Organ 59, 75-78.

Perales, B., de la Luna, S., Palacios, I., Ortin, J., 1996, Mutational analysis identifies functional domains in the influenza A virus PB2 polymerase subunit. J Virol 70, 1678-1686.

Perales, B., Ortin, J., 1997, The influenza A virus PB2 polymerase subunit is required for the replication of viral RNA. J Virol 71, 1381-1385.

Pielak, R.M., Schnell, J.R., Chou, J.J., 2009, Mechanism of drug inhibition and drug resistance of influenza A M2 channel. Proc Natl Acad Sci U S A 106, 7379-7384.

Pina-Martins, F., Paulo, O.S., 2008, concatenator: sequence data matrices handling made easy. Mol Ecol Resour 8, 1254-1255.

Pinto, L.H., Holsinger, L.J., Lamb, R.A., 1992, Influenza virus M2 protein has ion channel activity. Cell 69, 517-528.

Poch, O., Sauvaget, I., Delarue, M., Tordo, N., 1989, Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J 8, 3867-3874.

Poole, E., Elton, D., Medcalf, L., Digard, P., 2004, Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. Virology 321, 120-133.

Portela, A., Digard, P., 2002, The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. J Gen Virol 83, 723-734.

Potdar, V.A., Chadha, M.S., Jadhav, S.M., Mullick, J., Cherian, S.S., Mishra, A.C., 2010, Genetic characterization of the influenza A pandemic (H1N1) 2009 virus isolates from India. PLoS ONE 5, e9693.

Pybus, O.G., Rambaut, A., 2009, Evolutionary analysis of the dynamics of viral infectious disease. Nat Rev Genet 10, 540-550.

Qu, Y., Zhang, R., Cui, P., Song, G., Duan, Z., Lei, F., 2011, Evolutionary genomics of the pandemic 2009 H1N1 influenza viruses (pH1N 1v). Virol J 8, 250.

Rambaut, A., Drummond, A.J. 2009. Tracer v1.5: MCMC trace analyses tool. .

Rambaut, A., Pybus, O.G., Nelson, M.I., Viboud, C., Taubenberger, J.K., Holmes, E.C., 2008, The genomic and epidemiological dynamics of human influenza A virus. Nature 453, 615-619.

Raymond, F.L., Caton, A.J., Cox, N.J., Kendal, A.P., Brownlee, G.G., 1986, The antigenicity and evolution of influenza H1 haemagglutinin, from 1950-1957 and 1977-1983: two pathways from one gene. Virology 148, 275-287.

Robb, N.C., Smith, M., Vreede, F.T., Fodor, E., 2009, NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. J Gen Virol 90, 1398-1407.

Ronquist, F., Huelsenbeck, J.P., 2003, MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Ruigrok, R.W., Crepin, T., Hart, D.J., Cusack, S., 2010, Towards an atomic resolution understanding of the influenza virus replication machinery. Curr Opin Struct Biol 20, 104-113.

Sanz-Ezquerro, J.J., Zurcher, T., de la Luna, S., Ortin, J., Nieto, A., 1996, The amino-terminal one-third of the influenza virus PA protein is responsible for the induction of proteolysis. J Virol 70, 1905-1911.

Scholtissek, C., Rohde, W., Von Hoyningen, V., Rott, R., 1978, On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology 87, 13-20.

Seo, S.H., Hoffmann, E., Webster, R.G., 2002, Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. Nat Med 8, 950-954.

Sharma, S., Joshi, G., Dash, P.K., Thomas, M., Athmaram, T.N., Kumar, J.S., Desai, A., Vasanthapuram, R., Patro, I.K., Rao, P.V., Parida, M., 2013, Molecular epidemiology and complete genome characterization of H1N1pdm virus from India. PLoS ONE 8, e56364.

Shi, L., Summers, D.F., Peng, Q., Galarz, J.M., 1995, Influenza A virus RNA polymerase subunit PB2 is the endonuclease which cleaves host cell mRNA and functions only as the trimeric enzyme. Virology 208, 38-47.

Shiino, T., Okabe, N., Yasui, Y., Sunagawa, T., Ujike, M., Obuchi, M., Kishida, N., Xu, H., Takashita, E., Anraku, A., Ito, R., Doi, T., Ejima, M., Sugawara, H., Horikawa, H., Yamazaki, S., Kato, Y., Oguchi, A., Fujita, N., Odagiri, T., Tashiro, M., Watanabe, H., 2010, Molecular evolutionary analysis of the influenza A(H1N1)pdm, May-September, 2009: temporal and spatial spreading profile of the viruses in Japan. PLoS ONE 5, e11057.

Shinde, V., Bridges, C.B., Uyeki, T.M., Shu, B., Balish, A., Xu, X., Lindstrom, S., Gubareva, L.V., Deyde, V., Garten, R.J., Harris, M., Gerber, S., Vagasky, S., Smith, F., Pascoe, N., Martin, K., Dufficy, D., Ritger, K., Conover, C., Quinlisk, P., Klimov, A., Bresee, J.S., Finelli, L., 2009, Triple-reassortant swine influenza A (H1) in humans in the United States, 2005-2009. N Engl J Med 360, 2616-2625.

Shope, R.E., 1931, The Etiology of Swine Influenza. Science 73, 214-215.

Skehel, J.J., Bayley, P.M., Brown, E.B., Martin, S.R., Waterfield, M.D., White, J.M., Wilson, I.A., Wiley, D.C., 1982, Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. Proc Natl Acad Sci U S A 79, 968-972.

Skehel, J.J., Wiley, D.C., 2000, Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69, 531-569.

Smith, G.J., Vijaykrishna, D., Bahl, J., Lycett, S.J., Worobey, M., Pybus, O.G., Ma, S.K., Cheung, C.L., Raghwani, J., Bhatt, S., Peiris, J.S., Guan, Y., Rambaut, A., 2009, Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459, 1122-1125.

Snyder, M.H., Buckler-White, A.J., London, W.T., Tierney, E.L., Murphy, B.R., 1987, The avian influenza virus nucleoprotein gene and a specific constellation of avian and human virus polymerase genes each specify attenuation of avian-human influenza A/Pintail/79 reassortant viruses for monkeys. J Virol 61, 2857-2863.

Song, J., Feng, H., Xu, J., Zhao, D., Shi, J., Li, Y., Deng, G., Jiang, Y., Li, X., Zhu, P., Guan, Y., Bu, Z., Kawaoka, Y., Chen, H., 2011, The PA protein directly contributes to the virulence of H5N1 avian influenza viruses in domestic ducks. J Virol 85, 2180-2188.

Steinhauer, D.A., Wharton, S.A., 1998, Structure and function of the haemagglutinin, In: Nicholson, K.G., Webster, R. G., Hay, A. J. (Ed.) Textbook of influenza. Blackwell Science Ltd, pp. 54-64.

Stevens, J., Blixt, O., Tumpey, T.M., Taubenberger, J.K., Paulson, J.C., Wilson, I.A., 2006, Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. Science 312, 404-410.

Stohr, K., 2003, The global agenda on influenza surveillance and control. Vaccine 21, 1744-1748.

Sugiyama, K., Obayashi, E., Kawaguchi, A., Suzuki, Y., Tame, J.R., Nagata, K., Park, S.Y., 2009, Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. EMBO J 28, 1803-1811.

Suzuki, T., Takahashi, T., Guo, C.T., Hidari, K.I., Miyamoto, D., Goto, H., Kawaoka, Y., Suzuki, Y., 2005, Sialidase activity of influenza A virus in an endocytic pathway enhances viral replication. J Virol 79, 11705-11715.

Takeuchi, K., Lamb, R.A., 1994, Influenza virus M2 protein ion channel activity stabilizes the native form of fowl plague virus hemagglutinin during intracellular transport. J Virol 68, 911-919.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28, 2731-2739.

Taubenberger, J.K., Kash, J.C., 2010, Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe 7, 440-451.

Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., Yang, H., Chen, X., Recuenco, S., Gomez, J., Chen, L.-M., Johnson, A., Tao, Y., Dreyfus, C., Yu, W., McBride, R., Carney, P.J., Gilbert, A.T., Chang, J., Guo, Z., Davis, C.T., Paulson, J.C., Stevens, J., Rupprecht, C.E., Holmes, E.C., Wilson, I.A., Donis, R.O., 2013, New World Bats Harbor Diverse Influenza A Viruses. PLoS Pathog 9, e1003657.

Toyoda, T., Adyshev, D.M., Kobayashi, M., Iwata, A., Ishihama, A., 1996, Molecular assembly of the influenza virus RNA polymerase: determination of the subunit-subunit contact sites. J Gen Virol 77 (Pt 9), 2149-2157.

Tscherne, D.M., Garcia-Sastre, A., 2011, Virulence determinants of pandemic influenza viruses. J Clin Invest 121, 6-13.

Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solorzano, A., Swayne, D.E., Cox, N.J., Katz, J.M., Taubenberger, J.K., Palese, P., Garcia-Sastre, A., 2005, Characterization of the reconstructed 1918 Spanish influenza pandemic virus. Science 310, 77-80.

Van Hoeven, N., Pappas, C., Belser, J.A., Maines, T.R., Zeng, H., Garcia-Sastre, A., Sasisekharan, R., Katz, J.M., Tumpey, T.M., 2009, Human HA and polymerase subunit PB2 proteins confer transmission of an avian influenza virus through the air. Proc Natl Acad Sci U S A 106, 3366-3371.

Varghese, J.N., Colman, P.M., 1991, Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 A resolution. J Mol Biol 221, 473-486.

Varghese, J.N., Laver, W.G., Colman, P.M., 1983, Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution. Nature 303, 35-40.

Venter, M., Naidoo, D., Pretorius, M., Buys, A., McAnerney, J., Blumberg, L., Madhi, S.A., Cohen, C., Schoub, B., 2012, Evolutionary dynamics of 2009 pandemic influenza A virus subtype H1N1 in South Africa during 2009-2010. J Infect Dis 206 Suppl 1, S166-172.

Viboud, C., Nelson, M.I., Tan, Y., Holmes, E.C., 2013, Contrasting the epidemiological and evolutionary dynamics of influenza spatial transmission. Philos Trans R Soc Lond B Biol Sci 368, 20120199.

Viboud, C., Simonsen, L., 2012, Global mortality of 2009 pandemic influenza A H1N1. Lancet Infect Dis 12, 651-653.

Wang, C., Takeuchi, K., Pinto, L.H., Lamb, R.A., 1993, Ion channel activity of influenza A virus M2 protein: characterization of the amantadine block. J Virol 67, 5585-5594.

Wang, P., Song, W., Mok, B.W., Zhao, P., Qin, K., Lai, A., Smith, G.J., Zhang, J., Lin, T., Guan, Y., Chen, H., 2009, Nuclear factor 90 negatively regulates influenza virus replication by interacting with viral nucleoprotein. J Virol 83, 7850-7861.

Watanabe, T., Watanabe, S., Kawaoka, Y., 2010, Cellular networks involved in the influenza virus life cycle. Cell Host Microbe 7, 427-439.

Webster, R.G., 1993, Emerging viruses, In: Morse, S. (Ed.) Influenza. Oxford University Press, New York.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992, Evolution and ecology of influenza A viruses. Microbiol Rev 56, 152-179.

Webster, R.G., Laver, W.G., 1975, Antigenic variation of influenza viruses, In: Kilbourne, E.D. (Ed.) The influenza viruses and influenza. Academic press, New York.

Whittaker, G., Bui, M., Helenius, A., 1996, Nuclear trafficking of influenza virus ribonuleoproteins in heterokaryons. J Virol 70, 2743-2756.

WHO 2002. WHO manual on animal influenza diagnosis and surveillance (Geneva, World Health Organization).

WHO 2009. World now at the start of 2009 influenza pandemic (WHO).

WHO 2013. Terms of reference for National Influenza Centres.

Wilson, I.A., Skehel, J.J., Wiley, D.C., 1981, Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature 289, 366-373.

Xu, R., Ekiert, D.C., Krause, J.C., Hai, R., Crowe, J.E., Jr., Wilson, I.A., 2010, Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. Science 328, 357-360.

Yamada, S., Hatta, M., Staker, B.L., Watanabe, S., Imai, M., Shinya, K., Sakai-Tagawa, Y., Ito, M., Ozawa, M., Watanabe, T., Sakabe, S., Li, C., Kim, J.H., Myler, P.J., Phan, I., Raymond, A., Smith, E., Stacy, R., Nidom, C.A., Lank, S.M., Wiseman, R.W., Bimber, B.N., O'Connor, D.H., Neumann, G., Stewart, L.J., Kawaoka, Y., 2010, Biological and structural characterization of a host-adapting amino acid in influenza virus. PLoS Pathog 6, e1001034.

Yang, J.R., Huang, Y.P., Chang, F.Y., Hsu, L.C., Lin, Y.C., Su, C.H., Chen, P.J., Wu, H.S., Liu, M.T., 2011, New variants and age shift to high fatality groups contribute to severe successive waves in the 2009 influenza pandemic in Taiwan. PLoS ONE 6, e28288.

Young, S.A., 2009, Hemadsorption and hemagglutination inhibition, In: Specter, S., Hodinka, R.L., Young, S.A. (Eds.) Clinical virology manual. ASM Press, Washington, DC, pp. 119-123.

Yuan, H.Y., Koelle, K., 2013, The evolutionary dynamics of receptor binding avidity in influenza A: a mathematical model for a new antigenic drift hypothesis. Philos Trans R Soc Lond B Biol Sci 368, 20120204.

Yuan, P., Bartlam, M., Lou, Z., Chen, S., Zhou, J., He, X., Lv, Z., Ge, R., Li, X., Deng, T., Fodor, E., Rao, Z., Liu, Y., 2009, Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. Nature 458, 909-913.

Zambon, M., Hayden, F.G., 2001, Position statement: global neuraminidase inhibitor susceptibility network. Antiviral Res 49, 147-156.

Zehender, G., Pariani, E., Piralla, A., Lai, A., Gabanelli, E., Ranghiero, A., Ebranati, E., Amendola, A., Campanini, G., Rovida, F., Ciccozzi, M., Galli, M., Baldanti, F., Zanetti, A.R., 2012, Reconstruction of the evolutionary dynamics of the A(H1N1)pdm09 influenza virus in Italy during the pandemic and post-pandemic phases. PLoS ONE 7, e47517.

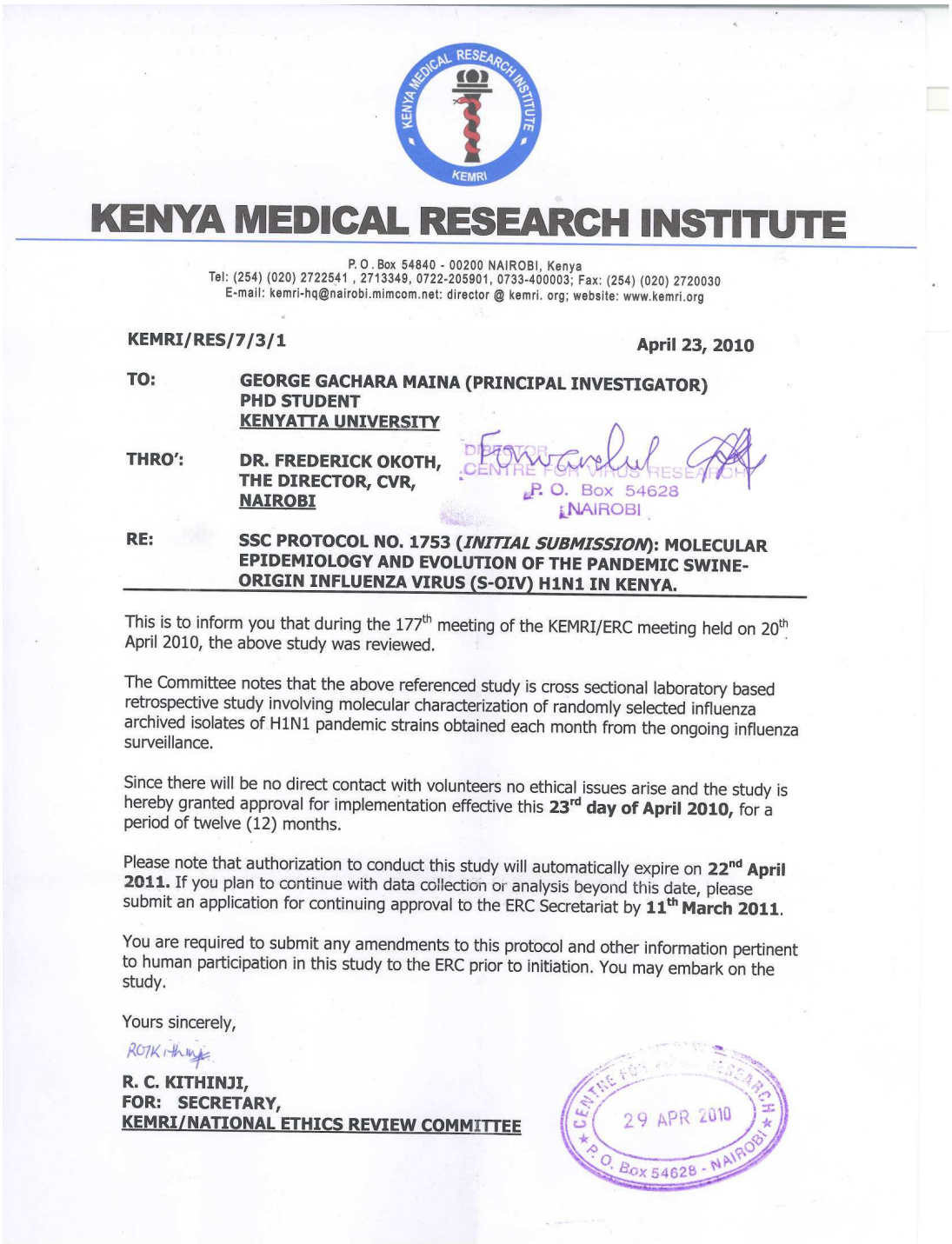
Zhang, W.D., Evans, D.H., 1991, Detection and identification of human influenza viruses by the polymerase chain reaction. J Virol Methods 33, 165-189.

Zhou, N.N., Senne, D.A., Landgraf, J.S., Swenson, S.L., Erickson, G., Rossow, K., Liu, L., Yoon, K., Krauss, S., Webster, R.G., 1999, Genetic reassortment of avian, swine, and human influenza A viruses in American pigs. J Virol 73, 8851-8856.

Ziegler, T., Hall, H., Sanchez-Fauquier, A., Gamble, W.C., Cox, N.J., 1995, Type and subtype-specific detection of influenza viruses in clinical specimens by rapid culture assay. Journal of Clinical Microbiology 33, 318-321.

APPENDICES

Appendix 1: Ethical approval for the study



KENYA MEDICAL RESEARCH INSTITUTE

P. O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541 , 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: kemri-hq@nairobi.mimcom.net: director @ kemri. org; website: www.kemri.org

KEMRI/RES/7/3/1 **April 23, 2010**

TO: **GEORGE GACHARA MAINA (PRINCIPAL INVESTIGATOR)**
PHD STUDENT
KENYATTA UNIVERSITY

THRO': **DR. FREDERICK OKOTH,**
THE DIRECTOR, CVR,
NAIROBI

DIRECTOR
CENTRE FOR VIRUS RESEARCH
P. O. Box 54628
NAIROBI

RE: **SSC PROTOCOL NO. 1753 (*INITIAL SUBMISSION*): MOLECULAR EPIDEMIOLOGY AND EVOLUTION OF THE PANDEMIC SWINE-ORIGIN INFLUENZA VIRUS (S-OIV) H1N1 IN KENYA.**

This is to inform you that during the 177th meeting of the KEMRI/ERC meeting held on 20th April 2010, the above study was reviewed.

The Committee notes that the above referenced study is cross sectional laboratory based retrospective study involving molecular characterization of randomly selected influenza archived isolates of H1N1 pandemic strains obtained each month from the ongoing influenza surveillance.

Since there will be no direct contact with volunteers no ethical issues arise and the study is hereby granted approval for implementation effective this **23rd day of April 2010,** for a period of twelve (12) months.

Please note that authorization to conduct this study will automatically expire on **22nd April 2011.** If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to the ERC Secretariat by **11th March 2011**.

You are required to submit any amendments to this protocol and other information pertinent to human participation in this study to the ERC prior to initiation. You may embark on the study.

Yours sincerely,

R. C. KITHINJI,
FOR: SECRETARY,
KEMRI/NATIONAL ETHICS REVIEW COMMITTEE

CENTRE FOR VIRUS RESEARCH
29 APR 2010
P. O. Box 54628 - NAIROBI

Appendix 2: Primers used in this study

Gene	Fragment	Primer name	Primer sequence 5'- 3'*#
PB2	1	PB2_F1	TGT AAA ACG ACG GCC AGT CTC GAG CAA AAG CAG GTC AA
		PB2_R2	CAG GAA ACA GCT ATG ACC GCT TTG RTC AAY ATC RTC ATT
	2	PB2_F3	TGT AAA ACG ACG GCC AGT CCT GGT CAY GCA GAC CTC AG
		PB2_R4	CAG GAA ACA GCT ATG ACC CCT CTA ACT GCT TTT A**Y**C ATG CAA T
	3	PB2_F5	TGT AAA ACG ACG GCC AGT CC**R** AC**W** GAA GAA CAA GCT GT
		PB2_R6	CAG GAA ACA GCT ATG ACC CTG AGA CCA YTG AAT TTT RAC A
	4	PB2_F7	TGT AAA ACG ACG GCC AGT CCA AG**Y** AC**M** GAG ATG TCA ATG AGA
		PB2_R8	CAG GAA ACA GCT ATG ACC AGT AGA AAC AAG GTC GTT TTT AAA C
PB1	1	PB1_F1	TGT AAA ACG ACG GCC AGT AGC AAA AGC AGG CAA ACC AT
		PB1_R2	CAG GAA ACA GCT ATG ACC GTT CAA GCT TTT C**R**C A**W**A TG
	2	PB1_F3	TGT AAA ACG ACG GCC AGT ACA AG**R** GTG GAC AAA T**R**A C
		PB1_R4	CAG GAA ACA GCT ATG ACC TTG AAC ATG CCC ATC ATC AT**Y** CCA GG
	3	PB1_F5	TGT AAA ACG ACG GCC AGT AAT CAA AA**Y** CCT **M**GA ATG TT
		PB1_R6	CAG GAA ACA GCT ATG ACC CCA AG**R** TCA TTG TTT ATC AT
	4	PB1_F7	TGT AAA ACG ACG GCC AGT ATG AG**Y** AAA AAG AAG TC**Y** TA
		PB1_R8	CAG GAA ACA GCT ATG ACC AGT AGA AAC AAG GCA TTT
PA	1	PA_F1	TGT AAA ACG ACG GCC AGT AGC AAA AGC AGG TAC TGA T
		PA_R2	CAG GAA ACA GCT ATG ACC TGA GAA AGC TTG CCC TCA ATG
	2	PA_F3	TGT AAA ACG ACG GCC AGT TAT GA**Y** TAC AA**R** GAG AA
		PA_R4	CAG GAA ACA GCT ATG ACC TC**R** CA**K** GCC TTG TTG AAC TCA TT
	3	PA_F5	TGT AAA ACG ACG GCC AGT AAT TGA ACC ATT CTT GAG GAC GA
		PA_R6	CAG GAA ACA GCT ATG ACC TCC GAT TGG CCA TGT TTC CGA T
	4	PA_F7	TGT AAA ACG ACG GCC AGT AAT GCA TCC TGT GCA GCA ATG GA
		PA_R8	CAG GAA ACA GCT ATG ACC AGT AGA AAC AAG GTA CCT TTT
HA	1	SwHAIMI3F	TGTAAAACGACGGCCAGTATGAAGGCAATACTAGTAG
		SwHAIMI3R	CAGGAAACAGCTATGACCGATCGGATGTATATTCTGAAATGG
	2	SwHAIIMI3F	TGTAAAACGACGGCCAGTGATTGCAATACAACTTGTC
		SwHAIIMI3R	CAGGAAACAGCTATGACCAATACATATTCTACACTGTAGAGACCCA
		NP_F1	TGT AAA ACG ACG GCC AGT CAG GGT AGA TAA TCA CTC AC

NP		NP_R5	CAG GAA ACA GCT ATG ACC AGT AGA AAC AAG GGT ATT TTT C
NA	1	SwNAIMI3F	TGTAAAACGACGGCCAGTATGAATCCAAAYCARAAGAT
		SwNAIMI3R	CAGGAAACAGCTATGACCGTGATAATTAGGGGCATTC
	2	SwNAIIMI3F	TGTAAAACGACGGCCAGTGACAGGCCTCATACAAGATCTTC
		SwNAIIMI3R	CAGGAAACAGCTATGACCTAAATGGMAACTCAGCACCG
MP		A-MP-M13F	TGTAAAACGACGGCCAGTAGCAAAAGCAGGTAG
		A-MP-M13R	CAGGAAACAGCTATGACCAGTAGAAACAAGGTAGT
NS		A-NS-M13F	TGTAAAACGACGGCCAGTAGCAAAAGCAGGGTGACAAAGACA
		A-NS-M13R	CAGGAAACAGCTATGACCAGTAGAAACAAGGGTGTTTTTTAT
Sequencing		M13F	TGTAAAACGACGGCCAGT
		M13R	CAGGAAACAGCTATGACC

* All primers bore the M13 tag

[#]Standard mixed oligonucleotide bases for primer sequences are listed in bold

Appendix 3: Epidemiological information about the cases and isolates used in this study

	Date of collection	Sample name	Site of collection	Age	Sex	Nationality
1.	02.07.2009	A/Kisumu/56/2009	CDC - Kisumu	18	Male	Kenyan
2.	02.07.2009	A/Kisumu/61/2009	CDC - Kisumu	22	Female	British
3.	03.07.2009	A/Nairobi/64/2009	Nairobi Hospital	5	Male	British
4.	05.07.2009	A/Garissa/78/2009	Garissa	8	Male	British
5.	11.08.2009	A/Keiyo/96/2009	Tabare	5	Male	Kenyan
6.	15.09.2009	A/Nairobi/58/2009	Kenya High	-	Female	Kenyan
7.	15.09.2009	A/Nairobi/59/2009	Kenya High	15	Female	Kenyan
8.	19.09.2009	A/Eldoret/119/2009	Sacred Heart Boys	15	Male	Kenyan
9.	19.09.2009	A/Eldoret/120/2009	Sacred Heart Boys	14	Male	Kenyan
10.	22.09.2009	A/Embu/169/2009	Mbeere District Hosp	13	Male	Kenyan
11.	24.09.2009	A/Malindi/238/2009	Malindi Dist Hospital	17	Male	Kenyan
12.	29.09.2009	A/Meru/467/2009	Materi Girls	18	Female	Kenyan
13.	29.09.2009	A/Nyeri/478/2009	Nyeri	13	Male	Kenyan
14.	29.09.2009	A/Mombasa/512/2009	Mombasa Hospital	12	Male	Kenyan
15.	30.09.2009	A/Kitale/531/2009	Weitaluu Sec Sch Kitale	20	Female	Kenyan
16.	30.09.2009	A/Kitale/532/2009	Bikere Pri-Kitale	15	Female	Kenyan
17.	02.10.2009	A/Mombasa/27/2009	AGH-Mombasa	11	Female	Kenyan
18.	03.10.2009	A/Nairobi/37/2009	Nairobi West	25	Male	Kenyan
19.	06.10.2009	A/Mombasa/91/2009	Mombasa	9	-	Kenyan
20.	12.10.2009	A/Kisii/115/2009	Kisii District Hospital	20	Female	Kenyan
21.	15.10.2009	A/Kisii/143/2009	Kisii District Hospital	11	Male	Kenyan
22.	14.10.2009	A/Mombasa/148/2009	Mombasa	6	Male	Kenyan
23.	22.10.2009	A/Transnzoia/168/2009	Kwanza Health Centre	16	Male	Kenyan
24.	27.10.2009	A/Mombasa/179/2009	Aga Khan Hospital Msa	10	Female	Kenyan
25.	27.10.2009	A/Kikuyu/184/2009	PCEA Kikuyu Hospital	12	Female	Kenyan
26.	27.10.2009	A/Kikuyu/185/2009	PCEA Kikuyu Hospital	11	Male	Kenyan
27.	27.10.2009	A/Nakuru/192/2009	Nakuru Prov Gen Hosp	19	Male	Kenyan
28.	29.10.2009	A/Kakamega/215/2009	Kakamega Prov G Hosp	16	Female	Kenyan
29.	29.10.2009	A/Kisii/205/2009	Kisii District Hospital	9	Female	Kenyan

30.	04.11.2009	A/Nairobi/3/2009	Mater Hospital	21	Female	Kenyan
31.	09.12.2009	A/Nairobi/16/2009	Kemri Staff Clinic	3	Male	American
32.	21/01/2010	A/Kijabe/16/2010	Kijabe	11	Female	-
33.	12.03.2010	A/Nairobi/11/2010	Kijabe	34	Female	-
34.	20.03.2010	A/Nairobi/20/2010	Nairobi	6	Female	Kenyan
35.	20.03.2010	A/Nairobi/21/2010	Nairobi	14	Male	Kenyan
36.	25.03.2010	A/Nairobi/24/2010	Nairobi	3	Female	Kenyan
37.	25.03.2010	A/Nairobi/25/2010	Nairobi	5	Female	Kenyan
38.	04.06.2010	A/Nairobi/80/2010	Nairobi	35	Female	Kenyan
39.	04.06.2010	A/Nairobi/72/2010	Nairobi	3	Male	Kenyan
40.	12.06.2010	A/Nairobi/97/2010	Nairobi	6	Male	Kenyan

Early pandemic samples are written in red, peak pandemic samples in black and late pandemic samples in blue

Appendix 4: Preparation of reagents

Virus transport medium (broth)

1. To a 500ml bottle add:
    a) 10 gm veal infusion broth (2.5% final conc.)
    b) 2.0 gm of BSA (0.5% ")
    c) 0.8 ml gentamicin sulfate (100 µg/ml ")
    d) 3.2 ml of Fungizone (2 µg/ml ")
    e) EASYpure water to 400 ml.

2. Swirl gently to dissolve or let stand at 4°C for 1 hour.
3. Sterile filter.
4. Label as "virus transport medium", with your name, today's date, and an expiration date of 3 months from the date prepared. Store at 4°C.

Preparation of Guinea pig Blood

1. Request 5 ml of guinea pig blood from animal resources at least 2 days before starting the assay.
2. Immediately add to 40 ml of Alsever's solution in a 50 ml sterile screw cap tube.
3. Centrifuge at 800-x g for 10 min and aspirate to remove <u>all</u> the supernatant. Add 35 ml of Alsever's and gently mix by inverting.
4. Repeat step 3 two additional times, but end by adding exactly 9 volumes of Alsever's to 1 volume of packed RBCs (to give a 10% solution). The volume of RBC can be checked by adding the RBCs to a 15 ml graduated centrifuge tube and centrifuging or allowing them to settle.
5. Label as "10% guinea pig RBC" with the day's date, name and an expiration date of 20 days.
6. Store at 4°C.
7. Before use, gently resuspend and add one ml of cells to nine ml of sterile PBS (to give a 1% final concentration). If there is clotting, the cells can be filtered through sterile gauze before use. This gives a 1% suspension ready for use.

1 % Agarose

1. Weigh 1g of agarose and add it onto a 250ml flask
2. Add 1x TAE slowly to the flask until the 100ml mark.
3. Microwave the flask for 1-2 min making sure that the mixture does not boil over.
4. Stop the microwave every 30 s and swirl the flask. Continue to heat until the agarose is completely dissolved
5. Add 5 µl of ethidium bromide to the dissolved agarose and mix.

Appendix 5: Accession numbers of nucleotide sequences of Influenza A(H1N1)pdm09 viruses from this study

POLYMERASE BASIC 2 (PB2) GENE

Isolate	Accession number
A/Kisumu/56/2009	EPI443014
A/Kisumu/61/2009	EPI443064
A/Nairobi/64/2009	EPI443103
A/Garissa/78/2009	EPI443156
A/Keiyo/96/2009	EPI443166
A/Nairobi/58/2009	EPI443173
A/Nairobi/59/2009	EPI443181
A/Malindi/238/2009	EPI443188
A/Eldoret/119/2009	EPI443195
A/Eldoret/120/2009	EPI443202
A/Embu/169/2009	EPI443209
A/Nyeri/478/2009	EPI443216
A/Meru/467/2009	EPI443223
A/Kitale/531/2009	EPI443230
A/Kitale/532/2009	EPI443238
A/Mombasa/512/2009	EPI443245
A/Nakuru/192/2009	EPI443252
A/Trans-Nzoia/168/2009	EPI443260
A/Mombasa/179/2009	EPI443268
A/Mombasa/148/2009	EPI443275
A/Mombasa/91/2009	EPI443282
A/Mombasa/27/2009	EPI446009
A/Nairobi/37/2009	EPI446017
A/Kakamega/215/2009	EPI446026
A/Kisii/205/2009	EPI446033
A/Kisii/143/2009	EPI446040
A/Kisii/115/2009	EPI446047
A/Kikuyu/184/2009	EPI446054
A/Kikuyu/185/2009	EPI446061
A/Nairobi/3/2009	EPI446068
A/Nairobi/16/2009	EPI446075
A/Nairobi/11/2010	EPI446083
A/Kijabe/16/2010	EPI446091
A/Nairobi/20/2010	EPI446099
A/Nairobi/21/2010	EPI446107
A/Nairobi/24/2010	EPI446115
A/Nairobi/25/2010	EPI446123
A/Nairobi/72/2010	EPI446131
A/Nairobi/80/2010	EPI446139
A/Nairobi/97/2010	EPI446152

POLYMERASE BASIC 1 (PB1) GENE

Isolate	Accession number
A/Kisumu/56/2009	EPI443015
A/Kisumu/61/2009	EPI443065
A/Nairobi/64/2009	EPI443112
A/Garissa/78/2009	EPI443157
A/Keiyo/96/2009	EPI443167
A/Nairobi/58/2009	EPI443174
A/Nairobi/59/2009	EPI443182
A/Malindi/238/2009	EPI443189
A/Eldoret/119/2009	EPI443196
A/Eldoret/120/2009	EPI443203
A/Embu/169/2009	EPI443210
A/Nyeri/478/2009	EPI443217
A/Meru/467/2009	EPI443224
A/Kitale/531/2009	EPI443231
A/Kitale/532/2009	EPI443239
A/Mombasa/512/2009	EPI443246
A/Nakuru/192/2009	EPI443253
A/Trans-Nzoia/168/2009	EPI443261
A/Mombasa/179/2009	EPI443269
A/Mombasa/148/2009	EPI443276
A/Mombasa/91/2009	EPI443283
A/Mombasa/27/2009	EPI446010
A/Nairobi/37/2009	EPI446018
A/Kakamega/215/2009	EPI446027
A/Kisii/205/2009	EPI446034
A/Kisii/143/2009	EPI446041
A/Kisii/115/2009	EPI446048
A/Kikuyu/184/2009	EPI446055
A/Kikuyu/185/2009	EPI446062
A/Nairobi/3/2009	EPI446069
A/Nairobi/16/2009	EPI446076
A/Nairobi/11/2010	EPI446084
A/Kijabe/16/2010	EPI446092
A/Nairobi/20/2010	EPI446100
A/Nairobi/21/2010	EPI446108
A/Nairobi/24/2010	EPI446116
A/Nairobi/25/2010	EPI446124
A/Nairobi/72/2010	EPI446132
A/Nairobi/80/2010	EPI446140
A/Nairobi/97/2010	EPI446153

POLYMERASE ACID (PA) GENE

Isolate	Accession number
A/Kisumu/56/2009	EPI443016
A/Kisumu/61/2009	EPI443066
A/Nairobi/64/2009	EPI443113
A/Garissa/78/2009	EPI443158
A/Keiyo/96/2009	EPI443168
A/Nairobi/58/2009	EPI443175
A/Nairobi/59/2009	EPI443183
A/Malindi/238/2009	EPI443190
A/Eldoret/119/2009	EPI443197
A/Eldoret/120/2009	EPI443204
A/Embu/169/2009	EPI443211
A/Meru/467/2009	EPI443225
A/Nyeri/478/2009	EPI443218
A/Kitale/531/2009	EPI443232
A/Kitale/532/2009	EPI443240
A/Nakuru/192/2009	EPI443254
A/Trans-Nzoia/168/2009	EPI443262
A/Mombasa/179/2009	EPI443270
A/Nairobi/37/2009	EPI446019
A/Mombasa/91/2009	EPI443284
A/Kisii/115/2009	EPI446049
A/Kisii/143/2009	EPI446042
A/Mombasa/148/2009	EPI443277
A/Mombasa/27/2009	EPI446011
A/Kikuyu/184/2009	EPI446056
A/Kikuyu/185/2009	EPI446063
A/Kisii/205/2009	EPI446035
A/Kakamega/215/2009	EPI446028
A/Mombasa/512/2009	EPI443247
A/Nairobi/16/2009	EPI446077
A/Nairobi/3/2009	EPI446070
A/Nairobi/20/2010	EPI446101
A/Nairobi/21/2010	EPI446109
A/Nairobi/24/2010	EPI446117
A/Nairobi/25/2010	EPI446125
A/Nairobi/72/2010	EPI446133
A/Nairobi/80/2010	EPI446141
A/Nairobi/97/2010	EPI446154
A/Nairobi/11/2010	EPI446085
A/Kijabe/16/2010	EPI446093

HAEMAGGLUTININ GENE

Isolate	Accession number
A/Kisumu/56/2009	EPI443017
A/Kisumu/61/2009	EPI443067
A/Nairobi/64/2009	EPI443120
A/Garissa/78/2009	EPI443159
A/Keiyo/96/2009	EPI443169
A/Nairobi/58/2009	EPI443176
A/Nairobi/59/2009	EPI443184
A/Malindi/238/2009	EPI443191
A/Eldoret/119/2009	EPI443198
A/Eldoret/120/2009	EPI443205
A/Embu/169/2009	EPI443212
A/Nyeri/478/2009	EPI443219
A/Meru/467/2009	EPI443226
A/Kitale/531/2009	EPI443233
A/Kitale/532/2009	EPI443241
A/Mombasa/512/2009	EPI443248
A/Nakuru/192/2009	EPI443255
A/Trans-Nzoia/168/2009	EPI443263
A/Mombasa/179/2009	EPI443271
A/Mombasa/148/2009	EPI443278
A/Mombasa/91/2009	EPI443285
A/Mombasa/27/2009	EPI446012
A/Nairobi/37/2009	EPI446020
A/Kakamega/215/2009	EPI446029
A/Kisii/205/2009	EPI446036
A/Kisii/143/2009	EPI446043
A/Kisii/115/2009	EPI446050
A/Kikuyu/184/2009	EPI446057
A/Kikuyu/185/2009	EPI446064
A/Nairobi/3/2009	EPI446071
A/Nairobi/16/2009	EPI446078
A/Kijabe/16/2010	EPI446094
A/Nairobi/11/2010	EPI446086
A/Nairobi/20/2010	EPI446102
A/Nairobi/21/2010	EPI446110
A/Nairobi/24/2010	EPI446118
A/Nairobi/25/2010	EPI446126
A/Nairobi/72/2010	EPI446134
A/Nairobi/80/2010	EPI446142
A/Nairobi/97/2010	EPI446155

NUCLEOPROTEIN GENE

Isolate	Accession number
A/Kisumu/56/2009	EPI443018
A/Kisumu/61/2009	EPI443068
A/Nairobi/64/2009	EPI443131
A/Garissa/78/2009	EPI443160
A/Keiyo/96/2009	EPI443170
A/Nairobi/58/2009	EPI443177
A/Nairobi/59/2009	EPI443185
A/Malindi/238/2009	EPI443192
A/Eldoret/119/2009	EPI443199
A/Eldoret/120/2009	EPI443206
A/Trans-Nzoia/168/2009	EPI443264
A/Embu/169/2009	EPI443213
A/Nakuru/192/2009	EPI443256
A/Nyeri/478/2009	EPI443220
A/Meru/467/2009	EPI443227
A/Kitale/531/2009	EPI443234
A/Kitale/532/2009	EPI443242
A/Mombasa/512/2009	EPI443249
A/Mombasa/179/2009	EPI443272
A/Mombasa/148/2009	EPI443279
A/Mombasa/91/2009	EPI443286
A/Mombasa/27/2009	EPI446013
A/Nairobi/37/2009	EPI446021
A/Kakamega/215/2009	EPI446030
A/Kisii/205/2009	EPI446037
A/Kisii/143/2009	EPI446044
A/Kisii/115/2009	EPI446051
A/Kikuyu/184/2009	EPI446058
A/Kikuyu/185/2009	EPI446065
A/Nairobi/3/2009	EPI446072
A/Nairobi/16/2009	EPI446079
A/Kijabe/16/2010	EPI446095
A/Nairobi/20/2010	EPI446103
A/Nairobi/21/2010	EPI446111
A/Nairobi/24/2010	EPI446119
A/Nairobi/25/2010	EPI446127
A/Nairobi/72/2010	EPI446135
A/Nairobi/80/2010	EPI446143
A/Nairobi/97/2010	EPI446156
A/Nairobi/11/2010	EPI446087

NEURAMINIDASE GENE

Isolate	Accession number
A/Kisumu/56/2009	EPI443019
A/Kisumu/61/2009	EPI443069
A/Nairobi/64/2009	EPI443141
A/Garissa/78/2009	EPI443161
A/Keiyo/96/2009	EPI443171
A/Nairobi/58/2009	EPI443178
A/Nairobi/59/2009	EPI443186
A/Malindi/238/2009	EPI443193
A/Eldoret/119/2009	EPI443200
A/Eldoret/120/2009	EPI443207
A/Embu/169/2009	EPI443214
A/Nyeri/478/2009	EPI443221
A/Meru/467/2009	EPI443228
A/Kitale/531/2009	EPI443235
A/Kitale/532/2009	EPI443243
A/Mombasa/512/2009	EPI443250
A/Nakuru/192/2009	EPI443257
A/Trans-Nzoia/168/2009	EPI443265
A/Mombasa/179/2009	EPI443273
A/Mombasa/148/2009	EPI443280
A/Mombasa/91/2009	EPI443287
A/Mombasa/27/2009	EPI446014
A/Nairobi/37/2009	EPI446022
A/Kakamega/215/2009	EPI446031
A/Kisii/205/2009	EPI446038
A/Kisii/143/2009	EPI446045
A/Kisii/115/2009	EPI446052
A/Kikuyu/184/2009	EPI446059
A/Kikuyu/185/2009	EPI446066
A/Nairobi/3/2009	EPI446073
A/Nairobi/16/2009	EPI446080
A/Nairobi/11/2010	EPI446088
A/Kijabe/16/2010	EPI446096
A/Nairobi/20/2010	EPI446104
A/Nairobi/21/2010	EPI446112
A/Nairobi/24/2010	EPI446120
A/Nairobi/25/2010	EPI446128
A/Nairobi/72/2010	EPI446136
A/Nairobi/80/2010	EPI446144
A/Nairobi/97/2010	EPI446157

MATRIX GENE

Isolate	Accession number
A/Kisumu/56/2009	EPI443020
A/Kisumu/61/2009	EPI443070
A/Nairobi/64/2009	EPI443144
A/Garissa/78/2009	EPI443163
A/Keiyo/96/2009	EPI443172
A/Nairobi/58/2009	EPI443179
A/Nairobi/59/2009	EPI443187
A/Malindi/238/2009	EPI443194
A/Eldoret/119/2009	EPI443201
A/Eldoret/120/2009	EPI443208
A/Embu/169/2009	EPI443215
A/Nyeri/478/2009	EPI443222
A/Meru/467/2009	EPI443229
A/Kitale/531/2009	EPI443236
A/Kitale/532/2009	EPI443244
A/Mombasa/512/2009	EPI443251
A/Nakuru/192/2009	EPI443258
A/Trans-Nzoia/168/2009	EPI443266
A/Mombasa/179/2009	EPI443274
A/Mombasa/148/2009	EPI443281
A/Mombasa/91/2009	EPI443288
A/Mombasa/27/2009	EPI446015
A/Nairobi/37/2009	EPI446023
A/Kakamega/215/2009	EPI446032
A/Kisii/205/2009	EPI446039
A/Kisii/143/2009	EPI446046
A/Kisii/115/2009	EPI446053
A/Kikuyu/184/2009	EPI446060
A/Kikuyu/185/2009	EPI446067
A/Nairobi/3/2009	EPI446074
A/Nairobi/16/2009	EPI446081
A/Kijabe/16/2010	EPI446097
A/Nairobi/11/2010	EPI446089
A/Nairobi/20/2010	EPI446105
A/Nairobi/21/2010	EPI446113
A/Nairobi/24/2010	EPI446121
A/Nairobi/25/2010	EPI446129
A/Nairobi/80/2010	EPI446145
A/Nairobi/72/2010	EPI446137
A/Nairobi/97/2010	EPI446158

NONSTRUCTURAL GENE

Isolate	Accession number
A/Kisumu/56/2009	HM855241
A/Kisumu/61/2009	HM855247
A/Nairobi/64/2009	HQ165791
A/Garissa/78/2009	HM855242
A/Keiyo/96/2009	HM855243
A/Nairobi/58/2009	HM855242
A/Nairobi/59/2009	EPI443180
A/Malindi/238/2009	HM855248
A/Eldoret/119/2009	HM855262
A/Eldoret/120/2009	HM855244
A/Embu/169/2009	HM855246
A/Nyeri/478/2009	HM855254
A/Meru/467/2009	HM855252
A/Kitale/531/2009	EPI443237
A/Kitale/532/2009	HM855256
A/Mombasa/512/2009	HM855255
A/Nakuru/192/2009	EPI443259
A/Trans-Nzoia/168/2009	EPI443267
A/Mombasa/179/2009	HQ165794
A/Mombasa/148/2009	HM855263
A/Mombasa/91/2009	HQ165792
A/Mombasa/27/2009	HQ165789
A/Nairobi/37/2009	HM855259
A/Kakamega/215/2009	HM855249
A/Kisii/205/2009	HQ165793
A/Kisii/143/2009	HM855257
A/Kisii/115/2009	HM855261
A/Kikuyu/184/2009	HM855258
A/Kikuyu/185/2009	HM855250
A/Nairobi/3/2009	HM855264
A/Nairobi/16/2009	EPI446082
A/Nairobi/11/2010	EPI446090
A/Kijabe/16/2010	EPI446098
A/Nairobi/20/2010	EPI446106
A/Nairobi/21/2010	EPI446114
A/Nairobi/24/2010	EPI446122
A/Nairobi/25/2010	EPI446130
A/Nairobi/80/2010	EPI446146
A/Nairobi/72/2010	EPI446138
A/Nairobi/97/2010	EPI446159

Appendix 6: Publications and conference abstracts from this study

Peer reviewed and submitted publications

Gachara, G., Samuel, S., John, M., James, S., Musa, N., Japheth, M., Wallace, B. 2011. Amino acid sequence analysis and identification of mutations in the NS gene of 2009 influenza A (H1N1) isolates from Kenya. *Virus Genes* **43**, 27-32.

Majanja, J., Njoroge, R.N., Achilla, R., Wurapa, E.K., Wadegu, M., Mukunzi, S., Mwangi, J., Njiri, J., **Gachara, G.,** Bulimo, W. 2013. Impact of Influenza A(H1N1)pdm09 Virus on Circulation Dynamics of Seasonal Influenza Strains in Kenya. *Am J Trop Med Hyg* **88**(5):940-5.

Gachara, G., Samuel S., J, Magana., B, Opot., M, Otieno., W, Bulimo. Molecular and Phylogenetic analysis of the surface glycoprotein genes of Influenza A(H1N1)pdm09 viruses isolated in Kenya. Submitted

Conference Abstracts

Gachara, G., Symekher, S, Mbithi, J, Simwa, J, and Bulimo, W. (2010). Mutational analysis of the NS gene of pandemic H1N1/09 influenza virus in Kenya. Kenya Young Scientists Seminar series 4. Kisumu, Kenya. (Oral Presentation) October 2010.

Gachara, G., Symekher, S, Mbithi, J, Simwa, J, and Bulimo, W. (2011). The molecular evolution of two internal genes, M & NS in pandemic H1N1 influenza A virus. 1st Medical and Veterinary Virus Research in Kenya (MVVR-K) symposium. Nairobi, Kenya. (Poster Presentation) September 2011.

Gachara, G., Symekher, S, Mbithi, J, Simwa, J, and Bulimo, W. (2012). Sequence analysis of the Neuraminidase (NA) gene of pandemic H1N1 influenza A virus in Kenya. XIV International Symposium on Respiratory Viral Infections. Istanbul, Turkey. (Poster Presentation). March 2012.

Gachara, G., Symekher, S., Otieno, M., Magana, J., Opot, B., Bulimo, W. (2012). Evolutionary analyses of influenza A/H1N1pdm virus during the early, peak and late phases of the pandemic in Kenya. 2nd Medical and Veterinary Virus Research in Kenya (MVVR-K) symposium. Nairobi, Kenya. (Oral Presentation). October 2012.